Statistics, Knowledge and Policy 2007

MEASURING AND FOSTERING THE PROGRESS OF SOCIETIES

OECD

T.R. PRIME MINISTRY SPO STATE PLANNING ORGANIZATION

TURKSTAT

Statistics, Knowledge and Policy 2007

MEASURING AND FOSTERING
THE PROGRESS OF SOCIETIES

OECD

T.R. PRIME MINISTRY
SPO
STATE PLANNING ORGANIZATION

TURKSTAT

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The OECD is a unique forum where the governments of 30 democracies work together to address the economic, social and environmental challenges of globalisation. The OECD is also at the forefront of efforts to understand and to help governments respond to new developments and concerns, such as corporate governance, the information economy and the challenges of an ageing population. The Organisation provides a setting where governments can compare policy experiences, seek answers to common problems, identify good practice and work to co-ordinate domestic and international policies.

The OECD member countries are: Australia, Austria, Belgium, Canada, the Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The Commission of the European Communities takes part in the work of the OECD.

OECD Publishing disseminates widely the results of the Organisation's statistics gathering and research on economic, social and environmental issues, as well as the conventions, guidelines and standards agreed by its members.

This work is published on the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the Organisation or of the governments of its member countries or those of Turkey's State Planning Office and the Turkish Statistical Office.

Corrigenda to OECD publications may be found on line at: *www.oecd.org/publishing/corrigenda*.

© OECD 2008

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to *rights@oecd.org*. Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at *info@copyright.com* or the Centre français d'exploitation du droit de copie (CFC) *contact@cfcopies.com*.

Foreword

The Organisation for Economic Co-operation and Development is called upon to improve the functioning of the world economy and to act as a hub for the discussions on global issues. To deliver our mandate, we decided to open our membership to five more countries, and to strengthen our links with major emerging economies. But we also decided to look at important issues such as the measurement of progress in our countries from a different standpoint. We should be able to answer key questions our societies are asking such as "How does social policy affect the way economies work?", or "How does globalisation change the functioning of our societies?" "How do we capture the essence of well-being in economic data?"

We depend on evidence to make decisions. And in today's world, where huge amounts of data are circulating, using information to make decisions has become more complex. But this greater flow of information also represents an opportunity to improve the functioning of our democracies. Increased access to relevant information for citizens will improve the quality of their decision-making and improve the accountability of their leaders. Better knowledge among citizens and leaders will translate into enhanced policies and a better future for the next generation. Knowledge about progress, about whether or not life is getting better in a society, is one of the most important ingredients in this process.

There is now a growing movement seeking to develop better measures of progress for societies, with hundreds of initiatives underway around the world. The OECD is not alien to this effort. In 2005 we launched an initiative to re-think the way we measure progress. To take this work forward, we organised, in 2007, the second OECD World Forum on "Statistics, Knowledge and Policy". With more than 1 200 participants from over 130 countries, 10 plenary sessions and 39 parallel sessions, 210 world class speakers from 50 countries and several international organisations, the Forum represented a key opportunity to debate crucial global policy issues and their measurement, including health, climate change, poverty and democracy.

The success of the World Forum, and the sheer scale of work underway in every continent, was a clear signal for the OECD to continue its leadership in this area. Participants called for a broad initiative to discuss and develop best practice around both the measurement of progress, and ways in which those measures can better improve democratic dialogue. The *Istanbul Declaration*, signed by international and national organisations during the Forum, reinforced the Global Project on "Measuring the Progress of Societies", hosted by the OECD and joined by an increasing number of international organisations, research institutes, universities and non-governmental organisations. This project will advance our understanding in this important area, and lead to conclusions that improve the quality of life around the world.

I am pleased to have the OECD among the many other institutions that want to get the measurement of progress right. Because measuring what matters to people is a basic component of good governance, a strong underpinning of democracy and of the design of better public policies.

Angel Gurría
Secretary-General of the OECD

Acknowledgements

These proceedings were produced by the OECD Statistics Directorate.

These proceedings were directed by Mr. Jon Hall, Project Manager for Global Project on Measuring the Progress of Societies, and co-ordinated by Ms. Yoon Jung Lee and Mrs. Suzanna Grant-Kejairi.

Special thanks are given to several individuals and the OECD for their hard work and advice, in particular, Mrs. Laura Belli, Ms. Theresa Cooke, Mrs. Katherine Kraig-Ernandes and Mrs. Eileen Capponi.

And of course we thank the authors for agreeing to let us reproduce their work.

Table of Contents

Foreword **3**

Introduction **13**

Istanbul Declaration 15

Part One
Measuring Progress: Does it Make a Difference for Policy Making and Democracy? 17

Chapter 1
Measuring Progress: Does it make a Difference for Policy Making and Democracy?
Angel Gurría 19

Chapter 2
How Key National Indicators Can Improve Policy Making and Strengthen Democracy
David Walker 23

Chapter 3
Promoting Progress: Making it Happen
Roy J. Romanow 27

Part Two
Measuring Progress and Political Processes to Foster Progress **37**

Chapter 4
Beyond Measuring: The Council of Europe's Instruments Contributing to the Progress of Societies
Andreas Siegel 39

Chapter 5
Global Public Goods for Local Decision Making: Empowerment through Evidence
Julio Frenk and Octavio Gómez-Dantes 51

Chapter 6
Evidence-based Policy Making: Just a Myth or a Must?
Cynthia A. Glassman 59

Chapter 7
Empowering and Challenging Voters through Governance Indicators: The Philippine Experience
Romulo Virola 65

Part Three
Accountability and Civic Engagement 79

Chapter 8
The Importance of Social Reality for Europe's Economy: An Application to Civil Participation
Marcel Canoy, Frédéric Lerais, Massimiliano Mascherini, Andrea Saltelli and Daniele Vidoni 81

Chapter 9
Power to the People: Indicators for Accountability
Shailaja Chandra..... 121

Chapter 10
Connecting Communities with Community Indicators
Alex Michalos 133

Part Four
Measuring Progress: People's Perceptions and Knowledge..... 147

Chapter 11
People's Perception vs. Reality: Drawing Lessons from the Use of Statistics in the European Union
Joaquín Almunia..... 149

Chapter 12
What U.S. Consumers Know About Economic Conditions
Richard T. Curtin..... 153

Chapter 13
Special Eurobarometer Europeans Knoweledge on Economical Indicators
Antonis Papacostas 177

Chapter 14
What Do Citizens Know about Statistics?: The Results of an OECD/ ISAE survey on Italian Consumers
Flora Fullone, Michela Gamba, Enrico Giovannini and Marco Malgarini 197

Part Five
Subjective Measures of Well-being..... 213

Chapter 15
A Short Introduction to Subjective Well-being: Its Measurement, Correlates and Policy Uses
André van Hoorn 215

Chapter 16
Measures of Gross National Happiness
Ruut Veenhoven 231

Part Six
Challenges for National Statistical Offices 255

Chapter 17
Innovative Statistics to Improve Our Notion of Reality
Henk K. van Tuinen 257

Chapter 18
Challenges of Data Collection: with Special Regard to Developing Countries
Asad Elahi 297

Part Seven
Measuring Progress: Economic, Social and Environmental Issues 305

Chapter 19
An Effective Global Statistical System: A Prerequisite for Consistency in Global Measurements
Dennis Trewin 307

Chapter 20
Policy and Statistical Issues Underpinning Financial Stability: The IMF Perspective
Robert W. Edwards 317

Chapter 21
International Integration and Societal Progress: A Critical Review of Globalisation Indicators
Philippe de Lombaerde and P. Lelio Iapadre 327

Chapter 22
Understanding Retirement Savings and Pensions
Len Cook 333

Chapter 23
Key Role of Cultural and Creative Industries in the Economy
Hendrik van der Pol 343

Chapter 24
Measuring and Improving Government Performance
Joe Grice 355

Chapter 25
Overcoming Indigenous Disadvantage in Australia
Gary Banks 361

Chapter 26
Trends in Worker Requirements and the Need for Better Information to Make More Informed Decisions in a Global Economy
Randall W. Eberts 387

Chapter 27
Child Development, Human Development and the Progress of Societies
Fiona Stanley and Stephen R. Zubrick 401

Chapter 28
Human Capital and the Older Worker: The Need for Solid Indicators
Peter Ester and Marcel Kerkhofs 411

Chapter 29
Societal Well-being after Experiencing Trauma at the Hand of "Others": The Intertwining of Political, Economic and Other Visible Factors with Hidden Psychological Processes Affecting Victimised Populations
Vamik Volkan 425

Chapter 30
The Living Planet Index and Ecological Footprint: Tracking the State of Global Biodiversity and Human Pressures on the Biosphere
Jonathan Loh 435

Chapter 31
The State of Ecosystems and Progress of Society
Odeh Rashed Al Jayyousi 441

Chapter 32
How Reliable are Climate Change Projections: A Statistician's Perspective
Dennis Trewin 453

Part Eight
Gender: A Key Dimension of Societal Progress 463

Chapter 33
Social Institutions and Gender Equality: Introducing the OECD Gender, Institutions and Development Data Base (GID-DB)
Denis Drescheler, Johannes Jütting and Louka T. Katseli 465

Part Nine
Measuring Societal Progress for Local Communities 479

Chapter 34
The Newfoundland and Labrador System of Community Accounts
Alton Holett, Douglas May and Cory Giles 481

Chapter 35
Developing Indicators for Local Communities: The New Zealand Experience
Abby Thornley 503

Chapter 36
Achieving Outcomes From Indicators: Community Assessment Projects in the United States, A Focus on Santa Cruz County
Susan Brutschy and Deanna Zachary 517

Part Ten
Using Indicators to Make Governments Accountable 529

Chapter 37
Measuring and Improving the Performance of Governments
Ömer Dinçer 531

Chapter 38
AfriMAP: On Not Using Indicators to Score Progress in Governance
Bronwen Manby 539

Chapter 39
National Progress and the Effectiveness of International Aid
Maarten Brouwer and Rob Swinkels 551

Chapter 40
Case Study: Introducing Standard Indicators in European Commission Funded Projects for Aggregation of Aid Effects
Androulla Kaminara 557

Introduction

Enrico Giovannini
Chief Statistician, OECD

Is life getting better? Are our societies making progress? Indeed, what does "progress" mean to the world's citizens? There can be few questions of greater importance in today's rapidly changing world. And yet how many of us have the evidence to answer these questions?

The OECD's Second OECD World Forum on "Statistics, Knowledge and Policy" held in Istanbul in June 2007, was organised by the OECD in partnership with Turkey's State Planning Office and the Turkish Statistical Office, and in co-operation with several international organisations. It followed on from the first World Forum held in Italy in 2004. Presidents and ministers rubbed shoulders with the leaders of civil society and senior statisticians. Captains of industry met the heads of charitable foundations and leading academics. They shared a common interest in wanting to develop better measures of how the world is progressing.

Crucial global policy issues were addressed and debated. The importance that statistical knowledge can play in increasing the transparency and accountability of public policies, and building democracies was also discussed, together with the role that different components of our society (government, NGOs, businesses, media, etc.) should play in promoting an evidence-based culture.

The World Forum led to the *Istanbul Declaration*, signed by the European Commission, the Organisation of the Islamic Countries, the OECD, the United Nations, the United Nations Development Programme, UNICEF, UNESCO, the United Nations Fund for Partnership, the World Bank, and several other organisations, which calls for action to identify what "progress" means in the 21st century and to stimulate international debate, based on solid statistical data and indicators, on both global issues of societal progress and how societies compare. In particular, the Declaration calls for actions to:

- Encourage communities to consider for themselves what "progress" means in the 21st century;
- Share best practices on the measurement of societal progress and increase the awareness of the need to do so using sound and reliable methodologies;
- Stimulate international debate, based on solid statistical data and indicators, on both global issues of societal progress and comparisons of such progress;
- Produce a broader, shared, public understanding of changing conditions, while highlighting areas of significant change or inadequate knowledge;

- Advocate appropriate investment in building statistical capacity, especially in developing countries, to improve the availability of data and indicators needed to guide development programs and report on progress toward international goals, such as the Millennium Development Goals.

The signing organisations commit themselves

> "to measuring and fostering the progress of societies in all their dimensions and to supporting initiatives at the country level."

urge

> "statistical offices, public and private organisations, and academic experts to work alongside representatives of their communities to produce high-quality, facts-based information that can be used by all of society to form a shared view of societal well-being and its evolution over time."

and invite

> "both public and private organisations to contribute to this ambitious effort to foster the world's progress and welcome initiatives at the local, regional, national and international levels."

The Forum was a milestone for the OECD's Global Project on "Measuring the Progress of Societies", whose characteristics were presented during the last session. The project's mission is "to foster the development of sets of key economic, social and environmental indicators and their use to inform and promote evidence-based decision-making, within and across the public and private sector and civil society". The project, hosted by the OECD, includes a growing list of partners.

This volume comprises some of the best material from the proceedings of the conference. The OECD is grateful to all speakers, chairmen and women, discussants and delegates for their contribution to the event. A special acknowledgement goes to the Turkish State Planning Organisation and Turkstat for co-hosting the conference, as well as to those international agencies who co-operated in its organisation: the European Commission, the Organisation of the Islamic Conference, the United Nations and the World Bank. A number of other organisations were associated with the event including the Central Bank of the Republic of Turkey, DevInfo, the International Statistical Institute, PARIS21 (Partnership for Statistics in the 21Century), the Statistical, Economic and Social Research and Training Centre for Islamic Countries. And of course the OECD thanks those corporations who sponsored the conference: IBM; Turk Telekom; ENEL; and the UniCredit Group.

These proceedings contain only some of the selected highlights of the conference. A wealth of other material, including webcasts of conference sessions and all papers and presentations, is available on line at www.oecd.org/progress.

ISTANBUL DECLARATION

We, the representatives of the European Commission, the Organisation for Economic Cooperation and Development, the Organisation of the Islamic Conference, the United Nations, the United Nations Development Programme and the World Bank,

recognise that while our societies have become more complex, they are more closely linked than ever. Yet they retain differences in history, culture, and in economic and social development.

We are encouraged that initiatives to measure societal progress through statistical indicators have been launched in several countries and on all continents. Although these initiatives are based on different methodologies, cultural and intellectual paradigms, and degrees of involvement of key stakeholders, they reveal an emerging consensus on the need to undertake the measurement of societal progress in every country, going beyond conventional economic measures such as GDP per capita. Indeed, the United Nation's system of indicators to measure progress towards the Millennium Development Goals (MDGs) is a step in that direction.

A culture of evidence-based decision making has to be promoted at all levels, to increase the welfare of societies. And in the "information age," welfare depends in part on transparent and accountable public policy making. The availability of statistical indicators of economic, social, and environmental outcomes and their dissemination to citizens can contribute to promoting good governance and the improvement of democratic processes. It can strengthen citizens' capacity to influence the goals of the societies they live in through debate and consensus building, and increase the accountability of public policies.

We affirm our commitment to measuring and fostering the progress of societies in all their dimensions and to supporting initiatives at the country level. We urge statistical offices, public and private organisations, and academic experts to work alongside representatives of their communities to produce high-quality, facts-based information that can be used by all of society to form a shared view of societal well-being and its evolution over time.

Official statistics are a key "public good" that foster the progress of societies. The development of indicators of societal progress offers an opportunity to reinforce the role of national statistical authorities as key providers of relevant, reliable, timely and comparable data and the indicators required for national and international reporting. We encourage governments to invest resources to develop reliable data and indicators according to the "Fundamental Principles of Official Statistics" adopted by the United Nations in 1994.

To take this work forward we need to:

- Encourage communities to consider for themselves what "progress" means in the 21st century;
- Share best practices on the measurement of societal progress and increase the awareness of the need to do so using sound and reliable methodologies;
- Stimulate international debate, based on solid statistical data and indicators, on both global issues of societal progress and comparisons of such progress;
- Produce a broader, shared, public understanding of changing conditions, while highlighting areas of significant change or inadequate knowledge;
- Advocate appropriate investment in building statistical capacity, especially in developing countries, to improve the availability of data and indicators needed to guide development programs and report on progress toward international goals, such as the Millennium Development Goals.

Much work remains to be done, and the commitment of all partners is essential if we are to meet the demand that is emerging from our societies. We recognise that efforts will be commensurate with the capacity of countries at different levels of development. We invite both public and private organisations to contribute to this ambitious effort to foster the world's progress and we welcome initiatives at the local, regional, national and international levels.

We would like to thank the Government of Turkey for hosting this second OECD World Forum on "Statistics, Knowledge and Policy." We also wish to thank all those from around the world who have contributed to, or attended, this World Forum, or followed the discussions over the Internet.

Istanbul, 30 June 2007

Organisations who signed the Istanbul Declaration

as of 30 June 2008

First organisations to sign the Istanbul Declaration:

- European Commission
- Organisation for Economic-cooperation and Development (OECD)
- Organisation of the Islamic Conference
- United Nations
- UN Development Programme
- World Bank

Other organisations who signed the Istanbul Declaration:

- Economic and Statistical Observatory for Sub-Saharan Africa (AFRISTAT)
- Atkinson Charitable Foundation, Toronto, Canada
- European Free Trade Association (EFTA)
- Gallup Organization Europe
- International Statistical Institute
- Lisbon Council
- McCaughey Centre
- Institute for Statistics and Economic Studies (STATEC)
- United Nations Educational, Scientific and Cultural Organization (UNESCO)
- United Nations Children's Fund (UNICEF)
- United Nations Office for Partnerships
- United Nations University - Comparative Regional Integration Studies (UNU-CRIS)

The complete list on organisations and individuals who signed the Istanbul Declaration can be found on our webpage, www.oecd.org/progress.

Part One
Measuring Progress:
Does it Make a Difference for Policy Making and Democracy?

Chapter 1
Measuring Progress: Does it make a Difference for Policy Making and Democracy?

Angel Gurría
Secretary-General, OECD

The topic of this particular session, "Does measuring progress make a difference for policy making and democracy?", touches on two very sensitive areas for our countries: governmental accountability and social participation. Measuring progress with reliable information is a key ingredient of the democratic process. On the one hand, it makes governments more accountable and trustworthy, and on the other, it encourages people to participate more actively.

This Forum will give us the opportunity to develop new, more reliable indicators of progress, based on concepts from different continents and cultures. Indicators that are more meaningful both to policy makers and to citizens. Indicators to calibrate the outcomes of public policies, and to evaluate the products of our democracies. But also indicators that allow us, as individuals, to make better decisions for our own lives and well-being.

Progress is a complex concept, because it means different things to different people, depending on their cultural background, history and personal beliefs; but also depending on the health of society, the environment and the economy. But if we agree that progress encompasses many elements, we therefore also have to agree that its measurement cannot be reduced to "growth in GDP per capita".

As Robert Kennedy once rightly put it: GDP "measures everything, in short, except that which makes life worthwhile".

It is time to call for a global effort to find better measurements of progress that will consider the multidimensional feature of societal well-being. In fact, in a rather bold and "risky" effort, the OECD has already produced some interesting statistical work on measuring happiness, which was discussed in April this year at an International Conference in Rome.

We are witnessing an explosion of initiatives to measure progress around the world. Statisticians, policy makers and civil society are discussing what progress really means and how it can be measured.

The magnitude and implications of this trend have not been fully recognised. And it isn't just being led by the public sector. Civil society is taking a lead through projects like the Genuine Progress Indicator, the Environmental Sustainability Index, the Canadian Index of Wellbeing or the State of the USA project, to name just a few.

Why are all these people and institutions spending so much time and resources on this? Why has this captured the interest of the world? Have societies seen this as a way to tackle a common challenge? These questions focus on the importance of measuring progress for democracy.

One of the main ingredients of a successful democracy is access to quality information. Reliable facts and figures help governments improve their policies by comparing them and measuring their impact. When societies can trust social and economic indicators, they can better assess a government's performance and put forward better proposals.

Globalisation has made our national realities more complex, as well as more sensitive to external actors and factors. The Information Age has made our daily lives more dynamic, more plural and more complicated. The mare magnum of available information makes it much more difficult to understand public affairs and develop a participative democratic culture. We must provide our societies with new, clear and reliable tools to form their opinions, to make their assessment of the effectiveness of their democracies in fostering social progress.

In many countries, there is distrust in public figures, in political parties and electoral campaigns. This scepticism affects the whole democratic process because it undermines accountability. Less than half of the economists polled (in preparation for this event) in Asia, Latin America and Africa, for example, thought the policy debates in their regions were based on sound statistical evidence.

If we want to improve the quality of public debate, the contribution of civil society to public policy, the transparency of governments and therefore the level of trust in democracy, we need to provide credible points of reference and reliable solid data. A set of progress indicators, supported by the joint expertise of international organisations, can provide this new reference.

In the 19th century, our societies established a new institution – the national central bank – to better manage the workings of our economies and help protect citizens from economic risk. In the 20th century, we built antitrust and audit institutions to improve the efficiency of markets and protect citizens. In the 21st century, it is time to build new schemes and institutions to empower our citizens to assess the quality of their governments and policies, but also to measure their own progress in a modern society.

What if we could build, in each and every country, an institution for progress? An institution where different parts of society (government, opposition, trade unions, business associations, NGOs, academia, media, statisticians and others) could discuss what progress means to them and the key indicators to measure it. An institution whose progress indicators are seen as having authority and legitimacy. Do we think this would significantly improve the quality of our political and social debates – the quality of our democracy?

I believe so. As I said before, better indicators of progress alone are not enough. They need to be trusted – to be seen as accurate and impartial. They need to be used and understood and become shared knowledge among citizens. It was Socrates who said "The only good is knowledge and the only evil is ignorance."

We have to ask ourselves what we can do, both as organisations and individuals. Let me begin with what I think the OECD can contribute.

Some of you might be asking why the OECD decided to get involved in this work. Our Organisation is well known worldwide for the quality of its statistics and has, since its inception, worked to provide the figures needed to explain and understand our social economic processes and improve our public policies. We have a wealth of expertise based on the experiences of our 30 members and nearly 70 other countries. We are also a well recognised source of publications and information delivery.

Based on this statistical supply and know-how, measuring whether and how life is getting better is one of the most important roles this Organisation can take on. This conference, which is part of a broader Global Project on Measuring the Progress of Societies, is a crucial step in that direction.

To succeed we need all of you to be involved. This conference has the power to improve policy making and breathe fresh life into democratic processes in each and every country.

We need to answer several questions:

How can we create a global catalyst for this work? Something that brings together the thousands of initiatives and the people working on them, share best practice, and discuss issues of common concern;

How can we support those countries that lack the resources and know-how to develop their own sets of progress indicators? How can we help strengthen the independence of statistical institutions in those countries where they are subject to political pressure? Can this work help to build a world progress monitoring system?

I hope that by Saturday lunchtime we will have some answers to these questions.

Measuring progress is instrumental for policy making and democracy. Countries that benefit from reliable statistics know where they stand, their capabilities and limitations; but most importantly, they know where they are going, and how to get there.

The OECD is ready to work with you to provide our societies with these tools of modernity. Only by measuring and comparing our realities will we be able to improve them. For never has a noble venture of human improvement come out of ignorance and obscurity.

Chapter 2
How Key National Indicators Can Improve Policy Making and Strengthen Democracy

Honourable David M. Walker
Comptroller General, United States

Abstract

The author advocates the development of key indicator systems, especially at the national level, because such systems have the potential to help prioritise resource allocation and improve government services, among other things. Key indicator systems pull together various measures, such as gross domestic product, unemployment levels, infant mortality rates and air quality indexes to tell a more complete story about how a city, region, state or nation is doing. The author describes the work of one SAI (supreme audit institution): the Government Accountability Office (GAO) of the United States in advancing the creation of a US key indicator system, and advocates that the International Organisation of Supreme Audit Institutions (INTOSAI)act as clearinghouse for SAIs developing such systems. Also discussed is the creation of the "State of the USA", a non-profit group whose mission is to provide Americans with quality information on key changes in societal, economic and environmental conditions

My topic today is one that's near and dear to me: key national indicators. I gave my first international speech on indicators at the OECD's World Indicators Forum in Palermo in 2004, and I've addressed indicators in many domestic speeches and congressional testimonies. This topic finally seems to be getting the attention it deserves. In fact, the International Organisation of Supreme Audit Institutions, commonly known as INTOSAI, has chosen key national indicators as one of two themes for its triennial congress in Mexico City later this year.

From the industrialised world to the developing world, all nations face a range of challenges. Some are long-standing and country specific. But increasingly, nations face common challenges that transcend national borders, economic sectors, and institutional divides. I'd include here vital issues like economic interdependence, environmental protection, and global pandemic preparedness. Most of these are long-term challenges, which can take years or even decades to address.

At the same time, nations face the reality of finite resources. The challenge before us is how we can stretch those resources and get the greatest value for the money we spend.

Fortunately, nations today have several tools at their disposal to help them achieve these goals. Examples of these tools include strategic planning, scenario modeling, and indicator systems. I've chosen to focus on key national indicators because of their powerful potential to help countries prioritise resource allocation and improve government services. With data from indicator systems, policymakers can better assess their current situation, make more informed decisions, and measure their progress over time and relative to other nations.

Gross domestic product, unemployment levels, infant mortality rates, and air quality indexes are all examples of commonly used indicators. As most of you are well aware, a key indicator system pulls together these various measures to tell a more complete story about how a city, region, state, or nation is doing. Used effectively, information from key national indicator systems can help clarify problems and reveal opportunities. Such data can also inform agenda setting, improve planning, and promote better decision making and oversight. It can enhance public understanding and citizen engagement as well.

Several nations around the world already have some form of an indicator system. Australia, Canada, Singapore, and the United Kingdom, as well as international organisations like the European Union, the U.N., and the OECD, all use indicators to measure economic, environmental, and social conditions over time. Clearly, indicator systems are helping to define what it means to be a leading democracy in the information age.

I know other SAIs around the world are pursuing the adoption of key national indicators in their countries. I am hopeful that these SAIs will share their knowledge, experiences, and lessons learned. INTOSAI comes to mind as a natural clearinghouse for this information.

Several U.S. cities and localities are also using indicator systems. For example, the city of Boston is involved in a public-private partnership that's measuring and evaluating various quality-of-life issues. An indicator system used by a county in Florida revealed a lack of basic information on its growing elderly population.

Despite these successes, the United States still lacks an indicator system at the national level. Every year, our federal government spends almost $3 trillion on a wide range of activities, provides hundreds of billions of dollars in tax preferences, and issues thousands of pages of regulations. Yet, what's astonishing is the federal government does all this without knowing which programmes and policies are making a real difference and which ones aren't. It's a little like an airplane pilot flying at night without an instrument panel. This must change.

The simple truth is it matters how a nation keeps score. Keeping score provides a clear sense of what a nation has achieved and what needs to be done. Indicators can reliably measure progress on a national level. With such fact-based information, public officials are more likely to ask well-framed questions and accurately analyse issues. They're also more likely to propose sound solutions and make wise decisions on appropriations, authorisation and oversight.

In countries that have used key national indicators, we've seen some improved government performance and better use of limited resources. In other words, we know that key national indicators can help a country better meet the needs of its citizens.

By educating policymakers and the public, key national indicator systems can also help to limit abuses of power. As the U.S. Supreme Court Justice Louis Brandeis once

said, "Sunshine is the best disinfectant." Indictors can shed much-needed light on the vast breadth of government operations today.

With more honest and transparent reporting, it's clearer how various government programmes and policies are working. Transparency has a remarkable ability to reduce waste, prevent corruption, and shift resources where they're truly needed. The data provided by indicator systems can help to ensure that no one is above the law and everyone is accountable for results.

Comprehensive, objective, and reliable information that's readily available to the public can also put pressure on politicians to make difficult but necessary policy choices. With greater public awareness, elected officials are more likely to consider the greater good and the longer term. With greater public awareness, elected officials are less likely to shirk their stewardship responsibilities to future generations.

Finally, the appropriate use of key national indicators can build public trust and confidence in government. Policy solutions backed by credible, objective information are more likely to gain public support. Indicators can help average individuals better understand complex issues and may encourage greater citizen engagement in the public policy process.

The increased transparency produced by indicators may even prompt voters to make better choices at the polls. In my view, an informed electorate is more likely to accept candidates who are prepared to make difficult choices. An informed and engaged electorate is also more likely to accept some degree of shared sacrifice today in order to help create a better tomorrow.

Governments in countries that lack key national indicators are more likely to remain inefficient and ineffective. Too many public officials in Washington and elsewhere suffer from myopia and tunnel vision. There's a tendency to focus on the problem of the moment. And all too often, the quick fix leads to over spending and over regulation.

For more than 80 years, my agency, the U.S. Government Accountability Office (GAO), has been promoting transparency and speaking truth to power. Some GAO reports look at whether taxpayer dollars are being spent appropriately. Other GAO reports examine whether government programmes are meeting their objectives and the needs of society. Increasingly, GAO has been seeking to alert policymakers to emerging trends and future problems, such as the rising cost of health care and funding shortfalls in the nation's pension system.

As professional services organisations with extensive experience in statistics, supreme audit institutions like GAO can play a key role in promoting the adoption of indicator systems. SAIs can also make a range of contributions to indicator system efforts, such as suggesting ways to ensure the reliability and reasonableness of the data being produced.

As many of you know, GAO has been leveraging its knowledge and credibility to promote an indicator system at the national level in the United States. Many of the key players in this initiative are either former GAO executives like Chris Hoenig and Jane Ross or members of one of GAO's advisory boards, like Harvey Feinberg.

In fact, the key national indicator effort in America really got its start at a forum GAO co-hosted with the National Academy of Sciences several years ago. This led to the creation of a non-profit group, the State of the USA, Inc., whose mission is to provide the American people with quality information on key changes in societal, economic, and

environmental conditions. State of the USA has received grants from several major foundations, and they have a prototype website under way. The State of the USA initiative has come a good distance, but there is a lot of work to do and the best is yet to come.

As I said earlier, I've been speaking out on key national indicators for some time. And not too long ago, GAO issued a comprehensive study comparing the use of indicator systems across various jurisdictions. The report assessed lessons learned and provided options for Congress to consider in establishing a key national indicator system for the United States.

GAO has also been working closely with Congress, federal agencies, and prominent professional associations, particularly our National Academies, to advance this issue.

I'm a big believer in partnering for progress by building bridges across government and among various sectors. In my experience, government, private industry, and nonprofit groups can all benefit from working together on projects of mutual interest and concern. That's certainly true in the case with key national indicators, whose benefits will be felt throughout American society.

While GAO remains a strong supporter of key national indicators for the United States, there has to be a limit to our advocacy efforts in order to preserve our institutional independence. Going forward, others will have to take the lead on this issue. This doesn't mean GAO won't be involved. We will. In this regard, GAO and a consortium of organisations from several sectors will continue to encourage the Congress to pass related legislation in order to help make this important concept a reality.

In closing, an African proverb says tomorrow belongs to the people who prepare for it today. With key national indicator systems, nations everywhere now have a powerful tool to provide their people with a better future. In my view, it's an opportunity that no government can afford to miss.

Chapter 3
Promoting Progress: Making it Happen

The Honourable Roy J. Romanow, P.C., O.C, Q.C.
Founding Chair, Canadian Index of Wellbeing (CIW) Network Board, Canada

Abstract

Measurements of societal progress, based on complex statistics and indicators, should first and foremost be tied to a society's social values, this paper postulates. The author considers the current trends toward decentralisation, individualism and privatisation, and questions whether these "new ways" really provide equitable societies that resonate with the values and concerns of their citizens. Using Canada as an example, the paper then outlines the creation of the Canadian Index of Wellbeing (CIW) and the five major lessons learned from its continued development, lessons that can help bridge the gap between government policy and societal values. The paper also outlines the eight domains of data gathered for the CIW: living standards; time use; healthy populations; ecosystem health; educated populace; community vitality; civic engagement; and arts and culture.

The Second Istnabul Forum succeeded in bringing together a remarkably impressive array of participants with the common goal of advancing our understanding and commitment to measure and foster genuine progress in our societies.

The speakers of the past three days have shown a vision for our shared destiny that gives hope for all our nations and our world. As we know, the genesis of why we are here today, is due to the OECD's groundbreaking **Measuring the Progress of Societies** Project. It has shined the world's spotlight on the central role that key indicators can play in underpinning democratic debate and promoting effective government policy making.

This is a vision that we all share. Everyone in this room understands that solid evidence, taking into account all aspects of societal progress, must play an increasing role in guiding us down a genuine path towards a better world.

The good news, as we have heard throughout this conference, is that there are already many initiatives that advance our overall societal progress.

Today I want to draw on my own personal experience with one such pioneering project that is trying to "connect the dots" between social aspirations, public policy, and hard evidence. This particular Canadian-based project is by no means alone in its aims.

But it is one example, perhaps, from which we may draw additional important lessons about the way forward.

Building on Values

Now this having been said, I would like to begin with the matter of values.

I realise that in a room full of scientists, researchers and academics who place a great emphasis on objectivity, this may sound like a perilous beginning. It may prove to be so.

But let me explain.

Looking at the world today, I sense that in many cases public policy is seemingly losing its ability to connect and resonate with the core values that are shared by its citizens.

Our nations face new challenges everyday that require us to clarify and question who we are, what we aspire to be and what we must do to build a better future.

Challenges such as climate change, health and well-being, diversity and inclusion, the social, cultural and economic impacts of globalisation.

What all these challenges have in common is that their resolution ultimately requires tough trade-offs and a reconciliation of the sometimes deeply held but competing values and aspirations of citizens and other actors in society. More to the point, they require values-based choices.

As indisputable as it is that solid evidence is a major component of good decision making, perhaps we have not taken careful enough stock, even at this Conference, of how that evidence must be linked to the values of our respective societies and how together these propel us forward.

Indeed, as someone whose own political career was galvanised serving on the front-lines of the battle for universal health care in my home province of Saskatchewan, I can attest first-hand why values matter in public policy. Why collective values, such as a belief in equity and social justice, must sometimes take precedence over individual values- such as limiting the ability of doctors to work outside the publicly-funded health care system.

Over the past three days, of this exhilarating Conference, we have debated and learned much about how to measure and foster the progress of societies.

And if the goal for us is to increase the impact that our research is going to have on public discourse – if we seek to make complex statistics, indicators, mathematical models and formulas meaningful – then we too must tie our efforts to social values.

I would venture to say that the cornerstone value that we share is the belief that our society is often best shaped through collective action. That there is a limit to how much we can achieve as individuals acting alone. That the sum of a good society and what it can achieve is greater than the remarkably diverse parts which constitute it.

I invite you for a moment to think about my own home, Canada. Two things that you might immediately think of are its sheer size on the world map along with its very long shared boundary with our American neighbours. And, secondly, perhaps, Canada's celebrated tradition of inclusion, diversity and peaceful co-existence.

Both are true. But building our nation was and remains a challenge. In a land so vast, in a terrain so rugged, in a climate so harsh, the very notion of Canada as one nation would not have lasted very long had the earliest settlers not embraced the idea of compromise and collective action.

I came of age in Canada's vast prairie lands where enormous grain farms grow a wide variety of cereal crops as well as hard wheat. But life on the prairies, as wonderful as it is, is not always easy. The harsh, often snow-blown conditions, droughts, distance from neighbours and isolation, the small populations forced us together, like poplar trees huddled on a windswept prairie plain.

And so it is with other regions in Canada, where the geography and demographics may vary, but where we all learned to see survival and progress as a test of our ongoing ability to organise collectively and to remain united around shared values.

In fact, I have argued that it is this notion of a "shared destiny" that is the essential narrative that has bound Canadians through years of nation-building across regional, economic, cultural and ideological divergences.

And I am also sure that this notion of "shared destiny" is not distinct to Canada.

If we reflect on the history of other nations, we would safely conclude that, in their unique ways, "shared destiny" – or some other such term – can be found in many other national narratives as well.

In recent times, however, we have all witnessed a noticeable drift away from this rich and successful legacy. Perhaps, new international trading arrangements have ushered in an era of Globalisation of commerce that has weakened our capacity and our appetite and resolve for progressive national enterprises (such as nation-building). Perhaps the recent era of fiscal challenges has sowed the seeds of doubt about our capacity to meet our nations' future challenges. Perhaps the contemporary challenges of diversity and inclusion have tested our commitment to a shared citizenship.

But one thing is clear: in Canada, as in many other societies, there is a palpable momentum toward decentralisation, individualism, and privatisation, all presenting themselves as the so-called "new ways" to deal with our most pressing common challenges.

But do these so-called "new ways" really provide answers that resonate with the values and concerns of citizens?

Consider that when the ordinary citizen looks around--whether in Toronto, Paris, Istanbul, Tokyo or Sydney – he or she is likely to the see the human face of poverty, disparity and inequality.

They might ask: how is it possible that in the midst of such economic progress, there are still so many children living in poverty? That the income gap continues to grow between the rich and the poor?

Why has the water we drink and the air we breathe been taken for granted for so long?

Why does it seem that so many poorer nations are still facing intolerable living conditions despite the abundance of wealth in others?

Why is it that the indigenous people of so many nations often face living conditions that are so much poorer than those of other citizens?

And faced with these fundamental questions that call out for decisive action from our leaders, why is there such a disconnect between the harsh reality and the vision to which we aspire as a society?

And so here we are, in Istanbul, asking fundamental questions about how we can bridge these gaps. How societies can better measure their progress in ways that tap into the values of their citizens and motivate decision-makers to action. How we can better use evidence to benefit democracy, build a stronger civil society, empower individual citizens and stimulate a global debate on what progress really means?

Well, in my view, we start by asking the right questions- those I raised earlier in my remarks: who are we, to what do we aspire and what must we do to build a better future?

And as we answer these questions, we can then begin to put practical policy in place that resonates with the values of citizens. From there, we must ask: how do we know if we are on the right track?

The Canadian Index of Wellbeing (CIW)

One Canadian project, which I have the privilege to Chair, may perhaps offer some insights how we can start to assess whether we are indeed on the right track.

It's called *The Canadian Index of Wellbeing*, or CIW, and it's based on the premise that what we count matters.

The CIW is an initiative that over the long haul has the potential to change the political discourse in our country, and to re-shape the direction of public policy – for the better.

In this room, we are all quite aware of the power of indicators. We know that what we count, measure, and report often drives our understanding of whether we are better off than we used to be, whether we are creating a better world for ourselves and future generations, and what we need to change.

But too often, we gauge our society's well-being according to a narrow set of strict economic indicators.

Everyday, we are bombarded with information about whether the stock markets have gone up or down. Every quarter we hear about shifts in our gross domestic product – the GDP. Even small changes in the GDP send an adrenaline rush down the veins of policy makers, and editorial writers who scurry to their keyboards.

The Institute for Economic Research and the International Chamber of Commerce recently surveyed 700 economic experts from over 80 countries and found that "while main economic indicators, such as GDP, inflation and public finance statistics are usually well established, the political debates on environmental aspects, research and education, as well as social conditions tend not to be well grounded on statistical evidence around the world."[1]

In the consequence, a balanced and objective point of view is understandably compromised when the single most influential lens that we use to measure our progress and well-being is confined to a narrow set of economic indicators – as important at these may be.

As we have learned over these past 3 days, our natural environment is becoming depleted, the gap between the rich and poor is growing, chronic disease rates are skyrocketing, the incidence of stress is driving us to distraction. Is it any wonder, then that so many of our citizens feel that the rosy economic picture they see in the news seems oddly disconnected from the reality of everyday life?

To this point, we haven't had a clear, coherent and compelling alternative to the dominance of economic measures. But what would happen if every time my fellow Canadians heard about their GDP, they also heard the results of another new and important index – an Index of Well-being?

A new Index that measures the variables that really contribute to, or subtract from, the health, well-being, and prosperity of Canadians.

An Index that actually links the economic reality and prosperity of our nation with the social, health and environmental conditions that defines our communities. An Index:

- That distinguishes between good things like good health and clean air, and bad things, like sickness and pollution;
- That recognises volunteer work and unpaid care-giving as social goods, and overwork and stress as social deficits;
- That attaches a value on educational achievement, early childhood development, economic and personal security, a clean environment, and social and health equity;
- That creates a better balance – as Dr. Chan noted in her remarks yesterday, between the imperative for investing more in prevention and health promotion in the upstream, rather than in treating often preventable illness downstream.

We believe that CIW can be such a tool. It is still under construction, but we are making significant progress.

We are currently assembling data that we hope will eventually be capable of regional and provincial disaggregation to best capture well-being trends in Canada.

More importantly, we have set out to monitor progress in eight areas, or domains. Each domain will be backed by rigorous Canadian and international peer review and many of the criteria to be included in the domains will be informed through consultations with ordinary citizens – to ensure we are measuring what really matters to them – on the ground and in their communities.

The domains are:

Living Standards, including data on secure and meaningful employment, adequate income, low-income rates, and related indicators;

Time Use, examining the balance between paid work, unpaid work, and free time;

Healthy Populations, such as self-rated health, life expectancy, overall satisfaction with health care services, and other factors;

Ecosystem Health, including good air and water quality, healthy forests and environmental sustainability;

Educated Populace, measuring literacy, numeracy, indicators of educational attainment, and quality of formal and informal learning;

Community Vitality, tracking community safety, cohesion, trust, sense of belonging, inclusion and identity;

Civic Engagement, including indicators of individual and collective actions to identify and address issues of public concern; and,

Arts and Culture, measuring activity in the broad area of culture, which covers all forms of human expression, and in the much more focused area of arts, including performing arts, visual arts, media arts and art facilities and institutes.

We believe that these domains are an appropriate starting point for building a robust CIW and represent a true expression of the kind of values that Canadians hold dear. Core values such as fairness, equity, diversity, inclusion, economic security, health, safety, democracy, and sustainability.

We know that the goal of creating a Composite Index is fraught with many hurdles, and there is heated debate about its attainability. But we are pursuing it because if we succeed, it can be the doorway to a new way of collecting, aggregating and communicating statistical knowledge about what matters to Canadians in a way that enables them to hold their decision-makers to account for why things are getting better or worse.

To say this is an ambitious project is to understate it. But learning from others as we have in this Conference, it's a worthy goal to pursue.

CIW: Lessons Learned

As I look back on all the work that we have done to get us to this stage, what key lessons might inform other like-minded projects?

I believe that there are five: A connection to values; Collaboration; Solid evidence; Public engagement; And, finally, communications.

I have already discussed the importance of being grounded, at all times, on the core values of citizens. So let me touch on the other four key themes briefly, starting with **collaboration**.

Collaboration

In this conference, we have a shining example of the kind of cross-sectoral partnerships that are needed to measure genuine progress. Similarly in Canada, collaboration has been to produce the right mix of people and partnerships to strive for our goals.

But, it has not been easy. Before the idea of the CIW came along, many Canadian indicator researchers and practitioners had been working in relative isolation. There were many excellent individual projects at both the national and regional levels, each doing their part to report on the dynamics of genuine progress. But for the most part, these initiatives remained detached from each other. The CIW has proven a catalyst and a conduit for much of this collaborative work. And just as you have done here at this Conference, we have brought together leading indicator practitioners with a cross-sectoral

partnership of leaders from government to business, social, health and community sectors.

Solid Evidence

The third point is that assembling measures based on solid evidence is essential to establishing credibility and progress. As this Conference has clearly demonstrated, the integrity of our work demands the greatest statistical and methodological rigour.

In Canada, fortunately along with our esteemed national treasure, Statistics Canada, we have some of the best thinkers and practitioners leading the technical development of the CIW around accurate indicators.

Public Engagement

Fourth, just as a rock-solid evidence-base is crucial to our credibility, so is public engagement. Tools that seek to inspire citizens cannot remain locked in the halls of academia or within the security of government agencies.

If we are going to be successful in ensuring our citizens, and in turn our governments, will pay attention to the latest research, then we must engage with the public at all stages of development.

So citizen engagement is an integral component of our work. Early on, when the experts were just beginning to conceptualise possible ways for moving forward, we decided to open up the discussion to ordinary Canadians. We asked Canadians questions about what mattered to them and their families. About their aspirations for their communities. About the types of information they felt would help them understand whether things were getting better or worse, and who to hold to account for this. In a world where we often seem to take the voices of common citizens for granted, what we heard was truly remarkable. We heard a deep desire by citizens to be engaged in a dialogue about what it means to have a good quality of life. We heard about the need for greater accountability and transparency – just as we have over these past 3 day – for the decisions and actions that are taken by policy makers.

And we heard about the imperative for new tools like the CIW that can provide policy shapers with a compass calibrated to Canadian values. We remain committed to public engagement because we have realised that this is key to the eventual acceptance and success of the CIW.

Communications

Which brings me to my final point, regarding the importance of **communications...**

Friends, a good statistical tool, even the most valid and reliable o ne in the world, will have no impact on decision making if it cannot speak to a variety of audiences.

So the CIW team has, from the beginning, applied a communications lens to everything that it does. Here too, there are challenges, as the spheres of academia and indicator development do not always mix easily with the lives of citizens and the world of everyday communications.

So we are working with communications experts and partners from across the country to till the soil for the growth of the CIW and to ensure its relevance to everyday issues. Only in this way can we ensure our impact on the policy making process. In time, we hope the CIW will be a major influence on daily discussions around the water cooler, in media coverage and on policy development and implementation.

So, the lessons? They are a combination of collaboration, solid evidence, public engagement and effective communications, all underpinned by values. The past 3 days have been spent collaborating, sharing and testing evidence, engaging with participants from a variety of sectors, and figuring out new ways to communicate so that our work can indeed become more relevant and so that we can all move closer to a vision of a better world.

If we all commit to carry on the work that we have accomplished here. If we all commit to engage our colleagues in our respective countries to keep talking to each other. If we all commit to hold each other accountable for advancing this cause. Then, we will all be closer to the goals of empowering citizens, stimulating a debate on genuine progress, and advancing democracy.

Conclusion: A Return to Values

Let me conclude with values and by quoting Judith Maxwell, former President of the Canadian Policy Research Networks, who earlier this year wrote in the British Medical Journal that:

"The legitimacy and sustainability of any major policy decision increasingly depends on how well it reflects the underlying values of the public."[2]

As one who has had the privilege of being elected to public office, of serving as a Premier of a Province of Canada, I have been truly blessed to have been involved in a good number of initiatives – large and small, and I have learned throughout that, as Judith Maxwell writes, being anchored on values is essential.

My friends, values provide the guideposts for how we can move forward, how we can orient ourselves during challenging times, how we can inspire our citizens and how we can be confident that the programmes we recommend and the path we choose will reflect the vision of our citizens.

My hope is that we can leave here today with a renewed commitment to produce many more improvements to the well-being index and thus continue to bridge the gap between policy and values.

With the right kind of leadership, success, for this is a daunting task, can be within our reach. But it requires, as I say, leadership; the kind of leadership assembled here:

- Leadership that reflects the core values of a nation and its people.
- Leadership that inspires diverse partners to work collaboratively.
- Leadership that is informed by the best available evidence.
- Leadership that embraces transparency and accountability.

- And, above all, leadership that dares to dream no little dream. That recognises the basic truth that economic progress and social justice go handin- hand. That you cannot have one without the other.

Good luck and all the best in your continuing contributions to building a more progressive and just world.

Note

1 Institute for Economic Research and International Chamber of Commerce. (February 2007). "Statistical information in public and political decision making," CESifo World Economic Survey. (WES 1/2007).

2 Judith Maxwell, Steven Rowell, and Pierre-Gerlier Forest. (2003) "Giving citizens a voice in healthcare policy in Cnanda," in British Medical Journal. May 10; 326(7397); 1031-1033.

Part Two
Measuring Progress and Political Processes to Foster Progress

Chapter 4
Beyond Measuring: The Council of Europe's Instruments Contributing to the Progress of Societies

Andreas Siegel
Director, Strategic Planning, Council of Europe

Abstract

This paper explains how the Council of Europe manages to observe and achieve progress via a unique mix of instruments, including Conventions, recommendations, monitoring mechanisms and other tools. First, the Council of Europe is described: its membership, mission and institutions. Then the benchmarking used to measure progress is outlined, including adherence to Conventions, such as the European Convention for the Protection of Human Rights and Fundamental Freedoms (ECHR) and other European Conventions. The instruments used to facilitate political action in member states are then listed, including examples of assistance programmes. Finally, some of the impacts achieved by Council of Europe Instruments, including the work of The European Court of Human Rights (ECHR), are presented

Introduction

Progress: a Chain of Instruments

This forum concentrates on the question of how we can improve the quality of life, individually and collectively. How can we recognise, measure and promote progress? Are there common denominators for progress in a society, in a region, in the world? In the past two centuries we have made tremendous progress in understanding that a society is not static but is constantly changing. And that we can influence change. With individual freedom and democratic rights the concept of choice has evolved. But on which basis can a consensus be reached between individual aspirations and the progress of societies?

I will attempt to demonstrate that in the Council of Europe, which brings together 47 European countries and 800 million people, we have found common ground. We have done so in spite of our cultural diversity and long history of conflicts all over the continent. Based on common values, we have established a unique system of institutions and a sophisticated arsenal of conventions and benchmarks, monitoring mechanisms and targeted assistance measures. The aim is to harmonise our methods to achieve and

monitor progress in the areas of Human Rights, Democracy and the Rule of Law, but also to reinforce social cohesion and the capacity to manage cultural diversity.

Democracy and Human Rights have always been closely linked with the notion of progress, and I believe we should add the concept of "good governance" because it describes the capacity of governments to transform their commitment to these common values and principles into action. Today, Democracy and Human Rights are still key factors of progress: they have a strong influence on our well-being and are a basic requirement to ensure a sustainable quality of life. Again, for the members of the Council of Europe member states, real progress can only be achieved if we also develop policy instruments for social cohesion, the protection of cultural and natural heritage and intercultural dialogue.

The main responsibility for achieving progress and for adjusting the tools on how to achieve it is in the hands of governments. And rightly so: they have to provide the conditions for sustainable development and should support the pursuit of happiness of the individual citizens. Civil society already contributes to this, but will have to participate more actively in measuring and promoting progress.

The underlying philosophy of the **Council of Europe** is based on the European Convention of Human Rights. Member states have agreed to create instruments and mechanisms which will ensure the respect of common standards, promote steady progress and assist those who have not yet arrived at the commonly agreed level of standards.

The Council of Europe has played an essential role in supporting the transition to and the consolidation of democracy from the mid-seventies, through the early nineties until today. Since its creation 58 years ago, the Council of Europe has established an important *acquis* – we have produced more than 200 conventions, benchmarks for human rights and democracy – a reference even beyond Europe.

Civil society and independent experts are constantly monitoring the quality of these common standards and make proposals for improvements.

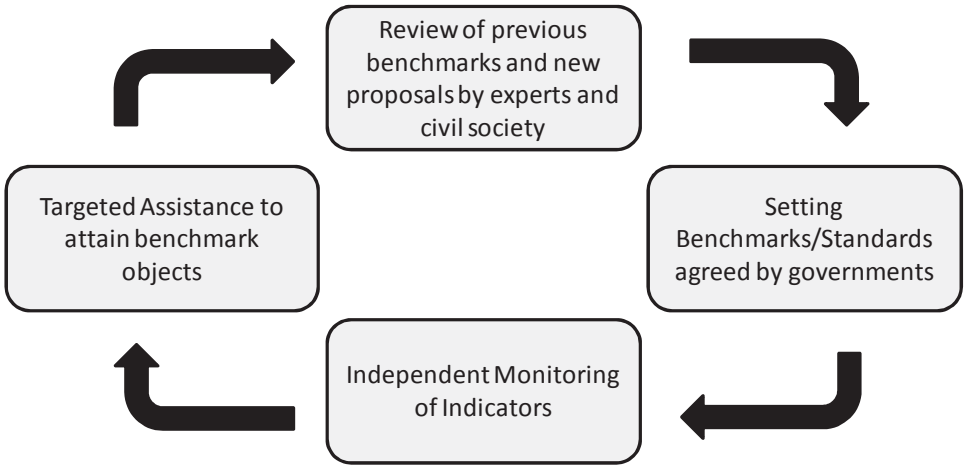
How do we manage to observe and achieve progress?

The **"chain of progress"** of the Council of Europe is the following:

- We develop common standards, indicators and benchmarks: by means of conventions, recommendations, resolutions and binding Court judgments as well as case law.
- We measure progress: by setting objectives, regular monitoring of indicators, results-based programming and evaluation.
- We foster progress:
  - by following–up member states' commitments and compliance with standards, court judgments and recommendations by monitoring bodies,
  - by adjusting and updating old standards and the creation of new ones,
  - by raising awareness about the evolving benchmarks, with target groups and the civil society,
  - by building, where necessary, national capacity to implement the agreed standards and

– by exercising collective political pressure whenever common standards are not respected, especially if basic human rights are violated.

Figure 4.1 The Council of Europe Chain of Progress



In a nutshell, with its **unique mix of instruments,** the Council of Europe is a key actor in promoting and monitoring progress, with targeted indicators, tools and procedures.

I would like to illustrate the impact of the work of the Council of Europe with a few examples.

But let me first introduce briefly the institutions of the Council of Europe.

The Council of Europe: Membership, Mission and Institutions

The Council of Europe, with its 47 member states and 5 observers[1] has a genuine pan-European dimension. It covers all major issues facing European society except for defense. Its work programme includes human rights, democracy (including local and regional democracy), legal cooperation, social cohesion, health, education, media, culture, heritage, sport, youth, and environment and regional planning. Its headquarters are located in Strasbourg, France.

The **institutional setup** of the organisation is the following:

The Committee of Ministers

The Committee of Ministers is the Council of Europe's decision making body and is composed of the Foreign Ministers of all member states or their Permanent Representatives.

The Parliamentary Assembly

The Parliamentary Assembly is the organisations' second statutory body. It deliberates on the preparatory work of the 10 committees *e.g.* in the areas of Political Affairs, Legal Affairs and Human Rights and the honouring of obligations and commitments by member states (Monitoring). The members of the Parliamentary

Assembly are appointed by national parliaments (318 members and the same number of substitutes).

The European Court of Human Rights

The European Court of Human Rights is directly accessible to the individual citizen and its jurisdiction is compulsory for all contracting parties. The Court's final judgments are legally binding on the states concerned. The execution of the Court's judgments is monitored by the Committee of Ministers.

The Congress

The Congress of Local and Regional Authorities of the Council of Europe is a consultative body representing local and regional authorities.

The Human Rights Commissioner

The post of Commissioner for Human Rights was created in 1999. The Commissioner is an independent institution responsible for promoting education, awareness and respect for human rights in member states and ensuring the full and effective compliance with the Council of Europe's normative instruments. He also provides advice, analyses and recommendations to member states to remedy any shortcomings and prevent human rights violations.

The Secretary General

The Secretary General directs and coordinates the organisation's activities. He is elected for a 5-year term. The current Secretary General, Terry Davis (United Kingdom), was elected by the Parliamentary Assembly in 2004.

Benchmarking to Measure Progress: A Vast Variety of Instruments to Set and Verify Standards (Conventions, Recommendations, Monitoring Mechanisms and Other "Tools")

The following instruments do not only set indicators and standards but mostly provide for regular scrutiny and monitoring on whether these standards have been implemented. Therefore, these instruments can be used to measure any progress in the specific fields they address. The list is not exhaustive and concentrates on the major areas of intervention:

The European Convention for the Protection of Human Rights and Fundamental Freedoms (ECHR)

This Convention was signed in 1950, and entered into force in 1953. It is the **"mother" of all Council of Europe conventions**. It sets forth a number of fundamental rights and freedoms, such as the right to life, prohibition of torture, prohibition of slavery and enforced labour, right to liberty and security, right to fair trial, no punishment without law, freedom of thought, conscience, religion, freedom of expression, freedom of assembly etc. More rights are granted by additional protocols.

Member states undertake to secure these rights and freedoms to everyone within their jurisdiction. The convention also establishes an international enforcement machinery to ensure its observance. The key role is being played by the *European Court of Human Rights* in Strasbourg. Individuals may refer alleged violations of human rights to the Court after they have exhausted legal remedies at national level. The parties to a case must abide by the judgment of the Court and take all necessary measures to comply with them. The ECHR is therefore a fundamental element of European democratic stability and of European cooperation and integration. For further information see: http://www.echr.cor.int/ECHR.

The European Social Charter

This Convention was signed in 1961 and entered into force in 1965. It is the counterpart of the European Convention on Human Rights in the sphere of economic and social rights. The Charter guarantees the enjoyment of fundamental social and economic rights by all appropriate means. Among those rights guaranteed by the Charter are the right to work, the right to organise, the right to bargain collectively, the right to social security, the right to social and medical assistance, the right to social, legal and economic protection of the family and the right to protection and assistance for migrant workers and their families.

The Charter has an international system of supervision of its application which is based on national reports and, increasingly, on collective complaints. Every year, the parties submit a report on some of the accepted provisions of the Charter, indicating how they implement it in law and in practice. If a party takes no action on a decision of non-conformity, the Committee of Ministers may address a recommendation and ask to change the situation in law and in practice. For further information see: http://www.coe.int/T/E/Human_Rights/Esc.

The Human Rights Commissioner

The Commissioner plays an essentially preventive role complementing the European Court of Human Rights and other Treaty based organs. He can submit reports, recommendations or opinions on specific matters to the Committee of Ministers and to the Parliamentary Assembly. For further information see: http://www.coe.int/t/commissioner.

The Parliamentary Assembly Monitoring

The work of the Parliamentary Assembly is prepared by specialist committees dealing, amongst others, with political affairs, legal affairs and human rights, social affairs, education and in particular the honouring of obligations and commitments by member states. In that context the Parliamentary Assembly has had a major role in preparing candidate countries for membership and it supports developing democracies and monitors closely the observance of their commitments when they are admitted to the Council of Europe. For further information see: http://assembly.coe.int.

The European Commission for Democracy through Law (Venice Commission)

Established in 1990, the Venice Commission is the Council of Europe's advisory body on constitutional matters and has played a leading role in the adoption of

constitutions in conformity with European standards. It contributes to the dissemination of the European Constitutional Heritage and provides constitutional first aid to individual states. The Venice Commission has also played an important role in crisis management and conflict prevention through constitution building and advice. The Venice Commission is a so-called enlarged partial agreement with originally 18 member states and has now 51 full members including Kyrgyzstan, Chile, the Republic of Korea and Morocco. For further information see: http://www.venice.coe.int.

The European Convention against Torture and Inhuman or Degrading Treatment or Punishment

This Convention was adopted in 1987 and entered into force in 1989. The Convention supplements the protection available under the European Convention on Human Rights by establishing a European Committee for the Prevention of Torture (CPT). CPT members are independent and impartial experts from a variety of backgrounds, for example lawyers, medical doctors and specialists in prison or police matters. The CPT visits places of detention to see how persons deprived of their liberty are treated. The committee has unlimited access to places of detention and may interview detainees in private. After each visit, the CPT draws up a report with recommendations when it considers that there is a need to improve the situation. The CPT's report and the replies from the states concerned are confidential, but it has become common practice for states to agree to make them public. For further information see: http://www.cpt.coe.int.

The Framework Convention for the Protection of National Minorities

This Convention was adopted in 1994 and entered into force in 1998. It is the first legally binding multilateral instrument to protect national minorities in general. Contracting parties undertake to pursue the Convention's objectives through national legislation and policies. These include equality before the law, preserving and developing cultures, safeguarding identities, religions, minority languages and traditions and guaranteeing access to the media. The implementation of the Convention is monitored by the Committee of Ministers and an advisory committee composed of 18 independent experts. The rules governing this monitoring procedure also allow non-governmental organisations and minority associations to submit alternative information or information reports. For further information see: http://www.coe.int/T/E/human_rights/minorities.

The European Commission against Racism and Intolerance (ECRI)

This mechanism was established at the 1993 Vienna Summit of the Council of Europe. ECRI is an independent monitoring mechanism, whose task is to combat racism, xenophobia, anti-semitism and intolerance in all member states. It covers all necessary measures to combat violence, discrimination and prejudice faced by persons, or groups of persons on the grounds of race, colour, language, religion, nationality and national or ethnic origin. ECRI works with 3 major methods: country reports, work on general themes and relations with civil society. ECRI then issues specific country related recommendations and thematic recommendations. For further information see: http://www.coe.int/T/E/human_rights/Ecri.

The Convention against Trafficking in Human Beings

This is one of the more recent Conventions which was signed in May 2005. The aim of this Convention is to prevent and combat trafficking in Human Beings in all its forms, national or international, whether or not it is linked with organised crime. The fundamental principle outlined in the Convention is that the protection and promotion of the rights of the victims shall be secured without discrimination on any grounds such as sex, race, colour, language, religion, political or other opinion etc. The focus of this Convention is on the protection of victims and on its independent monitoring mechanism guaranteeing the parties' compliance with its provisions. For further information see: http://www.coe.int/t/dg2/trafficking/campaign.

The Group of States against Corruption (GRECO)

GRECO (an enlarged partial agreement of currently 44 member states, with the OECD and the UN as observers) was established in 1999 in order to monitor Council of Europe member states' compliance with the organisation's anti-corruption standards. It helps identify deficiencies in national anti-corruption policies through mutual evaluation and peer pressure. This frequently leads to legislative, institutional and practical reforms. For further information see: http.//www.coe.int/t/dg1/Greco.

The Select Committee of Experts on the Evaluation of Anti-Money-Laundering Measures (MONEYVAL)

The aim of MONEYVAL is to ensure effective national systems to counter money-laundering and terrorist financing in compliance with relevant international standards. Its evaluation reports provide highly detailed recommendations on actions required to effectively combat money-laundering. MONEYVAL was established in 1997 and has currently 28 member states and a number of observers, for example the Secretariat of the Financial Action Task Force on Money Laundering (FATF) and the United Nations Counter-Terrorism Committee (CTC). Moneyval reports and recommendations are taken into account by the IMF and the World Bank for their overall financial assessment report. For further information see: http://www.coe.int/t/e/legal_affairs/legal_co-operation/combating_economic_crime.

Fighting Drug Abuse and Illicit Trafficking in Drugs: the POMPIDOU Group

This intergovernmental body was founded in 1971 and presently comprises 35 member countries. A key feature of its work is compiling and harmonising information on drug abuse and trafficking. The group contributes to the development of multidisciplinary, innovative and evidence-based drug policies in member states which are in line with basic European standards in the areas of health, ethics, social cohesion and penal policy. For further information see: http://www.coe.int/t/dg3/pompidou.

The Anti-doping Convention

The aim of the anti-doping Convention of 1989 is to combat doping in sport in an active and coordinated way. A monitoring group supervises its implementation, which aims at restricting trafficking in doping substances, stepping up dope testing and improving dope detection, supporting education and information programmes and ensuring that the penalties imposed on offenders are effective. Together with the

European Commission and the member states of the European Union, the Council of Europe helped set up the World Anti-doping Agency (WADA). The Council of Europe has 2 seats on the WADA Foundation Board and therefore helped to lay down a number of common rules regarding legislative, financial, technical and educational measures. For further information see: http://www.coe.int/t/dg4/sport/Doping.

The Compendium of Cultural Policies and Trends in Europe (COMPENDIUM)

The Compendium is a Europe-wide information and "clearing house" system on cultural policy measures, instruments, debates and trends providing also statistical data, indicators, cross-country comparisons and good practices. It is a unique reference tool pooling knowledge, available online (http://www.culturalpolicies.net) and regularly updated.

Beyond Measuring: Instruments to Facilitate Political Action in Member States

In order to complement the legal standards and monitoring mechanisms, the Council of Europe organises a vast variety of targeted assistance activities, which mainly consist of policy advice, capacity building and awareness raising in 5 key areas relevant to the progress of societies: human rights, rule of law, democracy and good governance, social cohesion and cultural and intercultural dialogue.

Currently 39 programmes, more than 100 projects and about 2 400 individual activities are carried out annually. The underlying principle is an objective-based approach: expected results are verified by means of performance indicators, predetermined sources of verification and assumptions about context factors. The results of these targeted assistance activities are regularly analysed in a progress review report in the first quarter of the following year, containing recommendations for further improvement.

A few examples of our assistance programmes:

- **European Programme of Human Rights Education for Legal Professionals** (HELP programme). The aim is to fully integrate the standards of the European Convention on Human Rights into national curricula of professional training for judges and prosecutors in schools of magistrates and similar institutions.

- The programme **"Making Democratic Institutions Work"** is a major programme in the field of democracy and good governance. This has, for example, led to the creation of a consultative forum for the future of democracy, to a project concerning good governance in the information society and a project supporting the right of the public to have access to official documents. All these projects contribute directly to improving the functioning of democratic institutions.

- Another important component in the field of democracy is the **strengthening in the role of civil society**. These initiatives are meant to encourage all groups of society to participate in democratic associative life in order to support the development and consolidation of a strong civil society attached to democratic values such as mutual tolerance, respect and understanding. In this context, the role of NGOs in a pluralist democracy is enhanced by promoting cooperation between NGOs and public authorities.

- A further important action of the Council of Europe is to promote **social cohesion** in Europe. This includes activities to avoid exclusion of vulnerable groups promoting the rights of the child and integration of migrants. Other special projects are dedicated to improving the quality of life of people with disabilities and the protection of Roma and Travellers against racism, intolerance and social exclusion.

- A number of **educational measures** also help to promote democratic culture in Europe, such as training of educational professionals, linking policy and practice in citizenship and human rights education and, of course, continuing to develop standards and materials for democratic citizenship and human rights education. Special attention is also given to **youth** and minority issues. New concepts are currently being developed in the area of **intercultural dialogue** and the **management of diversity** in pluralist societies. A white paper on intercultural dialogue will be finished in 2007. The links between human rights, democracy and **sport** are underlined by the activities around the Anti-doping Convention and the Convention against Spectator Violence.

Making Change Happen: the Impact Achieved by Council of Europe Iinstruments

The institutions, mechanisms and targeted assistance activities described above have contributed to a number of significant policy changes in the member states of the Council of Europe. Numerous legislative changes were adopted following judgments of the European Court of Human Rights. Changes in the case law of domestic courts are also quite frequent as a consequence of judgments of the European Court of Human Rights. General policy recommendations have also had a significant influence and have led national governments to adopt new policies in various specific sectors. The Council of Europe mechanisms have, for example, led to the following measurable progress:

The European Court of Human Rights (ECHR) Court judgments have led

- to legislative reform prohibiting aliens' expulsion to countries where they risk ill-treatment (Austria);

- to decriminalisation of conscientious objectors (Bulgaria);

- to legislative reform prohibiting the use of evidence obtained under compulsion (UK).

One of the greatest achievements of the Council of Europe is the *de facto* **abolition of the death penalty in peace time in all member states**, following Protocol No. 6 to the ECHR, which entered into force in 1998.

European Social Charter

The Charter's mechanisms have had the following effects:

- New legislation was adopted in the Czech Republic concerning the ban on employment of minors in hazardous occupations;

- New legislation was adopted in Greece, banning corporal punishment;

- Child pornography is introduced as an offense in the Penal Code (Belgium);
- In Cyprus, recent legislation provides for equal treatment of people with disabilities as compared to other employees;
- New legislation in the United Kingdom introduced the offense of trafficking of people, in particular children, for the purposes of sexual exploitation.

Framework Convention for the Protection of National Minorities

- New legislation was adopted in Cyprus permitting mixed marriages between different communities;
- New legislation was adopted in Russia allowing the use of minority languages on federal television;
- Armenia has set up a new department dealing with minority issues and has introduced legislation to guarantee the right to use minority languages when communicating with the administration.

European Convention for the Prevention of Torture (CPT)

- Finland set up a special centre for persons detained under alien's legislation;
- The French government has issued instructions on the dignity of persons in police custody (with express reference to CPT standards);
- In Lithuania, the legal standards for the provision of living space for prisoners were increased to 5 m2 per person in multi-occupancy cells.

European Commission against Racism and Intolerance (ECRI)

ECRI's country-monitoring reports have triggered post-publication debates in many countries and have had, amongst others, the following impact:

- Adoption of anti-discrimination legislation (Bulgaria);
- Reinforcement of criminal law provisions to combat racism (Croatia);
- Adoption of the Criminal Code, which now provides that racial and ethnic motivation of a crime are considered as aggravating circumstances (Denmark).

European Commission for Democracy through Law (Venice Commission)

The Venice Commission has strongly influenced constitutional improvements in member states. Examples:

- The **Constitution of Albania** was prepared throughout 1998 in close co-operation with the Venice Commission. The text of the Constitution **reflects to a large extent**

Venice Commission input. This was a decisive contribution to build stability with functioning institutions in Albania.

- The **Constitution of Armenia** was revised in close co-operation with the Venice Commission in order to remove obstacles to **complying with commitments to the Council of Europe** and bring the country closer to European standards. As a result, there are now **sufficient checks and balances** to the powers of the President.

- In December 2005 the Venice Commission adopted a **legal opinion** on European standards to be respected when carrying out a **referendum** on the independence of Montenegro. On the basis of this opinion, an agreement between the main political forces in Montenegro was achieved on the rules applicable to this referendum. As a **result**, **all political forces, participated in the referendum** and its result was recognised both at the domestic and international level. In this way the country achieved independence without destabilising the sensitive regional context.

Council of Europe Commissioner for Human Rights

- Slovakia: An Ombudsman Institution was created; the crime of defamation of the Republic and its representatives was eliminated from the Penal Code;

- Latvia: An Ethics Code on the Conduct of State Police Officers was introduced;

- Portugal: the maximum time limits for pre-trial detention in the Code of Criminal Procedure were reduced.

Parliamentary Assembly Monitoring

The Parliamentary Assembly monitoring procedure has been instrumental in accompanying 20 member states on their path to democratic reform. Some examples:

- Turkey has implemented several changes to the Constitution, seven reform packages and numerous other laws and decrees related to Council of Europe accession requirements. Since 2004, the monitoring procedure has therefore given way to a post-monitoring dialogue.

- In Armenia, monitoring has contributed to the abolition of the death penalty and a comprehensive constitutional reform.

- In Azerbaijan, the Electoral Code and the law on Freedom of Assembly was amended.

- In Ukraine, the monitoring procedure has closely accompanied and promoted democratic reforms since 1999.

Conclusions: the Council of Europe's Contribution to the Progress of Societies

The list of examples of specific impacts achieved through Council of Europe's instruments and mechanisms is far from exhaustive. You will find additional

documentation on your particular fields of interest at our information stand at the exhibit, and of course on our web site: www.coe.int.

We are well aware at the Council of Europe that we are not the only organisation dealing with these key topics of progress, namely democracy and human rights. That is why we have been striving to establish cooperation arrangements with other international partners, including international NGOs. It is also a factor of progress that international organisations respect the tax payers and avoid, as far as possible, any unnecessary duplication. Therefore, we should redouble our efforts to work together and use the respective added value of each organisation in order to achieve common goals. Very recently, the Council of Europe has concluded a memorandum of understanding with the European Union: The Council of Europe and the European Union have been actively involved in promoting human rights and democracy in Europe and have organised a number of joint programmes.

This forum is a good opportunity to compare sectoral and regional models and try to build on these in order to promote progress worldwide. We should pursue the fascinating debate of this forum in order to further develop our common views and indicators to measure progress in these areas. Progress by definition is a dynamic process. We need to be innovative and have to adjust to constantly evolving needs. However, we must, first and foremost, monitor the respect of our common standards. In a globalised world we must attempt to adopt a common approach on how to define the quality of life, how we protect basic human rights and promote a democratic culture with informed and active citizens.

The notion of progress is intrinsically linked to the mission of the Council of Europe: It has a unique institutional and instrumental mix, bringing together all the key actors – governments, NGOs, independent monitoring mechanisms and the enforcement mechanism of the Court. Under one roof we can set standards, give assistance to help with their implementation, monitor progress, and where necessary, ensure that commitments undertaken by member states are respected in practice.

The Council of Europe is ready to contribute to further promoting progress in the areas where it can provide a measurable added value: human rights, democracy and the rule of law, supported by the fundamental role of social cohesion and the management of diversity.

Notes

1 Canada, Holy See, Japan, Mexico and the United States of America

Chapter 5
Global Public Goods for Local Decision Making: Empowerment through Evidence

Julio Frenk, MD, PhD
Senior Fellow, Global Health Program, Bill and Melinda Gates Foundation, United States
Chairman of the Board, Institute for Health Metrics and Evaluation, United States
President, Carso Health Institute, Mexico
Former Minister of Health, Mexico (2000-2006)

Octavio Gómez-Dantés, MD, MPH
Director of Analysis and Evaluation, Carso Health Institute, Mexico

Abstract

The focus of this paper is how scientifically derived evidence can – and must – guide the design, implementation and evaluation of public policy programmes, at the national, bilateral and multinational levels. The case study presented is Mexico's 2001 healthcare reform. The author first examines the main challenges faced by health officials worldwide. Then the impact of the two main sources of knowledge and evidence that lead to Mexico's healthcare reform is discussed: (1) the WHO framework for the assessment of health-systems performance, launched in 2000 as part of the World Health Report; and (2) an analysis of the National Income and Expenditure Surveys for Mexico. The author then presents five major lessons learned from the Mexican experience. Finally, the creation of Seattle-based Institute for Health Metrics and Evaluation is used to illustrate the ongoing effort to create knowledge-related global public goods.

The OECD Forum focuses on the interplay between knowledge and policy. In order to contribute to the understanding of this crucial interaction, in this paper we present a real case of national policy making that made extensive use of knowledge-related global public goods. Its main message is that evidence-based policy is no longer a buzz word but a real possibility. In our contradictory and often contentious world, we require, more than ever before, the power of knowledge to illuminate the arduous process of purposeful social change.

First, we briefly examine the main challenges in global health, which provide the context for the evidence-based reform experience that took place in Mexico and that we describe in this paper. We then discuss the background, contents, and initial results of this experience.

More than ever before, today we are all keenly aware that in health matters the world has become a single neghbourhood. This awareness comes at a time of unprecedented change. We are in the midst of a **tense** and **intense** health transition unlike anything the world has seen before, which is linked to broader demographic, social, and economic transformations. The most dramatic expression of the health transition is the fundamental shift in the patterns of disease, disability, and death. Most countries in the world are facing a **triple** burden of ill health: first, the unfinished agenda of infections, malnutrition, and reproductive health problems; second, the emerging challenges represented by non-communicable diseases and injury; third, the health risks associated with Globalisation, including the threat of pandemics like AIDS and influenza, the trade in harmful products like tobacco and other drugs, the health consequences of climate change, and the dissemination of harmful lifestyles leading to the silent epidemic of obesity (which someone has called "globesity," precisely to underscore its link with Globalisation).

Our biggest challenge today is that most health systems in the world simply have not kept up with the pressures derived from this complex transition. As a result, we are facing a number of unacceptable paradoxes. We will mention four. First, never before has the power of science been greater, yet millions die unnecessarily from diseases whose prevention and treatment were solved decades ago. Second, countless countries simultaneously have rural communities without doctors and urban doctors without jobs. Third, while unprecedented sums of financial capital in the form of aid are flowing from North to South, intellectual capital moves in the opposite direction through the migration of health personnel, thus rendering much of that aid ineffective. Fourth, whereas health is a key factor in the fight against poverty, health care itself becomes a cause of impoverishment when hundreds of millions of uninsured families have to pay out of pocket for services and drugs.

These are indeed perplexing paradoxes. And they can only be solved with decisive action. But action must be guided. Fortunately, during the past few years we have witnessed the emergence of a new consensus in global health.

The first part of the consensus is based on the increasing evidence that health is not only the result of economic growth, but also one of its major determinants. The WHO, the OECD, the World Bank, and other multilateral institutions have made lasting contributions to this line of reasoning. As a result of this first consensus, today we have a greater opportunity than ever before to overcome the barriers to global health equity. While resources remain short of need, there has been a major increase in funding in recent years, and renewed efforts to mobilise additional funds hold the promise of even greater resource capacity. New actors, such as philanthropic foundations, global NGOs, and multilateral initiatives, have brought increased human and financial capacity and public attention to the field.

The second part of the emerging consensus is that scientific knowledge represents the driving force for health progress. We all agree that research is a value in itself, an essential part of human culture. At the same time, knowledge has an instrumental value as a means to improve health. This is achieved through three mechanisms.

First, knowledge gets translated into new and better technologies, such as drugs, vaccines, and diagnostic methods. This is the best-known mechanism through which it improves health. But, second, knowledge is also internalised by individuals, who use it to structure their everyday behaviour in key domains like personal hygiene, feeding habits,

sexuality, and child-rearing practices. In this way, knowledge can empower people to modify their lifestyles in order to promote their own health and also to become informed users of services and citizens conscious of their rights. Finally, knowledge becomes translated into evidence that provides a scientific foundation for decision making both in the delivery of health services and in the formulation of public policies.

Recent developments in Mexico illustrate this last point. Thanks to the cooperation among several academic and international organisations, notably the WHO, the OECD, and the World Bank, the analytical armamentarium for health policy has been greatly enriched during the past few years to include such robust tools as the measurement of burden of disease, cost-effectiveness analysis, national health accounts, and standardised surveys. The rigorous application of these knowledge-related global public goods, coupled with excellent country-specific data and long-term investment in capacity building, helped to catalyse a structural reform of the Mexican health system.

We will not go into the details of this reform, which has been the subject of a series of seven articles in *The Lancet* (2006, p.368). In addition, the OECD published in 2005 an excellent study of the health sector in Mexico, which included an initial appraisal of the reform (OECD 2005). For the purpose of this presentation, it will suffice to mention that the Mexican reform is probably a textbook case of evidence-based policy, since it was designed and implemented making use of the best available knowledge. Because of this feature, the Mexican experience can hold interesting lessons for other countries. Indeed, policy makers all over the world require sound evidence in order to understand and act upon the complexities of our times.

In the Mexican case, solid analysis made decision makers and the public aware of critical realities that required solution. Thus, the careful calculation of national health accounts revealed that more than half of total expenditure in Mexico was out-of-pocket. This proved to be a direct result of the fact that approximately half of the population lacked health insurance.

These findings were unexpected as it was generally believed that the Mexican health system was based on public funding. Instead, the analysis revealed one of the unacceptable paradoxes that was mentioned before: we know that health is one of the most effective ways of fighting poverty, but medical care can itself become an impoverishing factor for families when a country does not have the social mechanisms to assure fair financing that protects the entire population.

The realisation that households had been paying catastrophic out-of-pocket sums generated a different perspective on the operation of the health system. Policy makers extended their focus to include financial issues that proved to have a great impact on the provision of health care and on levels of poverty among Mexican households.

Another global public good that helped to make the local case for reform was the WHO framework for the assessment of health systems performance. This framework, launched in 2000 as part of the *World Health Report*, highlighted fairness of financing as one of the intrinsic goals of health systems (WHO 2000).

As a direct result of its high levels of out-of-pocket spending, Mexico performed very poorly on the international comparative analysis of fair financing. Instead of generating a defensive reaction, this poor result spurred detailed country-level analysis in 2001 that showed that impoverishing health expenditures were concentrated among poor and uninsured households. The analysis was undertaken jointly by the Ministry of Health of Mexico, WHO, and the Mexican Health Foundation, an example of how national

governments, international organisations, and non-governmental institutions can join forces.

The country-level analysis was based on data from the National Income and Expenditure Surveys for Mexico, yet another global public good. These surveys are produced by many countries in the world and provide homogenous data sets that are key for cross-national comparisons, yet are seldom used for health policy formulation and evaluation.

The careful interplay between national and international analyses generated the advocacy tools to promote a major legislative reform establishing a system of social protection in health, which was approved by a large majority of the Mexican Congress in 2003.

This system is reorganising and increasing public funding by a full percentage point of GDP over seven years in order to provide universal health insurance, including the 50 million Mexicans, most of them poor, who had been excluded until now from formal social insurance schemes because they are self-employed, are out of the labor market or work in the informal sector of the economy.

The increased funding is spearheading a major effort to realign incentives throughout the health system. Poor families can now enroll in a new public insurance scheme called *Seguro Popular*. Enrolment is the basis for allocating federal funds to states, which are responsible for the delivery of services. In this way, the old model of "bureaucratic budgeting," which subsidised providers without regard to performance, is being replaced by "democratic budgeting," whereby money follows people in order to assure the best balance between quality and efficiency.

To achieve this aim, the macro-level financial reform is being complemented by a micro-level management reform, which is strengthening delivery capacity through a series of specific interventions, such as long-term planning of new facilities, technology assessment, efficient schemes for drug supply and rational prescription practices, human resource development including managerial training, outcome-oriented information systems, facility accreditation, provider certification, quality improvement in the technical and the interpersonal dimensions of care, and performance benchmarking among states and organisations. These are all critical components of the stewardship role that ministries of health must fulfill with increasing proficiency, especially to deal with the problems of quality in both the public and the private sectors.

The element that articulates the financial and the managerial reforms is an explicit package of benefits, which was designed using cost, effectiveness, and social acceptability as the guiding criteria. Apart from serving as a planning and priority-setting tool, the package is a means of empowering people by making them aware of their entitlements.

A hallmark of the Mexican experience has been a substantial investment in research to design the reform, monitor progress towards its implementation, and evaluate its results. This is a clear example of the possibility of harmonising two core values of research in health: **scientific excellence** and **relevance to decision making**.

The value of research for enlightened decision making is underscored by the worldwide search for better ways of strengthening health systems. Because of the gaps in our current knowledge, every reform initiative should be seen as an experiment, the effects of which must be documented for the benefit of every other initiative, both present

and future. This requires a solid investment in research on health systems. Each innovation constitutes a learning opportunity. Not to take advantage of it condemns us to rediscover at great cost what is already known or to repeat past mistakes. To **reform** it is necessary to **inform**, or else one is likely to **deform**.

Let us conclude by drawing the global lessons from the Mexican reform experience as was done in *The Lancet* series. We will summarise those lessons as the ABCDE of successful reform (Frenk 2006).

A stands for **agenda**. The first ingredient for success is to link health to the broader agenda of development and security. Public health experts must learn to address the larger concerns of heads of government, legislators, ministers of finance, and other policy makers who have to balance the claims of many different sectors. In this advocacy effort we can make use of global evidence showing that, in addition to its intrinsic value, a well-performing health system contributes to the overall welfare of society by relieving poverty, improving productivity, increasing educational abilities, developing human capital, protecting savings and assets, and directly stimulating economic growth with a fairer distribution of wealth. In other words, a well-performing health system is a crucial determinant of the progress of societies. Because it is concerned with economic development, the OECD is particularly well positioned to advance the centrality of health in this broader agenda. A decisive step was taken when the OECD organised in May 2004 its first-ever meeting of ministers of health, which one of us was honored to chair and which included an exemplary encounter with ministers of finance.

Recalling that historic encounter leads us naturally to the **B**, which stands for **budget**. By placing health at the centre of the development agenda of a country it is possible to endow it with the degree of priority that it deserves. These arguments enhance the negotiating power of ministers of health, who can then convince decision makers to allocate **more money for health**. But we should also assure, in the words of the legendary Professor Ramalingaswami of India, that we develop the capacity to deliver **more health for the money** (Ramalingaswami n.d.).

And this takes us to the **C**, which stands precisely for **capacity**. There is no substitute for long-term investment in capacity building in two main areas. The first refers to health-service delivery, through investments in physical infrastructure and, most importantly, in human resources. The second has to do with the development of institutions that can undertake the necessary research and analysis to generate sound evidence for policy. In the case of Mexico, the current reform has reaped the benefits of 20 years of sustained efforts to establish and nurture organisations such as the National Institute of Public Health and the Mexican Health Foundation. These centres of excellence have produced relevant research and policy analysis, trained researchers who occupy key policy making positions, carried out independent and credible evaluations, and greatly enriched the quality of information.

With this capacity, we can then move to the **D**, which stands for **deliverables**. A key ingredient to gain public support for a reform is to identify and communicate its specific benefits. The best way to do so is to focus on priority diseases and risk factors. In this way, the public can link abstract financial and managerial notions to concrete deliverables. This is also the way to bridge the divide between two public-health traditions: on the one hand, the "vertical" approach, focusing on specific disease priorities, and, on the other, the "horizontal" approach, aimed at strengthening the overall structure and functions of the health system. In order to go beyond this false dilemma, it is necessary to extend the geometry metaphor and develop what Jaime Sepúlveda calls

the "diagonal" approach (Sepúlveda 2006), namely, a strategy in which explicit intervention priorities are used to drive the required improvements into the health system. A fundamental lesson from the Mexican experience is that health-system capacity can be built up through the scale-up of effective preventive and therapeutic interventions against specific priority problems grouped in an explicit package of benefits.

Finally, **E** stands for **evidence**. The Mexican experience confirms what several researchers (Deaton 2004; Jamison 2006) have pointed out: that the health of people in rich and poor countries alike is depending more and more on their ability to locally adopt knowledge that has been generated as a global public good.

In this respect, the Mexican case shows that the dilemma between local and global research is a false one. The process of Globalisation can turn knowledge into an international public good that can then be brought to the centre of the domestic policy agenda in order to address a local problem. Such application, in turn, feeds back into the global pool of experience, thus generating a process of shared learning among countries. This virtuous process can only take place if we mobilise international collective action for the common good of all countries (Jamison *et al.* 1998).

It is this enlightened conception of international cooperation to generate knowledge-related global public goods that has inspired the very recent establishment of the Institute for Health Metrics and Evaluation. Affiliated to the University of Washington in Seattle and supported initially by the Bill and Melinda Gates Foundation, this institute will conduct independent evaluation of health initiatives worldwide, develop new methods and tools, disseminate its products through multiple publication outlets, and offer free access to an electronic data base on global health. The Institute for Health Metrics and Evaluation will complement existing efforts by developing innovative indicators and carrying out rigorous analyses (Lancet 2007). Indeed, the Institute may contribute to the larger endeavor of measuring the progress of societies, as stated in the goals of this conference.

The new institute represents a constructive piece of the institutional architecture for global health, because it aspires to provide an independent assessment of performance. There are three main reasons that make such an initiative needed at this particular juncture: first, to sustain interest in global health by demonstrating results from the increased investments in this field; second, to enhance efficiency by building a solid knowledge base of what really works and may be transferred across countries when it is culturally, politically, and financially reasonable; third, to promote the values of transparency and accountability as essential ingredients of democratic governance both at the national and global levels.

Those of us involved in the development of this exciting new institution are determined to make it an excellent example of what global health requires today: global partnerships for the creation of global public goods that will foster global understanding to help us address common global problems by empowering local actors with the formidable tools of evidence.

Fortunately, the topic of the OECD event that generated the papers presented in this report – the value of knowledge to inform policy – involves at its essence the possibility of sharing. It is a topic in which we can all participate and from which we can all benefit. It is a topic where the **self interest** of each country coincides with the **common interest** of all nations.

One of the thinkers who best captured the sharing character of knowledge was Thomas Jefferson, who almost two centuries ago stated:

"He who receives an idea from me, receives instruction himself without lessening mine; as he who lights his taper at mine, receives light without darkening me."

We are certain that through our common commitment, many candles will be lit. The path is clear: scientifically derived evidence must be the guiding light for designing, implementing, and evaluating programmes in national governments, bilateral aid agencies, and multilateral organisations. This is the path that will lead to more equitable development through better policy making for health and that will make an enduring contribution to the progress of societies.

References

Deaton A. Health in an age of globalisation. Draft prepared for the Brookings Trade Forum, Brookings Institution, Washington, DC, May 13-14, 2004.

Frenk J. Bridging the divide: global lesson from evidence-based health policy in Mexico. Lancet 2006;368: 954-61.

Frenk J, González-Pier E, Gómez-Dantés O, Lezana MA, Knaul FM. Comprehensive reform to improve health system performance in Mexico. Lancet 2006;368:1524-34.

Gakidou E, Lozano R, González-Pier E *et al.* Assessing the effect of the 2001-06 Mexican health reform: an interim report card. Lancet 2006;368:1920-35.

González-Pier E, Gutiérrez-Delgado C, Stevens G *et al.* Priority setting for health interventions in Mexico´s System of Social Protection in Health. Lancet 2006;368:1608-18.

Knaul FM, Arreola-Ornelas H, Méndez-Carniado O *et al.* Evidence is good for your health system: policy reform to remedy catastrophic and impoverishing health spending in Mexico. Lancet 2006;368:1828-41.

Lozano R, Soliz P, Gakidou E *et al.* Benchmaking performance of Mexican states with effective coverage. Lancet 2006;368:1729-41.

OECD. OECD Reviews of Health Systems. Mexico. Paris: Organisation for Economic Co-operation and Development, 2005.

World Health Organization. World Health Report 2000. Health systems: Improving performance. Geneva: WHO, 2000.

Ramalingaswami V. Personal communication.

Sepúlveda J. Foreword. In: Jamison DT, Breman JG, Measham AR *et al.*, editors. Disease control priorities in developing countries [2nd edition]. New York: Oxford University Press for The World Bank, 2006: xiii-xv.

Sepúlveda J, Bustreo F, Tapia R *et al.* Improvement of child survival in Mexico: the diagonal approach. Lancet 2006;368:2017-27.

Jamison DT. Investing in health. In: Jamison DT, Breman JG, Measham AR *et al.*, editors. Disease control priorities in developing countries [2nd edition]. New York: Oxford University Press for The World Bank, 2006: 3-34.

Jamison DT, Frenk J, Knaul F. International collective action in health: Objectives, functions and rationale. Lancet 1998;351: 514-7.

Editorial. A new institute for global health evaluations. Lancet 2007;369: 1902.

Chapter 6
Evidence-based Policy Making: Just a Myth or a Must?

Cynthia A. Glassman
Under Secretary for Economic Affairs, Department of Commerce, United States

Abstract

Using experiences gleaned from her years as a Commissioner at the United States Securities and Exchange Commission and from her present position overseeing two of the major U.S. statistical agencies, economist Cynthia Glassman explores the importance of seeking and using evidence in policy making. She illustrates the real and serious consequences that can occur when policy makers do not have a solid, evidence-based foundation on which to base their decisions, and she outlines some current efforts to expand the information frontier in the United States.

When I first learned the topic of this session, I thought to myself how obvious. The answer to the question "Is evidence-based policy making a myth or a must?" is clear; that is, it is a must. This view no doubt reflects my background and training. As we say in the United States, where you stand usually depends on where you sit. At present, I sit at the head of two of the key federal statistical agencies in the United States – places whose sole purpose is to gather and compile data about the U.S. economy and the American people.

I am also an economist by training. These days, the profession of economics is largely – if not entirely – about evidence-based analysis. We have come a long way from the era when economists, like Thomas Malthus, theorised about the growth of population and the growth of the food supply. Today's economists go out and get the data, then test the hypothesis.

I have also learned through years of working in policy and economics that there can be real and serious consequences when policy makers do not have a solid, evidence-based foundation when they make policy. I would like to illustrate this with three stories, one drawn from the history of the Bureau of Economic Analysis, one of the agencies that I oversee, and the two others from experiences from my prior role as a Commissioner at the U.S. Securities and Exchange Commission.

For those of you who might not know, the Bureau of Economic Analysis (BEA) is the agency in the United States that produces the National Income and Product Accounts. The most well known number from these accounts is the gross domestic product or GDP. The accounts produced by BEA are among the most important tools that we have for understanding the mechanics of the U.S. economy. However, it is easy to forget how

recently this body of evidence came into existence and what it was like before we got monthly and quarterly statistical updates on the state of the economy.

BEA recently published a fascinating account of its history and the birth of the National Income and Product Accounts in the United States (Marcuss & Kane 2007). The desperate need to understand the workings of the U.S. economy during the Great Depression was the motivating force in producing these statistics. Before then, well-meaning policy makers did make economic policy, but there was no systematic means of knowing whether the policies made matters better or worse for the economy as a whole.

As Rosemary Marcuss and Richard Kane aptly described it in their article, during the early 1930s, "[n]either the public nor elected officials understood the workings of the economy that seemed to be perpetuating the crisis, nor did they know quantitatively its scale and scope." In hindsight, we now know that "Asset values had plummeted, the banking system was breaking down, deflation was reversing the gears of the economy, and sales were insufficient to keep businesses going. Farm income, on which one-fourth of the population depended had fallen by a half." The income statistics that the Department of Commerce began to produce in 1934 were a huge step forward in providing evidence on which to base policy. Imagine how much better policy could have been if decision makers had had data to help them disentangle the various economic forces at work in a timely manner.

This story illustrates the consequences of realising in the middle of a crisis that you need more information than you have at hand. My second story illustrates some of the same principles on a more micro scale. It also dramatically demonstrates how all of us – personally and professionally – make evidence-based decisions, and just how problematic it can be when we find out too late that we did not have enough of the right information when we made choices.

Prior to my appointment as Under Secretary for Economic Affairs, I was one of the five commissioners at the U.S. Securities and Exchange Commission (SEC). While I was at the SEC, there was a series of corporate reporting scandals that highlighted how important accurate financial reporting is to our capital markets. Investors, creditors, regulators, and other market participants rely on getting accurate, timely and comparable financial information from public companies. The efficient allocation of capital depends on financial reports that provide a realistic picture of firms' past performance and future prospects.

When information provides a misleading picture, the results can be devastating, as we saw in those corporate reporting scandals. The misleading information reported by those companies resulted both directly and indirectly in a serious misallocation of resources.

- Investors in those companies paid a huge opportunity cost by investing in companies with unrealistic, inflated values;
- Competitors made business decisions on a distorted playing field;
- Creditors did not price credit appropriately for the real risk taken; and
- Employees tragically made career and retirement investment decisions based on a false picture of their employer's financial prospects.

The impact did not end there. As I learned in my new position, the inflated profits contributed to a level of Gross Domestic Income calculated by BEA that was higher than Gross Domestic Product. The difference in these two measures raised questions by policy makers about why the GDI numbers were growing faster than the GDP numbers. Not knowing the answer led to challenges in determining monetary and fiscal policy at the time. In retrospect, now we know that the discrepancy was due, in part, to the same misleading financial information.

My final story illustrates the consequences of choosing not to look for evidence. Does this happen? I suspect it happens more often than I would hope. In fact, I have some evidence. Two economists, Alan Blinder and Alan Krueger, conducted a study entitled "What Does The Public Know About Economic Policy, And How Does It Know It?" They found that "On a variety of major policy issues (*e.g.*, taxes, social security, health insurance), ideology is the most important determinant of public opinion, while measures of self-interest are the least important. Knowledge about the economy ranks somewhere in between" (Blinder & Krueger 2004). This suggests that many people are simply not out there seeking the facts before making important decisions.

While at the SEC, I made it my mission to try to instill more economic analysis into the Commission's rulemaking and enforcement initiatives and to show the value of empirical data in framing rules designed to protect investors and maintain the integrity of the markets in the most efficient, effective way. Routinely I would ask: What are the objectives of the rule? Will the rule meet the objectives? Does it go far enough – or does it go too far? Does it meet the spirit as well as the letter of the law? Does it make sense? Are there likely to be unintended consequences? Are the costs commensurate with the benefits? Does it create unrealistic expectations? What are the regulatory alternatives? Within the Commission ranks, I became famous – or perhaps infamous – for consistently asking these questions. But they were, and they are, important. Too often, regulators – who, particularly at the SEC, are more likely to be lawyers than economists – tend to promulgate a rule and see how it works rather than try to ascertain its effects in advance. To me, that is putting the cart before the horse.

I asked all these questions when we were first presented with a mutual fund governance proposal requiring funds to have at least 75% independent directors and an independent chair. I particularly wanted to know how the performance and costs of funds with inside chairmen compared to that of funds with independent chairman. But, at least at that time, no internal studies were forthcoming, and the rule was promulgated over my dissent. It was, however, struck down by a unanimous court. Why? Because the Commission had failed to give adequate consideration to the costs of, and alternatives to, the measure. Recently, the agency sought public comment on two internal studies. As a result, the agency now has the benefit of public input on the studies. I do not know what the outcome will be, but I am encouraged that any decision will be informed by the evidence.

What are the lessons that I draw from these three stories? First, there are plenty of examples of situations where policy is made without the benefit of evidence, or enough evidence. There are always reasons why this happens. Sometimes you do not know what information you need until you are in the middle of an event. Other times you rely on evidence only to find out that it was a lot less reliable than you thought. Sometimes you know exactly what information you need, but it is not available and you must make do with the incomplete information that is available.

The primary lesson that I take away from these stories is that we must always strive to find the best evidence we can in order to make informed decisions. All of these stories reinforce my opening presumption. Evidence based policy is a must because making policy with no evidence or bad evidence has real and serious consequences. These stories show how serious the consequences can be if decision makers – from people making economic policy to ordinary people planning their retirement – do not have complete information. This is also why I find it troubling when smart people – people who should know better – choose not to go out and look for the evidence when it is in their power to do so.

This search for more and better evidence about how the U.S. economy works is something I deal with every day. The core mission of the bureaus that I oversee – the Bureau of the Census and the Bureau of Economic Analysis – is to produce solid, impartial statistics. And, it is data like that produced by these two agencies that forms the foundation on which analysis that informs policy can be built.

I would like to tell you about three areas where we at the U.S. Department of Commerce are currently working to push against the information frontier and improve our understanding of the economy. Two of these are at the Bureau of Economic Analysis and the third is a cross-cutting effort at the Department level.

Last September, BEA, with support from the National Science Foundation, published a preliminary satellite account on investment in research and development. Satellite accounts are supplementary estimates that do not change the official national accounts. However, by remaining consistent with the broader economic accounts, while adding new information and formats better suited to answering particular analytical questions, these accounts provide a "laboratory" for economic accounting research into specific industries and markets.

This particular satellite account was developed to explore the effect of investment in research and development on U.S. economic growth. What we have learned from this preliminary account is fascinating. It appears that R&D may have accounted for a substantial share of the resurgence in U.S. growth in recent years (BEA 2006).

A vital next step in this effort to explore the role of intangibles in the U.S. economy is currently underway at the Bureau. They are working on a prototype health satellite account, which we hope will provide insight into the role of health care spending in the U.S. economy.

This is a really big deal. The increasing cost of health care is the number-one long-term budget issue confronting the U.S. government. Health care spending currently represents 16% of U.S. GDP. It has doubled as a share of GDP over the last ten years and it is expected to continue to increase. We know that some of the data that we are now using to measure health care is inadequate. Preliminary research indicates that our measures of health care inflation are likely overstated. If this is the case, we could be looking at an overstatement of overall inflation and an understatement of overall economic growth and productivity because the sector represents such a large share of our economy. Getting the best evidence we can in this area is critical not only for health care policy, but also for monetary and fiscal policy.

Another policy-sensitive area where we are pushing against the frontiers is in the area of innovation measurement. This is an area that I know is also of high importance at the OECD and within the European Union.

We have begun by looking at what evidence is already available. We have examined metrics across our entire statistical system. We have talked to colleagues at the European Union and the OECD to see what they have discovered as they struggle with similar questions. But, we are also taking a next step and asking what data do we not collect that we should collect if we want to really understand the impact of innovation in the U.S. economy.

In February the Secretary of Commerce convened an Advisory Committee of CEOs and academics to "advise the Secretary on new or improved metrics on innovation in the economy." I must point out that it is very rare for a member of the President's Cabinet – the most senior group of policy makers in the United States – to take an interest in data collection at the technical level that this exercise requires. But, Secretary Gutierrez personally recruited CEOs from America's top-tier companies for this effort. As decision makers, they believe – and this is a quotation from the first meeting of the Advisory Committee – that "If you cannot measure it, you cannot manage it."

This is the type of person who believes that evidence-based policy is a must.

Ultimately, I believe that this forum is about finding the evidence. We are all here to compare notes about how we go about measuring progress in our economies and our societies. I look forward to our discussion this afternoon.

References

Alan S. Blinder and Alan B. Krueger, "What Does the Public Know about Economic Policy, and How Does It Know It?," NBER Working Paper No. 10787, September 2004.

Bureau of Economic Analysis, "Preliminary Study Shows Research And Development Contribution To Economic Growth," Press Release, September 28, 2006 (http://www.bea.gov/newsreleases/general/rd/rdspend06.htm).

Rosemary D. Marcuss and Richard E. Kane, "U.S. National Income and Product Statistics: Born of the Great Depression and World War II," Survey of Current Business, Bureau of Economic Analysis, Vol. 87, no. 2, February 2007 (http://www.bea.gov/scb/index.htm).

Chapter 7
Empowering and Challenging Voters through Governance Indicators: The Philippine Experience

Romulo A. Virola[i]
Secretary General, National Statistical Coordination Board, Philippines

Abstract

This paper describes the Philippine experience in creating and using indicator sets to measure not only the progress and performance of a society but to also strengthen its democratic processes. The focus is the results obtained, and lessons learned, from applying a Good Governance Index (CGI) to election results. The CGI, created by the National Coordination Statistical Board (NCSB) of the Philippines, grades each province on three aspects: its economic, political and administrative governance. The CGI was computed in 2000 and 2003, placing provinces in one of four categories: "best provinces" and "best-performing provinces", as well as "worst provinces" and "worst-performing provinces". During 2004, the election results for provincial governors were correlated to the CGI Index results. The aim was to show to what extent GDI results affected election results, if any. A Voters' Index for the 2004 Philippine Elections was then created, and the lessons learned from that are presented.

Introduction

Existing Indicator Sets To Measure Progress

Various indicators have been developed to assess and measure progress. The most widely used are the macroeconomic aggregates like the Gross Domestic Product (GDP) and the Gross National Product/Income (GNP/GNI) in the System of National Accounts (Commission of European Commission *et al.* 1993). The criticism against the GDP/GNI is that they do not capture dimensions of development like improvement in knowledge, health and the quality of life, in general. They also cannot explicitly describe the progress of societies towards good governance. Inevitably, the search for better indicators continues, a reflection of the multifarious development initiatives, the varying priorities of different groups and the multidimensional nature of many concerns of society.

i The author thanks Kristine Faith S. Agtarap, Noel S. Nepomuceno, Severa B. de Costo, Ma. Ivy T. Querubin and Mai Lin C. Villaruel for the assistance in the preparation of this paper.

The Human Development Index or HDI conceptualised by the United Nations Development Programme (UNDP) in 1990 attempts to measure human development as a process of enlarging people's choices (UNDP 1990). More specifically, it covers three human development goals: to lead a long and healthy life, to be educated and to enjoy a decent standard of living. But aside from the usual indexing issues that have been raised about the HDI, its value is diminished by the data constraints faced by the indicators comprising the index. In developing countries for example, life expectancy and literacy[1] are available very infrequently rendering the monitoring of the HDI ineffectual.

In September 2000, under the United Nations Millennium Declaration, 192 United Nations member states agreed to try to achieve eight goals by the year 2015: (a) Eradicate extreme poverty and hunger; (b) Achieve universal primary education; (c) Promote gender equality and empower women; (d) Reduce child mortality; (e) Improve maternal health; (f) Combat HIV/AIDS, malaria, and other diseases; (g) Ensure environmental sustainability; and (h) Develop a global partnership for development. These goals are to be assessed based on 18 targets and 48 indicators (UNDP 2003). Like the HDI, the MDG indicators face data availability constraints[2]; in addition questions have been raised on the relevance of some of the indicators or at least the lower priority considerations given to them in some countries.

The World Economic Forum, on the other hand, developed the Global Competitiveness Index (GCI) which focuses on the macroeconomic environment, the quality of the country's institutions and the state of the country's technology and supporting infrastructure to "explain why some countries are much more successful than others in raising income levels and opportunities for their respective populations (WEF 2007). The GCI is composed of nine pillars: institutions, infrastructure, macro economy, health and primary education, higher education and training, market efficiency (goods, labour, financial), technological readiness, business sophistication and innovation. Data come from publicly available administrative data and from the World Economic Forum Executive Opinion Survey.[3] The GCI has similar limitations as the HDI and is based on expert opinion, which has been found out by DIAL[4] through a mirror survey conducted for METAGORA, to systematically overestimate the level of corruption suffered by the citizens (OECD 2006).

Since 1995, Transparency International (TI) has been publishing an annual Corruption Perceptions Index (CPI) providing data on "corruption in international business transactions". It uses data from international surveys[5] on perceptions because "unbiased, hard data continue to be difficult to obtain and usually raise problematic questions with respect to validity" (Lambsdorff 2006, 2007). TI ranks countries according to the degree of corruption defined as "misuse of public power for private benefit, for example, bribing of public officials, kickbacks in public procurement, or embezzlement of public funds". Since the index is about relative positions of countries, it "cannot be used as a measure of national performance in combating corruption". Moreover, it suffers from comparability issues because "year-to-year comparisons of a country's score do not only result from a changing perception of a country's performance but also from a changing sample and methodology."

In 1972, Bhutan offered the Gross National Happiness[6] (GNH) index as an alternative to GDP in measuring progress. The four pillars of GNH are the promotion of equitable and sustainable socio-economic development, preservation and promotion of cultural values, conservation of the natural environment and establishment of good governance. Although two international conferences (Center for Bhutan Studies 2007) have been held

and a third is planned late this year[7] to discuss the GNH and interest in measuring happiness has increased through the conduct of various surveys[8], the idea has not caught up significantly enough, much less so with national statistical systems.

An inventory[9] of sources of governance indicators at the international level is provided among other places, in (UNDP 2006).

Likewise, in the Philippines, several initiatives have been undertaken to come up with tools for monitoring progress. Some of these tools have been used not only at the national level but they have been extended to the subnational[10] level where development planning has received greater attention than in the past with the passage of the Local Government Code in 1991. For instance, the National Statistical Coordination Board (NSCB) now regularly compiles the GDP for the regions (the Gross Regional Domestic Product or GRDP); but efforts are also ongoing towards the compilation of provincial product accounts. Similarly, provincial HDIs are being generated by the NSCB in collaboration with the Human Development Network (HDN) of the Philippines and pilot municipal HDIs are now being conducted.

As with the HDI, the MDGs have also gone down to the subnational level with the NSCB as one of the movers in the localisation of the MDGs.

Since 1998, the NSCB has been publishing annually the Countryside in Figures (CIF), which ranks the performance of the provinces in various aspects of governance but does not come up with an overall ranking (NSCB 1998). The CIF contains provincial data on the following: governance, income and expenditures, labor and employment, prices, finance, agriculture, health, education, communications, tourism, transportation and public order, safety and justice. The CIF also provides rankings of provinces by major sectors/categories as well as listings of local officials by province and city. In the CIF rankings, some provinces rank very high in some areas but very low in others, making it difficult to assess the overall performance of the provinces.

In 2004, the author presented a paper (Virola *et al.* 2004) during the 8th National Convention on Statistics (NCS)[11] which ranked provinces according to a governance framework developed by the NSCB in 1998 (NSCB 2003).

In addition, an earlier version of this paper was presented (Virola 2007) during the OECD World Forum on Statistics, Knowledge and Policy: Measuring Democracy and Human Rights Regional Conference for Asia in Seoul, South Korea on 8-9 February 2007.[12]

Obviously, the proliferation of the initiatives mentioned and other similar efforts has not solved the problems of society. Obviously, too, such initiatives will continue to proliferate from all directions, many of which will not necessarily make meaningful contributions towards progress and development. It is likewise obvious that much of the work is done by research institutions and private organisations outside of the national statistical systems. Thus, the initiatives are generally disconnected and many are *ad hoc* efforts which have not been mainstreamed in the realm of official statistics.

Nonetheless, it is worthwhile to mention that there are initiatives being pursued by various organisations that can lead towards better coordination and connectivity of the work being done in this area. This includes the work of the UN and its many agencies like the UN Statistics Division, the UNDP Oslo Governance Center and UNESCO, the World Bank, specifically its advocacy on Managing for Development Results under the Marrakech Action Plan for Statistics or MAPS and the OECD through its Statistics

Directorate, the Partnership in Statistics for Development in the 21st Century or PARIS21 and METAGORA, among others.

Survey-based initiatives, such as some of those mentioned that are used to generate governance indicators are of course not without flaws and weaknesses; however, it cannot be denied that they contribute to public debate but at the same time, they raise questions on the overall usefulness of the proliferation of indicator sets, nationally and internationally.

But one area that is worth exploring is the direct and proactive involvement of national statistical agencies in the construction of indicator sets that can help build democracies, particularly in the area of human rights and democratic governance. While this raises questions on the scope of work that statistical offices can do, it also offers golden opportunities to mainstream statistics in policy formulation and decision–making thereby enhancing the relevance of official statistics to society and its progress.

Involvement of National Statistical Agencies (NSAs) in Measuring Governance

Necessarily albeit not quite sufficiently, NSAs have been involved in the construction of indicator sets to monitor progress. Not quite obvious is the role of the NSAs in building indicators to strengthen the democratic processes of our societies. In fact, there may be official statisticians who still do not see their role expanding beyond the generation and dissemination of indicators on prices, employment, health, education, poverty and other traditional social and economic variables. And there may be others who feel some apprehension getting involved in possibly controversial, intensely political and high-profile issues such as democracy, human rights and governance.

The question therefore is, should NSAs get involved in the construction of indicator sets on these burning issues? Does such involvement contribute towards strengthening democratic processes?

In 2004, the Philippines thru the Commission on Human Rights, in collaboration with the NSCB and the National Commission of Indigenous Peoples agreed to participate in a world-wide project on "Democracy, Human Rights and Governance" or METAGORA. The project was hosted by the OECD through PARIS21 (Partnership in Statistics for Development in the 21st Century. METAGORA intends (OECD 2006) to measure good governance in the context of human rights and democracy and attempts to address the needs for evidence-based policy making, impact studies, aid delivery, empowerment of informed civil society and support to democratic dialogue, in particular in the context of distributive development. The Philippines is one of seven countries where pilot activities focusing on various themes were implemented. The Philippine component entitled "Development of a Pilot Survey Methodology for a Diagnosis of Rights of Indigenous Peoples in the Philippines" aimed to develop instruments for monitoring the rights to ancestral domains of indigenous peoples. The survey dealt with awareness/perception, realisations and violations, as well as mechanisms available to redress violations of these rights. The Project proved to be an eye-opener for the statistical agencies to get involved in the measurement of these concerns. The NSCB staff, who were traditionally more involved in statistical planning and coordination activities, gained conceptual knowledge of DHRG which gave them better insights in their planning and coordination function. It also enhanced the appreciation of DHRG institutions for statistics and provided the groundwork for closer collaboration and cooperation among the concerned institutions who, in fact, have not previously worked together. Aside from the survey, project

activities included capacity building of the partner organisations and focus group discussions to enrich the quantitative data with qualitative observations.

The METAGORA project therefore served as an eye-opener and catalyst that strengthened the resolve of the NSCB to get involved in the measurement of democratic governance.

At present, a number of institutions in the Philippines, both public and private, are generating information on governance.

In 2006, the NSCB collaborated with an academic institution[13] on a project that surveyed the users of governance indicators in the Philippines. In a way, this survey seeks to probe the usefulness of the indicators and the nature of their uses if any. One of the curious results of this survey is that neither one of the two respondents from the judiciary branch of government said that they are using governance indicators in the performance of their functions. The NSCB also worked with the government planning agency[14] on the development of a statistical framework to monitor and evaluate governance reform initiatives of the government.

In addition, private survey firms[15] periodically conduct surveys monitoring public opinion on corruption and governance issues.

The Construction of Useful Indicator Sets

The literature teems with ideas and principles on the construction of indicators (Tilbury *et al.* 2006) but some common characteristics of useful indicators are the following:

1. the indicators must be policy relevant and have wide application.
2. the indicators must be understandable and comparable, nationally and internationally
3. the indicators combine top-down and bottom-up approaches to enhance country/subnational ownership; unfortunately some initiatives have been basically imposed on countries/localities that sometimes render the indicators useless.
4. the indicators supplement quantitative with qualitative data.
5. the development and use of indicators are participatory and transparent, engaging and seeking wide consultation with many stakeholders.
6. the indicators must have timely data support and should be regularly monitored.

In addition, the successful use of the indicators will require capacity building (Virola 2007) on the part of both the producers and the users of the indicators. At present, official statisticians generally do not have conceptual knowledge of areas like democracy, human rights and governance and would need to be capacitated in these areas. Similarly, the target users of these indicators are generally ill-equipped with the statistical knowledge to be able to use them effectively, and correspondingly would need capacity building in statistics and measurement.

As tools for evidence-based decisions in managing for development results, let it be emphasised however, that the power of statistics lies not so much in its capacity to tell the stories of yesterday and today as in its use to shape and influence the stories that will be told tomorrow!

This paper is not about how to construct indicators. Rather, it shares the Philippine experience in the involvement of national statistical agencies like the NSCB in the measurement of progress and performance, hopefully, contributing in the debate towards answering questions on how best to construct and use indicator sets to strengthen the democratic processes in our societies.

Philippine Experience: Good Governance Index (GGI)

The start of election campaign season is the time when voters feel very important and powerful. This is the time when political candidates in the Philippines do everything they can to woo the voters through advertisements, insurance policies, cash, food, and other freebies. Appropriate or not, this is deeply imbedded in the culture of Philippine elections. All seasoned candidates know they ought to do more to entice the public to vote for them, beyond experience and performance. The excitement of the election thus draws on the decision of the voters.

In theory, there are so many areas to look into before deciding who to vote for. Philippine reality however is that for a greater majority of the voting public, the spirit of elections is lost on the colourful streamers, posters, and campaign jingles of political candidates. Surely, democratic processes cannot be strengthened if voters' decisions continue to be guided by questionable motives; when candidates who present strong credentials lose to platformless candidates who nevertheless exude almost unbelievable appeal to the masa.[16]

The challenge therefore, is to develop instruments that can guide and educate the voters to vote for the right candidates, based on performance. Toward this end, national statistical systems can play a significant role. The rest of the paper dwells on the efforts of the NSCB in that direction.

Methodology Used

The governance framework for the GGI as shown in Annex 7.1 was developed by NSCB (Virola 2004) and covers three areas: economic, political and administrative governance. It has sub-themes for each area with 40 core indicators for the three areas. It utilises mostly administrative-based data. Under this framework, governance is the manner in which power is exercised in the management of the country's economic and social resources for development. It also refers to the exercise of economic, political and administrative authority to manage the nation's affairs at all levels. Good governance promotes the collective responsibility of the government, civil society and private sector for improving the lives of all Filipinos, particularly the poor. The government, therefore, needs to efficiently direct scarce public resources in order to provide goods and services and protection to the poor.

While this paper uses the governance framework in Annex 7.1, it proposes a different set of indicators for the different subthemes. Some subthemes for which no indicators with appropriate data support could be identified have been excluded; while for others, some indicators have not been included because either the data are not timely or the desired provincial disaggregation has not been generated, or data quality is questionable like the data on crime rate. Some Millennium Development Goals indicators were also added.

For each province, the GGI is computed as the unweighted arithmetic average of the Economic Governance Index (EGI), the Political Governance Index (PGI), and the Administrative Governance Index (AGI). The EGI, PGI, and AGI are computed as the unweighted averages of the indexes corresponding to their subthemes. However, for this paper, the PGI was excluded due to unavailability of data.

Similarly, the above sub-indexes are computed as the unweighted arithmetic averages of the indicator-indexes under each subtheme. These indicator-indexes will themselves be unweighted arithmetic averages of another set of indicator-indexes. To introduce some benchmarking of the index, the index for the Philippines in 2000 is set at 100. At the lowest level of indexing, for indicators for which high values are desirable (positive indicators), the index for a province is obtained by dividing the value of the indicator for the province by the value of the indicator for the Philippines in 2000. On the other hand, for negative indicators, the index for a province is obtained by dividing the value of the indicator for the Philippines in 2000 by the value of the indicator for the province

In the choice of the indicators in this paper, aside from the availability of provincial level data and availability of more timely data, since the purpose is to assess the performance of the local government unit (LGU) executive, only those indicators which are more or less within the control of the executive are included. Examples of indicators excluded due to data constraints are court case disposition, cases resolved on graft and corruption and infant mortality rate. Examples of indicators excluded because they are not clearly within the control of the LGU executives are population growth rate, number of banks and number of pawnshops. Also excluded are indicators which may not be clearly reflective of good governance or which may not be uniformly relevant across all provinces. Some of these are police to population ratio, total bank deposits and agricultural production.

In addition, the GGI is compiled at the beginning and at the end of the term of office of the LGU executives in order to assess their performance. The difference in the GGI between the beginning and the end of the term is used as a measure of improvement, if any.

In the actual indexes computed in this paper, some indicators are missing for some subthemes for reasons mentioned earlier and the formula for the indexes have to be adjusted accordingly. In cases where indicators are not available for a certain year such as those on health, housing and poverty, the indicators are not included in the index computation for that year. In addition, in order to prevent an indicator from unduly influencing the index, such as when the value of the subindex for an indicator is over 1000, maximum limits are set for each subindex: 500 at the lowest level, 400 at the second level and 300 for the third level. The limits have been arbitrarily set; an alternative would be to use a certain number of standard deviations from the mean. The methodology is discussed in detail in NSCB (Viola 2004) and Virola, et.al. (NSBC 1998).

Obviously, the GGI presented in the paper has limitations. The methodology can be refined to enhance its robustness and the validity of the GGI as an indicator of governance. Toward this end, there is a need to generate timely data with the appropriate geographical disaggregation. More generally, there is a need to mainstream the measurement of DHRG in national statistical systems. It would also be useful to seek consultations with various stakeholders including the politicians.[17]

GGI Results: Voters Rewards & Punishment (2004 and 2007 elections)

As illustration, the GGI was computed for 2000, 2003 and 2005.[18] Best provinces are those with the highest GGIs and Best Performing Provinces are those with the highest differences between the 2003 GGI and the 2000 GGI and between the 2005 GGI and the 2003 GGI. For the 2004 and 2007 elections, the results on the performance of the governors are summarised as follows:

- 30 Best Provinces:
  - 2004 3 out of the 30 best provinces lost, 25 won! (2 did not run)
  - 2007 7 out of the 30 best provinces lost, 18 won! (1 did not run,4 no results yet)
- 30 Worst Provinces:
  - 2004 23 out of the 30 worst still won, 7 lost! (all still ran!)
  - 2007 19 out of the 30 worst still won, 4 lost! (2 did not run, 5 no results yet)
- 30 Best Performing Provinces:
  - 2004 4 of the 30 best performing lost, 24 won! (2 did not run)
  - 2007 7 of the 30 best performing lost, 18 won! (2 did not run, 3 no results yet)
- 30 Worst Performing Provinces:
  - 2004 22 of the 30 worst performing still won, 7 lost! (1 still ran!)
  - 2007 19 of the 30 worst performing still won, 4 lost! (1 did not run, 6 no results yet)

Voters' Index

How do voters assess their candidates? Is it based on the candidates' performance? Or is it based on other factors? The increasing number of actors/actresses/sports personalities-turned politicians shows the power of the media in influencing the decision of Filipino voters.[19] Also, cultural values such as kinship ties and sense of gratitude (repayment of debts and favors received) are known to influence voters' decisions.

The results of the 2004 and 2007 elections clearly demonstrate the need to advocate for better behaviour among voters. The Voters' Index is intended as a tool for this advocacy. The Voters' Index aims to measure the "wisdom" of the voters in selecting their candidates. A value of "0" or "1" is assigned to a province, depending on the results of the election and their GGI: 1 for best or best performing provinces whose governor[20] won and 0 otherwise; 1 for worst or worst performing provinces whose governor lost and 0 otherwise.

The table below shows the Voters' Index for the 2004 and 2007 Philippine Elections, using 10, 20 and 30 provinces in the computations. The table shows that good performance is not sufficient for the governor to win, neither is bad performance

sufficient for a governor to lose; worse, the Voters' Index has deteriorated between 2004 and 2007. This is alarming and it would be interesting to know public reaction to this; certainly, it is a challenge that civil society and all stakeholders concerned about voters' decisions during elections must address!

Table 7.1 Voters' Index for the 2004 and 2007 Philippine Elections

Voters' Index									
	10 Provinces			20 Provinces			30 Provinces		
	2004	2007	Change	2004	2007	Change	2004	2007	Change
Best Provinces 2005	0.89	0.78	-0.11	0.89	0.80	-0.09	0.89	0.72	-0.17
Worst Provinces 2005	0.20	0.13	-0.07	0.15	0.13	-0.02	0.23	0.17	-0.06
Best Performing Provinces 2003-05	0.90	0.67	-0.23	0.85	0.78	-0.07	0.86	0.72	-0.14
Worst Performing Provinces 2003-05	0.22	0.13	-0.09	0.26	0.07	-0.19	0.24	0.17	-0.07

Lessons Learned

1. Obviously, there is increasing interest worldwide in governance indicators. Interestingly, local interest has also been generated among various stakeholders – the governed as well as the governors, the media, civil society , government and the private sector. Participants in various fora have shown great interest in the presentation of governance indicators, with interest extending beyond the National Capital Region.
2. Reaction to the involvement of NSCB in governance indicators has been overwhelmingly positive and encouraging. Even a governor whose province did not do well in the GGI and who initially had questions about the GGI, eventually recognised the merits of the GGI.
3. It is imperative that methodologies be transparent, especially with their limitations.
4. The media, as usual, could spark greater interest in the indicators and is an indispensable partner in advocating the use of governance indicators to guide voters during elections.
5. Governance indicators are used. Politicians whose provinces do well would use the GGI to advance their political causes.
6. Somehow, rankings give especial attraction to the indicator sets, particularly from the media and the subjects of the rankings themselves. In the Philippines, while we have tried recently to focus attention on the best performers, there would always be probing, curious questions on the worst performers.
7. The timing and extent of dissemination of the results are important for maximum impact. For example, advocacy for voters to use information on the GGI is needed so that the Voters' Index will increase over time; this requires that the GGI is released strategically around election time. That the Voters' Index did not increase between 2004 and 2007 could mean that much more advocacy work needs to be done to guide voters towards better decisions during elections. Likewise, the compilation of the governance index down to lower administrative

levels such as the municipalities, where attribution of governance results can be more properly pinpointed, should be considered.

8. The relevance and usefulness of the GGI depend obviously on the capability of the Philippine Statistical System (PSS) to provide the necessary data support, which, as might be expected, the PSS is currently unable to do.

Concluding Remarks/Future Directions

While some indicator sets may not be totally useful to all countries and while these indicator sets have various limitations, they have their uses, both positive and negative.

It is thus important that compilers of indicator sets recognise their responsibility to stakeholders, in the case of governance indicators, the governed, the governors and the media. Wide consultations should be done and best practices in information dissemination should be followed. The methodologies should continuously undergo refinements and greater advocacy for the use of indicator sets should be pursued. For example, it is important that the voters appreciate that the Voters' Index increases over time.

More importantly, it is now becoming fairly evident that it would be desirable for national statistical agencies to get involved in the construction of indicators sets to monitor the progress of societies beyond the traditional GDP/GNP(GNI), HDI or the MDGs; more specifically towards the measurement of democracy, human rights and governance. Official statisticians should not be afraid to come up with information, no matter how controversial. As long as the appropriate metadata and methodological limitations are shared with the users, some information is always better than no information. Taking this view will promote statistics and the use of information towards evidence-based decisions. In return, the role of statistics and official statisticians in society will be appreciated better.

Finally, with the increasing demand for statistics in various areas, government as well as the private sector must be ready to invest more resources in statistics.

Notes

1 In the Philippines, the latest available data on life expectancy and literacy are for 2005-2010 and 2000, respectively, based on the 2000 Census of Population and Housing .

2 In the Philippines, as of August 2006, out of the 48 MDG indicators, 29 were available, 8 were not available and 11 were not applicable.

3 For The Global Competitiveness Report 2006-2007, over 11 000 business leaders were polled in a record 125 economies worldwide.

4 Developpement, Institutions et Analyses de Long terme (Paris)

5 For the CPI 2006, 12 sources from 9 independent institutions covering 163 countries were included.

6 The term was coined by Bhutan's King Jigme Singye Wangchuck in 1972.

7 The first International Conference "Operationalizing Gross National Happiness" was convened in Thimphu, Bhutan on 18-20 February 2004 by the Center for Bhutan Studies. The second international conference was organised in Nova Scotia, Canada by GPI Atlantic on 20-24 June 2005. The third is scheduled in Thailand on 22-28 November 2007 with the Sathirakoses Nagapradipa Foundation and the Center for Bhutan Studies as co-organisers.

8 An example is the World Values Survey whose Philippine member is the Social Weather Station which first incorporated the question "Taking all things together, would you say you are Very happy, Fairly happy, Not very happy, or Not at all happy?" in July 1991 (http://www.sws.org.ph/pr050314.htm

9 A similar inventory entitled "Governance Indicators in the Philippines: A Users' Guide" has been compiled in the Philippines by the National College of Public Administration and Governance jointly with the National Statistical Coordination Board of the Republic of the Philippines with assistance from the UNDP and the Oslo Governance Centrre.

10 As of 31 March 2007, the Philippines has 17 regions, 81 provinces, 118 cities, 1 510 municipalities and 41 994 barangays.

11 The NCS started in 1978 and is now held every three years as a collaborative effort of the Philippine Statistical System under the stewardship of the NSCB.

12 Other presentations by the author on similar topics are a) "Governance Frameworks and Indicator Systems: Some Experiences in the Philippines" during the UNDP Oslo Governance Center-Indian Council of Social Science Research International Workshop on Governance Indicators for Pro-Poor and Gender-Sensitive Policy Reforms, in New Delhi, India, 20-22 April 2005; b) "Measuring Democracy, Human Rights and Governance: Should Official Statisticians Be Involved? " to be presented during the 56th Session of the International Statistical Institute in Lisboa, Portugal on 22-29 August 2007; and c) "Governance Statistics: Did Performance Matter in the 2007 Elections?" to be presented during the 10th National Convention on Statistics, EDSA Shangri-La, 1-2 October 2007 He also wrote an article on 12 February 2007 for Statistically Speaking - "Governance Statistics: Who Should We Vote For in the May 14 Elections?", http://www.nscb.gov.ph

13 The National College of Public Administration and Governance (NCPAG) of the University of the Philippines.

14 The National Economic and Development Authority (NEDA)

15 The two most visible and credible polling firms in the Philippines are the Social Weather Station and Pulse Asia.

16 The greater majority of the public, connoting low social and economic standing.

17 The principal author made a presentation entitled "Towards Informed Decisions of Voters During Elections: Measuring Governance of Local Officials" on 13 April 2007, during a forum "2007 Elections: The Good, The Bad and the Ugly in the Local Government" sponsored by the Mayors Development Center, Inc. Philippines.

18 Elections were held in 2001, 2004 and 2007 and the relevant GGIs to compute would be for 2000, 2003 and 2006. However, due to data constraints, the GGI for 2006 could not yet be computed. Also, not all the results of the 2007 elections have been released so that the list of best/worst provinces is based only on those provinces for which the election results are already known.

19 Interestingly, majority of the actors/actresses/sports figures/media personalities, who ran during the 2007 elections lost. The casualties include two popular actors who lost badly in the senatorial race and a very popular, well-loved and well-liked sports figure who ran for a seat in Congress.

20 Or a candidate related to the governor.

References

Center for Bhutan Sudies (2007), *3rd International Conference on Gross National Happiness*, http://www.bhutanstudies.org.bt/annnounce.htm ,

Commission of the European Communities *et al.* (1993), *System of National Accounts*.

Lambsdorff, Johann Graf (2002), *Background Paper to the 2002 Corruption Perceptions Index*, Transparency International (TI) and Gottingen University.

Lambsdorff, Johann Graf (2006), *The Methodology of the Corruption Perceptions Index 2006.* Transparency International (TI) and University of Passau.

National Statistical Coordination Board (1998-), Countryside in Figures, various issues.

National Statistical Coordination Board (2003), *Report on the Development of Indicators and Design of a Database and Information Network of the Governance Statistics Project*.

OECD (2006), Metagora, *Measuring Human Rights and Democratic Governance to Inform Key Policies*. Light version of the synthesis report on results and lessons from the two-year pilot phase of the Metagora project (February 2004-February 2006).

Tilbury, D. *et al.* (2006), *Monitoring and Assessing Progress during the UN DESD in the Asia-Pacific Region: A Quick Guide to DevelopingESD Indicators*. UNESCO Asia and Pacific Regional Bureau for Education, Bangkok.

Virola, Romulo A. *et al.* (2004), "The Best and Worst Provinces in the Philippines: What Happened to their Leaders in the 2004 Elections?", paper presented at the 9th National Convention on Statistics, Mandaluyong City, Philippines.

Virola, Romulo A. (2007) "Indicators of Democratic Governance as Guide to Voters in Philippine Elections: Inspiration from METAGORA", paper presented at the Regional Conference for Asia, Measuring and Fostering the Progress of Societies, Measuring Democracy and Human Rights, OECD and the Korean National Statistics Office, Seoul, South Korea.

Virola, Romulo A. (2007) "Challenges and Opportunities in the Coordination of a Decentralised Statistical System". Paper presented at the Panel Discussion on Prospects and Risks in the Future: How to Manage Uncertainties of the Seminar on the Evolution of National Statistical Systems, United Nations Statistics Division, New York, United States.

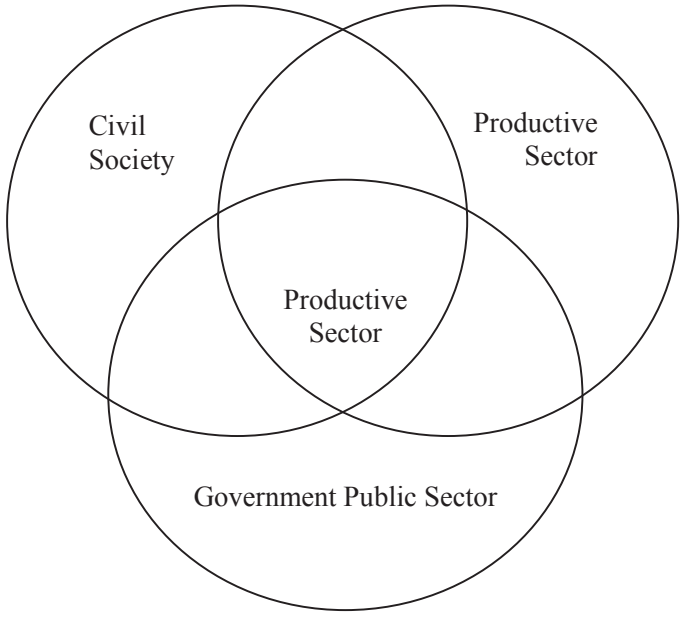
UNDP (1990) Human Development Report and other issues.

United Nations Development Group (2003) Indicators for Monitoring the Millennium Development Goals.

UNDP (2006), Oslo Governance Centre Governance Indicators: A Users' Guide, Second Edition.

World Economic Forum (2007), The Global Competitiveness Report 2006-2007. http://www.weforum.org

Annex 7.1 Framework for the Development of Governance Indicators



Economic

1. Sustainable Manager of Resources
2. Enhanced Governme Responsiveness to th poor

Political

1. Improvement of internal and external security
2. Law enforcement and Administration of justice
3. Elimination of graft and corruption

Administrative

1. Efficiency in the delivery of services
2. Improved transparency and accountability
3. Continuous building of capacities
4. Expanded use of ICT

Source: Report on the Development of Indicators and Design of a Database and Information Network of Governance Statistics Project, NSCB.

Part Three
Accountability and Civic Engagement

Chapter 8
The Importance of Social Reality for Europe's Economy: An Application to Civil Participation[i]

Marcel Canoy; Frédéric Lerais
Bureau of European Policy Advisers (BEPA)[ii], European Commission

Massimiliano Mascherini; Andrea Saltelli; Daniele Vidoni
Centre for Research on Lifelong Learning (CRELL, JRC)[iii], European Commission

Abstract

This paper attempts to gauge what variables affect participation in civil society, arguing that full participation is key to the future of Europe. The first part of the paper provides an overview of Europe's social conditions, including discussion of why globalisation triggers the need for social change, the fate of trade-replaced workers, and European social models. Part two focuses on the Active Citizenship Composite Indicator, which draws data from the European Social Survey of 2002. The authors create participation models correlating socioeconomic and demographic variables – income levels, age, education, employment, etc. – to participation in private and/or social organisations. Results are then grouped into four regional country clusters: Nordic, Continental; Mediterranean; Eastern. Also considered is the relationship between people's social participation and how they perceive the world – from their trust in others to their assessment of the economy. Conclusions and further research suggestions are presented.

i The analysis contained in this report is personal to the authors and does not necessarily reflect the views of the European Commission.

ii The mission of the Bureau of European Policy Advisers (BEPA) is to provide timely, informed, policy and political advice to the President and Commission Services on issues relevant to the President's agenda and the future of policies in the Union. Due to its special position, working directly to the President, BEPA can lead inter-service groups on specific policy issues and participates in horizontal work within the Commission. In order to achieve its mission, BEPA aims to produce research and policy analysis up to high professional standards. It is on the basis of this strong conceptual and empirical work that BEPA contributes to effective communication not only within the Commission and the EU Institutions but also with academia, markets and the public in general. The activity of BEPA is complementary to those of other Commission Services since it concentrates on the early (strategic) stage of the policy cycle, thereby contributing to shaping policy options in the medium and long run. BEPA interacts with academia, research institutes and, in general, outside professionals in order to ensure that in making policy the President and through him, the Commission are informed by the best analysis available.

iii The mission of the Joint Research Centre is to provide customer-driven scientific and technical support for the conception, development, implementation and monitoring of European Union policies. As a service of the European Commission, the JRC functions as a reference centre of science and technology for the Union. Close to the policy making process, it serves the common interest of the Member States, while being independent of special interests, whether private or national.

Introduction

In the course of history the European Union is perceived as a level of policy making that is largely economically driven. The Single Market, the stability and growth pact, the introduction of the euro, competition policy and trade policy are all key European level competences. But ever since Jacques Delors there has been an 'uncomfortable feeling' by many about the dominance of economic issues. Should the Union show its 'social face', too? Is economics 'all that matters'?

For a while 'showing a social face' or quid pro quo social policies to 'compensate for the (alleged) losses of economic policies' could be seen as the driving political forces of social policies at the European level, exemplified by the original Lisbon agenda of 2000, where social policies and economic policies where identified as separate pillars, arguably in part for those reasons.

One of the risks of such an approach is that it puts economic and social policies as antagonistic entities, often leading to polarised political discussions as if countries have to choose between free markets or social societies. And even if that is not the case it erroneously displays economic and social policies as disjoint entities that do not interact. In reality they do interact intensively either in a mutually reinforcing way or in a way that trades off one against the other.

At the top of the Danes' wish list as to how the EU should prioritise its tasks are: protecting the environment, peace and security, as well as fighting terrorism and reducing the gap between rich and poor. Only 19% of the Dutch and Danes think cost of living is a priority. And even in a liberal country such as Ireland the most popular answer to the question 'which actions should the European Union prioritise?' is a surprising 'fight poverty/social exclusion'.[1]

Gradually, the European Commission has tried to integrate social policies into their core business without doing injustice to relative responsibilities by other policy levels. To do so effectively, since the revision of the Lisbon strategy in 2005 into a growth and jobs agenda and particularly after the informal Summit in Hampton Court in the end of 2005, a new logic prevailed.

The conventional wisdom that greater economic prosperity drives jobs growth that in turn drives well-being in general can be challenged (Liddle & Lerais 2007). Not only is such logic not perceived as such by many citizens, it is also still based on a pillared view of economic and social policies. The logic is that economic activities create the size of the pie while social policies deal with the division of it.

If we look at the characteristics of our current and likely future societies such logic falls apart. Europe of today and tomorrow is more diverse than yesterday, people face different social risks, the world moves at a higher speed, is more knowledge driven and the population is ageing rapidly. The consequence of the combination of these characteristics is that future economic prosperity does not only depend on economic activities. Tomorrows Europe needs inter alia: (i) the highest possible education for all; (ii) the highest labour participation possible; (iii) well integrated migrants that add value to European societies; (iv) healthy citizens; (v) citizens fully participating in civil society.

These requirements are not needed (just) for social reasons but for economic reasons too. They do not relate (solely) to division of the pie but have clear implications for creation of a larger pie. More, healthier, better educated and participating citizens create higher growth.

One of the consequences of the different nature of interactions between economic and social policies is that there is a shift in responsibilities. There is more need for policy coherence across layers and actors within society, and for a greater role of private participation – such as that of NGO's. More than in the past an increased ownership and shared responsibility of social issues is seen as beneficial for policy, '*e.g.*' via an increased capacity for anticipation and prevention.

This paper underlines the importance of social and economic policies as reinforcing each other. The example which we consider in detail is civil participation. In other publications we have stressed the link between labour market and economic policies (Canoy & Smith 2006), on youth policies shaping future growth (Barrington-Leach *et al.* 2007), we have provided a snapshot of social reality (Liddle & Lerais 2007) and stressed the importance of a coherent EU migration policy as shaping Europe's future labour markets (Canoy *et al.* 2006).

Civil society[2] is both the substratum that will 'eat the pie', and a primary stakeholder of the policy making process. It is also clear that policies will only work if civil society is ready and able to take its responsibilities. This necessitates a process of co-ordination and valorisation of different forms of interactions (both formal and informal) between state and society. The existence of different European social models with different structures and rules prevents the implementation of one-size-fits-all solutions.

In order to determine how civil participation in these different models contributes to Europe's future, we first have to measure the current situation. Subsequently, we ask the question: 'What are the determinants of social involvement and how can involvement enhance the results of the coordination process between State and stakeholders in an increasingly diverse society?' The paper investigates this question on two levels. First, it studies the macro-level characteristics of Europe's social reality and identifies the leading avenues for change. Second, it dives deeper in the issue of 'participation' and proposes a model for investigating the issue across Europe and the relationship between participation and policy.

Our paper is divided up in two main parts. In the first (Sections "A New Paradigm: Adjusting to Change", Responsibilities of Governments Levels and Other Stakeholders") we provide an overview of Europe's social reality, develop a new paradigm of equipping people for change, and discuss the various responsibilities of government levels and other stakeholders.

As previously stated, tomorrows Europe needs citizens fully participating in civil society. Hence, the second part of the paper (Sections from "A Renewed Role for Individual Participation" to "Participation and Perception of Europe's Social Reality") is a case-study that investigates the characteristics of the shares of individuals that are more likely to be involved with different kinds of formal organisations and shows how each of these groups perceives the political, economical, and relational spheres of social reality. The last section concludes the paper with conclusions and further research.

A New Paradigm: Adjusting to Change

The red thread within the social and economic issues in Europe today is how to make sure that people can adapt to – and benefit from – change. The need for adaptation and change is accepted by EU citizens: 76% of Europeans agree that lifetime jobs with the same employer are a thing of the past; 76% also consider that being able to change easily

from one job to another is a useful asset to find a job nowadays; 72% of people said work contracts should become more flexible to encourage job creation. Finally, 88% of citizens said that regular training improves one's job opportunities.[3] This new paradigm provides a natural viewpoint to reunite economic and social goals, and to design institutions that aim at realigning them. To explain this paradigm we have to discuss first the nature of the relationship between economic and social policies. We approach the issue from four different angles: *i)* Why globalisation triggers the need for social change; *ii)* how are European social models performing; *iii)* which trends reinforce the need for change; and *iv)* what are the consequences for policy.

Why Globalisation Triggers the Need for Social Change[4]

In many Member States there are discussions and fears that a stronger internal market, trade, competition policy and globalisation will jeopardise national 'social models'. If we analyse the impact of globalisation on social realities, we can better understand how social and economic policies interact.

A trade-off between economic and social policies is a variant of the well-known efficiency equity trade-off that dates back to 1975 when Arthur Okun wrote that 'the pursuit of efficiency necessarily creates inequalities. And hence society faces a trade-off between equality and efficiency' (Okun 1975). The textbook trade-off emerges because protection and taxes reduce incentives to work. It also assumes countries are at the efficiency frontier. There is a growing empirical literature that denies the general validity of the claim that there exists such a trade-off. However, the literature is less precise on the conditions under which the trade-off can emerge (Esping-Andersen 2002; Jackson 2000).

In a general sense increased international competition strengthens the mobile factor (capital) potentially at the expense of the immobile factor (labour). In addition the demand for skilled workers rises relative to unskilled workers. However, neither of those effects implies that countries have difficulties of redistributing income should they wish to do so. Empirically, there are many examples of countries that have succeeded in redistributing income despite increased global competition (for example the Nordic countries). Theoretically, there are a number of factors that moderate, compensate or possibly reverse the trade-off.

First, it matters how one defines equity. It is one thing to find one self temporarily in shortened straights and quite another thing to be permanently excluded from improving one's position. Opening up opportunities for those that are permanently excluded is also a crucial dimension of equity, not to be confused with a point measure such as those below a specific poverty line. Second, an effective way of reducing inequality is to provide jobs for low income groups. There are many efficiency enhancing strategies available that do just that. Finally, there are political economy reasons whereby it is easier to redistribute when there is more to share.

If the evidence for trade-offs at the macro level is mixed, the obvious way to get closer to the conditions under which there exists a trade-off is to verify on the micro level where open market and social policies 'meet'. The central meeting place is the labour market. An important consequence of policies that open up markets to competition is that companies will have to restructure their business more often or faster than before.

A first implication of business restructuring is that it may 'punish' inappropriate labour market institutions and 'reward' appropriate ones. Open market policies expose countries and their enterprises to outside forces, which can reveal certain institutional

weaknesses of a country that would otherwise have been concealed (or revealed at a later stage).

Think of the debate on relocation in contemporary policy making. As a result of open market policies firms restructure their business, relocating certain activities to low wage countries and potentially getting some high-skilled jobs in return. What might happen with the people who lose their jobs in countries with rigid labour markets? The people who lose their job are not easily re-employable, either because they have obsolete skills, they make use of early retirement or other exit schemes, there are no appropriate jobs available for them or they are not prepared to move.

Table 8.1 Fate of Trade Displaced Workers

	High international competition manufacturing	Medium international competition, manufacturing	Low international competition, manufacturing	All manufacturing	Services	All sectors
Characteristics of workers						
Age at displacement						
15-24	10.4	13.1	11.6	11.8	12.2	11.4
25-54	75.1	75.8	78.1	76.4	78.0	76.9
55-64	14.5	11.2	10.3	11.9	9.8	11.7
Mean	40.9	38.8	39.4	39.7	37.9	39.2
Share women	31.7	44.9	26.2	34.8	43.2	38.2
Predisplacement occupation						
White collar	31.9	20.0	27.1	25.9	73.3	48.5
Blue collar	68.1	80.0	72.9	74.1	26.7	51.5
Job tenure at time of displacement						
More than 10 years	32.1	30.4	27.7	30.0	18.6	21.5
Mean job tenure	7	6.6	6.2	6.3	4.7	5.0
Hourly earnings in old job						
Mean (€)	9.51	9.15	9.08	9.43	9.15	9.08
Adjustment costs						
Share reemployed after two years	51.8	58.7	59.6	57.0	57.2	57.3
For reemployed						
Share with no earnings loss or earning more	44.0	45.7	47.3	45.8	49.6	47.1
Share with earnings losses greater than 30%	5.4	7.0	6.8	6.5	8.4	7.5

Countries: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain and the United Kingdom. Source: European Household Panel, waves 1 to 8, April 2003, as quoted in OECD (2005).

Empirically, negative effects can translate in different ways. First, there are significant divergences between displaced workers. While some fare well, others have significant problems finding a new or equally well-paid job. One reason could be that displaced workers are on average older, less educated, or possess inadequate skills. Second, there are also differences between the US and Europe. In the US, displaced workers tend to find a new job, but often take a significant wage cut. In Europe, displaced

workers tend to face long-term unemployment or complete withdrawal from the labour market.

Table 8.1 shows how displaced workers fare in 14 European countries. It shows that blue-collar workers are more exposed to adjustment than white collar ones. Moreover, these workers tend to have been employed for a long time. It also shows that a large part of those employed in manufacturing sectors facing high international competition fail to find a new job within two years. However, if they find a job, they do not have to take too much of a cut in wages.

Depending on the conditions, relocation can be positive for both countries involved when complementarities are strengthened or negative for the sending country when it is losing production because of an unfavourable domestic business climate. In the second case, not only do firms decide to find their luck elsewhere, but also the impact of relocation on the labour market is visibly negative, both in real terms and in the perception of the people involved. So the economic and social costs of relocation are high and the benefits low. However, in this particular case the problems are not created by open market policies but revealed by it.

There is a clear parallel with government subsidies of obsolete industries. On the surface such subsidies may seem 'social', but in the long run they are wasteful. The social and economic impact of delaying restructuring is that the problems emerge later anyway and will be bigger.

So far the link from open market policies to the labour market has been considered. What about the reverse? Consider two polar cases. Suppose that as a result of institutional changes a substantial percentage of low-skilled people, men, women, ethnic minorities, young and old, who previously relied on social benefits have now found a new enthusiasm in finding a job. What might happen (Solow 1998)?

In a world characterised by low trade, lenient competition policy and lack of innovation the labour market is like a game of musical chairs. There are a limited number of jobs out there, so adding a large quantity of low skilled people on the job market only has reshuffling consequences. Some of the new ones might find a job, but at the expense of others. And wages may drop too as a result of the increased competition for the chairs. Ineffective competition policy favours incumbent companies and workers, and mobility is lacking.

A dynamic economy will generate new employment opportunities and mobility is high. The rewards of modernising the labour market will be much higher, since more people will improve job matches. In a world with appropriate competition policy, innovation and trade, new opportunities are taken either by start-ups, by FDI or incremental innovations by multinationals. Competition policy helps by not favouring the incumbents and jobs are created.

But a flexible labour market alone is not enough. The challenge of adaptation also requires satisfactory levels of security to respond simultaneously to the new needs of employers and employees. Europe must be able to avoid more segmented labour markets with the risk of more precarious jobs, less sustainable integration and less accumulation of skills. Women, the young and migrants are already over-represented among labour market outsiders and older workers face a number of difficulties keeping or finding a new job.[5] Even insiders are threatened: their protection becomes less real as they also face the same difficulties moving on to good quality jobs if they experience redundancy. New forms of flexibility and security are needed, for individuals, companies, Member States

and the Union as a whole. Individuals increasingly need employment security rather than job security, as fewer have a job for life. Companies, especially SMEs, need to be able to recruit staff with a better skills match, who will be more productive and adaptable leading to greater innovation and competitiveness.

To conclude, economic policies that open up market can only be effective if labour markets are sufficiently flexible and adequate levels of security are in place. But for labour market reform to be successful open market policies should be in place. Labour and product markets are communicating vessels. One explanation for the mixed evidence at the macro level is that some countries (Nordic countries) are more successful in letting the vessels communicate than others (Germany, France). Successful countries design policies that make vessels communicate, and policies that combine flexibility on the labour market with security for the individual affected (protecting the person not the job).

The second way in which open market policies affect equity is by their distributional effects. Here one must be careful how to measure distribution. Empirical studies generally show that the magnitude of macro distributive effects remains limited as compared with the potential gains (Baldwin & Martin 1999).

However, the political impact can be much greater for several reasons. The effects are generally concentrated in certain sectors or regions, therefore their political impact can be much greater than would be the case for effects uniformly distributed across the whole population. Costs tend to be more significant in the first years, while most gains require more time to take full effect. Costs and benefits are typically dissociated in space. Some regions suffer from the adjustments induced by trade opening in certain sectors, while others benefit from the expansion of sectors based in their area. The spatial dimension is thus critical: in declining regions, job losses constitute a clear cost; benefits derived from comparative advantages are only apparent in dynamic regions.

It follows that standard ways of measuring distribution such as the Gini coefficient are too aggregated to pick up distributional effects in time and space, but those effects can be politically vital. Notice also that there is a link between the communicating vessels argument and redistribution. Lack of labour mobility and inappropriate labour market institutions aggravate the spatial dimension of the distributive effects.

Concluding, economic policies and social policies can go hand in hand when a country realises that they are communicating vessels and designs policies to enable the vessels to communicate. But there remain distributional effects in time and space, which may need flanking policies.

Underperformance of European Social Models

Apart from the need to accommodate to the changes triggered by globalisation it is also the social models themselves that need modernisation irrespective of globalisation. European social models do not meet modern social needs. The shortcomings of social models are often attributed to ageing and globalisation. Yet, both phenomena play a somewhat different role than the one which they are credited for. Ageing has a direct impact on the financing of health care and pensions, but has also more subtle consequences for the labour market and intergenerational solidarity. However, even in the absence of ageing, problems would arise to finance pensions and health care at existing levels of provision. Problems are aggravated not caused by ageing (cf. Canoy &Smith 2006). As already argued above, globalisation exposes rather than causes problems.

Social performance can be analysed in a number of ways, with regard to the situation in different countries, with regard to the degree to which social policy meets its current objectives or the extent to which new social needs are catered for. All of these are useful and provide insights into the question 'how well is the European social model meeting needs'?

Many social systems have failed to respond to new social demands and risks arising from major changes in social, economic and cultural parameters since the inception of the social systems themselves. Welfare state institutions today were designed to deal with the challenges of growing industrial economies. Yet risks have changed. The transformation of modes of production in the move to a post industrial society yield faster obsolescence of skills, discontinuity in professional careers, labour market demands for women's participation, part time work and short term contracts. Greater competition creates a need for more mobility and flexibility of workers and firms. Technological advances imply improved productivity, instant communication, the rise of the service economy, development of knowledge society, new demands on democratic systems. Demographic evolutions lead to ageing and a multicultural composition of societies. Roles of women and men have changed, both in their private and their public life (two-wage earner families, family breakdowns, single mothers, higher cost of children).

It is necessary to adapt the existing social institutions by considering the actual risks people face today and will face in future. The most striking example concerns the lack of consideration given to changes in family functions and structures. As a result, most of the European Social Systems inherited from the post war era failed to account for the home and care work done by women at the time when women were becoming actors in the labour market. Hence fertility rates have moved below the renewal level in most Member States as parents cannot afford and/or cope with having children, while labour participation of women still lags at the edge of the labour market.

One of the main challenges for today's societies is to find quality and affordable solutions to the new 'caring' needs faced by families. Answers to demand for childcare when children have become rare and care of the elderly in a rapidly ageing society are the two wheels to reinvent: we need more children, educated to perform in a knowledge society. These children will keep the machine going for, inter alia, a very old baby boom generation with high caring needs.

In that context it is worthwhile mentioning that any remedy that is being put in place after early childhood tends to be much more expensive than early childhood policies (Heckman & Lochner 2000). In several Member States childhood poverty is rising. Female employment is the key to mitigate child poverty (Esping-Andersen 2002).

The conclusion therefore is that while welfare states do a good job in meeting the risks that they were designed for, they are much less well attuned to those of today. In addition, they are increasingly failing to meet the challenge of social exclusion, which is dominated by changes in the demographic composition of the population, of the particular needs of particular segments and a jobs market which, in many countries, is very far from full employment.

Trends Reinforce the Need for Change[6]

Above we showed that current social systems are built on social risks that are in part obsolete. Several trends reinforce the need for modernisation of social systems.

Globalisation is the frame of reference that defines the modern purpose of the European Union. It focuses attention on critical issues such as trade policy, the rise of Asia, and climate change, as well as how in a world of mass migration, cross border crime and terrorism, Europe can combine openness with security. These are issues where the European Union has a key role to play if Europe's response is to be effective.

The European Union has a long experience in dealing with these types of challenges. Besides economic policies the EU has made a substantial difference to the European 'quality of life'. The EU's social commitments to employment rights, social dialogue, gender equality and action against discrimination together with the leading role it has played on environmental questions and consumer protection are both consequence and cause of social transformation in Europe. The EU has consolidated democracy across Europe through enlargement, as dramatic a transformation in our time as Franco-German reconciliation for a previous generation. It has greatly expanded the scope of individual freedom to live, work and travel.

Yet, several trends add to a need to step up a gear. Europe's rapid trajectory to a post-industrial knowledge and service economy is transforming the nature of work and social class divisions as well as the conditions of access to economic opportunity, the extent of social mobility and the incidence of poverty and inequality.

The achievement of mass affluence and the process of economic modernisation have profoundly shaped values, as we see in the trend to individualisation, to the new patterns of family life, in the changing position of women in society, the rise of post-material concerns and the new challenges of democratic engagement. But this individualisation often goes side by side with an increased yearning for a more socially cohesive and socially responsible society, as well as with citizens' increased awareness and concern about quality of life issues such as clean air, clean water and environmental security which are seen as fundamentals that matter as much, if not more to people than material affluence.

Radical demographic change is seen in declining fertility and longer life expectancy raising issues of social and economic sustainability, new and diverse housing demands, work life balance in dual earner households, gender equality in who shares the burden of care and equity between the generations.

The rise of the citizen as consumer is changing the way we think about issues such as choice in housing and public services, individual responsibility for one's health, the management of waste and recycling, food safety and quality, while other concerns such as ethical consumption and community empowerment emerge. There is no evidence of a lessening of citizens' concerns with public issues, but participation and trust in traditional forms of politics and citizen engagement is generally in decline.

The development of the welfare state, itself a social and political response to the industrial age, has in turn reshaped life opportunities for tens of millions in our societies, as well as arguably changing incentives and creating new dependencies from which individuals can find it difficult to escape. Welfare states have been shown to contain complex dynamics of their own that are slow to work through our societies and difficult to change.

Policy Consequences for the Need for Adapting to Change

Above we showed that economic policies and social policies could often be designed to reinforce each other. We showed that European social systems need modernisation, and existing trends reinforce this need. As said, all these elements combine to the need for people to adapt to change. In terms of policies at the European level the most important reactions to the need for change are the following:

Mid-term Review of Lisbon

The European Council in March 2008 re-launched the Lisbon strategy in refocusing on growth and jobs. The Commission's mid-term review had three main objectives: i) more focus with ''rigorous prioritisation": the Commission proposes to focus on economic growth and employment by launching the idea of a ''Partnership for Growth and Jobs", which would be supported by an action plan at Union level and national action plans in the Member States; ii) mobilise support for reforms: national Member States, social partners and even citizens need to take 'ownership' of the necessary reform processes. Lisbon "must become part of national political debate"; iii) simplification, clarification and simpler reporting: instead of the myriad reports, there will be a single Lisbon report at EU level and one at national level.

There are new guidelines for 2005-2008. For 2005-2008, there are three priorities: 1) Attract more people on the labour market and modernise social protection system; 2) improve the adaptability of workers and firms and increase flexibility of the labour market; 3) invest more in human capital by improving educations and qualifications. They offer a consistent framework for reforms. Each Member State has to present a country specific report to identify concrete measure that should be taken into account to foster growth and jobs. One of the consequences is that the guidelines are reshaped. All these guidelines are monitored by quantitative indicators.

In recent years, there have been very favourable developments in terms of employment, unemployment, and specific issues such as employment of women and older workers.[7] The key question is to what extent these favourable developments can be attributed to the business cycle. The coming years will prove whether structural reforms have made European labour markets more resilient to less favourable general economic conditions.

Social Agenda

As a complement to the mid-term review, the Commission has committed itself to the modernisation of the European social model(s) and to the promotion of social cohesion.[8] One of the main elements of this five-year strategy plan is to strengthen citizens' confidence in accepting the changes needed to cope with new challenges such as increased global competition and Europe's ageing population. The Commission sees three 'key conditions for success': an intergenerational approach with a strong emphasis on youth, a partnership for change with social partners and civil society, and a strong focus on the social dimension of globalisation. The communication sets out two priority areas: i) achieving full employment; ii) a more cohesive society with equal opportunities for all. In all these inter-related domains there is an important role of monitoring the progress that involves a bundles of quantitative indicators. In a recent package, the European

Commission spells out its social vision, for a rapidly changing 21st century Europe.[9] It sets out how in a globalised age, all Europe's citizens must have access to resources improving their "life chances" and enabling them to share in rising prosperity. There is a need of investment in a number of areas: youth; career opportunities; longer and healthier lives; gender equality; inclusion and non-discrimination; mobility and integration; culture, participation and dialogue.

Migration

The EU needs to deal with migration in the overall socio-economic context of Europe with the trends as underlined above. At the same time efforts aimed at integration of migrants need to be stepped up to ensure peaceful and stable societies, while the EU must also react to increasing migration pressure of illegal migrants trying to make their way into Europe, not only from neighbouring countries but increasingly from countries further afield.

The so-called 'comprehensive EU migration policy', as defined by the European Council, provides a coherent and efficient manner to respond to the challenges and opportunities related to migration. It builds on the conclusions of the Tampere European Council in 1999, the Hague Programme of 2004 and the Global Approach to Migration adopted by the European Councils in 2005 and 2006.

The comprehensive approach deals with many different aspects of migration policy, each with a different role of the EU. It deals with negotiations with third countries, asylum policy, integration, common border patrols, fighting illegal employment and circular migration.

Other Policy Areas

There are many areas where the European Commission is not the main responsible policy level, but where nevertheless important initiatives have been taken. In particular on education, *e.g.* the Bologna Process and the Erasmus programs stand out as triggers for reform and mobility, and health where the EU has been active in transmitting best practices on how to deal with diseases such as obesity, alcohol and mental health. Also in the area of employment the EU has been active in areas such as health and safety at work, and more recently on flexicurity.[10]

While these EU policies are pertinent, still most of social policies are Member State or local responsibilities. The next section will elaborate on this latter issue in more detail.

Responsibilities of Governments Levels and Other Stakeholders

Member States-European Union

Most of the policy areas previously mentioned to support/promote citizen's responsibility are, however, in the hand of the Member States. This said, that does not mean that EU does not have an important role to play. For instance, in re-launching Lisbon agenda, the Commission has proposed to focus on economic growth and employment by promoting the idea of a 'Partnership for Growth and Jobs', supported by an action plan at Union level and national action plans in the Member States.

As said the Member States have a clear responsibility for the design of their labour market institutions and other aspects of their social model. Europe has a responsibility essentially for product and financial market policies.

European economic policies can only be successful if the social models have the capacity to absorb the consequences of economic policies such as corporate restructuring. The reverse is also true. Member States have been active, some more than other, in modernising their social systems. But it is of no use to push large quantities of benefit recipients to the labour market if there are no jobs. People in difficult circumstances do not like to play games of musical chairs with their future as the stake.

Innovation, internal market and trade policies are needed to create new opportunities for start-ups or FDI. Competition policy helps by not favouring the incumbents. So for economic policies to be effective job markets need to be sufficiently flexible. But for labour market reform to be successful economic policies should be in place.

The symbiotic relationship between economic and social policies creates a subsidiarity trap. One cannot live without the other, but decisions are not made on the same level. This does not mean that all decision power should be shifted to one or the other.

In a happy marriage both partners will find ways to reconcile conflicting interests, sometimes it implies that ones gives in sometimes the other. A strong marriage is vital when the going gets tough. The Member States and the Commission must realise that they are in this together and that only a strong marriage can succeed in reconciling economic, financial and social goals.

However, the marriage of tomorrow will not be the same as the marriage of today. In the post war period the male breadwinner brought in the money, while the women took care of the house and children. Today, the world is more complicated. Women pursue careers, such that housekeeping and children raising has become a shared responsibility. Shared responsibilities are more complicated since it requires coordination and communication.

For a long time the Commission was – so to speak - in a position similar of that of the male breadwinner in a couple. Internal markets and competition law created prosperity, and the role of the Commission was clear. The Member States took care of their house and children. They redistributed the wealth that was created to make sure that everybody could profit from it. So what happened?

In a gradual process the male breadwinner role of the Commission has been put under pressure. It had to undertake caring activities such as setting labour standards. The actions by the Commission to generate wealth became more demanding for the Member States, who have to permanently adapt their households to make the wealth creation effective. Meanwhile the children of the Member States became more demanding. There were some small fights between the partners and the children were blaming the father from becoming too dominant. The father responded by ignoring the children and telling the wife to get the house done.

At some stage the partners have to come to terms with this. They need to reallocate some of their tasks to maintain a happy marriage. The open method of coordination (OMC) was a heroic attempt to avoid the subsidiarity trap. OMC is a specific tool that helps to reconciliate Member State competences and the fact that there is a need of exchange of experiences and need of convergence of objectives in Europe. The method

consists of five elements: 1) Agreeing on common objectives for the Union; 2) establishing common indicators as a means of comparing best practice and measuring progress; 3) translating the EU objectives into national/regional policies on the basis of national reports in strategies for social protection and social inclusion; 4) publishing reports analysing and assessing the National Reports; 5) Establishing a Community Action Programme to promote policy cooperation and transnational exchange of learning and good practice. This process helps Member States to benefit from experiences of others. However this process and the way the EU is leading it, are not without shortcomings.[11] In particular, one can note that the process that consists in converging objectives are not always known or shared by the various stakeholders who are in the meantime responsible for action, and the problem of citizens lacking ownership of the reform process has been identified by many as one of the obstacles to the process itself (Sapir & Pisani-Ferry 2006).[12] This could partly explain that at the beginning, the Lisbon agenda did not seem to deliver the expected results.

The constitution was another one attempt to get to grips with coherence across policy layers. But neither open methods of coordination nor constitutions will be sufficient to avoid future competence clashes on the economic and social policies. There are elements of these clashes already present today and there will be more tomorrow and even more the day after tomorrow. Questions are more easily posed than answered. But what is clear is that a lacking European narrative – already underlined in the Kok report, could also be attributed to a scarce dialogue between institutions and large segments of the EU society and the question whether citizen participation could enhance such a dialogue is one of the motivations of the present work.

Citizens and State Responsibilities

Globalisation, modernisation of social models and ongoing trends imply that people need to adapt to change and that we do the most with human resources. This requires citizens to assume ownership of problems and to foster their responsibility, as well as policies have to be designed to facilitate that. A number of examples can be mentioned:

- Citizens are responsible for updating their human capital at all stages of their life. Governments are responsible for designing an education and training system that allows a smooth and efficient upgrading with a specific focus on a strong start. The nature of the labour market and the service economy will make such investment more important in the future.

- Citizens are responsible for investing in good health by appropriate food and physical activities. Governments are responsible for ensuring Member States have high quality and affordable health care systems as well as healthy environment;

- Citizens are responsible for their own social lives and for treating fellow citizens in a respectful way. Governments are responsible for facilitating social inclusion and providing adequate housing for lower income groups.

- Citizens are responsible for raising adequately their children. Governments are responsible for ensuring an adequate system of child care is put in place.

- Migrants are responsible for integrating in the host societies. Host societies are responsible for treating migrants on the same footing as their native citizens and to promote equal opportunity in education.

In this perspective, the challenge concerns the setting up of strategies that would enable people to take their responsibilities and seize opportunities, but would also give second chances when needed.

To conclude, realigning social and economic policies is a real need that creates challenges for all levels of policy making and also at the individual level. It also creates a perceived democratic gap. Often decisions are made at a central level in a diverse Europe of 27. National politician do not always have the right incentives to defend Europe level policies in their own country. The European Commission does not find it easy to fill this gap. Therefore civic participation is an important step to reduce the perceived democratic gap. Hence this is the topic for our case study in the next sections.

A Renewed Role for Individual Participation

As detailed in the previous sketch, the success of a paradigm by which European Citizens 'adjust to change' rests on a series of structural solutions to be implemented on the European social model. In structural terms, the countries must design policies that combine flexibility of the labour market with security for the individuals (protecting the person not the job). However, these designs must be referred to a social woven dependant on a population whose demographic composition is rapidly changing. The situation is also complicated by new and renewed challenges – social exclusion, unemployment, security, etc. – that pose a further burden on the evolving dimension of a sustainable 'European quality of life'.

Alternative Focuses

Traditional welfare states are not well equipped in general for facing these challenges. Hence, 'innovative social policies' tend to delegate power for local implementation to collaborative public-private arrangements and even to wholly private actors. We see more and more decentralised co-ordination. This applies also to policy experiments at the EU level (Esping-Andersen *et al.*). *Vertical subsidiarity* becomes the necessary tool for *horizontal subsidiarity*, as it opens the public service sector to the market, civil society and social entrepreneurship.[13] In social policies developed along these lines, the voice of the citizen needs to be heard and taken into account at the outset of the public-value creation process and not merely when protests or complaints are made. This means ensuring that, via the appropriate social channels and institutions, citizens have access to and be part of the policy design and implementation.

As indicated before, citizens' concerns with public issues have always been there, but the traditional forms of participation and policy engagement show signs of fatigue. This suggests that individuals are more and more inclined to use alternative architectures – such as cohesive local networks (civil society, private sector) and global networks (international organisations and NGOs) – for advocating their needs. As Dobell and Mitchell (1997, p. 27-28) suggest, "analytic formulation of policy within an agency has been replaced to a considerable extent by organic, unpredictable and often highly conflictual public multi-stakeholder processes of formation of policy, not only through executive or legislative stages of policy adoption, but through multi-agency and private

sector partnerships for implementation and co-management." Issues such as the adoption of genetically modified food, nuclear power, liberalisation of services industry and reform of the common agricultural policy are all examples of settings where private stakeholders often set the agenda within a polarised context which policy must strive to mediate.

For optimising their responsiveness, the institutions, on one side, must learn *how* these structures are built and interact and, on the other side, must understand *who* the individuals that participate in the process are.

Individual Participation and Contribution to Social Capital

The scope of our analysis is to identify the characteristics of the shares of individuals that are more likely to be involved with different kinds of formal organisations. In particular, to investigate how each of these groups perceives the political, economic, and relational spheres of social reality.

As Putnam (1993, 2000) and other researchers (Woolcock 2001; Adler & Kwon 2002; van Oorschot & Arts 2005, for effects on the European welfare states) suggest, not all participation is conducive to positive results in terms of attentiveness to social and economic challenges. The classical Putnam's distinction between *bridging* (or inclusive) and *bonding* (or exclusive) social capital suggests that:

> *'Some forms of social capital are, by choice or necessity, inward looking and tend to reinforce exclusive identities and homogeneous groups. Examples of bonding social capital include ethnic fraternal organizations, church-based women's reading groups, and fashionable country clubs. Other networks are outward looking and encompass people across diverse social cleavages. Examples of bridging social capital include the civil rights movement, many youth service groups, and ecumenical religious organizations' (Putnam 2000, p. 22).*

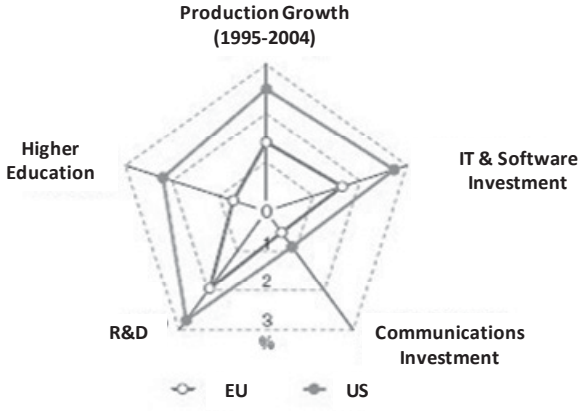
Previous work (*e.g.* Beugelsdijk & Smulders, 2003; Sabatini, 2005; Beyerlein & Hipp, 2005) focuses on the distinction between *bridging* and *bonding* social capital, eventually to categorise the various networks and organisations to which people belong. However, as Woolcock (2000, p. 80) suggests '[t]he policy response to reading the social capital literature should not be a call for more choirs and soccer clubs, as writers satirising Putnam have tended to infer.' In fact, as suggested earlier, we are interested in the characteristics and development of the European social reality as a whole. Specifically, we are trying to understand what the determinants of social involvement are and whether involvement can enhance the results of the coordination process between State and stakeholders in an increasingly diverse society.

A considerable body of literature (see for example: Dasgupta 2003; Völker & Derk 2004) highlights the positive link between *bridging* and *linkage* capital and economic prosperity in the sense that 'bridging social ties (sometimes called 'weak' ties) are more likely to be drivers of economic growth than bonding social ties' (Putnam 2002). One possible explanation of such parallelisms is given by Sapir and highlights the need of the 'adaptation to change paradigm' proposed in Section "A New Paradigm: Adjusting to Change":

> *'Instead of fostering the necessary adaptation and flexible responses to increasingly rapid changes, modern European welfare states … [] now often protect the status quo. And as James Heckman (the Nobel Prize winner) rightly states in his insightful analysis of Europe, 'The opportunity cost of security and*

preservation of the status quo – whether it is the status quo technology, the status quo trading partner, or the status quo job – has risen greatly in recent times' (Heckman 2002).

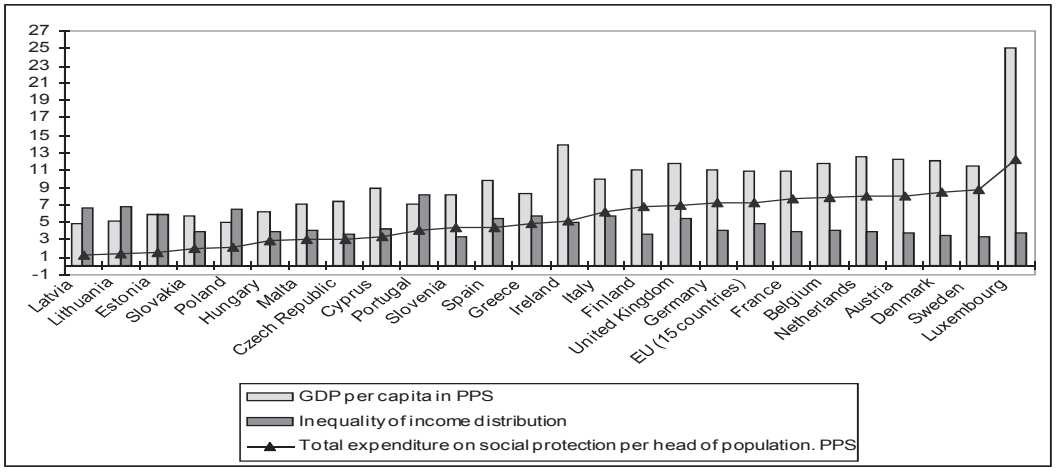
Figure 8.1 Investments in Various Forms of Knowledge and Productivity Growth (1995-2004): Comparison between EU and United States



Source: Indepen, 2006

Indeed, in a comparison with the United States (Figure 8.1), Europe as a whole lags behind in all aspects of investment in R&D, Communications, IT & Software, and Higher Education. The result is a substantial slowness in productivity growth (1995-2004). Hence, recalling the initial argument, this result (depicted in Figure 8.1) could suggest a predominance of bonding ties that anchor the EU to maintain the status-quo, rather than invest in equipping for change.

Figure 8.2 GDP per capita, Social Protection and Inequality of Income Distribution: An European Overview (2004-2005)

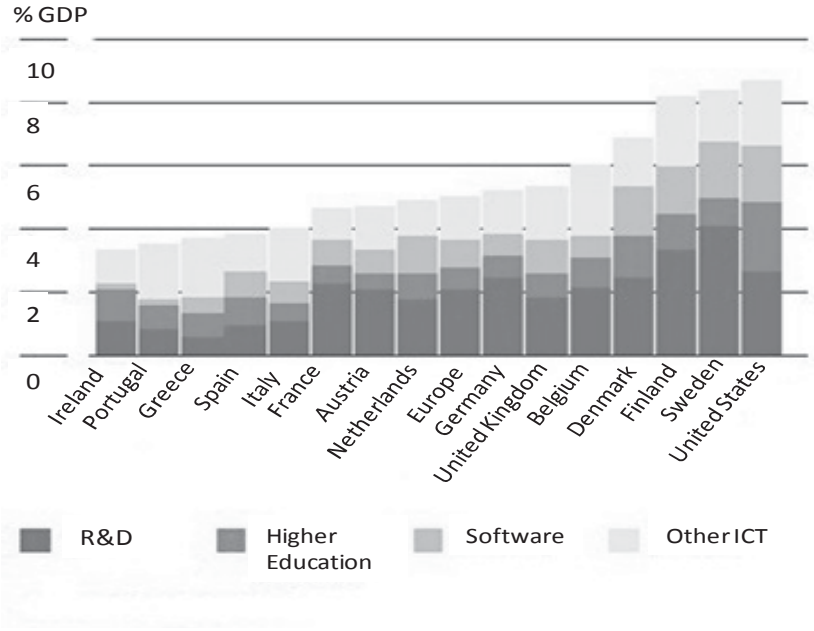


Source: JRC elaboration of Eurostat data.

However, if we break-up these effects at country level, we can notice that the intensity of the investments tends to co-vary with welfare state characteristics, especially on a North-South axis. For instance, high welfare spending in the Scandinavian countries goes hand in hand with a relatively higher level of wealth (GDP), a relatively smaller income inequality. Low welfare spending in the Mediterranean countries goes together with a lower level of wealth, a larger income inequality (Figure 8.2).

The picture does not change when looking at investments in R&D, Higher Education, and ICT (Figure 8.3) whereby Finland and Sweden match closely the US figure of 8% of GDP invested in the afore specified chapters with a specific focus on R&D. Southern European countries, on the other hand stop half way through at about 4%.

Figure 8.3 Investment in Knowledge: An European Overview (2005)



Source: OECD & GGDC (2006)

Referring again to the initial paradigm, these figures make us wonder whether there is any parallelism between such diversified investment strategies and the prevalent type of social capital in the various European countries.

As Liddle and Lerais (2007) suggest, however, Europe's social reality is a rather fragmented picture. Many people now 'see their lives in terms of a personal biography they write for themselves. This is not to say that people are no longer concerned about 'social cohesion': on the contrary surveys suggest that many are, but many have no desire to return to the old conformities of family, class or religion.' Social cohesion, individualism, participation, acceptance, disinterest; making sense of all these apparently diverse indicators requires a 'dashboard' that recovers basic message behind all these signals. One such measure is the Active Citizenship Composite Indicator (ACCI). This aggregated index has been built by the European Commission – Centre for Research on Lifelong Learning to measure the European active and democratic citizenship, defined as:

'Participation in civil society, community and/or political life, characterized by mutual respect and non-violence and in accordance with human rights and democracy' (Hoskins 2006).

The Active Citizenship Composite Indicator covers 19 European countries and is based on a list of 63 basic indicators for which the data has been principally drawn from the European Social Survey of 2002. This composite indicator shows a heterogeneous Europe where Nordic countries lead and southern European countries perform well in Values and Political Life but lag behind in Civil Society and Community Life dimensions (Hoskins *et al.* 2006, Figure 8.4).[14]

Figure 8.4 The Active Citizenship Composite Indicator

Source: Measuring Active Citizenship in Europe, Hoskins *et al.* 2006

The overall picture produced by the indicator supports the hypothesis by which the intensity of change follows a North-South axis. However, the indicator cannot tell us much with respect to the *role of the different types of social capital* for the progress of European societies. In fact, as depicted in Figure 8.4 the structure of the indicator considers positively *all* kinds of participation. Still, as suggested earlier, the prevalence of different forms of participation may imply a trade-off between behaviours aimed at maintaining the status quo or at fostering progress. Hence, it is necessary to investigate whether such trade off does exist and whether the behaviour follows a mere North-South axis, or whether more articulated structures are required.

Participation and European Social Models

Box 8.1 The European Social Survey

Favouring the fine-tuning of the process of institutional responsiveness requires identifying the objective characteristics and the perceptions of the actively engaged individuals.

To achieve this aim, we investigated the European Social Survey (http://www.europeansocialsurvey.org/), which ran a specific module on citizenship in 2002. This data is more up-to-date than that which is available from alternative sources, such as the World Values Survey and IEA's CIVED, which is currently only available from 1999.

The European Social Survey (the ESS) is an academically-driven social survey designed to chart and explain the interaction between Europe's changing institutions and the attitudes, the beliefs, and the behavioural patterns of its diverse populations. The survey is administered every two years (the fourth wave is scheduled for 2008) and consists in a core questionnaire and two rotating modules. In the first wave of 2002 one of these modules was on citizenship; such data was not available in the 2004 issue and the 2006 data hasn't been disclosed yet.

The European Social Survey (ESS) aims to be representative of all residents among the population aged 15 years and above in each participating country. The size and the quality of the sample make the European country coverage in the ESS data reasonably good, with 18 EU Member States providing data of sufficient quality. Among the 18 investigated Member States, Austria, France, Hungary, and Ireland have been excluded from the analysis because some variables were completely missing in the dataset as some questions were not asked in the national versions of the questionnaire. Although data on Poland has revealed anomalous patterns (see Section 4.6) we have decided to keep the country in the dataset for the analysis.

The dataset under investigation is therefore composed by 26.491 observations. The sample is representative at the national level for the 14 European countries considered, which are: Belgium, Germany, Denmark, Spain, Finland, United Kingdom, Greece, Italy, Luxemburg, the Netherlands, Poland, Portugal, Slovenia and Sweden. The number of observations was subsequently reduced to 24.023 to exclude people below 18 or above 80 years of age.

The Various Ways of Participation

In the European Social Survey 2002, individuals were asked about their *ways of participating* to several types of organisations (sport, cultural, trade union, business, human rights, environment/peace, religious, politics, social, teacher/parents organisations), and they could choose among four different behaviours (membership, donating money, participation and voluntary work) recorded as dichotomous variables (values 0/1). For each type of organisation, the *strength* of individual participation amounts to the sum of the scores that the individual has obtained with respect to the foresaid possible behaviours.

Specifically, we created a new variable equal to the sum of the four binary variables describing the possible action of engagements in that organisation

$$Y_h = \sum_{i=1}^{4} X_{h,i}$$

where $h=1..10,$ is the type of organisation and $i=1..4,$ are the four different ways of engagement. Each variable can assume a value from 0 (no action taken) to 4 (the person is engaged in all possible ways).

On the basis of the definitions of *bridging* and *bonding* social capital, we divided individual participation into two categories:

- Social engagement, individuals participating to organisations that are outward looking and aim at improving the society at large (Cultural, Human Rights, Social, Religious, Environmental/Peace organisation).

- Private engagement, encompasses the organisations that work closer to the private interest of the respondent (Sport, Trade Union, Business, Teacher/Parents, Political Party).

A factor analysis using polychoric correlations for ordered-category data[15] was carried out to validate the consistency of this theoretical grouping. The results of the analysis confirm the hypothesis of the existence of two main groups (social engagement and private engagement) and explain approximately 40% of the variance.

The two groups were used as the basis for constructing two variables – SOCIAL ENGAGEMENT and PRIVATE ENGAGEMENT – given respectively by the aggregation of the variables reporting higher factor loadings in factor 1 (social engagement) and in factor 2 (private engagement), respectively. On the basis of the 10 categories of organisations previously identified (sport, cultural, trade union, business, human rights, environment/peace, religious, politics, social, teacher/parents organisations), the variables *social engagement* and *private engagement* ended up having the following structure:

$$\text{SOCIAL ENGAGEMENT} = Y_{Cultural} + Y_{Social} + Y_{Env/Peace}\ Y_{HumanRights} + Y_{Religious}$$

$$\text{PRIVATE ENGAGEMENT} = Y_{Sport} + Y_{TradeUnion} + Y_{Business} + Y_{Teacher/Parents} + Y_{Political}$$

Subsequently, we created the variable ENGAGEMENT that will constitute the basis of our analysis. ENGAGEMENT is a multinomial variable that takes values from 0 to 3 defined as follows:

ENGAGEMENT = 0 - individuals not participating in any organisation;

ENGAGEMENT = 1 - individuals participating just in social organisations;

ENGAGEMENT = 2 - individuals participating just in private organisations;

ENGAGEMENT = 3 - individuals engaged in both social and private organisations.

To characterise the factors that can determine the decision of an individual to participate in formal organisations, the variable ENGAGEMENT was considered as the dependent variable, and a set of objective variables reported in the European Social Survey was included as explanatory variables. The complete list of the explanatory variables considered is reported in Table 8.2.

Explanation of the Various Patterns of Participation

The variable objects of the analysis are qualitative, hence the application of classical regression models (OLS) is not allowed due to the violation of the basic assumptions underlying the model. Modelling qualitative variables with two or more categories requires the use of models based on different assumptions that can produce reliable estimates and allow for the correct application of standard statistical techniques. If the adoption of Logistic Regression could solve the problem of modelling dichotomous variables, in the case of variables that can assume more than two categories the use of the Ordinal Logit or the Multinomial Logistic regression model can face the challenge of providing correct and reliable estimation of the parameters.

Table 8.2 List of Explanatory Variables

Type	Description
Family Income	Family income categorised in 12 leves following the ESS scale: Less than 1 800; 1 800-3 600; 3 600 – 6 000; 6 000 – 12 000; 12 000 – 18 000; 18 000 – 24 000; 24 000 – 30 000; 30 000 – 36 000 – 36 000 – 60 000; 60 000 – 90 000; 90 000 – 120 000; more than 120 000. The last category "'120 000 or "more" has been choosen as reference.
Age	Age of the respondent divided in class: 18 - 24; 24 -35; 35 - 44; 44 - 55; 55 - 65; 65 or more. The last category "'65 or "more" has been choosen as reference.
Individual Human Capital	expressed as years of education.
Sex	0 = Male; 1 = Female
Citizenship	0 = No; 1 = Citizen of the country where interviewed
Number of persons	Size of the Family
Declared Religious	0 = No; 1 = Yes
Place where the respondent lives	Place where the respondent lives categorised in classes following the ESS scale: 0 = A big city; 1 = The suburbs of a big city; 2 = A town or a small city; 3 = A country village; 4 = A farm or a home in the countryside. The category "' 0 = A big "city" has been chooses as reference.
Time spent in watching TV	Time spent during the day in watching TV following the ESS scale: 0 = No time at all; 1 = Less than 1/2 hour; 2 = 0.5 -1 hour; 1 - 1.5 hour; 1.5 - 2 hours; 2 -2.5 hours; 2.5 - 3 hours;more than 3 hours.
Time spent in listening to the Radio	Time spent during the day in listening to the radio, following the ESS scale: 0 = No time at all; 1 = Less than 1/2 hour; 2 = 0.5 -1 hour; 1 - 1.5 hour; 1.5 - 2 hours; 2 -2.5 hours; 2.5 - 3 hours;more than 3 hours.
Time spent in reading newspapers	Time spent during the day in reading newspapers, following the ESS scale: 0 = No time at all; 1 = Less than 1/2 hour; 2 = 0.5 -1 hour; 1 - 1.5 hour; 1.5 - 2 hours; 2 -2.5 hours; 2.5 - 3 hours;more than 3 hours.
Time spent in surfing the Internet	Time spent during the day in surfing the internet following the ESS scale: 0 = No time at all; 1 = Less than 1/2 hour; 2 = 0.5 -1 hour; 1 - 1.5 hour; 1.5 - 2 hours; 2 -2.5 hours; 2.5 - 3 hours;more than 3 hours.
Main Activity	Main Activity performed by the respondent during the seven days before the interview, coded following the ESS scale: 1 = paid work; 2 = In Education; 3 = unemployed and actively looking for a job; 4 = unemployed but not actively looking for a job; 5 = permanent sick or disable; 6= retired; 7 = in community or military service; 8 = doing houseworks; 9 = other. The category "' 1 = in paid "work" has been choosen has reference.
Country Dummies	A set of dummy variables representing the country effect was included in the model

The model we decided to adopt to analyse the variable ENGAGEMENT is the Multinomial Logit (mLogit). The mLogit is a straightforward extension of the classical Logistic model and it is a strategy often used in the literature when categories are unordered (as the case of the variable ENGAGEMENT). In the mLogit model a category (Non Participant, in our case) is designated as reference category. The probability of membership in other categories (Social, Private and Full Engagement) is compared to the probability of membership in the reference category. In general for a dependent variable with M categories, the mLogit requires the calculation of M-1 equations, one for each category relative to the reference category to describe the relationship between the dependent variable and the explanatory variables. In our case, three different equations have been computed.

Although possible violations of the Independence from Irrelevant Alternatives (IIA) assumptions need to be further investigated, similar results have been obtained by repeating the analysis and applying several logistic regressions that perform pairwise comparisons between the reference category and the other categories.

The Clustering: Partition of Countries

We find a similar partition to many authors[16]but we group Anglo Saxon and Continental countries because Great Britain was left alone, and we include an extra group – Eastern – composed by Poland and Slovenia, which were not considered in other taxonomies (Table 8.3).

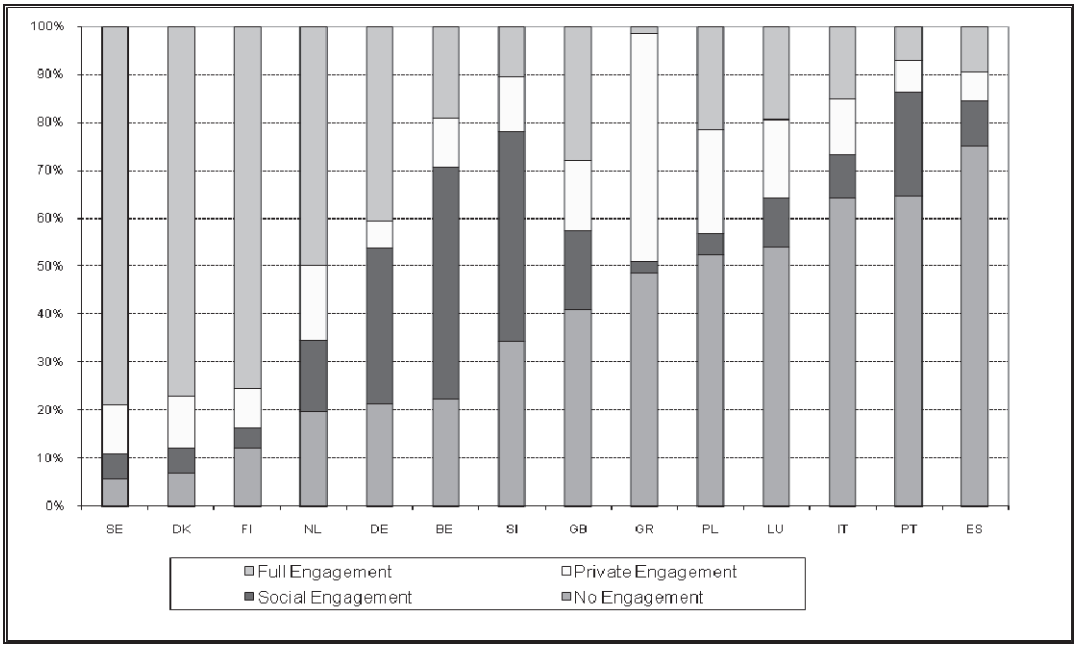
Table 8.3 Clusters of Countries

Cluster	Countries
1	Denmark, Sweden, Finland, The Netherlands
2	Belgium, Germany, Luxemburg, Great Britain
3	Greece, Italy, Portugal, Spain
4	Poland, Slovenia

Full Engagement: More than the Sum of the Parts

An initial analysis of the participatory patterns at country level (Figure 8.5) suggests that individuals in all Nordic countries are preponderantly engaged in both private and social organisations, while more than 50% of the individuals in Mediterranean countries – with the exception of Greece – are not engaged. Greece is an interesting outlier because it shows a strikingly high 45% of people that are involved in private organisations, a result partly due to the peculiar status of Greek business organisations.[17] In Continental and Anglo-Saxon countries the rates of participation are generally above 50%; Slovenia, Germany, and Belgium show high percentages of people involved solely with social activities.

Figure 8.5 Engagement: How Is It Distributed among European Countries



Source: JRC elaboration of ESS 2002 data

Various authors have criticised the use of micro-data to account for the behaviour of social structures that exist only at aggregate level such as social capital.[18] Although such critiques are legitimate, the variables derived from the ESS dataset provide valuable information in understanding the *character* and *nature* of the social connections in a place and the identification and strengthening of the social channels through which citizens have access to those institutional conditions that will allow them to be a part of the intervention programming process and of the ensuing strategy adoption phase.

Specifically, if the dimensional identification is correct, the level and type of participation should reveal a driver of Europe's social reality whereby social and full engagement (which correspond to bridging and linking social capital) should be correlated with positive social and economic dynamics. The preponderance of private engagement could, on the other hand, be correlated with a more stagnant situation because '[b]onding social capital is, as Xavier de Souza Briggs puts it, good for 'getting by' … [*while*] bridging social capital is crucial for 'getting' ahead" (Putnam 2000, p. 22).

If we sketch the situation at country level with respect to the trade off between social engagement and private engagement, we are presented with an interesting situation whereby we register a clear polarisation between Nordic countries (most fully engaged) and Mediterranean countries (most non participants), while the other countries present trade-offs in the patterns of participation.

When considering the variable ENGAGEMENT as the dependent variable and the variables listed above as explanatory variables, the analysis of the results the mLogit shows an interesting picture, highlighting different profiles of the participants. The result of the analysis and the complete list of the coefficients are shown in Table 8.4.

As suggested earlier, we submit that the *social engagement* groups are *bridging* and bring citizens into contact with people from a cross-section of society, while the *private engagement* organisations are *bonding* and tend to aggregate people who have mainly the same background. Consistently, the depicted profiles indicate that the *income level* and the *main activity of the previous week* are not relevant explanatory variables for social participation, while it's highly so for *private participation*. More specifically, lower income and non-active participation to the job market are associated with a substantial reduction of the probability of being privately engaged. The age and the place of residence have opposite effects whereby older people living in suburbs are more likely to participate in social, and working-age individuals living in small cities are more likely to participate in private groups. In both cases, males are more likely to participate than women, but this distinction is much more evident with respect to private participation and it is likely to indicate a *bias* due to a non-yet-closed gender gap. Level of individual human capital (proxied by the years of education), access to the media – and in particular listening to the radio – and being natives of the specific country of reference increases chances of participation in both types of organisations. The same is true for religious affiliation, although this characteristic has a much stronger effect on social participation.

As expected, these profiles show consistent differences between the two populations, but present also some relevant similarities. Direct analysis of the profile of the fully engaged individuals allows to further elaborating on this issue.

As reported in the Table 8.4, individuals that are fully engaged (participate both in social and private organisations) have a peculiar profile that seems to be a composition of the previous two identikits. First, income plays a big role as the probability of being fully engaged decreases for the individuals with a yearly family income below 30 000 euros.

Table 8.4 Coefficients of the Model for the Whole Europe

Depedent Variable: ENGAGEMENT	Comparison between Non Engaged and Socially Engaged		Comparison between Non Engaged and Privatelly Engaged		Comparison between Non Engaged and Fully Engaged	
Explanatory Variables	Coef.	P>t	Coef.	P>t	Coef.	P>t
Income: Less than 1 800Euros	0.703	0.511	**-2.670**	**0.001**	**-2.527**	**0.004**
Income: Less than 3 600Euros	0.786	0.444	**-1.939**	**0.001**	**-2.882**	**0.000**
Income: Less than 6 000Euros	0.618	0.547	-0.351	0.529	**-0.496**	**0.350**
Income: Less than 12 000Euros	1.141	0.256	**-1.529**	**0.007**	**-1.754**	**0.001**
Income: Less than 18 000Euros	1.878	0.061	**-1.399**	**0.012**	**-1.587**	**0.002**
Income: Less than 24 000Euros	1.783	0.076	-0.723	0.190	-1.005	0.052
Income: Less than 30 000Euros	0.884	0.384	-1.021	0.069	**-1.045**	**0.043**
Income: Less than 36 000Euros	1.240	0.224	-1.140	0.440	-0.800	0.124
Income: Less than 60 000Euros	1.274	0.208	-0.601	0.283	-0.434	0.398
Income: Less than 90 000Euros	1.470	0.156	-0.217	0.714	-0.389	0.478
Income: Less than 120 000Euros	0.690	0.602	-0.936	0.157	-0.694	0.280
Age (18-24)	**-1.997**	**0.000**	-0.333	0.313	**0.712**	**0.010**
Age (24-35)	**-1.940**	**0.000**	-0.181	0.539	**0.875**	**0.000**
Age (35-44)	**-1.726**	**0.000**	**1.001**	**0.000**	**1.228**	**0.000**
Age (44-55)	**-1.391**	**0.000**	0.340	0.223	**1.231**	**0.000**
Age (55-65)	**-0.960**	**0.000**	**0.648**	**0.011**	**1.600**	**0.000**
Sex	**-0.389**	**0.000**	**-0.815**	**0.000**	**-0.741**	**0.000**
Watching TV	**-0.113**	**0.000**	0.035	0.153	**-0.253**	**0.000**
Listening to the Radio	**0.060**	**0.001**	**0.048**	**0.006**	**0.077**	**0.000**
Reading Newspaper	**-0.268**	**0.000**	**0.174**	**0.000**	**0.177**	**0.000**
Surfing the Web	-0.006	0.779	**0.042**	**0.000**	**0.042**	**0.000**
Individual Human Capital	**0.196**	**0.000**	**0.098**	**0.000**	**0.178**	**0.000**
Number of persons	0.060	0.145	0.057	0.132	**-0.181**	**0.000**
Residence: suburbs of a big city	**0.506**	**0.007**	**-0.495**	**0.002**	**0.447**	**0.009**
Residence: small town	**0.630**	**0.000**	**-0.625**	**0.000**	**0.492**	**0.000**
Residence: country village	0.005	0.979	**-0.862**	**0.000**	**0.397**	**0.006**
Residence: home in the countryside	0.118	0.735	**-0.634**	**0.041**	0.406	0.116
Citizenship	**0.672**	**0.003**	**0.610**	**0.004**	**0.669**	**0.000**
Declared Religious	**0.946**	**0.000**	**0.348**	**0.004**	**1.027**	**0.000**
Main Activity: In education	0.548	0.057	-0.059	0.834	0.227	0.382
Main Activity: Unemployed, looking for job	0.042	0.877	**-1.159**	**0.000**	**-0.981**	**0.000**
Main Activity: Unemployed, not looking for job	-0.519	0.238	**-1.086**	**0.003**	-0.415	0.275
Main Activity: permanently sick	-0.040	0.891	**-2.015**	**0.000**	-0.582	0.063
Main Activity:retired	0.098	0.592	**-1.249**	**0.000**	**0.706**	**0.000**
Main Activity:community or military service	1.437	0.155	0.389	0.745	-1.745	0.070
Main Activity:houseworker	0.203	0.217	**-0.995**	**0.000**	-0.308	0.054
Main Activity:other	-0.365	0.403	**-1.081**	**0.083**	0.091	0.823
Constant	-3.606	0.115	**-6.198**	**0.000**	**-7.110**	**0.000**

Source: JRC elaboration of ESS2002 data

In line with the previous comparison, individuals in the 'productive age' are more likely to be part of the fully engaged group, as well as men have a higher probability than women. Moreover, Individual Human Capital plays a significant positive role so that individuals with a higher lever of ICH have a higher probability to be members of the Fully Engaged group. Citizenship and religiosity yield the same effect detected in the previous two groups.

The place where people are living gives an interesting picture; individuals that live outside big cities but not completely isolated in the countryside have a higher probability to be part of the fully engaged group. Conversely, the effect of the media is controversial;

if people who spend more time watching TV have a lower probability to be part of the group, increasing the time spent listening to the Radio, reading Newspapers and surfing the Web increases the probability to being fully engaged.

A model for Northern and Continental Countries

In order to deepen the investigation, we repeated the analysis for the four clusters identified in Section "Full Engagement: More than the Sum of the Parts". The comparison of the results shows a peculiar and picture of the Europe.

Table 8.5 Coefficients of the Model for the Nordic Countries

Dependent Variable: **ENGAGEMENT** - Cluster 1: **Denmark, Finland, Sweden, Netherlands**	Comparison between Non Engaged and Socially Engaged		Comparison between Non Engaged and Privately Engaged		Comparison between Non Engaged and Fully Engaged	
Explanatory Variables	Coef.	P>t	Coef.	P>t	Coef.	P>t
Income: Less than 1 800Euros	-0.065	0.967	-1.489	0.349	-0.609	0.677
Income: Less than 3 600Euros	0.522	0.669	0.485	0.675	0.151	0.885
Income: Less than 6 000Euros	-0.312	0.757	-0.366	0.719	-1.482	0.089
Income: Less than 12 000Euros	0.547	0.562	-0.346	0.714	-0.589	0.454
Income: Less than 18 000Euros	0.725	0.440	0.184	0.845	-0.037	0.963
Income: Less than 24 000Euros	0.735	0.431	0.434	0.641	-0.130	0.867
Income: Less than 30 000Euros	0.503	0.594	0.478	0.609	0.388	0.618
Income: Less than 36 000Euros	0.523	0.582	0.324	0.733	0.575	0.464
Income: Less than 60 000Euros	0.741	0.439	1.287	0.171	1.434	0.068
Income: Less than 90 000Euros	0.889	0.360	0.188	0.846	0.198	0.805
Income: Less than 120 000Euros	0.019	0.988	1.205	0.283	0.891	0.355
Age (18-24)	**-1.022**	**0.016**	0.130	0.764	-0.004	0.992
Age (24-35)	**-1.136**	**0.002**	0.203	0.613	**1.251**	**0.000**
Age (35-44)	**-1.134**	**0.003**	**0.807**	**0.037**	**0.866**	**0.007**
Age (44-55)	-0.544	0.106	0.511	0.156	0.542	0.066
Age (55-65)	**-0.779**	**0.005**	**0.696**	**0.024**	**0.806**	**0.000**
Sex	0.002	0.990	**-0.489**	**0.001**	-0.113	0.403
Watching TV	**-0.103**	**0.023**	-0.036	0.378	**-0.194**	**0.000**
Listening to the Radio	0.018	0.565	0.046	0.085	**0.080**	**0.001**
Reading Newspaper	0.045	0.576	**0.210**	**0.010**	**0.346**	**0.000**
Surfing the Web	0.008	0.861	0.008	0.858	0.040	0.358
Individual Human Capital	**0.099**	**0.000**	0.030	0.279	**0.123**	**0.000**
Number of persons	-0.054	0.567	**-0.162**	**0.019**	**-0.257**	**0.000**
Residence: suburbs of a big city	0.405	0.150	-0.022	0.928	**0.853**	**0.000**
Residence: small town	**0.492**	**0.037**	0.212	0.314	**0.556**	**0.006**
Residence: country village	**0.520**	**0.038**	0.211	0.362	**0.954**	**0.000**
Residence: home in the countryside	0.026	0.943	-0.095	0.783	**1.308**	**0.000**
Citizenship	0.462	0.079	**0.883**	**0.000**	**1.171**	**0.000**
Declared Religious	**0.842**	**0.000**	-0.159	0.305	**0.982**	**0.000**
Main Activity: In education	0.592	0.131	0.024	0.943	0.298	0.315
Main Activity: Unemployed, looking for job	-0.058	0.919	-0.617	0.114	**-1.065**	**0.011**
Main Activity: Unemployed, not looking for job	-0.404	0.492	**-0.933**	**0.037**	**-1.465**	**0.000**
Main Activity: permanently sick	-0.274	0.513	**-1.315**	**0.000**	**-0.843**	**0.009**
Main Activity:retired	-0.039	0.899	**-1.593**	**0.000**	-0.431	0.081
Main Activity:community or military service	…	…	…	…	…	….
Main Activity:houseworker	-0.192	0.493	**-1.237**	**0.000**	**-0.887**	**0.000**
Main Activity:other	0.107	0.823	**-1.805**	**0.001**	**-0.785**	**0.044**
Constant	-2.071	0.069	-1.106	0.316	-2.260	**0.014**

Source: JRC elaboration of ESS 2002 data

The analysis of cluster 1 (Nordic countries) gives a clear picture of the dynamics related to participation and Table 8.5 shows the results in detail. The comparison of the three models resembles the differences identified in the overall results, but not quite. First of all, it is worth noticing that income variables are not significant in any of the three models, showing that – in the Nordic cluster – income does not have any effect on the probability of participating. Subsequently and unlike in the overall picture, females do not have a significantly lower probability of being socially of fully engaged, while people outside the job market are less likely to be fully engaged.

Table 8.6 Coefficients of the Model for the Continental Countries

Depedent Variable: ENGAGEMENT - Cluster 2: Belgium, Germany, Great Britain, Luxemburg	Comparison between Non Engaged and Socially Engaged		Comparison between Non Engaged and Privately Engaged		Comparison between Non Engaged and Fully Engaged	
Explanatory Variables	Coef.	P>t	Coef.	P>t	Coef.	P>t
Income: Less than 1 800Euros	-0.379	0.792	**-2.169**	**0.045**	**-4.971**	**0.000**
Income: Less than 3 600Euros	0.620	0.639	**-2.245**	**0.017**	**-3.240**	**0.000**
Income: Less than 6 000Euros	0.010	0.994	**-2.936**	**0.004**	**-1.833**	**0.012**
Income: Less than 12 000Euros	0.290	0.812	**-2.815**	**0.000**	**-2.356**	**0.000**
Income: Less than 18 000Euros	1.441	0.233	**-2.277**	**0.002**	**-2.208**	**0.001**
Income: Less than 24 000Euros	0.982	0.421	**-1.764**	**0.012**	**-1.543**	**0.014**
Income: Less than 30 000Euros	0.142	0.908	**-1.785**	**0.011**	**-1.625**	**0.010**
Income: Less than 36 000Euros	0.606	0.623	**-1.645**	**0.020**	**-1.591**	**0.012**
Income: Less than 60 000Euros	0.694	0.570	**-1.505**	**0.032**	**-1.310**	**0.036**
Income: Less than 90 000Euros	0.602	0.634	-0.926	0.204	-1.232	0.063
Income: Less than 120 000Euros	-1.425	0.294	-0.002	0.914	-1.693	0.057
Age (18-24)	**-2.389**	**0.000**	0.050	0.919	0.156	0.677
Age (24-35)	**-2.457**	**0.000**	0.238	0.568	-0.009	0.976
Age (35-44)	**-2.065**	**0.000**	0.303	0.464	0.101	0.746
Age (44-55)	**-1.451**	**0.000**	**0.806**	**0.046**	**0.688**	**0.021**
Age (55-65)	**-1.081**	**0.000**	**0.900**	**0.014**	0.484	0.052
Sex	**-0.549**	**0.000**	**-0.381**	**0.020**	**-0.631**	**0.000**
Watching TV	**-0.118**	**0.004**	-0.074	0.077	**-0.176**	**0.000**
Listening to the Radio	**0.071**	**0.007**	**0.056**	**0.045**	0.001	0.979
Reading Newspaper	**-0.351**	**0.000**	0.053	0.380	**0.215**	**0.000**
Surfing the Web	**-0.003**	**0.039**	**-0.005**	**0.004**	-0.002	0.083
Individual Human Capital	**0.229**	**0.000**	**0.119**	**0.000**	**0.210**	**0.000**
Number of persons	0.089	0.217	0.082	0.248	-0.018	0.753
Residence: suburbs of a big city	**0.685**	**0.020**	0.201	0.472	0.429	0.078
Residence: small town	**0.607**	**0.017**	-0.122	0.607	0.081	0.689
Residence: country village	0.206	0.484	0.110	0.677	**0.461**	**0.035**
Residence: home in the countryside	0.649	0.244	0.366	0.476	0.594	0.183
Citizenship	**0.609**	**0.044**	0.281	0.311	**0.579**	**0.012**
Declared Religious	**0.993**	**0.000**	**0.364**	**0.029**	**1.059**	**0.000**
Main Activity: In education	**1.285**	**0.010**	0.172	0.749	0.491	0.225
Main Activity: Unemployed, looking for job	-0.018	0.963	0.017	0.959	**-0.873**	**0.021**
Main Activity: Unemployed, not looking for job	0.273	0.665	-0.120	0.840	0.210	0.610
Main Activity: permanently sick	-0.055	0.887	**-1.819**	**0.000**	-0.570	0.140
Main Activity:retired	0.075	0.789	-0.481	0.152	-0.060	0.819
Main Activity:community or military service	0.817	0.415	1.715	0.212	-1.413	0.087
Main Activity:houseworker	**0.573**	**0.019**	-0.215	0.421	-0.205	0.361
Main Activity:other	-0.499	0.522	-0.359	0.764	0.815	0.193
Constant	-3.547	0.008	-1.117	0.250	**-1.765**	**0.032**

Source: JRC Elaboration of ESS2002 Data

The analysis of the results of cluster 2 (Belgium, Germany, Great Britain and Luxemburg) are shown in Table 8.6 and gives a slightly different picture. In contrast with the results obtained with the Nordic countries, in this cluster the income play a major role in the probability to be privately and fully engaged where people with an income lower than 60 000 € have a decreasing probability of participate. Nevertheless the income turns out to be non significant for the probability to be part of the socially engaged group.

Older people are still more likely to be part of the socially engaged group, but the effect is much clearer in this cluster than the previous analysis. The same consideration holds for men, who have a much higher probability of being part of any group.

Table 8.7 Coefficients of the Model for the Mediterranean Countries

Dependent Variable: ENGAGEMENT Cluster 3: Greece, Italy, Spain, Portugal	Comparison between Non-engaged and Socially-engaged		Comparison between Non-engaged and Privately-engaged		Comparison between Non-engaged and Fully-engaged	
Explanatory Variables	Coef.	P>t	Coef.	P>t	Coef.	P>t
Income: Less than 1 800Euros	-1.435	0.226	-1.788	0.064	1.020	0.223
Income: Less than 3 600Euros	**-3.047**	**0.001**	**-1.994**	**0.010**	**-1.919**	**0.025**
Income: Less than 6 000Euros	-1.341	0.093	-1.017	0.132	0.001	0.999
Income: Less than 12 000Euros	-1.037	0.183	-0.405	0.416	-0.725	0.241
Income: Less than 18 000Euros	-1.047	0.168	-0.652	0.192	-0.180	0.756
Income: Less than 24 000Euros	-0.026	0.972	0.073	0.879	-0.097	0.870
Income: Less than 30 000Euros	-0.993	0.217	-0.428	0.440	-0.521	0.401
Income: Less than 36 000Euros	-0.504	0.530	**-1.254**	**0.036**	0.244	0.698
Income: Less than 60 000Euros	-1.398	0.112	0.251	0.655	0.295	0.641
Income: Less than 90 000Euros*	…	….	…	….	…	….
Income: Less than 120 000Euros*	…	….	…	….	…	….
Age (18-24)	**-1.025**	**0.042**	-0.858	0.294	-0.961	0.163
Age (24-35)	**-0.921**	**0.023**	-0.303	0.690	-0.829	0.177
Age (35-44)	**-1.294**	**0.003**	0.692	0.347	-0.282	0.611
Age (44-55)	**-0.778**	**0.050**	-0.072	0.922	-0.406	0.463
Age (55-65)	**-0.647**	**0.050**	0.395	0.559	0.064	0.894
Sex	0.347	0.142	**-0.909**	**0.000**	**-0.614**	**0.008**
Watching TV	-0.070	0.161	0.020	0.717	-0.095	0.119
Listening to the Radio	-0.039	0.329	0.028	0.518	**0.083**	**0.046**
Reading Newspaper	0.036	0.723	0.060	0.486	**0.196**	**0.030**
Surfing the Web	-0.051	0.309	0.019	0.118	**0.026**	**0.005**
Individual Human Capital	**0.120**	**0.000**	**0.102**	**0.000**	**0.127**	**0.000**
Number of persons	-0.040	0.626	0.067	0.427	0.084	0.373
Residence: suburbs of a big city	**1.039**	**0.006**	0.476	0.231	0.139	0.758
Residence: small town	**0.994**	**0.001**	0.545	0.068	**0.769**	**0.022**
Residence: country village	0.564	0.067	0.273	0.426	0.483	0.184
Residence: home in the countryside	0.482	0.450	0.526	0.372	-0.864	0.302
Citizenship	0.968	0.078	0.618	0.228	0.632	0.222
Declared Religious	0.094	0.759	0.162	0.575	0.455	0.107
Main Activity: In education	0.014	0.981	0.717	0.206	0.470	0.375
Main Activity: Unemployed, looking for job	0.469	0.328	-0.696	0.251	-0.887	0.126
Main Activity: Unemployed, not looking for job	-0.340	0.445	-0.199	0.765	0.334	0.432
Main Activity: permanently sick	0.685	0.274	**-3.734**	**0.001**	0.009	0.993
Main Activity:retired	-0.169	0.657	**-1.259**	**0.044**	-0.350	0.468
Main Activity:community or military service*	…	…	…	…	-1.413	0.087
Main Activity:houseworker	0.237	0.234	**-2.560**	**0.000**	**-1.667**	**0.000**
Main Activity:other*	…	…	**0.572**	**0.017**	0.429	0.116
Constant	**-3.547**	**0.008**	**-3.424**	**0.001**	**-3.915**	**0.000**

* variable excluded for scarcity of data
Source: JRC Elaboration of ESS2002 Data

On the other hand, the effect of the place of domicile maintains its direction but is weaker than in the Nordic countries, and people outside the job market seem to have a higher probability to be part of the socially engaged groups whereas they are less likely to be member of the privately or fully engaged groups.

The story changes altogether when looking at the Eastern and the Mediterranean clusters. In fact, the analysis of the Mediterranean countries (Italy, Greece, Portugal and Spain, Table 8.7) and of the Eastern countries (Poland and Slovenia, Table 8.8) triggers concerns on the validity of the proposed model for these clusters.

Table 8.8 Coefficients of the Model for the Eastern Countries

Depedent Variable: ENGAGEMENT - Cluster 3: Poland, Slovenia	Comparison between Non Engaged and Socially Engaged		Comparison between Non Engaged and Privately Engaged		Comparison between Non Engaged and Fully Engaged	
Explanatory Variables	Coef.	P>t	Coef.	P>t	Coef.	P>t
Income: Less than 1 800Euros*	…	…	…	…	…	…
Income: Less than 3 600Euros	-0.357	0.387	0.707	0.250	0.839	0.304
Income: Less than 6 000Euros	-0.509	0.209	0.780	0.102	**2.463**	**0.001**
Income: Less than 12 000Euros	-0.094	0.824	0.355	0.562	**2.078**	**0.010**
Income: Less than 18 000Euros	0.256	0.626	0.011	0.988	**2.302**	**0.010**
Income: Less than 24 000Euros	-0.157	0.811	0.275	0.753	**1.929**	**0.041**
Income: Less than 30 000Euros	0.605	0.369	1.361	0.123	**3.478**	**0.002**
Income: Less than 36 000Euros	-0.534	0.647	0.129	0.922	0.582	0.613
Income: Less than 60 000Euros*	…	….	…	….	…	….
Income: Less than 90 000Euros*	…	….	…	….	…	….
Income: Less than 120 000Euros*	…	….	…	….	…	….
Age (18-24)	**-1.450**	**0.012**	0.015	0.983	0.633	0.430
Age (24-35)	-0.937	0.054	-0.184	0.763	0.342	0.649
Age (35-44)	**-0.943**	**0.048**	**1.952**	**0.000**	1.197	0.073
Age (44-55)	**-0.997**	**0.024**	0.512	0.314	**1.483**	**0.016**
Age (55-65)	-0.252	0.484	0.346	0.501	**2.609**	**0.000**
Sex	-0.179	0.422	**-0.696**	**0.001**	**-1.055**	**0.000**
Watching TV	-0.074	0.106	0.069	0.186	**-0.471**	**0.000**
Listening to the Radio	0.033	0.385	0.019	0.636	**0.197**	**0.000**
Reading Newspaper	0.084	0.375	**0.284**	**0.001**	0.215	0.113
Surfing the Web	0.059	0.274	**0.280**	**0.000**	0.148	0.048
Individual Human Capital	**0.083**	**0.032**	-0.036	0.467	**0.144**	**0.001**
Number of persons	0.081	0.250	0.077	0.256	**-0.440**	**0.000**
Residence: suburbs of a big city	-0.837	0.114	-0.724	0.127	0.976	0.116
Residence: small town	-0.022	0.932	**-1.072**	**0.000**	**1.414**	**0.001**
Residence: country village	**-0.622**	**0.042**	**-1.683**	**0.000**	0.403	0.403
Residence: home in the countryside	-0.280	0.665	-0.863	0.212	**2.442**	**0.000**
Citizenship	**2.103**	**0.002**	-0.683	0.314	**1.094**	**0.040**
Declared Religious	0.360	0.261	0.278	0.433	0.716	0.083
Main Activity: In education	0.499	0.354	-0.671	0.249	0.662	0.324
Main Activity: Unemployed, looking for job	-0.393	0.481	-0.177	0.435	0.613	0.320
Main Activity: Unemployed, not looking for job	-0.145	0.556	-0.987	0.643	0.099	0.543
Main Activity: permanently sick	-0.553	0.516	**-4.279**	**0.000**	-2.228	0.052
Main Activity:retired	0.017	0.965	-0.674	0.091	**0.959**	**0.022**
Main Activity:community or military service*	…	…	…	…	…	…
Main Activity:houseworker	-0.360	0.352	-0.698	0.078	-0.581	0.372
Main Activity:other	0.411	0.665	0.133	0.897	-1.267	0.228
Constant	-4.664	0.000	**-2.616**	**0.026**	**-8.079**	**0.000**

* variable excluded for scarcity of data
Source: JRC Elaboration of ESS2002 Data

On one side, the analysis confirms, for both the clusters, some common trends appeared in the previous groups of countries:

- Older people are more likely to participate in the socially engaged group (such as in the Nordic and continental countries);
- Individual human capital has a persistently positive effect;
- People outside the job market that looks less likely to be part of the privately engaged group;
- Women still appear to be less likely to be engaged in all the groups.

However, the limited number of significant coefficients coming out from the model suggests that, in these clusters, participation is driven by variables not considered in the model and not collected through the survey. This result is specifically relevant for policy because it suggests that the drivers of participation are different in Northern-Continental and Southern-Eastern Europe. Hence, the actions put in place to enhance participation and the social dialogue must be evaluated against different benchmarks along the different models.

Moreover, the variables used for the analysis are *self reported* and doubts persist on the reliability of the variable *income*. For example, almost nobody declared a yearly family income above €90 000.00 in any of the two clusters. With respect to the Eastern cluster, as mentioned before, doubts persist on the reliability of the Polish data. In fact, the Polish dataset shows anomalous patterns of response by which 80% of the individuals answer consistently in the same manner. Although we haven't reported the detailed table, we have also re-run the model on the Slovenian sample alone, and the results are quite similar to those of the Continental countries.

The adoption of a different model exploring new dimensions and variables appear necessary for modelling the participation in these parts of Europe.

Participation and Perception of Europe's Social Reality: The Differences.

The institutional slowness in adapting to change shown by some countries in a situation of incremental rise of the cost of the status-quo has been identified by many (Esping-Andersen 1999; Bison & Esping-Andersen 2000; Kohli 1999; Easterlin 1987) as one of the key reasons for the decline of the traditional welfare state, especially in terms of sustainability. Moreover, some authors (see for example Wolfe 1989; Fukuyama 2000; Putnam 2000; Beugelsdijk & Smulders 2003)[19] suggest the possibility of a kind of 'crowding out effect'. The focus on the sole social expenditures and social programmes in the traditional welfare state can undermine informal caring relations, social networks, and systems of reciprocity. There is then a risk of leading to a decline of commitment to civil norms, of participation in civil society, and trust in fellow citizens and social institutions.

A situation whereby individuals rely only on institutions to solve problems is a structural problem. As Delors foresaw, 'the problem of how we finance the welfare state should not obscure a separate issue: if each person thinks he has an inalienable right to welfare, no matter what happens to the world, that's not equity, it's just creating a society where you can't ask anything of people'. As mentioned already in Section "Citizens and

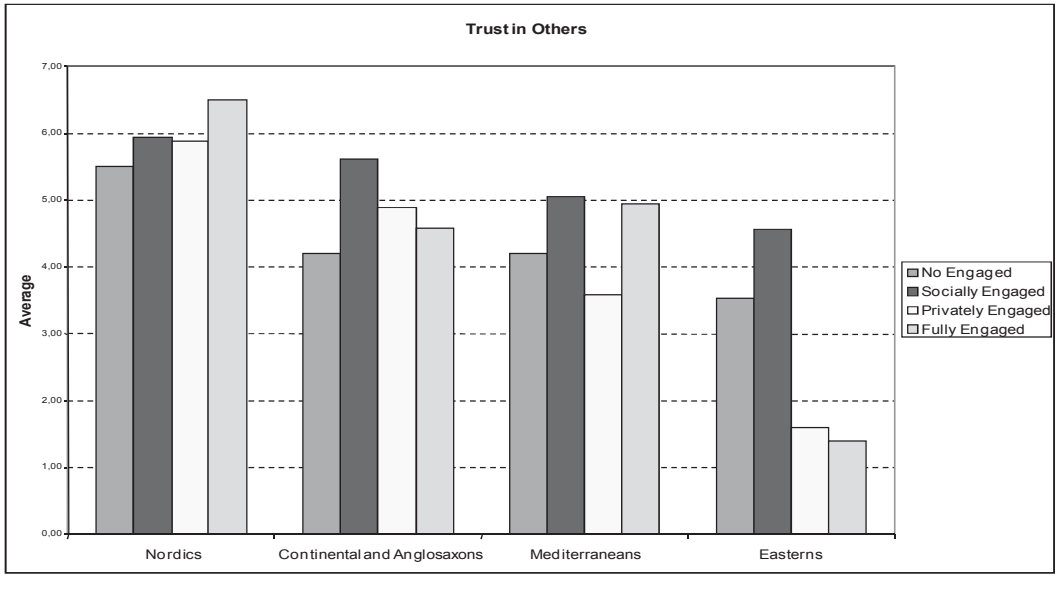
State Responsibilities", and to avoid the 'crowding out effect', modern welfare states need to be built on joint responsibilities between all stakeholders.

Part of the Nordic success is due to the ability of these countries to invest in the creation of a renewed form of welfare state in *closer dialogue* with civil society (Esping-Andersen 1990, 2002, 2006). Such political architecture creates the structural and cultural conditions for a welfare mix operating in a social environment of high 'generalised trust' between the populations and state-systems (Rothstein & Stolle 2003) and where voluntary activity works complementary to the state, not a substitute for it (Bartkowski & Jasinka-Kania 2004).

Although the levels of participation are different, these considerations suggest the need of investigating whether such feeling of 'generalised trust' holds systematically in case of individual participation, whether different participatory styles suggest substantial differences in attitudes towards policy, society, and the economy, or whether other latent dimensions must be investigated before making any policy suggestion.

The first set of graphs investigates individual perceptions towards a number of social dimensions. What is apparent is that people who participate perceive the world in a more positive way than those who don't participate. In this sense, the first remark to be made is that the greater the participation, the higher the trust in others. 'Trust lubricates cooperation. The greater the level of trust within a community, the greater the likelihood of cooperation. And cooperation itself breeds trust' (Putnam 1993, p.171). Considering people's trust in the others (Figure 8.6), it is apparent that people in Nordic and Continental Europe engaged in participative activities has a more positive approach to the society. On other hands, this trend is not completely confirmed in Southern and Eastern Europe.[20] By the way, Nordic fully engaged people report the highest mark with 6.57 while eastern European fully engaged report the lowest score: 1.56.

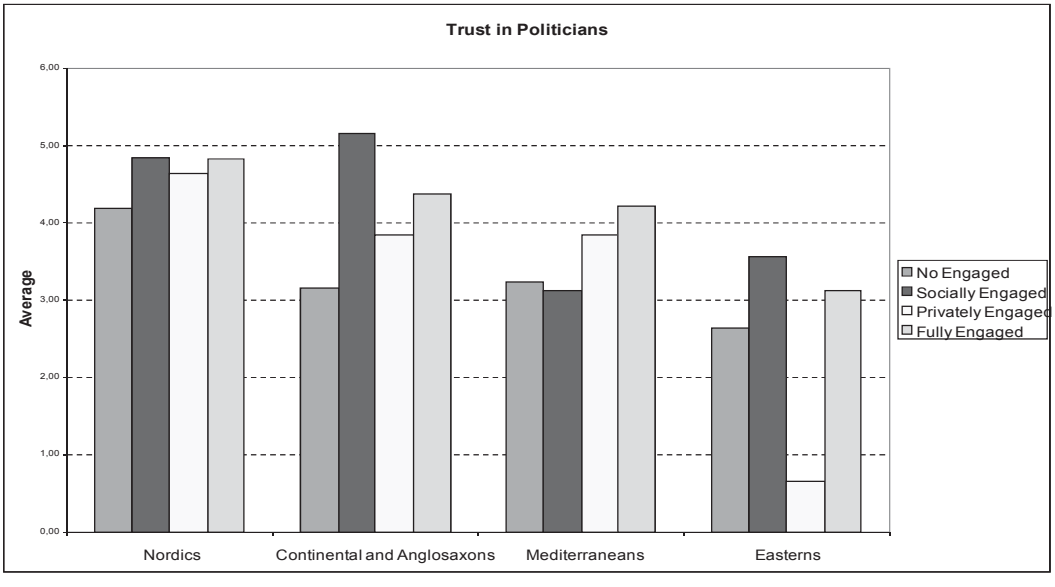
Figure 8.6 Trust in Others



Source: JRC elaboration of ESS2002 data

The general picture shown in the previous graph is also confirmed when analysing the 'Trust in Politicians' (Figure 8.7). It is also interesting to notice that the average score assigned in the different clusters shows the same trend, whereby people in the Nordic countries – on average – trust much more than those in the Southern and Eastern countries. The anomaly of Polish data is clear also in this question, with an peculiar behaviour of privately engaged people.

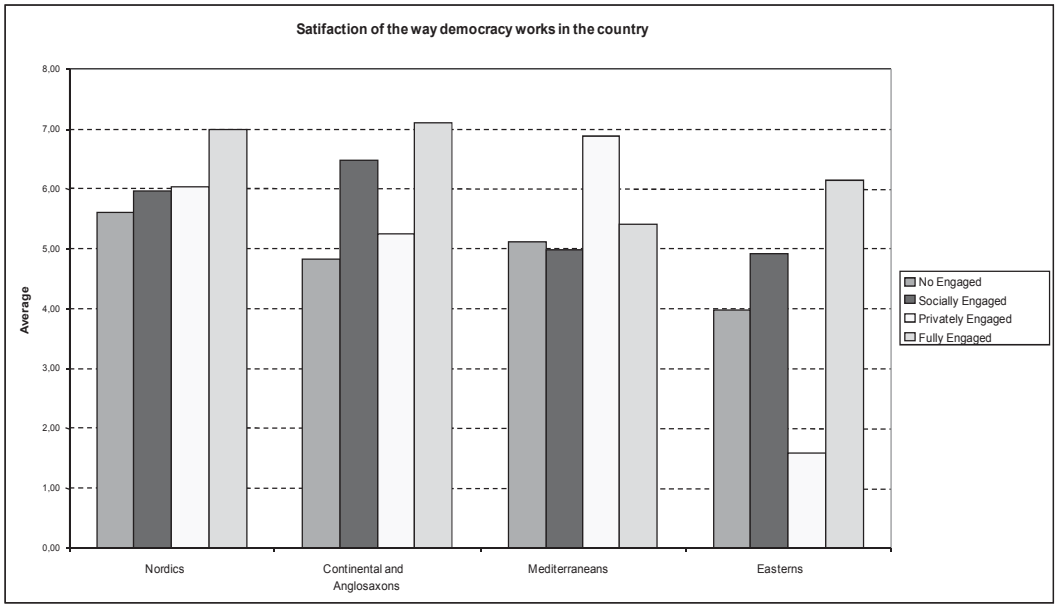
Figure 8.7 Trust in Politicians



Source: JRC elaboration of ESS2002 data

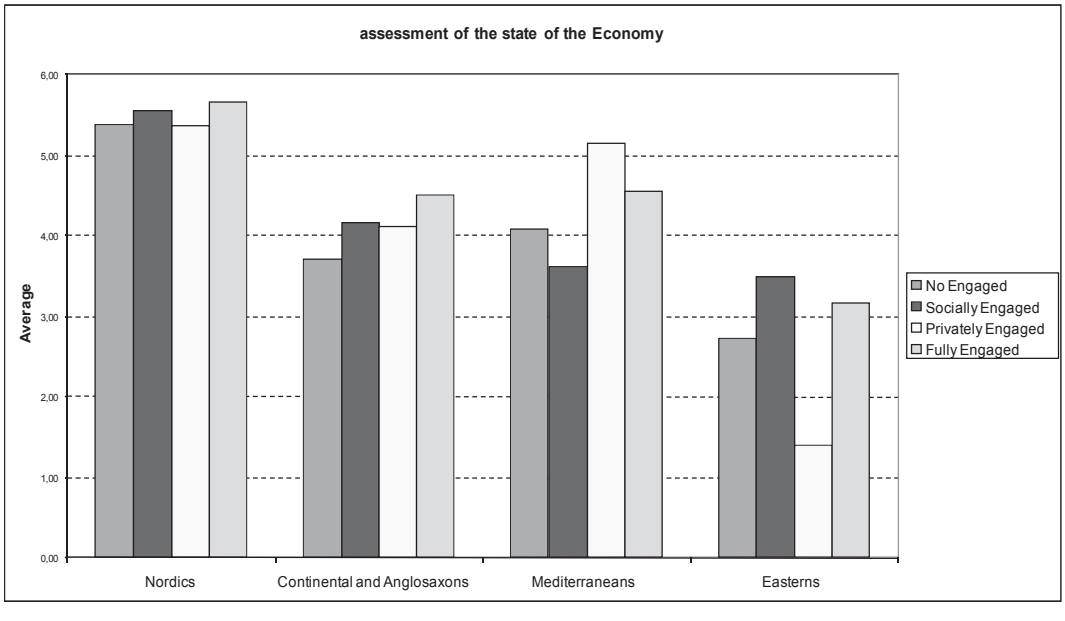
The overall trend is still confirmed for the assessment of the satisfaction of the way democracy works in the country (Figure 8.8) and the perception of the state of the economy (Figure 8.9), where participants are more positive than non participants. The trend is uniform for Nordic and continental countries and more peculiar for the rest of Europe. Still present is the Polish anomaly.

Figure 8.8 Satisfaction of the Way Democracy Works in the Country



Source: JRC elaboration of ESS2002 data

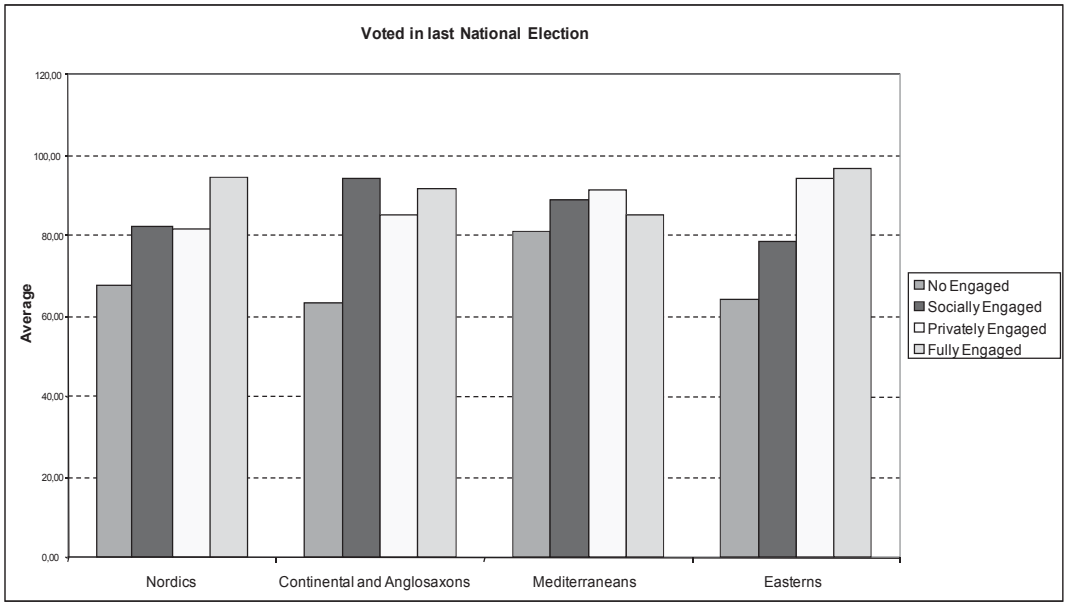
Figure 8.9 Assessment of the State of the Economy



Source: JRC elaboration of ESS2002 data

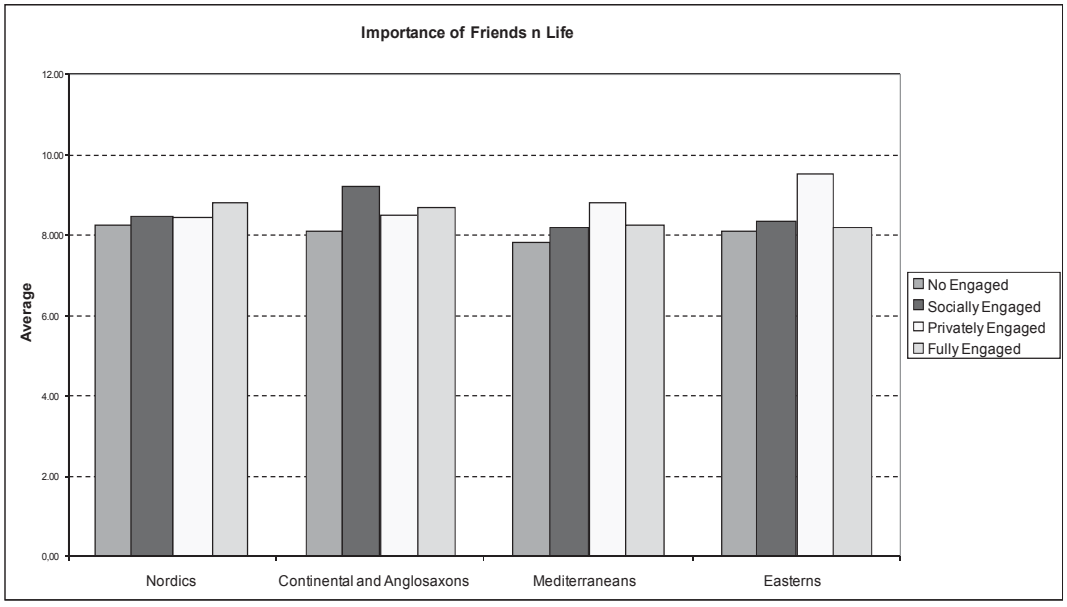
One final remark regards the existence of some transversal dimensions such as 'voted in the last elections' (Figure 8.10) or 'importance of friend in life' (Figure 8.11) that show common trends shared all across Europe.

Figure 8.10 Voted in the Last Elections



Source: JRC elaboration of ESS2002 data

Figure 8.11 Importance of friends in life



Source: JRC elaboration of ESS2002 data

The analyses presented do not advocate for the development of specific organisation or structures. In fact, as suggested earlier, we are interested in the characteristics and development of the European social reality as a whole. Specifically, we are trying to understand what the determinants of social involvement are and whether involvement can enhance the results of the coordination process between State and stakeholders in an increasingly diverse society.

Of course, we have sketched a picture that depends on the *indicators chosen*, which determine the representation of reality that we are using. A set of indicators is not reality itself, but rather a descriptive model of reality. This caveat explains the limited validity of the model for Eastern and Mediterranean countries and suggests that, in these countries, participation is driven by variables not considered in the model and not collected through the survey. Moreover the data used are from a survey undertaken in 2002, meaning that the picture today might be different. The possibility of drafting policy measures that can effectively support the close coordination between the stakeholders depend on the availability of adequate data and – in the present situation – we must highlight a *problem with the quality of the available information* as important variables, like *informal participation*, which is a typical asset of the Mediterranean region, are poorly or not at all represented.

One point to highlight is the *consistently positive effect of education for all kinds of participation in all the clusters*. Such result identifies for looking at education as one of the possibly most transversal policy actions to undertake and suggests the importance of working on the quality of education systems both at basic and higher level.

This empirical analysis characterises some of the dimensions along which European social models differ while also suggesting that a situation whereby individuals expect solutions from institutions that are not able to keep up with the development of society is a structural problem." Part of the Nordic success is due to the ability of these countries to invest in the creation of a renewed form of welfare state in *closer dialogue* with civil society.

The existence of such strong structural differences impedes the application of one-size-fits-all policies and suggests the need of cooking *ad hoc* recipes closely agreed upon with the stakeholders, who take on the independent and legitimate role of interpreting social needs and of launching cooperation strategies at territorial level. In this paradigm, favouring individual participation to the public-value creation process is a key element and reminds us that '. . . how we associate with each other, and on what terms, has enormous implications for our well-being. . .' (Woolcock 2001, p.15).

The empirical results of our paper also feed the general thesis of the paper: future success depends on realigning economic and social models. That does not imply that Mediterranean countries should convert into Nordic models. But it does show that within their own systematic idiosyncrasies they need to find ways to achieve civil participation. In this sense, the actions put in place to enhance participation and the social dialogue must be evaluated against different benchmarks along the different models. The existing measurement instruments are adequate for investigating the results of the Northern regions of the continent, but different recipes require different sorts of leaven; the objective of 'increasing the size of the pie' rests also on the capability of identifying, measuring, and supporting the adequate raising agents also in Mediterranean and Eastern Europe, otherwise economic reform will not be socially sustainable and thereby not politically feasible.

Conclusion

The paper shows that new realities call for a better involvement of social dimensions to foster economic dynamism. Despite the rise of an affluent society, social difficulties persist and these difficulties are likely to hamper future economic dynamism. Child poverty is rising, migrants' integration is problematic and the performance in terms of

education is poor in certain countries. Given the need for high skilled labour and given demographic trends, the social problems will need to be dealt with now.

To meet the challenges of adapting to change and to reduce the democratic gap of citizens that become less interested in politics while politics becomes ever more complex, civil participation becomes key. That is why we have illustrated our general thesis by an empirical study on participation in society.

Modernisation of social model may be easier in a society characterised by bridging social ties, linked to a higher level of trust and a greater disposition of citizens to negotiate and own the reform process. Not by chance, 'Learning to live together' is one of the four pillars advocated by Jacques Delors (1996) for a lifelong learning society.

The empirical analysis of participation clearly identifies specific patterns according to the distinction between private and public participation. The analysis shows that involvement of citizens is key to social and economic performances, but not all kind of involvements lead to valuable results. The way that people involved in one part of Europe can be very different from other parts of Europe, but different approaches could yield similar results. The variety of participation patterns reflects the variety of (the quality of) coordination between state and stakeholders when running policies and different traditions and cultures. The common element is that participation by itself, in whichever form – is a crucial input for economic and social performance.

In a sense, our empirical analysis underlines the role of another dimension than the usual dimension of the welfare state to explain the variety of social models in Europe. The existence of such differences in internal coordination implies that one-size-fits-all policies do not work and suggests the need of specific methods closely agreed upon with the stakeholders, who take on the independent and legitimate role of interpreting social needs and of launching cooperation strategies at territorial level.

Our empirical study also underlines the importance the active involvement of non-profit organisations and NGO's. And we do not mean this is the traditional way of 'getting all stakeholders on board', but non-profit organisations as essential players in civil participation, both as suppliers of information, intermediaries between government and citizens and platforms for social interactions.

Another aspect underlined by this paper is that there is growing interaction between economic and social dimension. This interaction as well as the (various) involvement of various stakeholders require a multidimensional approach to assess and undertake policies in Europe. This is crucial since policies are increasingly cross-cutting in nature. A future challenge will be to develop a well-being approach that goes beyond traditional indicators such as GDP per head.

Notes

1 http://ec.europa.eu/public_opinion/archives/eb_special_en.htm#273

2 As the CCS (Centre for Civil Society at the London School of Economics) indicates, 'Civil society commonly embraces a diversity of spaces, actors and institutional forms, varying in their degree of formality, autonomy and power. Civil societies are often populated by organisations such as registered charities, development non-governmental organisations, community groups, women's organisations, faith-based organisations, professional associations, trades unions, self-help groups, social movements, business associations, coalitions and advocacy groups'.

3 European Employment and Social Policy, Special Eurobarometer 261, October 2006.

4 This subsection and the next borrow from Canoy and Smith, 2006.

5 On average across the OECD the hiring rate of workers aged 50 and over was less than half the rate for workers aged 25-49, see 'Live longer, work longer', OECD 2006

6 This subsection borrows from Liddle and Lerais 2007.

7 http://ec.europa.eu/employment_social/employment_analysis/employ_2007_en.htm

8 Communication on the social agenda 2005-2010. COM(2005) 33 final

9 The Commission adopted on 20 November 2007 a Communication on "Opportunities, access and solidarity: towards a new social vision for 21st century Europe" [COM(2007)726]

10 See eg http://ec.europa.eu/education/index_en.html, http://ec.europa.eu/health/dyna/press_rel/ph_press_en.cfm, and http://ec.europa.eu/employment_social/employment_strategy/index_en.htm

11 See Atkinson, A B, et al. (2005), Taking Forward the EU Social Inclusion Process, Luxembourg: Luxembourg Presidency of the European Union.

12 For a critique of the OMC see also Altomonte and Nava 2006.

13 As hinted, the concept of horizontal subsidiarity addresses the fundamental role of the social partners in the implementation of the social dimension of the EU. In its first Communication concerning the application of the Agreement on Social Policy (COM (93) 600 final, Brussels, 14 December 1993), the Commission acknowledged: 'a dual form of subsidiarity in the social field: on the one hand, subsidiarity regarding regulation at national and Community level; on the other, subsidiarity as regards the choice, at Community level, between the legislative approach and the agreement-based approach.' (http://www.eurofound.europa.eu/areas/industrialrelations/dictionary/definitions/horizontalsubsidiarity.htm).

14 The full report can be downloaded from: http://crell.jrc.ec.europa.eu/ActiveCitizenship/AC-Final%20Report-December%202006/measuring%20AC.pdf

15 Factor Analysis (FA) can be used to group the information contained in the indicators. The aim of the methodology is to explore whether the different dimensions of the phenomenon are statistically well balanced in the composite indicator. The higher the correlation between the indicators, the fewer the statistical dimensions that will be present in the dataset. For continuos variables the standard Pearson correlation coefficient can be computed, but in case of categorical variables the literature suggests strongly the adoption of polychoric correlation.

16 For instance, Gosta Esping-Andersen 1990.

17 See for example: Aranitou, V., (2003), 'The strengthening of the employers' organisations representation and social dialogue', paper presented at the S. Karagiorgas Conference on Social Change in Contemporary Greece (1980-2001), Athens. http://www.lse.ac.uk/collections/hellenicObservatory/pdf/symposiumPapersonline/Aranitou.pdf

18 For example, Putnam's analyses have been criticised because his aggregation of individual-level data up to collectivities does not capture the social structural nature of social capital (see Skocpol 1996, or Sampson 2001).

19 The 'crowding out" hypothesis must not be considered valid altogether. In fact some scholars reject the crowding out hypothesis by arguing that a well developed welfare state creates the structural and cultural conditions for a thriving and multiform civil society. The point could be certainly shared.

20 Again, the results of Eastern Europe are affected by the Polish data, which shows an anomalous concentration of answers.

References

Altomonte, C., Nava, M.(2006), *Economics and Policies of an Enlarged Europe,* Edward Elgar, Cheltenham Glos.

Atkinson, A B, *et al.* (2005), *Taking Forward the EU Social Inclusion Process,* Luxembourg Presidency of the European Union, Luxembourg.

Baldwin and Martin (1999), "Two waves of globalisation: fundamental differences", *NBER Working paper* 6904.

Barrington-Leach Leanda, *et al.* (2007), "Investing in Youth: an empowerment strategy", BEPA paper, http://ec.europa.eu/dgs/policy_advisers/publications/docs/Investing_in_Youth_25_April_fin.pdf

Beugelsdijk, S., Smulders, S. (2003), "Bridging and Bonding Social Capital: which type is good for economic growth?," ERSA conference papers ersa03p517, European Regional Science Association. http://www.ersa.org/ersaconfs/ersa03/cdrom/papers/517.pdf

Bison, I., Esping-Andersen, G. (2000), "Income packaging, poverty and unemployment in Europe", In D. Gallie and S. Paugham, eds. *The Experience of Unemployment in Europe*. Oxford University Press, Oxford.

Canoy M., P.M. Smith (2006), "A social and economic model for Europe", Intereconomics - Review of European Economic Policy, Volume 41, Number 6/ November, 2006.

Canoy Marcel, *et al.* (2006), "Migration and public perception", BEPA paper http://ec.europa.eu/dgs/policy_advisers/publications/docs/bepa_migration_final_09_10_006_en.pdf

Centre for Civil Society, London School of Economics (2004), "What is civil society?", http://www.lse.ac.uk/collections/CCS/what_is_civil_society.htm

Dasgupta, P. (2003) "Social Capital and Economic Performance: Analytics" in: E. Ostrom and T.K. Ahn, eds., (2003). *Critical Writings in Economic Institutions: Foundations of Social Capital,* Edward Elgar, Cheltenham, UK.

Delors, J. *et al*. (1996), "Learning: The Treasure Within- Report to UNESCO of the International Commission on Education for the Twenty-first Century". UNESCO Publishing, Paris, France.

Dobell, A.R. and D. Mitchell, (1997), "The British Columbia Case: Natural Resource Management and Sustainable Development", Cited in: Dobell, R. and Bunton, M. (2001), "Sound Governance: The emergence of collaborative networks and new institutions in the Clayoquot Sound region", Background paper for Clayoquot Sound

Regional Workshop, Tofino.
http://www.clayoquotalliance.uvic.ca/PDFs/SOUND_GOV_2.pdf

Easterlin, R. (1987), "The new age structure of poverty in America", Population and Development Review, Vol. 13 No. 2:p. 195-208.

Esping-Andersen G. (2002), *Why we need a new welfare state*, Oxford University Press, Oxford.

Esping-Andersen,G. (1990), *The three Worlds of Welfare Capitalism,* Oxford University Press, Oxford.

Esping-Andersen,G. (1999), *Social Foundations of Postindustrial Economies ,* Oxford University Press, Oxford.

Esping-Andersen,G. (2006), "An Equitable Social Model", *Samfundsøkonomen* 1.

Esping-Andersen, G., Gallie, D., Hemerijck, A. and J. Myles (2002), *Why We Need a New Welfare State*, Oxford University Press, Oxford.

Eurobarometer http://ec.europa.eu/public_opinion/index_en.htm

European Commission (2005a), "Commission Staff Working Document in support of the report from the Commission to the Spring European Council", 22-23 March 2005, on the Lisbon strategy of economic, social and environmental renewal.

Fukuyama, F. (1995), *Trust: The social virtues and the creation of prosperity*. The Free Press, New York.

Fukuyama, Francis (1999), "Social Capital and Civil Society", IMF Conference on Second Generation Reforms
/www.imf.org/external/pubs/ft/seminar/1999/reforms/fukuyama.htm).

Granovetter, Mark (1973), "The Strength of Weak Ties", American Journal of Sociology, Vol. 78, No. 6, p.1360-1380.

Hayek, Friedrich (1945), "The Use of Knowledge in Society", The American Economic Review, Vol. 4, p.519-530.

Heckman, James J. and Lance Lochner (2000), "Rethinking Education and Training Policy: Understanding the Sources of Skill Formation in a Modern Economy," in S. Danziger and J. Waldfogel (eds.), Securing the Future: Investing in Children from Birth to College, Russell Sage Foundation, New York, p. 47-83.

Hoskins B., *et al*. (2006), *Measuring Active Citizenship in Europe*, EUR 22530 EN – DG Joint Research Centre, Luxemburg.

Indepen (2006), "Restoring European economic and social progress: unleashing the potential of ICT", http://www.indepen.co.uk/panda/docs/brt-summary.pdf

Jackson, Andrew (2000), *Why we don't have to choose between social justice and economic growth: the myth of the equity/efficiency trade off*, Canadian council on social development

Jerzy Bartkowski and Aleksandra Jasinka-Kania, (2004), 'Voluntary Organisations and the Development of Civil Society' in Arts and Halman, eds. , *European Values at the Turn of the Millennium,* Brill, Leiden/Boston:

Kohli, M. (1999). "Private and public transfers between generations", European Societies, *Vol.*1, No. 1, p. 81-104

Kok, (2004), "The High Level Group on Lisbon Strategy (chaired by Wim Kok)", *Facing the Challenge*, European Communities, Luxembourg.

Liddle R. and F. Lerais (2007) 'Europe's social reality' a consultation paper from BEPA, European Commission', http://ec.europa.eu/citizens_agenda/social_reality_stocktaking/docs/background_document_en.pdf

OECD (2005), *Society at a Glance*: *OECD Social Indicators*, OECD, Paris.

OECD (2006), *Live longer, work longer*, OECD, Paris.

Okun A. (1975), *Equality and Efficiency: The Big Trade-off*', Brookings Institution, Washington.

Pisani-Ferry, J. and Sapir, A. (2006), *Last exit to Lisbon*, Bruegel Policy Brief, Bruxelles.

Putnam, R. (2002). 'Robert Putnam's response'. Austin American Statesman, Special Report – Cities of Ideas: Prosperity and its price, April 28, 2002. (http://www.statesman.com/specialreports/content/specialreports/citiesofideas/0428putnam.html.)

Putnam, Robert D. (1993a). *Making Democracy Work: Civic Traditions in Modern Italy,* Princeton University Press, Princeton, NJ.

Putnam, Robert D. (1993b). "The Prosperous Community: Social Capital and Public Life", American Prospect, Vol. 13, p. 35-42.

Putnam, Robert D. (1995), "Bowling Alone: America's Declining Social Capital", *Journal of Democracy,* Vol.6, p. 65-78.

Putnam, Robert D. (2000). *Bowling Alone: The Collapse and Revival of American Community,* Simon and Schuster, New York, NY.

Sapir, A. (2005), *Globalisation and the Reform of European Social Models*, Bruegel policy brief, 1/2005 Bruxelles.

Solow R. (1998), *Work and Welfare,* Princeton University Press.

UNCTAD (2004) *World Development Report*, United Nations, New York and Geneva.

Van Oorschot & Arts (2005), "The social capital of European welfare states: the crowding out hypothesis revisited",. Journal of European Social Policy, Vol.15, p. 5-26.

Völker, B., Derk, H. (2004), *Creation and Returns of Social Capital: A New Research Program*. Routledge, London.

Wolfe, A. (1989), *Whose keeper? Social science and moral obligation*. University of California Press, San Francisco.

Woolcock M. (1998), "Social capital and economic development: towards a synthesis and policy framework", Theory and Society, Vol. 27, p. 151-208

Chapter 9
Power to the People: Indicators for Accountability

Shailaja Chandra
Executive Director, National Population Stabilisation Fund, India

Abstract

In this article a In this article a definition of what constitutes "indicators" and "information" in the context of civil societies understanding has been delineated. A description of how a government-citizen partnership called Bhagidari coupled with the Delhi Right to Information Act 2001 empowered the citizenry of New Delhi, the capital of India and promoted a climate of participatory democracy has been presented. The paper describes how the power of information could alter citizen's lives and give them a public voice.

The article goes onto explore how in nearly half the 600 districts in India, the amalgamation of GIS maps and census data has enabled disparities and iniquities in the provision of health services to be placed in the public domain, in a visual, easily downloadable fashion. This internet based indicator tool is aimed at enhancing the understanding of all stakeholders viz., the affected public, media, academic bodies and non-governmental organisations to enable them to use the information to augur change. As the mapping exercise represents almost 94% of a country of 1.2 billion people, the visual data depiction is expected to widen understanding about the extent of inequity and to increase public accountability..

Section I

Participatory Democracy: How Indicators Gave Power to the People

Statistical data, surveys, projections, extrapolations, portrayal of inequities through inter and intra regional mapping, are all vitally important as **indicators**. Generally however international organisations, national Governments and statistical organisations provide the status of developments and the result of analysis through the release of reports, political speeches, press releases, seminars and discussions in which improvements (or lack of them) are projected. Couched in statistical terms they do not create an impact on the wider general public, because in their very nature such reports do not appear to relate directly to their concerns. Instead they are of use only to planners and

academics working within specific sectors. **Information** constitutes material which is generally derived using the outcome of research, media sources or increasingly by the public using the Freedom of Information Acts. This has the shortcoming of presenting a misleading picture unless it is qualified properly. The general public is not well versed in appreciating nuances or using the information to demand better performance.

Against these presumptions, Section I of this paper expounds on three important developments which took place in New Delhi, the capital of India, which depict how indicators and information empowered the public to hold government departments and even the political executive accountable.

Delhi is mega-city with a population of more than 14 million. Since, the city is divided into 7 parliamentary seats, 70 assembly constituencies and 272 municipal wards, political conflict and contending priorities are inherent in the situation. Being the seat of the central and state governments and with three levels of political representation, with 16 newspapers published from the city and a score of TV channels devoted to its affairs, it is a daunting situation for both political and administrative management. Arriving at a consensus is extremely difficult while pursuing government policies and in the absence of public participation, an unforeseen backlash can occur at any time, which can derail even good schemes and projects.

Since the political colour of the Members of Parliament, Members of Legislative Assembly and Municipal Councillors in New Delhi is generally different, the Government of Delhi needed to identify a way of involving the public directly, to create an understanding about who was responsible for what aspect of governance and service delivery. The government also realised the importance of giving the public a route to measure the responsiveness of the government to their concerns. Educating and involving the public was therefore of paramount importance for which they had to be first equipped with facts and data, information about processes, reporting levels and where responsibility for action finally lay. Likewise localised information on the performance of utilities, schools, hospitals and dispensaries, municipal services, upkeep of roads, parks etc. had to be first made available to evoke interest and involvement.

Bhagidari: A Government Citizen Partnership

Bhagidari in Hindi literally means partnership. This is an example how people's power catapulted over 3 levels of political representation and was responsible for the creation of more than 1 800 Resident Welfare Associations in New Delhi in a span of 3 years from 2000. The movement grew in strength as the public became aware of the systems and processes through which Government organisations could be accessed and also became conversant with how the provision of services could be demanded, as due. Collaboration with the citizenry was fundamental to address the existing and future issues affecting their quality of life. Thus the partnership was of mutual interest.

The *Bhagidari* process began with the organisation of large group meetings with a wide range of stake-holders – resident's associations, market associations, school standards committees, environment friendly groups held at district and sub-district levels as citizen's problems are generally localised. At the Apex level, meetings were organised by rotation, district wise every 2 months for 3 days at a time, in a large hall where 36 tables were occupied by 12 persons at each table – 6 government representatives and 6 resident's representatives. During the 3 day conclave, representatives at each table were give an opportunity to exchange lists of deficiencies and to listen to the response of the

public servants about the theory of resolving civic problems. By the third day the macro picture emerged and senior officers of the government responded to the main issues. This exposed the Residents Welfare Committees representatives to the working of the government departments and also gave them information about the intra-colony status of services, the intra-district funding available, future priorities, programmes on the anvil and the basis on which decisions had been taken.

In response the citizen's representatives could differ with the organisation, question arbitrary decisions and also highlight instances of "pick-and-choose" and favouritism. They could raise issues of immediate concern to the neighbourhood and expect a positive response in the presence of the organisational hierarchy of officials.

The *Bhagidari* movement used newspapers, documentation, reports, awards and incentives and introduced systems for sharing lessons from failures and successes as instruments for generating momentum, expanding the support base and encouraging collaboration. The movement has now covered more than 3 million people and has helped create a network of associations which represent a large spectrum of citizens needs. Although they have no elected status, the Resident Welfare Associations have gained prominence and credibility because of their numbers and ability to leverage attention. From early initiatives which were localised, the movement spread to include more important subjects like electricity and water distribution, solid waste management. It jolted the ordinary citizen out from a state of helplessness to which he had become a accustomed and created a public awakening that public services that needed to be demanded and obtained through collective action and effective interface with Government agencies. The fulcrum on which this movement rested was the provision of **information** which led to empowerment.

The *Bhagidari* initiative received the United Nation's public service award for improving transparency, accountability and responsiveness to public service categories in 2005.

Human Development Report for Delhi

When the *Bhagidari* movement was in full swing it was decided to collect information for a Human Development Report (HDR) for Delhi city (the first city HDR in the world). An independent survey was commissioned which extended over 14 000 households in Delhi in 2003-04. The respondents were asked to provide their assessment of twelve different services which included education, health, water supply, power supply, sanitation, roads, transport, environment, livelihood and housing. Women's status and safety and overall governance also featured in the survey. The methodology of the survey gave due weightage to residents of eleven types of colonies which exist in the city ranging from posh up market colonies to slum clusters from rural and urban villages to the old walled city of Delhi, from Government housing, to self-promoted colonies on private land. The respondents had to comment on the performance of various sectors in terms of services offered to residents and also identify the most pressing problems that needed attention. While 82% of the residents wished to continue living in the city, their main complaints related to water supply, garbage disposal, traffic, high levels of crime and public transportation deficiencies. The collection of these indicators and zone wise display of survey results gave credibility to the process and provided a direct voice to the citizens in an organised way. It also provided a basis for making demands and holding the political representatives as well as the bureaucracy accountable. The Report was prepared by academics and experts in operational research which fortified the sincerity of approach

and reliability about the results projected in the HDR. The results of the surveys are available on the website http://data.undp.org.in/shdr/delhi/completereport.pdf.

The HDR survey results gave credibility to the Bhagidari initiative and increased public participation.

The Right to Information Act in Delhi

Side by side, the Delhi Right to Information Act was introduced in 2001 which further gave people the right to ask questions about governance, systems, processes and decisions. The introduction of the Right to Information law coupled with the *Bhagidari* movement brought to the fore the reluctance within the bureaucracy to open files to public scrutiny. Since the Act allowed citizens to inspect public records, documents and works and even take samples of work, NGOs in particular became quick to use the legislation in creative ways. Data on the public distribution system, the food stock position in warehouses and its distribution showed that there was large scale diversion of essential commodities into the open market in some parts of the city. The Indian Express group of newspapers conducted a series of camps along with a local Non-government organisation based upon the information received. The Right to Information Act created an enabling environment because people were empowered to expose wrong doing and keep public servants under check and accountable. It gave people the direct authority to question what Government's organisations were expected to achieve and why there was an apparent shortfall in the outcomes.

Table 9.1 Status of Applications received under the Delhi Right to Information Act, 2001 and Disposed off Upto September, 2007)

	Number of Applications		Information	
	Received	Disposed Off	Given	Not Given
Total	11 557	11 036	10 067	969

Source : Department of Administrative Reforms, Government of National Capital Territory of Delhi.

How Giving Information can also Derail Decisions

The *Bhagidari* movement and the avenue of Right to Information worked as a bulwark for participatory democracy. The government in power was re-elected on the strength of giving the public a meaningful place in governance, but the outcomes have to be viewed in totality.

The Government had privatised electricity at a distribution end by 2002. The expectation was that this would improve efficiency. In 2005 when the public realised that electricity meters being supplied through private companies were running fast, and the electricity bills were spiraling, they rejected the electronic meters and began questioning the privatisation process itself. Using the platform of *Bhagidari* and by demanding information under the Right to Information Act, non-government organisations used the information made available under the Act, to motivate the media and the public to raise issues which forced the government to abandon the proposed privatisation of water distribution in the city. Information relating to power privatisation was used to draw attention to how expensive the decision to privatise water distribution could eventually become, thereby using the power of information to forestall a government decision that had already been taken. By exposing the inter and intra-city inequities in the supply of

water and government's intention to hand over the problem to private distributors an uproar was created, involving the media, academic bodies and NGOs. The result has been that the privatisation of water has been put off indefinitely because of people's capacity to impede the process, backed by data and information derived from knowledge of the system. This too was an outcome of the power gained under *Bhagidari* and the Right to Information Act.

Comment

For participatory democracy to be sustainable, there is no doubt the power has ultimately to be exercised through the ballot. Information can imbue a non-elected group of people with temporary power and authority to question government policies, schemes and the intention behind policies. The dilemma is that while civil society can exert counterveiling pressure on the domination of the state and an exploitative market, once it begins to exercise authority it can also challenge democratically established institutions. If the arbitration of what is right and wrong, important or unimportant, cost effective or sustainable, is to be determined by the public at large, it could have consequences on long term sustainability. Such civil society movements though powerful, have not been tested in the history of democratic countries. Never-the-less the power of information can make authoritarian systems accountable which point is underscored through the preceding examples.

Section II

Enhancing Accountability through G. I. S. Mapping and Census Data

Population growth is one of India's biggest challenges. India is a second most populous country in the world sustaining almost 16% of the world population on 2.4% of the land surface. The population of Indian states can be compared to the population of many countries. Unlike Europe and North-America, Australia, New Zealand, Japan, Thailand and Korea, India has the distinction of having an enormous growth of a young population for the next few decades. In comparative terms India is facing fewer challenges of dealing with an ageing population that many countries in the world face today. It can look forward to high productivity on account of a high proportion of the population belonging to the young cohort. By 2026, the population of India will rise by 371 million and share of 15-59 age group would be 64%. However, this cannot be an asset unless social indices like health, education, drinking water, sanitation improve side-by-side.

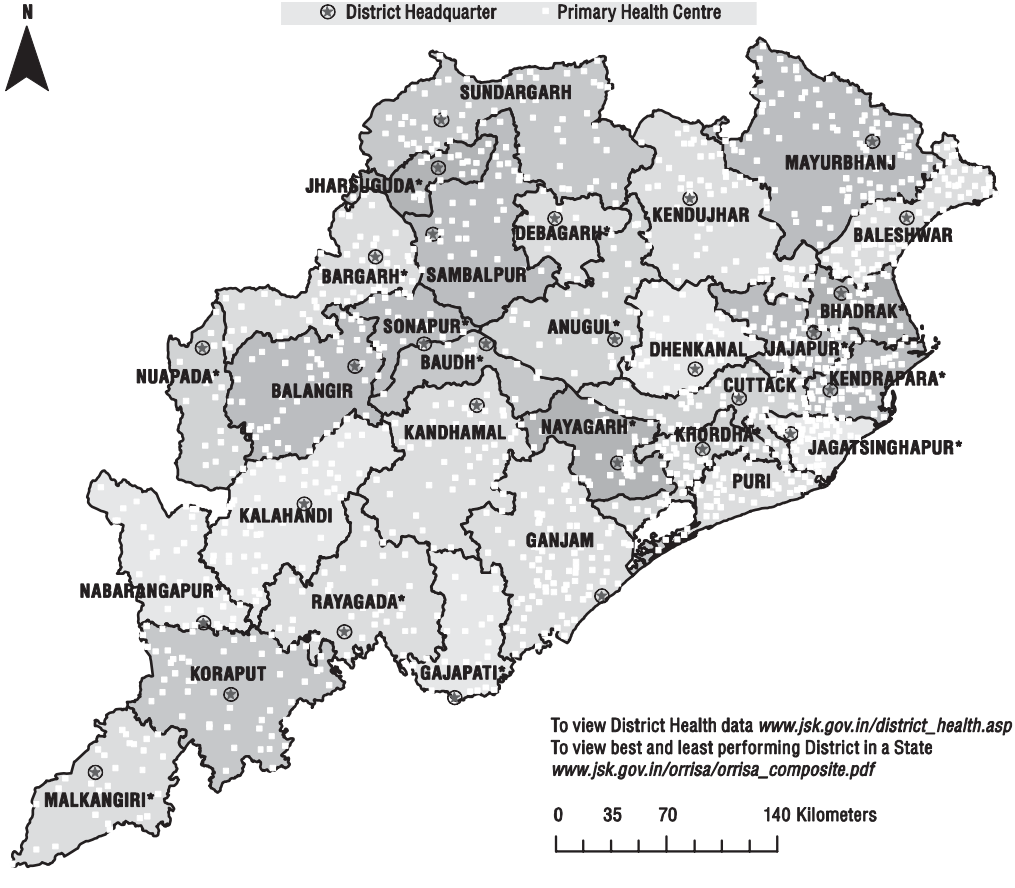
It was therefore felt that the provision of localised indicators and information could highlight larger issues and motivate Government authorities, the media and Non-Governmental Organisations to understand the bigger picture and what lies in store. There is every need to project statistical data in simple visual ways to empower the public. One of the most effective ways is to display inter and intra regional disparities is through maps, graphs and charts in a way that can create direct interest at a local level.

Drawing Attention to Disparity - The Power of Comparative Analysis

India is made of 28 States administered by elected governments and 7 Union Territories (administered through the Central Government). In an effort to provide information to the public, the National Population Stabilisation Fund of India (called Jansankhya Sthirata Kosh (JSK) in Hindi), prepared **State Level Health Facility** maps for 19 states (*One northern state (Himachal Pradesh) and seven North Eastern States were excluded as data was not available from the source – Survey of India)*.

The GIS based State Health Facility maps created by JSK cover 482 out of 600 districts in India.[1] Figure 9.1 gives the picture in respect of just one state of India (Orissa) in terms of spatial distribution of Primary Health Centres.

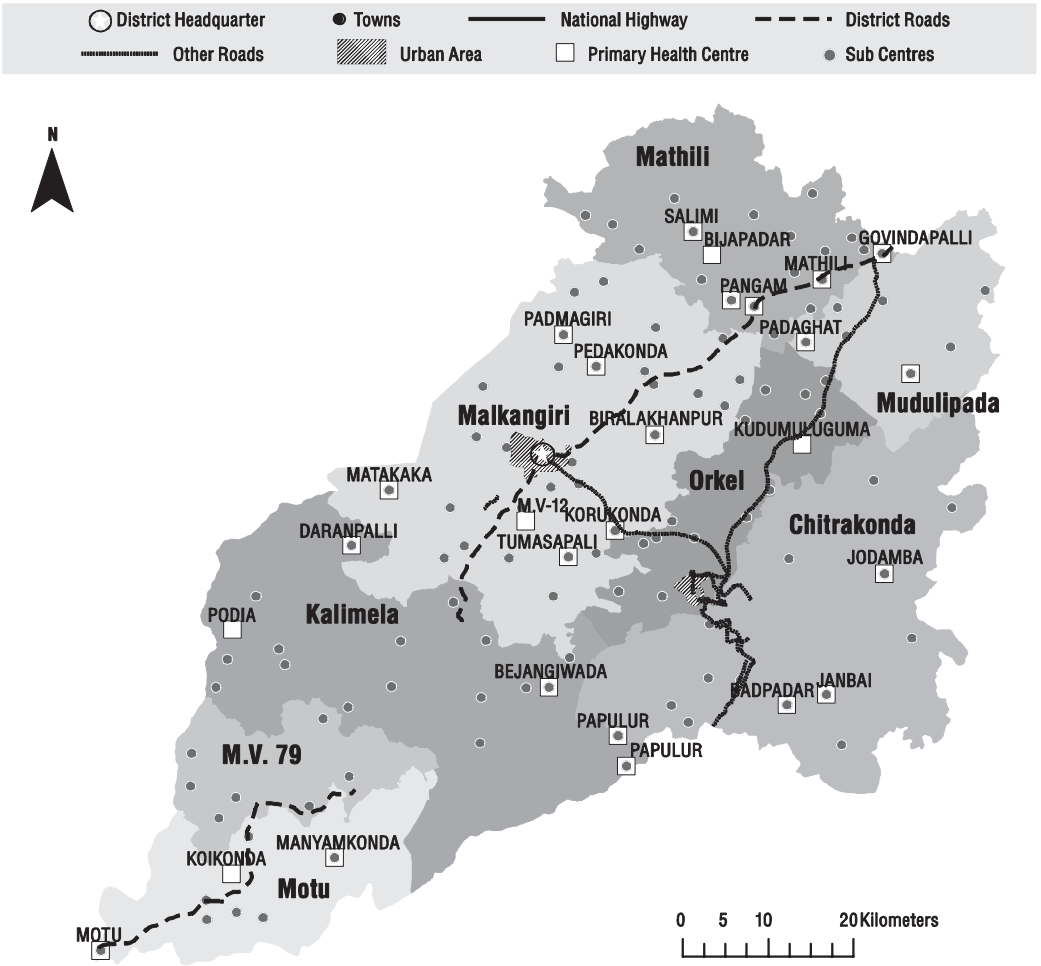
Figure 9.1 Spatial Distribution of PHC's in Orissa



From the state map one can move to the district map and view the picture of each district, its sub-divisions and the population of every village along with its distance from the nearest primary health centre. The maps highlight inequities in health facility coverage down to every village to enable resources to be targeted to underserved areas. The national, state district and other roads further exhibit the proximity not only to Primary Health Centre but also to surface communication in general. All this can be viewed alongwith rank of every district on JSK's website (www.jsk.gov.in).

The maps are an amalgation of census data and GIS mapping and they depict not only the disparities in access to health facilities in respect of every sub-division of 485 districts but also provide information on the distance people have to travel to reach a health facility. This information was sent on CD to the senior most administrator of each district (called Collector or District Magistrate), the Chairman of the District Development Committee, the member of the Legislative Assembly among others.

Figure 9.2 Health Facilities in District Malkangiri, Orissa



The information highlights the inter and intra district differences and disparities in coverage by health facilities. The census data has been used to show the distance of a village from a primary health centre alongwith the population of each village living 5 kms. away, 5-10 Kms and more than 10 Kms. away from the health facilities. The GIS mapping has been used to display the clustering of health facilities and the presence of large underserved areas. The map shown in Figure 9.2 indicates the clustering of facilities in Malkangiri District of Orissa State. The Distance chart at Figure 9.3 shows the distance to be traveled to reach a Primary Health Centre represents just one sub-division (taluka) out of eight sub-divisions of that particular district. There are 600 districts in the whole country.

Table 9.1 State - Orrisa: District – Malkangiri: Range from PHC's

	0-5 Kms.		5-10 Kms.		More than 10Kms	
Sub District	Village Name	Total Population	Village Name	Total Population	Village Name	Total Population
Mathili	Samarathaguda	471	Katepali	273	Saradhaput	631
	Badaguda	553	Bimanpali	405	Kianga	1 241
	Kansariput	826	Pedaguda	98	Jhulanibahal	124
	Badarengabeda	457	Kamarpali	1069	Badserpali	460
	Sunapadar	295	Bedurpali	461	Katapadar	112
	Saragiguda	713	Katapali	432	Kutunipali	923
	Brahmanguda	44	Galiaguda	438	Gojiagoda	297
	Padiarasi	185	Pakhanaguda	425	Kutilaguda	231
	Jamuguda	57	Daniguda	231	Biranpali	540
	Bariabahal	824	Meka	2015	Kamalapadar	199
	Dangarkhali	357	Salapapadar	255	Bengapani	33
	Daiguda	415	Anlajodi	48	Debhaguda	128
	Dhaudaguda	674	Talapadar	459	Rukniatal	104
	Sarangapali	353	Katrimajhiguda	534	Disariguda	98
	Bandhaguda	33	Sananayakguda	245	Kadhuapada	40
	Bhatiguda	192	Dangariguda	162	Bachalmeta	91
	Matiguda	38	Ramaguda	234	Sindhiguda	294
	Nuaguda	385	Kenduguda	363	Baghaangar	41
	Mundaguda	747	Duraguda	682	Netalgandhi	543
	Patraput	941	Karapali	771	Andhari	46
	Dokarijodi	34	Raniguda	63	Koprakut	23
	Durkajodi	2	Bhainsaghat	408	Tangirigadon	133
	Mendukuli	1 191	Purunaguda	666	Badaghat	48
	Dhungiaput	1 265	Sorgiguda	74	Bara	643
	Ambaguda	283	Singalbecha	279	Kerangajaba	64
	Patraguda	244	Godiali	531	Telekabecha	126
	Burudiguda	217	Kantabaunsaguda	15	Kartanpali	781
	Gundenguda	99	Gaudaguda	31	Dandipadar	299
	Madkamiguda	291	Durukajodi	540	Udegiri	779
	Kadamguda	501	Kadapadar	1034	Kopara	506
	Khadimati	562	Katiguda	367	Dangaras	127
	Bandamajhiguda	252	Samakunda-1	22	Sarangapali	832
	Podaguda	343	Timasput	893	Kosalamundaguda	75
	Siriguda	344	Sindhabeda	1369	Kangarabeda	210
	Baliamba	55	Sana Jhulanibahal	50	Samakunda-II	39
	Dumaliguda	145	Chedenga	1321	Nuaguda-II	42
	Kaudiguda	322			Chalanguda	466
					Udulibeda	569
					Patraguda	166
					Luler	668
					Sanoguma	301
					Durkaguda	197
					Dalapatiguda	843
					Tulasi	665
					Jigidipali	163
					Baliguda	252
					Kurumpali	1169
					Dombabeda	53
					Kandikiguda	98
					Madhiguda	25
					Damapada	1206
					Dalapatiputa	503
					Dusanadu	409
					Mahupadar	2549
					Adarladi	73
					Kukurkund	381
					Temurupali	1528
					Chamandaras	640
					Barubeda	504
					Kulabeda	423
					Dhakadarasi	325
					Muchukiras	210
					Chaulmendi	2682

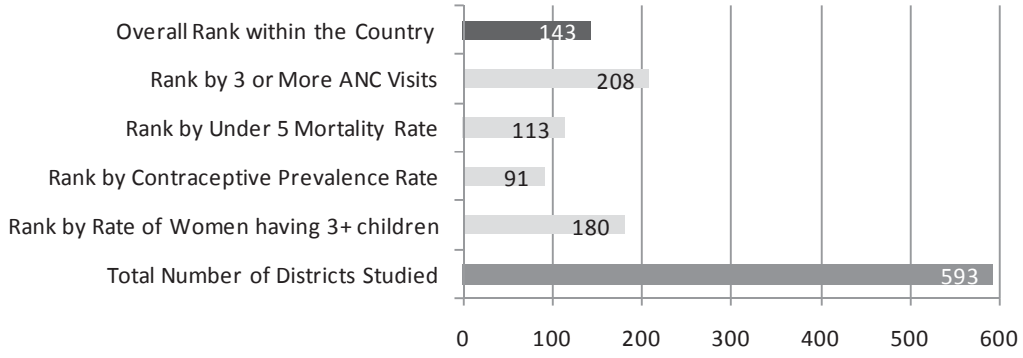
Table 9.2 Indices Covered in Study "Ranking and Mapping of Districts" for Composite Ranking and Composite Index

	Indices
1	Percentage of Population 0-6 Years
2	Birth Order Three And Above
3	Birth below Age 20
4	Complete Immunisation Coverage
5	Dropout from Full Immunisation
6	Female Literacy Rate
7	Households Using Safe Drinking Water
8	Households with Toilet Facility
9	Percentage of Electrified Households
10	Women Receiving 2 TT Injections
11	Women Receiving 3 or More ANC Visits
12	Under 5 Mortality Rate
13	Contraceptive Prevalence Rate
	Composite Index
	(Composite Index is the average of the above 13 indices)

Source: Ranking and Mapping of Districts, IIPS 2006.

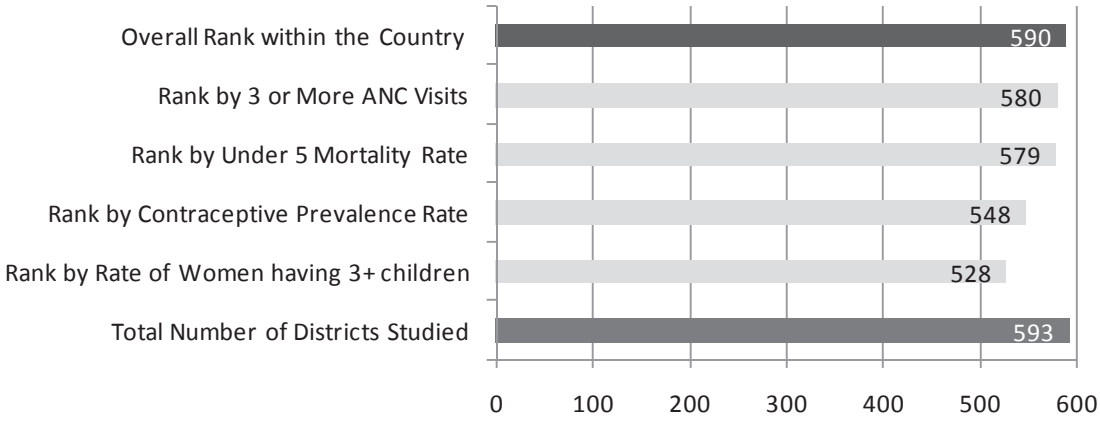
In 2006, the International Institute for Population Sciences India released a report titled "Ranking and Mapping of Districts – Based on Socio-economic and Demographic Indicators". The Report had given a composite Ranking to every district (Table 9.2). This information was converted into bar charts and the ranking based of the district in terms of the composite index was made available for scrutiny on the CD sent to the districts. The bar-charts shown in Figure 9.3 and 9.4 indicate the difference between the best performing and a lowest performing district within the same state. This is something people can relate to and raise questions about.

Figure 9.3 Best Performing District of the State according to the Composite Index : District of Indore (State – Madhya Pradesh)



All districts have not been covered due to unavailability of data for some districts.
Source: Ranking and Mapping of Districts. IIPS 2006 and DLHS-RCH State and Nutritional Report, 2006.
*IIPS: International Institute for Population Sciences, Deonar, Mumbai, is an autonomous insititution under the administrative control of the Ministry of Health and Family Welfare. Government of India. It Offers academic courses in the area of population studies.

Figure 9.4 Lowest Performing District of the State according to the Composite Index : Distric of Sidhi (State – Madhya Pradesh)



All districts have not been covered due to unavailability of data for some districts.
Source: Ranking and Mapping of Districts. IIPS 2006 and DLHS-RCH State and Nutritional Report, 2006.
*IIPS: International Institute for Population Sciences, Deonar, Mumbai, is an autonomous insititution under the administrative control of the Ministry of Health and Family Welfare. Government of India. It Offers academic courses in the area of population studies.

The National Population Stabilisation Fund has placed all this data on its website **http://jsk.gov.in** which can be accessed by public representatives, administrators, media and research bodies. The organisation has publicised the availability of this information through national newspapers and magazines so that a wider group of stakeholders can access it. An average of 100-150 persons a day access the website every working day.

The point to be underscored is that until now the availability of the data on health and social indices as well as the resident population was merely seen as statistical inputs. These have now been converted into visual form and made easily accessible on the website. This is expected to empower the public to question why the inter-district and intra district indices show so much variation and why availability of health services and accompanying health indices are better in certain districts as compared to others. It is expected that in planning new health facilities the underserved areas will now get prominence.

Conclusion

It can be concluded that indicators and information can be used to empower the public to hold their policy makers accountable. Were it not for access to information and indicators, the *Bhagidari* initiative may not have been able to gain the leverage it gained in the context of empowering citizens of New Delhi.

It is also expected that the exercise undertaken by the National Population Stabilisation Fund of India (called Jansankhya Sthirata Kosh in Hindi) to provide inter-state and intra-state, inter-district and intra district variations which highlight underserved districts, sub-districts and villages will lead to questions about low performance. Both the political executive as well as civil servants would be answerable for allowing apparent inequity to continue. It would necessarily require infusion of funds for upgrading

facilities and services since disparities would be evident down to every village. Since data as compiled by national organisations like the Survey of India and the Registrar General of India has been used, complaints that it has been selectively compiled to credit or discredit any particular segment would not hold true.

Although such data has been available for decades, the difference now is that technology has enabled it to be displayed in visual form capable of being down loaded and lending itself to analysis even by a layperson. By improving access to information, leaders can keep themselves abreast of developments and have a basis to profess how they are promoting the quality of citizens life and how progress can be measured.

The process would also help international donors and external agencies to evaluate where critical gaps exist. The data can be used to focus on under-served areas so enabling the funding to be targeted properly.

Notes

1 JSK has completed the exercise for 19 States in the country out of a total of 29 States which is available on the website www.jsk.gov.in.

The exercise includes a population of 960 million (out of a total of 1.2 billion) from the States of Bihar, Chhatisgarh, Jharkhand, Orissa, Madhya Pradesh, Rajasthan, Uttar Pradesh, Uttarakhand, Haryana, Jammu & Kashmir, Punjab, West Bengal, Goa, Gujarat, Maharashtra, Andhra Pradesh, Karnataka, Kerala and Tamil Nadu covering 482 districts, 4 645 sub-division and 579 779 villages in India.

Ten States have been omitted because village boundary data is presently not available with the Survey of India. The excluded states constitute 111 districts and a population of 68 million which is relatively small (6% only), given the size of the country.

Accordingly almost 94% of a large country of 1.2 billion people has been covered through JSK's exercise, giving a visual idea of the location of health facilities, alongwith names and the population of every village alongwith the distance to be covered to health facilities.

References

Annual Report of Public Grievances Commission, Government of National Capital Territory of Delhi

Bhagidari Cell, Office of the Chief Minister, Government of National Capital Territory of Delhi.

Delhi Human Development Report 2006: Partnership for Progress, Oxford University Press, New Delhi, 2006

International Institute for Population Sciences, *Nutritional Status of Children and Prevalence of Anaemia among Children, Adolescent Girls and Pregnant women*, International Institute for Population Sciences, Mumbai, 2006

National Informatics Centre, Ministry of Communications and Information Technology, Department of Information Technology, Government of India

National Population Stabilisation Fund, Ministry of Health & Family Welfare, Government of India.

Ram, F and Chander Shekhar, *Ranking and Mapping of Districts: Based on Socio-economic and Demographic Indicators*, International Institute for Population Sciences, Mumbai, 2006

Registrar General of India, Census Volumes, *Census of India 2001*, New Delhi

Survey of India and Registrar General of India data

World Bank, *Reforming Public Services in India: Drawing Lessons from Success,* 2006

Chapter 10
Connecting Communities with Community Indicators

Alex C. Michalos, FRSC[i]
Chancellor, University of Northern British Columbia, Canada

Abstract

This paper presents an initial attempt to answer the question: how can statistical indicators that define "the good life" and well-being be used to connect communities of diverse sizes, in order to build solidarity while embracing diversity? Using approaches and structures of the Canadian Index of Wellbeing System, the Community Indicators Consortium and the National Neighborhood Network of Canada, the author defines 12 qualities that indicators of well-being must possess to be effective. The author also identifies 21 critical issues that must be addressed in order to assemble a set of indicators or indexes that are coherent and comprehensive. Finally, a multi-level nesting of indicators is presented, starting with the individual and ending with the world community.

Introduction

The question I would like to begin answering in this essay is this. How can one connect communities of diverse sizes by means of a system of statistical indicators in the interests of building solidarity and embracing diversity? I say "begin answering" because I have only been able to see some salient aspects of a beginning. A complete and logically tidy vision is not yet clear.

The general structure and approach to answering the question before us is that used by the team of researchers building the Canadian Index of Wellbeing System (CIWS). The main sponsor of the CIWS work is the Atkinson Charitable Foundation. Since the team is fairly large (a few dozen people), we do not have complete agreement about everything, but we generally aim for consensus as our work progresses.

Besides following the structure and approach of the CIWS, my remarks are significantly informed by discussions with the governing Board and participants at meetings of the Community Indicators Consortium and the newly formed National Neighbourhood Network in Canada. There are levels of theoretical and practical

i E-mail: Michalos@unbc.ca.

expertise, and of enthusiasm and energy in the CIWS, CIC and NNN groups that are quite remarkable. Indeed, adding the new initiatives of the OECD to these grass roots organisations and those of the International Society for Quality of Life Studies promises a very bright future for work in the general field of quality of life, well-being and progress measurement.

Because our time is limited, I will make several assumptions immediately.

First, as explained in detail elsewhere (*e.g.*, Michalos 2003, 2005), it is assumed that a good life or a life that is qualitatively good may be measured in quantitative terms with statistical time series data, broadly referred to as social, economic and environmental indicators.

Second, it is assumed that the progress referred to in the title of this conference is progress toward precisely such a life.

Third, the communities referred to in our research question are primarily place-based, *i.e.,* communities of place, and secondarily communities of interest. Although there are frequently interactions among these two kinds of communities, it seems more likely that people sharing the same space would have common interests than that people having common interests would share the same space. The people assembled at this conference to discuss issues of common interests illustrate my point precisely, since we come from all parts of the globe. What is perhaps extraordinary about our meeting here is the fact that our particular common interest concerns issues related to connecting the earth's great variety of place-based communities by identifying additional common interests.

Fourth, it may be useful to distinguish what we might call *fundamental definitions* of 'community' from *derived definitions*. Fundamental definitions would have individuals (people) as their constituents, while derived definitions would have groups (collectivities) as their constituents. A fundamental definition of a 'place-based community' would be a network of people sharing some common place or space (*e.g.*, the city of Prince George, the Carney Hill Neighbourhood in Prince George, the province of British Columbia, or Canada), while a fundamental definition of an 'interest-based community' would be a network of people sharing some common interests (*e.g.*, dentists, city mayors, sociologists, book publishers). A derived definition of a 'place-based 3 community' would be a network of place-based communities (*e.g.*, the network of municipalities of British Columbia, the network of Canadian provinces, the network of countries in the European Union or the United Nations), while a derived definition of an 'interest-based community' would be a network of interest-based communities (*e.g.*, the Canadian Labour Congress or the Social Sciences and Humanities Federation of Canada).

Fifth, following Aristotle (1999) in the fourth century BCE, I think that a good life for an individual or community is a function of the actual conditions of that life and what an individual or community makes of those conditions. What a person or community makes of those conditions is in turn a function of how the conditions are perceived, what is thought and felt about those conditions, what is done and, finally, what consequences follow from all these things. People's perceptions, thoughts, feelings and actions, then, have an impact on their own and others' living conditions, as illustrated in Exhibit 10.1.

Sixth, the basic components of a good life from the time of Aristotle to today, across all cultures, are familiar to most people. In the ancient philosopher's terms, they included goods of the body (*e.g.*, good physical health, strength, vitality, bodily pleasures), goods of the mind (*e.g.*, wisdom, moral virtue and mental pleasures) and external goods (*e.g.*, wealth, security, good friends, loved ones, beautiful built and natural environments, and

good communities). People whose lives are characterised by such features have generally been regarded as happy, enjoying *eudaimonia* or "living well and doing well" in Aristotle's phrase, or briefly in contemporary terms, enjoying well-being.

Seventh, and finally, at least since the report of the Brundtland Commission (1987), the idea of a good quality of life has been essentially connected to that of sustainable development.

> *"Sustainable development", we are told in the Report, "is development that meets the needs of the present without compromising the ability of future generations to meet their own needs. . .The satisfaction of human needs and aspirations is the major objective of development. The essential needs of vast numbers of people in developing countries – for food, clothing, shelter, jobs – are not being met, and beyond their basic needs these people have legitimate aspirations for an improved quality of life. . .Sustainable development requires meeting the basic needs of all and extending to all the opportunity to satisfy their aspirations for a better life" (p.43-44).*

Acceptability Criteria and Critical Issues

Over some years of work on indicators of a good life, I have gradually produced lists of Acceptability Criteria and Critical Issues that must be addressed by anyone interested in constructing such indicators. Briefly, acceptable indicators and/or indexes should be:

Acceptability Criteria

1. Relevant to the concerns of our main target audiences
2. Easy to understand
3. Reliable and valid
4. Politically unbiased
5. Easy to obtain and periodically update
6. Comparable across jurisdictions and groups
7. Objective or subjective, or both
8. Positive or negative, or both
9. A constituent or determinant of well-being, or both
10. Attributable to individuals or groups of animate or inanimate objects, or all of these
11. Obtained through an open, transparent and consultative review process
12. Going to contribute to a coherent and comprehensive view of a good life or human well-being

Acceptability Criteria are not usually specified with great precision (*e.g.*, see Hagerty *et al*. 2001), but they provide useful guidelines for discussions and negotiations over particular indicators and indexes. Some judgment is required in their application in order to prevent the achievement of some goals at the expense of others, *e.g.*, measures that are easy to obtain and update must not be allowed to arrest the development of new measures

that might have greater validity and make a greater contribution to a comprehensive view of well-being.

Additional complications arise when one considers the array of Critical Issues that have to be settled in order to assemble a set of indicators or indexes satisfying the final criterion in the list, *i.e.,* contributing to a coherent and comprehensive set. Explicitly or implicitly, every indicator or index must be specified by addressing the 21 issues listed below. The issues are not categorised or presented in any particular order of priority or importance because no single principle or set of principles seems sufficient to provide a single rank order of the total set. Some are clearly logically prior to others in the sense that answers to some must be obtained before others, *e.g.*, one cannot assess causal relations (#17) or benefits and costs (#9) without knowing what sorts of things (#5) have been selected for measurement or evaluation. Other ordering principles would lead to different orders, *e.g.*, one might suppose that the most important question in the lot is who gets to assess the adequacy of answers to all other questions (#20), or perhaps, who is entitled to any consideration at all (#4). As well, it is unlikely that the issues in this list are mutually exclusive in pairs and collectively exhaustive. Nevertheless, for all their limitations, in order to identify and measure *a life, way of life or quality of life worthy of pursuit*, an individual or community must make decisions concerning the following issues.

Critical Issues

1. Individual well-being, group well-being or both: e.g., per capita incomes are inferred attributes applying to individuals, while unemployment rates are inferred attributes applying to groups.
2. Spatial coordinates: e.g., the best size to understand air pollution may be different from the best size to understand crime.
3. Temporal coordinates: e.g., the optimal duration to understand resource depletion may be different from the optimal duration to understand the impact of sanitation changes.
4. Population composition: e.g., analyses by language, sex, age, education, ethnic background, income, etc. may reveal or conceal different things.
5. Domains of life composition: e.g., different domains like health, job, family life, housing, etc. give different views and suggest different agendas for action. 5
6. Objective versus subjective indicators: e.g., relatively subjective appraisals of housing and neighbourhoods by actual dwellers may be very different from relatively objective appraisals by “experts”.
7. Positive versus negative indicators: negative indicators seem to be easier to craft for some domains, which may create a biased assessment, e.g., in the health domain measures of morbidity and mortality may crowd out positive measures of well-being.
8. Input versus output indicators: e.g., expenditures on teachers and school facilities may give a very different view of the quality of an education system from that based on student performance on standardised tests, and both may be very different from assessing whether the populace at large is becoming more literate, knowledgeable, educated and wise.

9. Benefits and costs: different measures of value or worth yield different overall evaluations as well as different evaluations for different people, e.g., the market value of child care is far below the personal, social or human value of having children well cared for.
10. Recipient populations: Who should be included as a recipient for particular benefits and costs?
11. Measurement scales: e.g., different measures of well-being provide different views of people's well-being and relate differently to other measures.
12. Research personnel: e.g., different stakeholders often have very different views about what is important to monitor and how to evaluate whatever is monitored.
13. Report readers: e.g., different target audiences need different reporting media and/or formats.
14. Aggregation function: e.g., once indicators are selected, they must be combined or aggregated somehow in order to get a coherent story or view.
15. Distributions: e.g., because average figures can conceal extraordinary and perhaps unacceptable variation, choices must be made about appropriate representations of distributions.
16. Distance impacts: e.g., people living in one place may access facilities (hospitals, schools, theatres, museums, libraries) in many other places at varying distances from their place of residence.
17. Causal relations: Prior to intervention, one must know what causes what (interaction effects), which requires relatively mainstream scientific research, which may not be available yet. At a minimum, correlations among variables should be explored with a view to discovering possible evidence of dependence or independence, redundancy and double-counting.
18. Discount rates: How much should one discount costs and benefits delivered some time in the future compared to those delivered today?
19. Confidence levels: What levels of confidence should one require to accept any particular claim or measure?
20. Auditors: Who should decide if any assessments are adequate or appropriate?
21. Auditing criteria: What criteria should be used to assess the adequacy of auditors' assessments, the adequacy of the procedures used for audits and even the adequacy of the answers to questions raised with the previous 20 issues?

The last question reveals the threat of either an infinite regress, a circular argument or an arbitrary end to analysis. Clearly, none of these options is very attractive, but it is in the very nature of foundational work that such a point must be reached. In any event, supposing that one had only two alternatives for each of the 21 Critical Issues (an absurdly conservative supposition), at least 2 097 152 different sets of indicators might be constructed. If nothing else, this suggests the size of the working space for indicator and index development.

Multi-level Nesting of Indicators

Given the particular question of this essay, our attention should be directed to the critical issues of individual versus group well-being (#1), spatial coordinates (#2), distance impacts (#16) and causal relations (#17). Assuming that we are seeking indicators in the form of statistical time series to connect diverse kinds of communities, Exhibit 10.2 illustrates an array of 8 kinds of entities that are usually nested within larger entities. The entities in the exhibit are not exhaustive of all those that might have been included and, since they are nested, they are not mutually exclusive. People often talk about places of work or places of worship, and statistical agencies routinely collect data on "establishments" of various kinds. Such places might be given an appropriate slot in a more robust exhibit. Apart from the smallest nest in the exhibit, that of individuals, all the other nests are place-based communities with diverse geographical, cultural and other component features.

As the exhibit clearly shows, individuals are multi-nested entities. Because each type of nest has some distinctive features and requirements for its own existence, multi-nested individuals typically have multiple roles to play throughout their own existence. Different nests produce different rights, privileges and obligations. For example, we are family members (fathers, mothers, sons, cousins), friends, neighbours, residents of some city, province and country. As fathers, for example, we have special obligations to our wives and children, as neighbours we have somewhat less distinct obligations, as residents of cities and so on, still other kinds of obligations as well as rights and privileges.

It seems to me that the most fundamental problem faced by people with an interest in community indicators is to precisely identify the particular nest or set of nests that they are thinking of when they think of community indicators, and then to clearly identify its or their essential features by working through the list of Critical Issues. Identification of the nest or nests is essential because of the diverse roles, rights, privileges and obligations that are attached to individuals and groups in the nests. It is impossible to begin taking stock of a community's strengths and weaknesses, opportunities and challenges, obligations and rights without community identification. Until there is common consent regarding community identification, there is no way to know who is or should be responsible or accountable for what, and no way to begin constructing a strategic plan for change or development. For example, homelessness will be addressed in different ways if it is regarded as a personal problem of some individuals, a problem of some neighbourhoods, some cities, provinces or the country as a whole. If people living in certain areas of a city do not perceive those areas to be neighbourhoods and do not think of the problem of homelessness as a neighbourhood problem, it will be impossible to address the problem at a neighbourhood level. Similar remarks apply to each of the kinds of nests named in Exhibit 10.2. At the broadest level, the so-called international or world community often fails to act (*e.g.*, to reduce global warming, the spread of AIDS or genocide) because the community and its interests are so vaguely defined. If these colossal problems are neglected for lack of a sufficient definition of a 'global community', it is difficult to be optimistic about the chances of global action for common measures of progress toward an economically, socially and environmentally sustainable future. Nevertheless, it is necessary to try to build a consensus for the development of such measures because, after all, we all inhabit a single planet with a limited capacity for growth and waste disposal.

In Exhibit 10.2, individuals are the smallest units of analysis. They are not places, certainly not in any geographical sense. Social scientists might suppose these "units" are

individual human beings, but biologists of various sorts might imagine individuals of other animal or plant species and other kinds of scientists might have still other paradigm cases of individuals. Regarding human beings, individual-level indicators for any place-based community would include things like simple census counts of total numbers of individuals, men, women, people of a certain age, born in a certain place, educated to a certain level, with an income of a certain amount and a Body Mass Index of a certain size. The place-based community might be as small as a single household or a remote village or a neighbourhood within a city, or as large as a province, a whole country or set of countries.

It is important to emphasise that even though social scientists might collect individual-level information about people (*e.g.*, Body Mass) without collecting any information about the place(s) occupied by those people, human beings are not disembodied spirits. So, they are always occupying some place(s) and the missing information would frequently, if not always, be relevant to the measures obtained. This happens all the time, for example, when psychologists gather information from students in diverse places. Depending on the issues being investigated, being in one class rather than another, or in one school or university rather than another, or in a particular family, neighbourhood and so on might be relevant, whether or not any such place-based information is collected. If such variables are just missing from the survey research instruments used, then it is impossible to test for place-based effects or interaction effects, or to undertake any kind of multi-level modeling involving place. Researchers' explanatory equations are then bound to be misspecified, as it so often happens, and their purported causal analyses are bound to be faulty.

Individual-level indicators are routinely aggregated in uncontroversial ways to form group-level indicators. For example, we might calculate the percentages of men and women in a certain place-based community, and find that 52% of the individuals are women and 48% are men. Since only individuals can be characterised as men or women while only groups can be characterised as having a certain percent of men or women, it is clear that the properties of being a man or women are distinctive of individuals while the properties of being a certain percent of men or women are distinctive of groups. There was a time when social scientists hotly debated the question of whether all social scientific data could be reduced to individual-level data or not, with those on the individualist side insisting on the possibility of reduction and those on the holist side insisting on its impossibility. We need not address this question here, but I have always been on the holist side (Michalos 1978).

Dwellings occupy the second level of Exhibit 10.2, typically "our places" for most people and most often family places. The tasks of taking account of people without dwellings (homeless people), in the simple sense of counting them as part of the official population and in the more profound sense of accounting for their very livelihood, usually do not fall on statistical agencies. Such tasks are routinely left to public or private welfare service agencies.

One might suppose that the aggregation of individuals into families is a relatively simple matter, but it is not. For example, for some years Statistics Canada had two definitions of a family. A *census family* was either a husband or wife couple, including common-law couples, with or without children, or a lone-parent with a child sharing a dwelling, while an *economic family* was any group of people, including all census families, sharing a dwelling and related by blood, marriage, common-law or adoption. Thus, the idea of an economic family was broader than that of a census family and both

kinds of families were essentially place-based communities of people, the places being the particular dwelling units occupied by those people. More recently, 'census families' have dropped out of usage, and we now have 'economic families' and 'unattached individuals'. For the purposes of their study of family violence, Poff and Michalos (1991) had to construct an even broader notion of family which they called a *socio-psychological family*. They defined such families as "economic or census families, or groups of individuals with self-avowed, personally and socially constructed, relatively intimate relationships without sexual or temporal boundaries" (p.155). Socio-psychological families must also occupy some place, which may differ from that of particular economic families.

If one went no farther down the pyramid of nested entities in Exhibit 10.2, one would still have to address the issues concerning the interactions and causal relations among individuals and groups occupying particular dwelling units. After all, individuals with different genetic and learned capabilities and attributes act and react in different ways in groups with different capabilities and attributes. So, what resources and constraints are present or absent, and what it might be necessary to provide or remove in the interests of improving the quality of the lives of individuals and groups can only be determined by individual and group-level research, with some notion of family or at least occupants of a single dwelling providing the group-level unit of analysis. Thus, if one went no farther down the pyramid of nested entities, it is obvious that the chances of appropriately measuring and understanding only the socio-psychological dynamics affecting the individuals and families would be very small unless one engaged the expertise of multi-disciplinary and trans-disciplinary teams of researchers. Nevertheless, because the basic building blocks of human communities are human beings, one can be sure that the Aristotelian core of internal and external goods will merit attention, be capable of measurement, periodic monitoring and some deliberate manipulation in the interests of improving the quality of life.

The term 'neighbourhood' in the third nest of Exhibit 10.2 is intended to do double duty, designating a set of contiguous streets and/or blocks with residents whose children often share a single public school as well as a set of contiguous rural properties. While the former idea and situation is probably familiar to more people because most people live in relatively urban areas, people living in sparsely populated rural and remote areas also think of themselves as having neighbours. That is, residents of urban and rural or remote areas all have neighbours sharing a common place, only those in urban areas might be just around the corner, across the street or upstairs while those in rural and remote areas might be several acres or even hundreds or thousands of acres away. That is why when we speak of neighbourhoods as place-based centres of community activities, we must be sensitive to the fact that the geography of the places occupied by some communities can be the most important feature of those communities, the greatest asset or liability for mobilising community action.

It is worth mentioning that there is no more agreement over definitions of 'urban' and 'rural' than there is over 'community' and 'neighbourhood'. Statistics Canada's concept of rurality is a function of population density and distance from an urban centre. So, for example, "one's measure of rurality would change, depending upon whether one was considering access to a monthly ballet performance or access to a curling rink or access to a sizable number of organic restaurants to market your organic farm products" (Bollman 2007, p.7). Since there is very little information about the size of most people's living space, how far they are willing to and actually do travel to access shops of various kinds (groceries, clothes, household appliances, automobile services), leisure activities

(theatres, restaurants, arenas, golf courses, parks, ballparks, museums), health care (clinics, hospitals, medical specialists and general practitioners), education and information (public and private trade and other kinds of schools, universities, libraries), it is very hazardous to assess the quality of people's lives by simply counting kinds of facilities and services available within a relatively well-defined political jurisdiction (*e.g.*, within a particular town or city). For example, when I moved from the city of Guelph, Ontario to Prince George, British Columbia, both cities had populations of 70 to 80 thousand people, a university, hospital and symphony. However, Guelph is about a one hour drive from downtown Toronto and the Hamilton-Niagara region, giving Guelph residents access to practically all the facilities and services of these two larger communities, while the closest relatively large city to Prince George is Edmonton, Alberta, at least a 7 hours drive away. Residents of Toronto can escape the city congestion and enjoy the theatres and more leisurely paced lives of the small towns of Stratford or Niagara on the Lake, Ontario by simply driving a couple hours west. Clearly, the so-called 'spillover problem' that researchers encounter when they try to assess neighbourhood facilities and services is a problem for researchers assessing the facilities and services of communities of all sizes in most places. Even in a relatively small city like Prince George, road maintenance is sometimes primarily the responsibility of the municipal government and sometimes the provincial government, depending on which road. A local survey respondent asked to approve or disapprove of the government's performance on road maintenance would seldom know which government is praise-worthy or blame-worthy.

The fourth nest in Exhibit 10.2 is set aside for villages, towns and cities, settlement areas which seem to have no universally agreed upon definitions. This is also true of metropolitan areas, our fifth nest. Although the diversity and complexity of facilities, services, institutions and organisations tends to increase with the populations of such settlement areas, there are no new Critical Issues so far as quality of life, well-being or progress measurement are concerned. I would not under-estimate the significance of the additional diversity and complexity introduced by great increases in population or population density in particular places. However, the changes that occur with such increases seem to me to be differences of degree rather than differences of kind. Villages, towns, cities and metropolitan areas are communities of communities insofar as each of the larger settlement areas contains smaller settlement areas. Each of the larger places contains smaller places, with diverse kinds of interactions and causal relations among the individuals and groups in the diverse places.

I mentioned the Community Indicators Consortium earlier and there will be a presentation at this conference focused specifically on the work of this group. However, Exhibit 10.3 has been included to briefly provide an overview of the sort of information, knowledge and wisdom generated at the most recent meeting of the group in Jacksonville, Florida in March 2007. Participants in the meetings of this group are primarily interested in the development of indicators and indexes, tools of measurement, advocacy and change for communities of the fourth and fifth levels of Exhibit 10.2. Hopefully, this essay will make it easier to see connections between these two levels and all the others.

It might be thought that differences of kind would appear at the sixth nest, involving relatively large political jurisdictions like provinces or states, or ecological water sheds or air sheds. But such a view does not seem warranted. The province of British Columbia, for example, has two large river basins, for the Columbia and Fraser Rivers. The largest of the two, the Fraser River runs approximately 1 400 kilometres from Mount Robson in the Rocky Mountains to the Pacific Ocean at the Straight of Georgia. About 2.8 million

people live in the basin, which is two-thirds of the province's population. According to the most recent "state of the basin" report of the Fraser Basin Council (2007, p.2), "The Fraser Basin is a special place". The Basin supports

> *". . .six salmon species, including steelhead, and 65 other species of fish. . .British Columbia's most productive waterfowl breeding area, home to hundreds of species of birds and mammals as well as reptiles, amphibians and insects. From Prince George to Williams Lake to Kamloops, and throughout the most populated stretches of the Fraser Valley and Greater Vancouver, communities depend on the Basin to support a range of economic activity –from natural resource industries, to agriculture to businesses of all types. This is the land where we live, work and play. Our well-being is so closely tied to the Fraser Basin that its future is our own".*

Such is the diversity and complexity of the Basin, that to make a reasonable assessment of the "state of the Basin", the total area covered has been divided into 5 regions with a common set of indicators applied to each region. Prince George is the largest city in the most northern region of the Basin, a region that enjoys the best water quality and the worst air quality of all 5 regions. Regardless of the fact that the Basin occupies much more space and has a considerably greater variety of natural species than, say, the metropolitan area of Toronto, measuring the quality of life in the Basin does not seem to introduce new Critical Issues. One would expect greater variability in the air and water quality across the length of the Basin than across the length of the metropolitan Toronto area, but not necessarily different kinds of indicators of air and water quality. Sampling techniques and quality standards that would apply to air and water in the Basin would apply in the metropolitan area. Measures of family violence, mortality and morbidity, educational attainment, employment, income and so on would be as relevant to the people in the smaller communities and neighbourhoods within communities within the "special place" known as the Fraser Basin as in the place known as metropolitan Toronto.

For a huge country like Canada, diversity and complexity significantly increase as one moves to the seventh level or nest, the whole country. Unlike the States of the United States of America, Canadian provinces have constitutionally guaranteed jurisdictional rights to the natural resources found in the provinces, *e.g.*, the oil and gas of Alberta is primarily a provincial resource and source of tax revenue. Apart from constitutional provisions, the geography of the country creates its own divisions. For example, to address the issue of a sustainable fishing industry in Canada, one must consider problems of the Pacific herring and salmon on the coast and inside of British Columbia as well as problems of the cod fish across the continent to the coasts of Newfoundland and Labrador. Forestry is vitally important to the economy of British Columbia but not to Manitoba. Air pollution is a significant problem in Toronto, but not a problem in Nova Scotia.

The community of communities that is Canada is not only vertically laced with many levels of nested communities, but horizontally laced with distinct geographic features that individually and collectively make distinct as well as causally interacting contributions to what we call 'the Canadian Mosaic'. A comprehensive assessment of the progress of the Canadian community over time must somehow take account of this great diversity and complexity, and so far we have not developed a vocabulary, much less a general theory to allow us to make such an assessment (Michalos 1997; Michalos *et al.* 2007) .

Additional complications arise at the eighth level, the level concerning international relations, whether they are bi-lateral (*e.g.*, Canada-United States Free Trade Agreement), tri-lateral (*e.g.*, North American Free Trade Agreement) or multi-lateral (*e.g.*, World Trade Organization). Such agreements create new opportunities and constraints on a country's capacity to act in the interests of national progress, and additional problems in measuring and accounting for progress (Michalos, in press). Then, of course, such problems are multiplied again when one moves to the global level, the nest of all nests. In a recent issue of *Prospect* magazine, Cooper (2007, p.24) wrote,

"In the 21st century, the new forms of communication have brought us a new world and we need a new constitutional form too. The big question is how to organise this world in which politics and identity are national, but we can survive and prosper only if we act internationally. It is fine to talk about "the international community," but who is it and how can it function?"

As already suggested, the form and functions of the international community are yet to be defined (constructed), but conferences like this one are helpful. Hopefully, so are observations on the multi-leveled forms and functions of communities of necessarily diverse and smaller sizes.

References

Aristotle: 1999, *Nicomachean Ethics*, T. Irwin, second edition trans., Hackett Pub. Co., Indianapolis.

Bollman, R.D.: 2007, *Factors Driving Canada's Rural Economy*, Agriculture and Rural Working Paper Series, Cat. No. 21-601-MIE-No.083, Statistics Canada, Ottawa.

Brundtland Commission: 1987, *Our Common Future*, World Commission on Environment and Development, Oxford University Press, New York.

Cooper, R.: 2007, Untitled, *Prospect*, March, p.24.

Fraser Basin Council: 2007, *2006 State of the Fraser Basin Report: Sustainability Snapshot, Inspiring Action 3*, Vancouver.

Hagerty, M.R.: 2001, "Quality of life indexes for national policy: review and agenda for research", *Social Indicators Research*, 55(1), p.1-96.

Michalos, A.C.: 1978, *Foundations of Decision-Making*, Canadian Library of Philosophy, Ottawa.

Michalos, A.C.: 1997, "Combining social, economic and environmental indicators to measure sustainable human wellbeing", *Social Indicators Research*, 40, p.221-258 (reprinted in Michalos, 2003).

Michalos, A.C.: 2003, *Essays on the Quality of Life*, Kluwer Academic Publishers, Dordrecht.

Michalos, A.C.(ed.): 2005, *Citation Classics from Social Indicators Research* (Springer, Dordrecht).

Michalos, A.C.: in press, *Trade Barriers to the Public Good: The NAFTA and AIT Cases Concerning MMT*, McGill-Queen's University Press, Montreal.

Michalos, A.C., A. Sharpe and N. Muharjarine: 2007, *An Approach to a Canadian Index of Wellbeing* (Working Paper prepared for the Atkinson Charitable Foundation, Toronto).

Poff, D.C. and A.C. Michalos: 1991, *Family Violence: A Literature Review*, Institute for the Study of Women, Mount Saint Vincent University, Halifax.

Exhibit 10.1 General Quality of Life Model

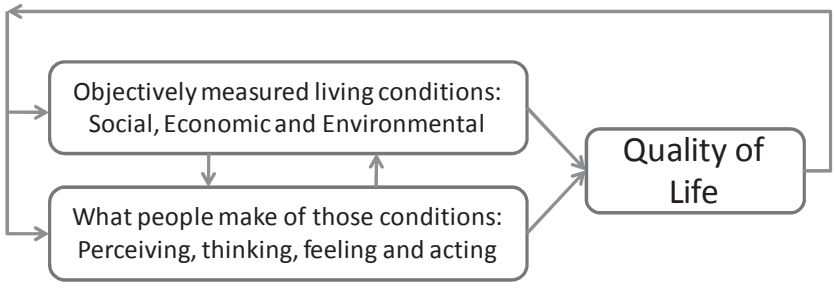


Exhibit 10.2 Multi-Level Nesting of Indicators

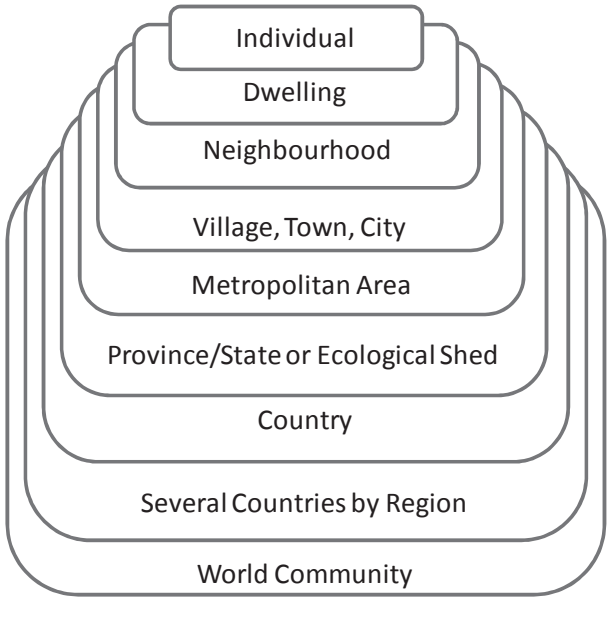


Exhibit 10.3 Some Results from the CIC Jacksonville Meeting: Lessons Learned, March 2007

Aim to build a culture of evidence-based policy making

Be as inclusive as possible and find the best talent

Inclusiveness is more important than speed

Expect mistakes, be patient, be very patient

The process is at least as important as the product

Let media and friends know how much they are appreciated

Agency responsible for reports needs recognised integrity

City level reports can influence state/provincial, federal gov.

QOL reports provide vision and allow goal setting

Benchmarking is important

Results-based accountability is important

QOL is connected to place and to heritage

Facts and perceptions are important

Different communities have different champions

Produce shorter and longer versions of reports

Use diverse media and means to communicate messages

Pictures are better than numbers

Provide user training sessions for diverse potential users

Identify key change agents for diverse kinds of changes

Allow elected and appointed officials to do their jobs

Undertake evaluations to find what works or does not work

Recognise that communities are not homogeneous

Don't try to do know everything or do everything

Focus on something, identify critical variables for action

From Bishop Tutu: Mobilise, mobilise, mobilise

From trade unions: analyse, organise, educate, agitate

Part Four
Measuring Progress: People's Perceptions and Knowledge

Chapter 11
People's Perception vs. Reality: Drawing Lessons from the Use of Statistics in the European Union

Joaquín Almunia
European Commissioner for Economic and Monetary Affairs

Ensuring that quality statistics feed into European policy making is, for me, a principal concern. Indeed, for a politician, the fundamental value of good statistics is clear. Sound statistics form the core of democracy. They justify our actions and document the results of these actions. They provide the transparency and accountability crucial for any democratic system.

Statistics are also vital decision making tools. We cannot consider what to do, how to act or take decisions without access to information that is as objective as possible.

In light of these considerations, it is clear that ensuring the quality of our statistical data is paramount. Accurate, reliable and timely statistics, produced by credible and independent institutions, underpin policy making. They support the legitimacy conferred by citizens upon their representatives.

In this way, high quality data is a public service. And as a public service, official statistics should be accessible and comprehensible. Good communication is crucial. And yet, I realise that the message statistical offices convey is sometimes differently perceived and interpreted by citizens. I will return to this topic in a moment, using the example of inflation misperceptions that followed the introduction of the euro in 2002.

The question of public perception is very relevant when we consider how to measure the progress of societies. Official statistics play a central role here and yet they are by no means perfect. Thus we are constantly developing the statistics we use to improve our understanding of this issue.

For instance, GDP was never intended to provide all-encompassing analysis of the progress of societies, rather to measure the economic wealth created by a market economy. It is a very important statistical indicator that is under permanent review, seeking always to more accurately capture economic reality. We are working alongside other major countries to further improve the GDP measure, in the framework of a future revision of the UN-based System of National Accounts.

But whether the statistical information we use is adequate to comprehend concepts such as economic and societal development, or social welfare, remains in question.

Combating climate change and striving for minimum standards of welfare for all are crucial objectives. And I stress that the use of indicators, such as those agreed as

Millennium Development Goals, have proved to be very powerful tools. They raise awareness and fuel public debate on how to address these global challenges.

However, GDP fails to cover more complex concepts, such as individual welfare. It brushes over important considerations like environmental damage and social inequality, and overlooks the long-term financial sustainability of pensions and liabilities.

Thus I share the objectives set down in the Istanbul Declaration and can agree on the need to build a more nuanced and therefore accurate understanding of economic and societal progress. This could involve enlarging the coverage of statistics to include indicators of social and environmental as well as economic outcomes.

But policy makers have to take important decisions on a daily basis. Therefore we need to strike a balance between improving measurements of societal progress in the long term, while continuing to make sure that we have appropriate, reliable and accurate statistical tools in the short term. This can be a challenging task when, as in the EU, decisions apply to a number of different countries.

And this task has become even more demanding because the role of European Statistics has been evolving in recent years. When Eurostat was set up in 1953, it collected, harmonised and disseminated statistics, providing information on the economy and societies of the six founding Member States. However, over time, new developments in EU policy and progressive enlargements to a more diverse set of countries meant that the Union required hard data, subject to ever more demanding quality standards.

The Treaty of Maastricht that paved the way for the introduction of the euro in 1999 is a prime example of this evolution in policy. To qualify for the euro, Member States had to converge towards agreed, quantified criteria in terms of inflation rates, interest rates and budgetary debt and deficit.

Until then there had been no common definition of a Consumer Price Index or of debt and deficit. So statisticians throughout Europe toiled long and hard to provide the agreed statistics and indicators that would underpin fundamental political decisions involved in setting up the Economic and Monetary Union.

As a result, the European System of National Accounts and the Harmonised Index of Consumer Prices have now developed into technically sound instruments that support important political decisions.

And the ambitious project of Economic and Monetary Union has proved a success, thanks partly to the effective management of quantified objectives, such as the reference values for inflation, government deficit and debt ratios.

It has helped us to create a strong and stable macroeconomic framework for EMU that shields members from external shocks, keeping inflation and interest rates at consistently low levels.

We are now extending the use of statistics to other important areas as a form of "soft" governance. For example, structural indicators covering domains like employment, research and social cohesion provide an objective instrument to measure progress towards the goals of the EU's Lisbon Agenda, our strategy for reforming the European economy. We are also developing a set of indicators to monitor, assess and review the EU's Sustainable Development Strategy.

Thus our use of statistics has evolved from an information tool that facilitates decision making, to become a key to the decision making process itself. Naturally, this

places increased scrutiny on their reliability and quality and brings the issue of public trust to the fore.

One visible example of public trust in official statistics being put into question can be taken from the euro area. Here the mismatch between consumer perceptions of inflation compared to the actual rate of inflation have become a concern.

After the changeover to euro notes and coins in January 2002, inflation perceptions rose sharply in all euro area countries. This increase was not in line with actual consumer price developments.

It is true that some unjustified increases in the price of certain frequently-purchased goods and services did take place around the changeover. But the largest part of the increases can be explained by normal inflation patterns and some special non-euro factors such as energy price developments.

The statistical data clearly indicate that the changeover effect on prices was rather limited. However, since 2002, the notion that the euro causes prices to increase has taken hold in many countries.

From a macroeconomic viewpoint, any overestimation of inflation by consumers can have a negative impact – for instance on consumption decisions and on inflation expectations – and may trigger unwarranted wage demands.

From a political viewpoint, high inflation perceptions can affect the support for the single currency among euro area citizens. But it can also affect public trust in our capacity to accurately measure consumer price developments and hence the credibility of monetary policy for the euro area. More generally, misperceptions such as these can lend weight to unfair criticisms that aim to attack the political legitimacy of public authorities and their ability to meet the expectations of citizens.

We constantly strive in Eurostat, together with the ECB, to ensure that the HICP accurately captures the complex reality of inflation in our economies. But this is only worthwhile if at the same time, we can make the HICP understood to every citizen as one of the most widely recognised macroeconomic statistics. Clear and concise communication on inflation measurement is essential and we are currently working to streamline communication and explanation of the HICP.

Euro area inflation is just one example that shows how important it is for people to understand the key economic facts that underpin their societies. This does not only help them take well informed decisions. But more generally, it is essential because in the absence of reliable statistics, genuine public debate cannot take place, and may instead fall victim to misinformation and ideology. Such an outcome can ultimately distort democracy and policy making. This is why it is important to monitor public perceptions of statistics and to improve our communication of data.

Another way to combat misperception and public mistrust is to ensure that our statistical systems remain fully credible, efficient and independent. With these lessons in mind, we are also working to reinforce the credibility and public acceptance of our European Statistical System.

Constructing good cross border statistics is particularly challenging in the EU. Our European Statistical System consists of 27 EU Member States. But their harmonisation extends well beyond EU borders, to include the four EFTA countries, three candidate

countries and five potential candidates in the Western Balkans, who are increasingly adopting European legislation in statistics.

Add to these the 16 members of the European Neighbourhood policy and one could argue that an enlarged European Statistical System consists of 55 countries – over one quarter of the total membership of the United Nations.

This stresses the importance of the European Statistical System as a key player in the global statistical system. It bears a heavy responsibility in setting standards and guidelines both on procedures and methods as well as on the quality of statistics.

Given the growing influence of the European Statistical system, I am committed to modernising and improving its governance structure and to enhancing its quality standards.

In 2005 we adopted the European Statistics Code of Practice which lays down 15 principles on the independence, integrity and accountability of national and Community authorities. This code is self regulatory: each member state has completed a self-assessment questionnaire on its implementation and this is currently being complemented by a round of peer reviews. For the sake of transparency, these results will be posted on Eurostat's web site.

In this way, the governance structure is better ensuring that EU statistics are independent and that the trust of EU citizens in these statistics is secured.

Let me conclude by stressing again that statistics are a fundamental instrument of policy making. But if statistics are to fulfil their democratic function, objective and high quality communication to the general public is needed. More than ever statistical information is accessible for all types of users. The growth of the internet is opening up the world of statistics to a broader public and is increasing mutual understanding. However, in this environment of greater transparency, agreed rules and standards on quality of data are crucial.

Chapter 12
What U.S. Consumers Know About Economic Conditions[i]

Richard Curtin
Research Professor, University of Michigan, United States

Abstract

This paper presents the findings of a University of Michigan survey on what Americans know about several key aspects of the US economy, and why, based on data collected in the United States in April and May 2007. This survey was part of the first co-ordinated international survey on what citizens know about key economic statistics (the other corresponding half of the survey was conducted in 29 European countries by Eurobarometer). The American survey gauges respondents' knowledge of three major economic statistics: unemployment, inflation rate and the growth rate of the overall economy. Results are then correlated to demographic data (including education, age, income and gender) via regression analysis. The sources of economic information are also correlated to responses, to ascertain the information determinants of knowledge and accuracy in people's reports about official economic statistics. Also analysed is coverage of official economic statistics in the mass media and people's informal knowledge of economic performance

Πάντες άνθρωποι του ειδέναι ορέγονται φύσει. Aristotle
(All men by nature desire knowledge.)

Introduction

The natural inquisitiveness of people was noted long ago when Aristotle began his book Metaphysics by saying "All men by nature desire knowledge." The acquisition of information about economic conditions has been a common facet of life since the dawn of civilisation. Even before governments devised economic accounts, people utilised information about economic conditions in their decisions about bartering and trade, borrowing and lending, and other economic decisions. Simply because it was not measured by an official statistical agency of the state, there is no reason to doubt that people sought out such information.

i E-mail correspondence to Curtin@umich.edu. Thanks to Phippe Blum and Alex Mirkin for reaserch assistance. This paper was originally presented at OECD World Forum on Statistics, Knowledge and Policy, June 2007.

In those earlier times, however, there was no reason to expect all people would arrive at the same estimate of relative prices or the same rate of economic growth. Information was scarce, costly to obtain and process, and there were no accepted paradigms to interpret the data. The significant innovations of national income and product accounts, including data on conditions in labor and product markets, reduced the cost of economic information and increased its quality and timeliness.[1] Information on economic conditions is now often viewed as nearly costless to obtain since people are constantly exposed to official government statistics by the mass media. Indeed, it is argued that exposure is so complete that people would have to actively ignore the available information to be unaware of economic data. Unlike in earlier times, there is reason to believe that everyone would now be fully aware of the official data on the key indicators of the performance of the national economy.

Unfortunately, that is not the case. Nearly every profession has been disappointed with the amount of knowledge ordinary citizens possess, whether they are political scientists, physicians, mathematicians, physicists, or economists. It is an all too frequent occurrence that some survey finds that a surprisingly high proportion of people could not name their representative in the legislature (Delli *et al.* 1996), have accurate knowledge about common medical conditions (Lucas 1987), correctly know about planetary orbits (Lucas 1988), how to do rather simple arithmetic operations (OECD 2006), or know the current rate of inflation or unemployment (Blendon *et al*. 1997; Blinder & Krueger 2004).

How can the often widespread lack of knowledge be reconciled with Aristotle's view that people naturally desire knowledge? Or was Aristotle simply wrong? Plato was sceptical about our ability to know the absolute truth, and suggested that people can only dimly perceive the truth from its shadows. Policy makers may find Plato's allegory compelling, as they often feel trapped in a cave making decisions based on shadows of truths that last only as long as the next data revision. The implications of Plato's views cut more deeply, however.

The assumption that people utilise official sources of economic information reflects the widespread tendency toward the reification of economic data – that is, treating conceptual measures as if they had a concrete existence.[2] It should be no surprise that few people think in terms of GDP in chained 2 000 dollars, or even know what that concept signifies. It does not follow, however, that people do not actively seek and use information about the performance of the national economy in making their economic decisions. While economists justifiably have strong preferences for the measures that they have developed, it does not follow that all people should adopt that same set of indicators. Indeed, economic theory suggests that people adopt whatever measures prove most useful, considering both the costs of obtaining the information and the benefits that can be achieved by more informed decisions.

For most consumers, the most relevant information is about their specific situation. In contrast, the concepts of Gross Domestic Product, the Consumer Price Index, and the unemployment rate, for example, were devised as appropriate measures of conditions in the macro economy. Rather than macro data, people's decisions are typically based on a different information set, namely the strength of the local economy, the change in the prices they actually face, and job prospects for people with their same skills and abilities. Of course, when aggregated across all communities and people, the data would approximate the national averages that economists favor.

The measurement of people's knowledge about the performance of the economy does not depend on a direct translation of the economist's definition into survey questions, but on the ability of the resulting measure to accurately reflect the underlying economic concept. This is more than simply rephrasing questions to use "population language" or decomposing complex economic variables into more manageable measurement objectives; it also means constructing questions that are based on the same conceptual framework used by ordinary people in making their economic decisions.

This paper will document what people know about several key aspects of the economy based on survey data collected in the United States in April and May of 2007. The data suggest that most people do not know the exact figures contained in the official releases on the rate of unemployment, the rate of change in consumer prices, or the rate of growth in the overall economy. To be sure, these results are quite sensitive to the wording and framing of the questions used to measure people's economic knowledge in population surveys. Moreover, unlike the assumptions of the standard economic model, a search of transcripts of television broadcasts and articles in newspapers could not confirm the universal dissemination of the regular releases of the official *rates* of unemployment, inflation, or economic growth. Media reports about the economy were more often qualitative rather than quantitative in content. As a result, the higher costs of acquiring economic information and the diminished benefits of *exact* figures compared with less precise information, means that consumers exhibit "rational inattention" and stagger their updates over time. Finally, the data suggest that consumers do possess substantial amounts of information about the economy, but that information was based more on private rather than public information sources and was more likely to be tailored to their own decision needs.

Theoretical Framework

The standard economic model has made several highly restrictive assumptions about people's knowledge about economic conditions. The conventional view is that people have full information about all relevant economic conditions. This information set is then updated with each release of official information by governmental agencies. In this simple formulation, the costs of acquiring, processing, and interpreting new information are ignored, as are the variations in potential benefits. The critical implication of these conventional assumptions is that all people are assumed to hold the same information about various aspects of the economy, and update that information at the same time from official sources. There is no information heterogeneity in the standard model.

Modern economic theory does not hold such strict views, however. There are two decisive departures from the standard model: first, rather than simultaneous, updating information occurs in a staggered pattern across individuals and over time, and second, the information that is relevant to people's economic decisions differs across people and over time depending on the characteristics of their situation.

Staggered Updating. As long as there is any positive cost involved in collecting and processing information, some agents will choose to sometimes hold views that are less accurate. The terms "sticky information" or "rational inattention" have been used to describe the impact of costs on the formation process (Mankiw & Reis 2002; Sims 2003; Bacchetta & Wincoop 2005; Reis 2007). These theories postulate that rational consumers may find the costs associated with updating their information on the economy to exceed the benefits. At any given time some people will find it worthwhile to incur the costs,

especially if that information is critical to a pending decision. Most of the time, however, rational inattention is the optimal course. Alternatively, agents may base their economic views on imperfect information, which can be conceptualised as less costly than perfect information. Whatever the cause, the process creates staggered changes, whereby at any given time people's views on economic conditions reflect a combination of current and past information.

Disagreement across people in their views of the economy at any given time is taken as an indication of such a process (Mankiw *et al.* 2004). Some have modelled the disagreements as the result of factors other than costs, such as an epidemiological process in which "expert opinion" spreads slowly through a population like the spread of a disease (Carroll 2003). Costs can also be assumed to vary across demographic subgroups, as some encounter lower costs for acquiring and using information, and other more economically active subgroups derive greater benefits from updating their expectations more frequently. This interpretation of disagreements stands in contrast to the older and still more common interpretation that the very existence of differences across demographic subgroups indicates non-rational processes (Bryan & Venkatu 2001; Souleles 2001).

Staggered changes could be created by a wide range of processes that either encourage or discourage agents from updating their information about the economy. A common hypothesis for staggered updating holds that it is due to asymmetric responses to economic information, with agents updating their expectations much more quickly in response to bad news about economic conditions. Akerloff, Dickens and Perry (2000) suggest that bad economic news is perceived by consumers to contain more potentially relevant information about their financial situation. The volume of news also matters, especially the volume of bad news, as well as news that represent a sharp and negative break from the past (Carroll 2003). Sims (2003) shows that based on information theory the tone and volume of economic reporting affects people's overall views of the economy beyond the information contained in the reports.

The same staggered information flows have been hypothesised to result from uncertainty about the correct structural model of the economy. Since model uncertainty is costly to resolve, it results in less frequent updating (Branch 2005). Although the data that indicates disagreement is similar to what could be expected to result from model uncertainty, these two concepts are distinct. More importantly, the prevalence of disagreement may be much more variable over time than uncertainty.

The models developed to capture the impact of staggered information are similar to consumption models that incorporate the division between "rule of thumb" and rational consumers. Mankiw and Reis (2003), Carroll (2003), Khan and Zhu (2006), and Curtin (2006) estimated that rather than continuously updating their expectations, most people update their expectations only a few times a year.

Relevant information. It is an artifact of the standard economic model that all agents are assumed to focus on the same definition of inflation, unemployment or economic growth. Nonetheless, it makes no economic sense to assume that people pay attention to an inflation rate that is higher or lower than the one they actually encounter. Empirical research has confirmed a good deal of variance in actual inflation rates across different demographic groups (Hobijn & Lagokos 2005; Hagemann 1982; Michael 1979). Most of the differentials are based on the differential inflation rates for specific products or services, such as higher health care costs among the elderly, or in specific areas of the country, such as price differentials between rural and urban areas. While many of the

price differentials do not persist over extended periods of time and people's circumstances change as they age, it is unreasonable to expect people would ignore these differences in prices. The same may be said for employment conditions, as people would naturally pay more attention to job opportunities that are relevant to their own skills and abilities. The national unemployment rate may be quite meaningless to workers living in areas or working in an industry that has a distinctly different outlook for employment, whether it was better or worse than the national average.

While relevant information about prices and employment conditions are in principle available from the state statistical agencies at no cost, in practice the cost to acquire and process this information is high. The cost to acquire and process private information, however, may be significantly lower and the potential benefit of that information may be significantly higher, leading people to prefer private over official data sources (Curtin 2003).

Survey Methodology

Reliable and valid measures of what people know about economic conditions are subject to all of the problems usually associated with sample surveys. Aside from the more general issues of survey methodology, the crucial measurement issue involves judgements about the capacity of individuals to provide meaningful responses. Questions that measure people's knowledge of economic conditions can be phrased in a number of ways, each differing in the cognitive burden placed on respondents. Increasing the precision of the measures also increases the extent of information that respondents must access from memory, the required computational skills, and the motivation of respondents to provide accurate responses.

This paper takes three approaches to the measurement of people's economic knowledge. The first set of questions asked about the respondents' knowledge of the official rates of unemployment, inflation, and economic growth. This set of questions specifically included an instruction that if the respondent had no information about the measure, they should simply tell the interviewer that they wanted to skip the question. These questions, however, did include a statement that it was important for the respondent to share whatever information they did know about the topic. The purpose was to identify people who had any information on the topic and would willingly undertake the cognitive burden of answering the questions. This group was expected to be the most knowledgeable about the official economic statistics.

The second set of questions were identical in every way to the first set with the only difference being that the questions did not tell respondents that they could opt-out or skip the questions if they chose. Respondents would need to volunteer that they did not know the information. These questions were designed to encourage all people to undertake the cognitive burdens of providing an answer.

After the first two sets of questions, everyone who replied they did not know the official rates of unemployment, inflation, or economic growth were asked if they had ever heard of the official measures. For an economist, it was hard to imagine anyone who had not heard of the unemployment rate or the inflation rate. The hypothesis was that people had simply not heard of the most recently announced rate, which is consistent with the notion of rational inattention and staggered updating.

The third set of questions was quite different in that the questions did not refer to an official rate produced by some government agency but asked about the same underlying economic concept. These questions were part of the regular monitoring of consumer expectations and were asked before the questions on their knowledge of the official rates. Perhaps the most striking difference was that these questions were phrased in "population" language and avoided any mention of official rates, the government agencies responsible, or how the measure should be defined. These questions were designed to capture public as well as private information about the economic measures. This third set of questions was asked before the other sets of questions as part of the regularly repeated standard questionnaire.

All of the questions were asked as part of the University of Michigan's Survey of Consumers, which are known worldwide for its measure of consumer sentiment. The monthly samples are representative of all households in the United States, with every adult given an equal probability of being selected for an interview. The data were collected in April and May of 2007 and included 1 008 cases.

Knowledge of Official Data on the Economy

The survey measured people's awareness of the official national unemployment rate, the Consumer Price Index (CPI), and the rate of growth in the Gross Domestic Product (GDP). Each question included three core elements: it defined the economic indicator, it identified the official governmental agency responsible for collecting the data, and asked for the most recently published figure. The wording of the questions was as follows:

> First, the Bureau of Labor Statistics counts people as unemployed if they are not currently working but have been actively looking for work during the prior four weeks. What was the most recent rate of unemployment published by this government agency?
>
> Another economic indicator published by the Bureau of Labor Statistics is the Consumer Price Index, or the CPI. Compared with a year ago, what was the percentage change in overall prices as measured by the Consumer Price Index, or CPI, published by this government agency?
>
> The Bureau of Economic Analysis regularly publishes data on the total amount of goods and services produced in the U.S. This figure is called the Gross Domestic Product and is often abbreviated as GDP. Compared to a year ago, what was the percentage change in the Gross Domestic Product, or GDP, published by this government agency?

Note that the first question asked was expected to be the easiest to answer: the unemployment rate is widely discussed in the media and the percentage is not a rate of change but a simple proportion. In contrast, the CPI, while widely publicised, is always expressed as a rate of change, and that rate is variously published as a simple month-to-month change, an annualised month-to-month change, or a year-to-year change. This implies that the information about the CPI may require more processing and calculation before it is useable as an answer to this question. The final question was the most difficult since it concerns a quantity that does not directly impinge on people's economic lives like inflation or unemployment, giving them less incentive to track the measure; moreover, the figure is repeatedly revised, and variously reported as an annualised quarter-to-quarter change or annual change, and adjusted for seasonal variations and inflation. It is also

important to note that unlike GDP, unemployment and inflation express negative outcomes, and people have been found to pay more attention to negative rather than positive economic developments.

Some people may not have answered the questions because they may not have had specific knowledge of all three required elements. People may not have known how the rate they knew was defined, may have never heard of that particular federal agency, or may not have heard an announcement for some time. While each of these possibilities was not investigated, a follow-up to each question asked to everyone who did not know the official figures:

Have you ever heard an announcement of the

- *unemployment rate by the Bureau of Labor Statistics?*
- *the Consumer Price Index, or CPI, by the Bureau of Labor Statistics?*
- *the Gross Domestic Product, or GDP, by the Bureau of Economic Analysis?*

Respondents could have indicated that they had heard of the economic indicator, but just didn't know the current figure, and may have simply engaged in rational inattention. Given that inflation and unemployment were relatively low at the time the data were collected, people's attention to the recent data may have been significantly reduced, and so engaged in staggered updating. With heightened salience, attention could quickly return – why people pay more attention to "bad" news in the media. While it could be expected that nearly everyone would have heard of all three indicators, a "no" answer could presumably mean either that they have not heard of the official indicator or that they did not know of the government agency, or both.

When respondents are faced with questions that require a high cognitive burden to answer, some respondents will simply say they don't know the answer when they simply do not want to exert the required effort to answer. Knowledge questions in general population surveys need special treatment given that asking the question could cause embarrassment and resentment and possibly cause the respondent to immediately end the interview. Typically, some effort is taken to diffuse the situation by framing the questions in an appropriate context. More importantly, the survey wanted to isolate the most informed people who were willing to undertake the required cognitive burden. As a result, the introduction to the questions was:

The next several questions are about the effectiveness of the mass media in communicating information from agencies of the federal government about the performance of the U.S. economy. If you do not have any information about one of these questions, please just say so, and I will go on to the next question. However, if you do not know the exact answer but just have a rough idea, it is important for you to tell me what you know.

The second form of the questions, asked in an independent random sample, did not provide an easy opt-out. It encouraged everyone to respond, although volunteered "don't know" answers were accepted. The introduction no longer told respondents that it was acceptable to skip the questions. This was done by deleting the sentence: "If you do not have any information about one of these questions, please just say so, and I will go on to the next question." The second form was:

The next several questions are about the effectiveness of the mass media in communicating information from agencies of the federal government about the performance of the U.S. economy. If you do not know the exact answer but just have a rough idea, it is important for you to tell me what you know.

In the analysis included in this paper, this difference in the introductions to the knowledge questions will be referred to as the "opt-out" option. In the tables, the opt-out option equals "yes" when the respondent was specifically given the option skipping the question, and equals "no" when there was no mention of skipping the question.

Reported Awareness of Official Economic Statistics

The data indicate that one-third of all respondents reported that they knew the most recently published official rate of unemployment, one-in-five reported knowledge of the most recently published rate of change in the Consumer Price Index, and about one-in-six knew the most recently announced official rate of growth in the Gross National Product (see Table 12.1). What was an even more dismal assessment of the public's knowledge of these official statistics was that one-fifth of all respondents reported that they had never heard of the official rate of unemployment published by the Bureau of Labor Statistics, one-third reported that they had never heard of the Consumer Price Index, and four-in-ten reported that they had never heard of the Gross Domestic Product reported by the Bureau of Economic Analysis. In between these extremes, just over four-in-ten respondents reported that they had heard of these official statistics but they did not know the most recently announced rate.

Based on the hypothesis of rational inattention, the results are not so dismal: more than three-in-four people knew of the official rate of unemployment, two-thirds knew about the CPI, and six-in-ten had heard about GDP. These estimates represent the maximum share of the population that could engage in rational inattention; in all likelihood the true proportion is much lower.

Even fewer people reported that they knew the official rates when the opt-out option was given to the respondent. About half as many respondents provided a "rate" answer when the opt-out option was given for the unemployment rate (26% versus 43%), the Consumer Price Index (13% versus 27%) and for the Gross Domestic Product (9% versus 23%). The data clearly indicate that people were quick to take advantage of the question skipping option.

The data generally confirmed that these differences were related to the cognitive burden of providing answers to these questions. The proportion that provide "rate" answers to these questions were 14 to 17% lower when given the option to not answer the question. For the unemployment rate and the CPI, the shift was mainly toward the response that they had heard of the official figure but didn't know its current level, but for GDP the shift was overwhelmingly toward the response that they had never heard of the figure. These results indicate that a combination of factors were at work, including the added cognitive burden of providing a "rate" answer, avoiding embarrassment from admitting that you never heard of these economic statistics, or simply from being a compliant respondent that provides answers to each question that was asked. Overall, these results suggest that responses to such knowledge questions are very sensitive to the amount of both formal and informal encouragement given to respondents to provide answers.

Table 12.1 People's Knowledge of Official Measures of Economic Performance

What is the official rate of . . .	Unemployment Rate			Consumer Price Index (CPI)			Gross Domestic Product (GDP)		
	All	Opt-out Option		All	Opt-out Option		All	Opt-out Option	
		Yes	No		Yes	No		Yes	No
Provided rate answer	34	26	43	20	13	27	17	9	23
Heard of, but didn't know current rate	42	46	37	44	48	40	42	44	41
Never heard of official rate or agency	23	27	20	34	37	31	40	46	34
DK; NA	1	1	0	2	2	2	1	1	2
Total	100%	100%	100%	100%	100%	100%	100%	100%	100%
Cases	1008	508	500	1008	508	500	1008	508	500
Addendum (medians)									
Actual Official Rate A_t	4.5%	4.4%	4.5%	2.4%	2.8%	2.6%		2.1%	3.1%
People's reports of official rates R_{it}	4.8%	4.7%	4.9%	3.1%	3.1%	3.1%	3.8%	3.3%	4.1%
Median absolute percentage point error in people's reports of official rates $\lvert R_{it} - A_t \rvert$	0.57	0.54	0.59	1.20	0.81	1.40	1.49	0.96	1.77

Note: The actual official rate (A_t) reflect the latest release available on the date the interview was conducted and do not reflect subsequent revisions.

The other critical information contained in Table 12.1 is the answers provided for the most recently announced rate of unemployment, inflation, or economic growth. For the question on the CPI, people reported an official rate of 3.1% whether the opt-out option was present or absent from the question. There was only a minor difference in the rate of unemployment, 4.7% with the opt-out option and 4.9% without. There was a larger difference in people's responses about the most recent GDP announcement, although given how few people reported any knowledge of GDP, the difference was not significant.

Given that the survey was conducted over a two month period, there were up to three official announcements for the statistic. Since the questions asked respondents for the most recent value of each statistic, it was assumed that people's answers would differ depending on the actual date of the interview (the interviews were spread rather evenly over the two months). It was assumed that respondents would become aware of the latest official statistic on the day it was released, so interviews conducted on the day after the official release would use the newly released number to determine the accuracy of the respondent's reports.

To assess errors in people's estimates, the absolute difference was calculated between what people reported (R_{it}) and the official statistic (A_t) in percentage points, so that both overestimates and underestimates are fairly treated. The median of the absolute differences ranged from one-half a percentage point to nearly two percentage points (see Table 12.1). In every case, the absolute errors were larger when respondents were not given the opt-out option, although for the question on the unemployment rate, the difference was quite small – just 0.05%.

The largest errors were for the question on GDP. For both the CPI and GDP, the percentages are rates of change, which makes these questions more difficult. Moreover, the official reports on GDP are variously stated as annualised quarterly rates of change or as year-to-year changes. These alternatives can complicate the message and depending on which version a person hears; it would require some calculation to convert to the form asked in the questionnaire. Although the question asked about year-on-year percentage changes, if people had these two rates confused, it would generally lead to less accurate

reports since the annualised quarterly GDP growth rates were considerably smaller than the year-on-year percentage changes at the time the data were collected.

More importantly, the calculation of the error rates ignored the hypothesis of rational inattention and staggered updating. People's answers presumably reflect a process of updating over time so that at any given moment their answers reflect the last update, which could have taken place at some time in the past and not immediately following the last data release.

Demographic Correlates of People's Knowledge of Official Statistics

It is of some interest to determine the demographic correlates of people's knowledge as they may indicate differences in costs or benefits of acquiring information. The most obvious variable to investigate for differences is the absolute error in people's perceptions of the official statistics. Given that so few people actually provided a numerical answer to that question, the analysis needs to be supplemented by investigation of those who reported that they knew of the statistic but didn't know the current rate, and an investigation of the demographic correlates of those that reported that they had never heard of the official statistics.

Specific Rate Answers. Regressions were performed to determine the net impact of the demographic variables using two dependent variables: the proportion who reported a rate and the size of the absolute errors. The demographic characteristics included in the analysis were education, income, age, sex, and whether the respondent was given the opt-out option (see Table 12.2). These variables are proxies for differences in the costs and benefits of more up-to-date information.

Table 12.2 Demographic Determinants of Knowledge and Accuracy of People's Reports about Official Statistics

	Unemployment Rate		Consumer Price Index		Gross Domestic Product	
	Rate Answer Givena	Absolute Errorb |Rit - At|	Rate Answer Givena	Absolute Errorb |Rit - At|	Rate Answer Givena	Absolute Errorb |Rit - At|
Opt-out given	-0.783*** (0.148)	-0.257 (0.352)	-0.922*** (0.175)	-1.308 (0.667)	-1.072*** (0.193)	-0.087 (0.872)
Income (log)	0.128 (0.096)	-0.308 (0.234)	0.188 (0.112)	-1.277** (0.414)	0.014 (0.119)	-0.338 (0.489)
Age in Years	0.125** (0.004)	-0.048*** (0.010)	-0.004 (0.005)	-0.053** (0.018)	-0.012* (0.005)	-0.047* (0.023)
Years of Education	0.188*** (0.038)	-0.242** (0.087)	0.104* (0.042)	-0.132 (0.139)	0.114* (0.046)	-0.313 (0.185)
Female	-0.889*** (0.149)	0.602 (0.358)	-0.851*** (0.173)	-0.502 (0.665)	-0.974*** (0.189)	2.035* (0.867)
R2 adjusted	0.169	0.092	0.128	0.132	0.137	0.104
Cases	925	357	911	202	919	165

Note: [a]Logistic regressions on the proportion that provided a rate answer. [b]Least-square regressions on the absolute percentage point errors in people's reports of the official statistics indicated by column headings. Intercept included but not shown. Standard errors in parentheses.
Significant levels: * < .05, ** < .01, *** < .001.

As expected, people who were offered the opt-out option were significantly less likely to report a figure for any of the official rates. Perhaps more noteworthy was that the size of the absolute errors in people's estimates of the official rates did not significantly vary

with the opt-out option. This means that although fewer people answered the question if given the opt-out option, errors in the estimated rates were not larger.[3] To be sure, when the option to skip the question was given to respondents it uniformly reduced errors among those who answered the question, consistent with the view that they were more informed respondents, but the reduction was not significant.

The capacity to handle the cognitive demands of the questions should have increased along with the formal education of respondents. Although high education did significantly increase the likelihood that respondents would provide a rate answer for each of the three statistics, high education only significantly reduced the size of the absolute errors on reports of the official unemployment rate but not for the CPI or GDP statistics.

The age of the respondent had a mixed impact. Older respondents were more likely to provide a rate when asked about unemployment, but less likely when asked about GDP. More importantly, the older the respondent, the smaller the size of the absolute errors on all three statistics. This is somewhat surprising since older respondents are more likely to have cognitive limitations; presumably, the experience that comes with older age dominated.

The income of the respondent's household had the least impact. Income only had significant impact on the size of errors on the CPI, with higher income associated with less error. The significant impact on CPI errors may reflect experience with a wider range of prices that usually accompanies higher incomes, whereas lower income households spend a much higher proportion of their incomes on food and energy, and these products have recently recorded higher rates of increase. Finally, female respondents had a much lower likelihood of reporting any specific figure for each of the three statistics, but were not significantly less accurate for the unemployment and inflation rates. Females did record a much larger error in their reports of GDP, an error of 2%.

Table 12.3 Demographic Determinants of Knowledge about Official Rates and Never Having Heard of the Statistic

	Unemployment Rate		Consumer Price Index (CPI)		Gross Domestic Product (GDP)	
	Heard of but DK rate	Never Heard of	Heard of but DK rate	Never Heard of	Heard of but DK rate	Never Heard of
Opt-out given	0.399** (0.136)	0.485** (0.172)	0.344* (0.138)	0.353* (0.154)	0.180 (0.139)	0.501*** (0.147)
Income (log)	0.280** (0.088)	-0.624*** (0.110)	0.223* (0.090)	-0.433*** (0.100)	0.293** (0.091)	-0.347*** (0.095)
Age in Years	0.007 (0.004)	-0.027*** (0.005)	0.016*** (0.004)	-0.019*** (0.005)	0.023*** (0.004)	-0.019*** (0.004)
Years of Education	-0.006 (0.037)	-0.206*** (0.039)	0.163*** (0.034)	-0.264*** (0.038)	0.150*** (0.035)	-0.255*** (0.036)
Female	0.698*** (0.138)	0.208 (0.174)	0.090 (0.139)	0.637*** (0.157)	-0.168 (0.140)	0.845*** (0.151)
Pseudo R^2 adjusted	0.064	0.207	0.094	0.222	0.110	0.208

Note: Logistic regressions on presence or absence of conditions indicated by column headings. Standard errors in parentheses. Intercept included by not shown. Significant levels: * < .05, ** < .01, *** < .001.

Lack of knowledge about economic statistics. Logistic regressions were performed to help sort out the impacts of the demographic variables when people professed to have no current knowledge of the official statistics and for those who reported that they had

never heard of the official statistics (see Table 12.3). As anticipated, the results of the logistic regressions indicate that whenever the opt-out option was given it generally had a significant impact.[4] Significantly higher proportions of respondents who were given the opt-out option professed a lack of knowledge of the most recently published rate or had no knowledge of the statistic.

Higher educated respondents were consistently less likely to report that they had never heard of the economic statistics, and generally more likely to report that they simply didn't know the most recent figure. Although education and income are correlated, income was still a significant predictor because it is a proxy not only for cognitive ability but also for the degree of engagement in the economy. Higher income respondents were more likely to simply not know the current figure for each statistic and significantly less likely to have never heard of these official economic measures. Even apart from education, the data indicate that higher income households may have a vested interest to give more attention to economic statistics, may face lower costs in acquiring economic information, or may garner greater economic benefits from updating their information more regularly.

The age of the respondent also had a widespread impact on these responses. Older respondents were significantly less likely to report having never heard of these economic statistics, and were generally more likely to report that they knew of them but didn't know their current levels. This probably reflects the greater life experiences of older adults, especially those experiences of several decades ago when inflation and unemployment were at double digit levels and GDP fell sharply. The results for gender were somewhat surprising, in that females were significantly more likely to report never having heard of the CPI or GDP.

Importance of Information on Official Economic Statistics

While economic models of consumer decision making usually include an assumption that all agents have full knowledge of official economic statistics, the data above indicate that official information is not widely known. It is useful to check the underlying assumption that people value exact information on official statistics. Following the questions on knowledge of official statistics, respondents were asked whether they thought it was important to know exact information on the performance of the economy. The question was worded as follows:

> How important is it for a person like you to have exact information about the rate of unemployment, the rate of change in prices, and the rate of change in the Gross Domestic Product — would you say it is extremely important, very important, somewhat important, not very important, or not important at all?

The responses were:

Extremely important	*6*
Very important	*20*
Somewhat important	*40*
Not very important	*24*
Not important at all	*9*
DK; NA	*1*
Total	*100%*

The results indicate that just one-in-four respondents thought it was extremely or very important to know exact information compared with one-in-three that thought it was not important for them to know the exact information on unemployment, inflation, or economic growth. It should be emphasised that the question asked about the importance of "exact" information not "any" information. It should not be so surprising that people do not think that their economic decisions would change if the exact figure was a few tenths of a percentage points higher or lower or even if it differed by a few percentage points. This result is what theories of "rational inattention" or "near rationality" imply.

Following this question, respondents were asked if they would like to be more informed about economic conditions. It would seem hard to say no to a question about whether you wanted to be better informed, especially in an economic survey conducted by a university. The primary cost would simply be the opportunity cost of devoting attention to economic matters rather than something else. Nonetheless, half of all respondents answered that they didn't want any more information when asked the following question:

> Would you like to be more informed about these topics or would you not want any more information about these topics?

Where do people actually get information on official economic statistics? The survey asked respondents the following questions to gain information about their top three news sources.

> We are interested in how people get official government information about the rate of unemployment, the rate of change in prices, and the rate of change in the Gross Domestic Product. Do you get **most** of this type of information from television, the radio, newspapers, magazines, the internet, your family, friends, or co-workers, your own personal experiences, from some other sources, or do you never get any official government information on these topics? What is your **second** most common source of official government information about these topics? What is your **third** most common source of official government information about these topics?

The dominant source of information on economic statistics was television, reported by nearly half of all people as their first choice, and by nearly three-in-four people among their top three choices (see Table 12.4). Newspapers came in second, with nearly one-in-five naming them their main source, and nearly six-in-ten reported newspapers among their top three choices. The Internet, radio, and personal contacts were each reported by about one-in-three people as among their top three choices. Nearly one-in-ten people volunteered that they never obtain information about the economic statistics.

It is difficult to determine the quality of the information people gain from these sources. Television encompasses a wide variety of reports on economic news, ranging from the in depth details of cable business channels to passing references that contain no details except a one-word summary that the news was "good" or "bad." The same can be said for newspapers, ranging from the detailed commentary included in the Wall Street Journal to the same one-word summaries of the latest rates included in many city newspapers. Of all the sources, only the Internet includes the possibility of a direct link to the official government agencies that produce the statistics, but it also includes access to the widest range of sites that can be expected to widely differ in terms of accuracy and quality.

Table 12.4 Sources of Information on Official Rates of Unemployment, Consumer Prices, and Gross Domestic Product

Sources of Information	Most Common Source of Information	Second Most Common Source of Information	Third Most Common Source of Information	Total Sources
Television	45	23	10	78%
Newspapers	18	25	15	58
Internet	10	15	12	37
Radio	11	11	12	34
Family/friends/coworker/personal experience	4	10	20	34
Magazines	2	4	8	14
Never obtain information	10	—	—	10
No other sources	—	12	23	
Total	100%	100%	100%	
Cases	1008	1008	1008	

The key issue is whether people who rely on any of these sources are more likely to report accurate information on the economic statistics in question. The regressions included in Table 12.2 were repeated with the addition of the sources people mentioned using for obtaining information. The data show a consistent lack of impact on reports of the unemployment and inflation rates for both dependent variables: the mere reporting of a rate and the accuracy of the reported rate (see Table 12.5). For GDP, the regressions indicate a significant impact from nearly all the news sources but no impact on accuracy.

Table 12.5 Informational Determinants of Knowledge and the Accuracy of People's Reports about Official Economic Statistics

	Unemployment Rate		Consumer Price Index (CPI)		Gross Domestic Product (GDP)	
	Rate Answer Given[a]	Absolute Error[b] $\lvert R_{it} - A_t \rvert$	Rate Answer Given[a]	Absolute Error[b] $\lvert R_{it} - A_t \rvert$	Rate Answer Given[a]	Absolute Error[b] $\lvert R_{it} - A_t \rvert$
Television	-0.072 (0.253)	0.221 (0.567)	0.066 (0.299)	-0.379 (1.000)	0.624 (0.392)	-0.938 (1.619)
Radio	0.304 (0.213)	0.256 (0.502)	0.201 (0.262)	-0.359 (1.064)	1.017** (0.368)	-0.680 (1.683)
Newspapers	0.271 (0.218)	0.743 (0.489)	0.596* (0.272)	0.488 (1.023)	1.321*** (0.381)	-1.408 (1.636)
Magazines	0.040 (0.264)	0.123 (0.601)	-0.040 (0.323)	-0.111 (1.215)	1.380*** (0.406)	-0.188 (1.773)
Internet	0.533* (0.229)	-0.635 (0.577)	0.354 (0.276)	0.902 (1.070)	1.171** (0.383)	-0.640 (1.667)
Friends, family, co-workers, personal experience	0.016 (0.222)	1.345** (0.520)	0.360 (0.274)	-1.521 (1.084)	0.987** (0.381)	-0.222 (1.688)
Never obtain information	-0.899 (0.600)	4.709** (1.547)	0.166 (0.712)	-0.464 (2.722)	2.448* (1.03)	-2.007 (4.507)

Note: a.Logistic regressions on the proportion that provided a rate answer.
b.Least-square regressions on the absolute percentage point errors in people's reports of the official statistics indicated by column headings. All regressions include the demographic controls listed in Table 12.2.
Standard errors in parentheses.
Significant levels: * < .05, ** < .01, *** < .001.

The impact of the media on the likelihood that the respondent had ever heard of the official statistics was also sparse. The logistic regressions reported in Table 12.6 indicate

only a few significant impacts: newspapers were more often associated with having heard of the unemployment rate, radio and the internet with having heard of the CPI. For the GDP equation, users of most all forms of the mass media were less likely to say that they never heard of the official rate of growth in the Gross Domestic Product. This may underscore the notion that unemployment and inflation consumers have personal experience, and require some information for knowledge about GDP.

Table 12.6 Informational Determinants of Knowledge about Official Rates and Never Having Heard of the Statistic

	Unemployment Rate		Consumer Price Index (CPI)		Gross Domestic Product (GDP)	
	Heard of but DK rate	Never Heard of	Heard of but DK rate	Never Heard of	Heard of but DK rate	Never Heard of
Television	0.203 (0.240)	-0.015 (0.322)	0.285 (0.240)	-0.212 (0.281)	0.296 (0.238)	-0.393 (0.262)
Radio	-0.030 (0.196)	-0.268 (0.237)	0.610** (0.200)	-0.851*** (0.224)	0.255 (0.200)	-0.648** (0.214)
Newspapers	0.425* (0.197)	-1.039*** (0.236)	0.212 (0.202)	-0.693** (0.221)	0.319 (0.200)	-0.919*** (0.214)
Magazines	0.109 (0.244)	-0.251 (0.335)	0.106 (0.246)	0.007 (0.273)	0.058 (0.247)	-0.733** (0.278)
Internet	-0.211 (0.211)	-0.392 (0.263)	0.560** (0.214)	-0.996*** (0.242)	0.206 (0.214)	-0.754*** (0.232)
Friends, family, co-workers, personal experience	0.241 (0.202)	-0.312 (0.244)	-0.053 (0.205)	-0.142 (0.218)	-0.299 (0.204)	-0.034 (0.211)
Never obtain information	-0.648 (0.514)	0.967 (0.552)	-0.962 (0.564)	0.412 (0.535)	-1.176 (0.567)	-0.040 (0.535)

Note: All logistic regressions include the demographic controls discussed earlier.
Standard errors in parentheses.
Significant levels: * < .05, ** < .01, *** < .001.

Coverage of Official Economic Statistics in the Mass Media

A critical assumption in testing whether people have accurate knowledge of the current official rates of unemployment, consumer prices, and economic growth is that those rates are communicated by the mass media. In an attempt to test this assumption, television transcripts and newspaper archives were searched to determine if they contained a report that cited a specific number for the official statistics on the day it was released by the government agency, or in the following three days for newspapers. A four day window was chosen because most U.S. newspapers are morning editions that are printed before the 8:30 release time of the agency, and since some releases occur on Fridays, a four day window was needed to include the following Monday. A report on the official statistic that did not mention the exact official rate was not counted even if it did occur in the four day window. The TV transcripts and newspapers were searched over the sixteen months from January 2006 to April 2007.[5]

News reports from the five major broadcast networks – ABC, NBC, CBS, CNN, and FOX – were searched.[6] All-business networks were not included, even if they were owned by one of the five networks that were included. The official release of the unemployment rate was reported every time on NBC and CNN and 94% of the time by FOX (see Table 12.7). The official unemployment rate was reported much more frequently than the CPI or GDP. Perhaps this is due to the fact that the percentage figure represents a proportion rather than the more complex concept of a rate of change. Reports

on the CPI and GDP were often given in qualitative rather than quantitative terms, such as "prices rose faster" or "the economy worsened." This tendency may also reflect the fact that the official releases of these figures are given both as monthly and annual rates of change for the CPI and as quarterly or annual rates of change for GDP. In any event, CBS reported the actual official rate of inflation for the CPI more than half the time, and only CNN and FOX reported the official GDP rate for half or more of the releases.

A total of 27 newspapers were searched, each having a circulation of more than 400 000 as of March 2006.[7] The total circulation of these newspapers was 21.1 million at the start of 2006. Newspapers are often read by more than one person, with a 2006 estimate that on average 2.3 persons read each copy (according to the Newspaper Association of America). Ignoring that some people read more than one paper each day, the gross number of people reading each copy in circulation was 48.6 million, or about 22% of all adults living in the U.S.

Four papers had a circulation of more than one million – USA Today, The Wall Street Journal, The New York Times, and the Los Angeles Times. None of these papers carried every release. Only the Washington Post, fifth highest in circulation with just below one million copies, carried all the official releases all the time, not surprising since all the agencies are headquartered in Washington, DC. The New York Times reported the official unemployment rate and the CPI for each of the sixteen months, and the GDP in fifteen months. The Wall Street Journal carried all the unemployment releases, and 80% to 90% of the CPI and GDP releases. The Los Angeles Times carried the reports between 50% of the time (CPI) and 75% of the time (GDP and unemployment). The paper with the highest circulation, USA Today, reported the official figures the least, reporting each official figure about half the time (see Table 12.7).

The publication of the official figures was even more dismal across all 27 newspapers searched. On average, just 39% of the official reports on GDP appeared, with the median rate just 19%. The CPI was reported on average 52% of the time across all 27 newspapers, with the median a dismal 38%. The official unemployment rate was the most likely to be cited, with an average of 52% of the time and a median of 44%.

The AP and UPI wire services, in contrast, carried reports on the latest official rates of unemployment, the CPI and GDP for every official release. If we presume that the 27 papers with the largest circulations all had access to the wire reports, the lack of complete coverage would be an active decision of the newspaper to not carry the report. It was likely to reflect a judgment about the newsworthiness of the latest figures given their subscribers' interests. There was a tendency for newspapers to more frequently report the latest official figures when it represented an unfavorable development, which may reflect the greater importance people place on the information content of "bad" news.

It is of some interest that the AP and UPI wires typically did not mention the specific government agencies (Bureau of Labor Statistics and Bureau of Economic Analysis) in their releases. They usually simply used the phrase that "the government reported..." or at most referred to the Labor or Commerce Department, the parent agencies for the Bureau of Labor Statistics and the Bureau of Economic Analysis. Presumably the inclusion of the agency names in the survey questions made the questions more difficult than warranted.

The functions of the AP and UPI wire services have been supplanted in recent years by the simultaneous Internet releases of the official statistics. The news wires (and others) still have the advantage of viewing the results early (in a locked room) so they can also provide commentary at the time of the release. Nonetheless, people from around the globe

can access the same data the instant it is released via the Internet. Data were provided by the Bureau of Labor Statistics on the number of times the full releases of the unemployment rate on May 4, 2007 and the CPI on May 15, 2007 were viewed on the Internet. For the unemployment rate it totaled 8 243 and for the CPI it was 11 959, with both accounting for about 1% of all the visits to their Internet sites on those days.

Table 12.7 Television and Newspaper Reports of Official Economic Statistics: Proportion of News Reports that Cited Official Rates from January 2006 to April 2007

Television Reports					
Unemployment Rate		**Consumer Price Index**		**Gross Domestic Product**	
CNN	100%	CBS	63%	CNN	81%
NBC	100%	CNN	50%	FOX	50%
FOX	94%	FOX	31%	ABC	44%
ABC	63%	ABC	19%	NBC	31%
CBS	56%	NBC	13%	CBS	25%
Mean	83%	Mean	35%	Mean	46%
Median	94%	Median	31%	Median	44%

Newspaper Reports					
Unemployment Rate		**Consumer Price Index (CPI)**		**Gross Domestic Product (GDP)**	
Wall Street Journal	100%	New York Times	100%	Washington Post	100%
New York Times	100%	Washington Post	100%	New York Times	94%
Washington Post	100%	Long Island Newsday	100%	Newark Star-Ledger	88%
Chicago Tribune	100%	Houston Chronicle	94%	Wall Street Journal	81%
Seattle Times	100%	Wall Street Journal	88%	Seattle Times	81%
Atlanta Journal Constitution	94%	Newark Star-Ledger	88%	Los Angeles Times	75%
Houston Chronicle	88%	Seattle Times	88%	Houston Chronicle	69%
Newark Star-Ledger	88%	St. Louis Post-Dispatch	81%	Long Island Newsday	69%
St. Louis Post-Dispatch	88%	Atlanta Journal Constitution	75%	Philadelphia Inquirer	63%
Los Angeles Times	75%	Chicago Tribune	69%	Atlanta Journal Constitution	56%
Long Island Newsday	56%	Philadelphia Inquirer	69%	St. Louis Post-Dispatch	56%
USA Today	44%	USA Today	63%	USA Today	50%
New York Post	44%	Los Angeles Times	56%	Chicago Tribune	31%
Minneapolis Star Tribune	44%	New York Post	38%	Boston Globe	19%
Baltimore Sun	44%	Dallas Morning News	38%	San Francisco Chronicle	19%
Dallas Morning News	38%	Minneapolis Star Tribune	38%	St. Petersburg Times	19%
Milwaukee Journal Sentinel	38%	San Francisco Chronicle	38%	Denver Post	13%
Philadelphia Inquirer	31%	Denver Post	31%	Detroit News	13%
Boston Globe	25%	St. Petersburg Times	31%	Minneapolis Star Tribune	13%
San Francisco Chronicle	25%	Boston Globe	25%	Arizona Republic	13%
Arizona Republic	19%	New York Daily News	19%	Milwaukee Journal Sentinel	13%
St. Petersburg Times	19%	Arizona Republic	19%	New York Post	6%
Denver Post	13%	San Diego Union-Tribune	19%	Dallas Morning News	6%
Detroit News	13%	Milwaukee Journal Sentinel	19%	Baltimore Sun	6%
San Diego Union-Tribune	13%	Detroit News	13%	New York Daily News	0%
New York Daily News	6%	Plain Dealer	13%	Plain Dealer	0%
Plain Dealer	0%	Baltimore Sun	6%	San Diego Union-Tribune	0%
Mean	52%	Mean	52%	Mean	39%
Median	44%	Median	38%	Median	19%

Overall, this review of the dissemination of official economic data suggests that people's lack of knowledge can be in part attributed to the inadequate communication of that information by the mass media. It was true that news on unemployment was more frequently reported in the media, and people's knowledge of the unemployment rate was more accurate in the survey. The coincidence is suggestive but does not prove causation.

Informal Knowledge of the Performance of the Economy

The questions about people's knowledge of official data by federal agencies can be compared with other questions that simply ask about unemployment, prices, and economic growth. Unlike the prior questions, which identified the official governmental agency responsible for collecting the data and asked for the most recently published figure, these alternate questions simply asked respondents about likely changes in unemployment, prices, and the economy. How each concept was defined also differed, especially for unemployment and GDP, with both questions using less technical jargon. Also note that these questions focus on the next twelve months rather than changes over the past twelve months. It should also be noted that these questions were asked prior to the questions on the official economic statistics, and the questions were separated in the questionnaire by dozens of other questions that took more than five minutes to ask. The wording of these questions is as follows:

> How about people out of work during the coming twelve months – do you think that there will be more unemployment than now, about the same, or less?
>
> During the next twelve months, do you think that prices in general will go up, or go down, or stay where they are now? By what percent do you expect prices to go (up/down) on the average during the next twelve months?
>
> Now turning to business conditions in the economy as a whole – do you think that during the next twelve months conditions will be better, or worse than they are at present, or just about the same?

The question on unemployment expectations and business conditions are measured using qualitative scales while the question on expected inflation is based on a quantitative response scale. For nearly all aspects of the analysis contained in this paper, the qualitative questions on unemployment and GDP cannot be directly compared with the earlier quantitative questions, except for one comparison: the percentage of people who responded that they did not know the answer. For both questions, the percentage of "don't know" responses totalled just 1% of the April and May 2006 surveys. Although not directly comparable the two prior questions, the questions about the unemployment rate and potential economic growth have an excellent track record in predicting changes in the unemployment rate (Curtin 1999, 2003) and GDP (the expectations index is part of the U.S. composite index of leading economic indicators).

It is the question on expected price changes that is most comparable to the question about the official CPI. This question has been analysed repeatedly over the past decades, and found to be predictive of the actual subsequent change in overall prices (Gramlich 1983; Grant & Thomas 1999; Thomas 1999; Mehra 2002).[8] Thomas (1999, p.141-142) summarized his findings by noting that "...consensus household inflation forecasts do surprisingly well relative to those of the presumably better-informed professional economists." Indeed, the median consumer forecasts of year-ahead inflation rates "...outperformed all other forecasts in the 1981-1997 period on simple tests of

accuracy as well as on tests for unbiasedness." Mehra (2002, p.35) also finds that Michigan's median inflation expectations outperforms the expectations of professional economists and forecasters: "They are more accurate, unbiased, have predictive content for future inflation, and are efficient with respect to economic variables generally considered pertinent to the behaviour of inflation."

Table 12.8 Responses to Expected Inflation Question by Response on Official CPI Question

	Provided Percentage Rate for Expected Inflation Rate		
	Opt-out Option		
Knowledge of Official CPI Rate	Yes	No	Total Sample
Provided answer for official CPI rate question	97%	92%	94%
Heard of, but didn't know current official CPI rate	93%	92%	92%
Never heard of official CPI rate or agency	85%	80%	83%
Total Sample	91%	88%	89%

What is of interest for the present analysis is the differences in responses to the two questions. One critical issue is to determine what exactly people did not know on the official question about the CPI: was it their knowledge about price changes or was it their knowledge of the official rate published by the Bureau of Labor Statistics. The biggest difference between the two questions was that one question was on the official rate published by a government agency and the other simply asked about the rate of change in prices. Whereas one form of the question required knowledge of a government announcement of an official rate, responses to the other question could be answered from more informal information, including personal experience. Unfortunately, no data were collected for a direct test of this hypothesis, although several interesting comparisons are possible.

Perhaps the most dramatic comparison is that 83% of those who said that they had never heard of the official CPI nonetheless provided a percentage inflation rate they expected during the year ahead (see Table 12.8). Among those who said that they had heard of the CPI but didn't know the current rate, 92% reported a percentage rate of expected inflation. A still higher proportion (94%) of those that reported knowledge of the official CPI, reported an expected inflation rate. Clearly, nearly all of the respondents knew something about trends in overall prices. It should again be stressed that the two questions are not perfectly comparable given the that the question on the official CPI asks about past changes in prices, and the informal question asks about potential future changes.

Table 12.9 Estimates of Expected Inflation by Response to the Question on the Official CPI

	Consumer Price Index (CPI)		
	Opt-out Option		
Knowledge of Official CPI Rate	Yes	No	Total
Provided answer for official CPI rate question	3.4%	3.1%	3.2%
Heard of, but didn't know current official CPI rate	3.1%	3.4%	3.3%
Never heard of official CPI rate or agency	3.4%	3.4%	3.4%
Total Sample	3.3%	3.3%	3.3%

Note: Table entries are the expected inflation rate based on people's responses to the informal question within the subgroups defined by people's responses to their knowledge of the official CPI.

Another indirect test is to compare the answers given by people on the informal question across the different responses to the official CPI question. Table 12.9 contains the inflation rate forecasts of people by their responses on the question on their knowledge of the official inflation rate. Across all of the table's cells, the data suggest that the median inflation rate showed only minor and statistically insignificant variations, ranging from 3.1% to 3.4%. The data also suggest that a major barrier confronted by people was that they did not know the official CPI rate published by the Bureau of Labor Statistics, but not that they did not have any knowledge of price trends. If the data were to be used in models describing economic behaviour, Occam's razor would recommend that extraneous details should be eliminated, and the question should focus on the central topic in a simple and straightforward manner.

Conclusions

What do people know about official economic statistics? Only a minority of people could report a specific figure for the current rates of unemployment, inflation, or economic growth, although the majority had heard about these official statistics. When they did report a specific figure, it typically differed from the official statistic by half a percentage point to one and a half percentage points, depending on the economic statistic. People reported more accurate figures for the unemployment rate, less accurate reports of the inflation rate, with reports of GDP growth rates being the least accurate. Errors were modestly related to the education level of the respondent, which can be interpreted as a proxy for the cognitive abilities required to process and recall information or potential costs. Other factors, such as age and income, were related to one statistic or another, and can be interpreted as proxies for economic experiences or potential benefits.

The tendency for people to overestimate the rates of inflation and unemployment could be conceptualised as a buffer against unexpected events. Bad economic news is perceived by people to contain more potentially relevant information about their future financial situation. As a result, people may adopt somewhat less favourable views about unemployment and inflation as a means to protect themselves against an unending stream of disruptive and costly small changes.

Importantly, the data provided considerable evidence that the questions about official economic statistics were viewed as burdensome, and when given the opportunity to skip the question, many respondents did opt-out of answering the question. Along with the cognitive burden, another motivation would be to avoid the embarrassment of an incorrect answer to a knowledge question where the respondent could assume that the interviewer knew the correct answer. It is impossible to disentangle these two hypotheses with the collected data. The rather large impact of a minor change in the question wording, however, indicates the high sensitivity of such questions, and indicates the important role of survey methodology in this research.

Conventional economic models assume that all economic agents always have full information on all relevant economic quantities. More recent theoretical advances have emphasised two departures from the standard model. First, rather than simultaneously, information updating occurs in a staggered pattern across individuals and over time. People make decisions about whether to update information depending on the costs of acquiring, processing, and interpreting new information compared with the potential benefits of the new information. Thus, the tests of "accuracy" included in this paper are too strict in that they implicitly assumed that all people update their economic

information immediately after its release. Rather than the standard of uniform views on unemployment, inflation, and economic growth, heterogeneity of beliefs can be expected across economic agents. While there is no universal standard to judge whether the current costs and expected benefits warrant updating economic information, it is nonetheless more likely when the inflation or unemployment rate is high and variable rather than low and stable. These data were collected when unemployment, inflation, and economic growth were relatively favourable and stable, which would imply little need for updating.

The second modification of the standard model is that the information that is relevant to people's economic decisions differs across people and over time depending on the characteristics of their situation. There is no reason to expect that people would seek out information about an inflation or unemployment rate that they did not face. Indeed, rather than economy-wide information, it is more likely that local information is more appropriate. Local unemployment rates for jobs that individuals are qualified for are more important than national unemployment rates, and people that consume a greater proportion of their incomes on certain products or services would naturally view the potential benefits of information on those products or services greater than information on overall inflation. The implication of the primacy of these more specific information needs increases the importance of what economists call "private" compared with "public" information.

A third modification of the standard model involves the cost of updating information on current economic statistics. It is typical that models assume that exact figures on the rates of unemployment, inflation, and growth in the economy are widely disseminated to the public in the mass media. Rather than universal reporting of these economic statistics, the record indicates more sporadic and incomplete reporting, with much greater use of qualitative terms than quantitative figures. This means the cost of acquiring information is much greater than usually assumed. Moreover, given the various ways rates of change for the CPI and GDP are reported, it also means greater computation and processing costs.

The survey included other economic measures that were more aligned with people's usual economic experiences. Answers to questions about trends in unemployment, inflation, and economic growth were nearly universal, standing in stark contrast to responses to the knowledge questions on the official rates. To be sure, there was one critical difference: these other questions did not ask what happened in the past, but asked people about their expectations of the future. The measure on the expected rate of inflation was asked using a percentage rate response scale comparable to the knowledge questions. The expected inflation responses were the same regardless of whether the respondents were explicitly given the option to skip the question, and if the respondents reported a figure for the knowledge question about the official CPI, or had heard of the official CPI but did not know the most recent figure, or had never heard of the official measure.

The lack of a relationship of the two questions on inflation indicates an independence between knowledge of the official CPI and the "private" information people possess on prospective trends in the inflation rate. The general lack of knowledge of the official CPI does not mean that people do not know about inflation, only that they do not know the official rate most recently published by a governmental agency. Private knowledge about expected price trends, as well as unemployment and economic growth, was widespread, and past analyses has shown those expectations to be relatively accurate.

Such a complex overall assessment of the public's knowledge of economic statistics is much less surprising than the premise that people would consistently use their scarce

resources to monitor official economic statistics published by government agencies. Consumers do desire knowledge about their economic situation, as Aristotle noted long ago. Nonetheless, just as Plato suggested, we all see reflections from our own perspective, and believe that these assessments best serve our own needs. It is within these shadows of diversity that economic theory and public policy will flourish.

Notes

1 National income accounts are a relatively recent innovation, with the first accounts published by William Petty in 1665, and modern national income accounts were devised by Simon Kuznets and others in the first half of the twentieth century (Vanoli 2005).

2 No economist would suggest that the GDP figure represents the absolute truth, but would readily admit that it is an estimate of economic performance that is subject to conceptual and measurement errors. The same is true of all other measures, including the CPI and the unemployment rate. Given whatever conceptual or measurement errors are present, as long as these errors remain constant over short periods of time, monthly or quarterly changes in the measure can reliably gauge economic trends.

3 Regressions of the absolute errors on just the opt-out variable gave similar results: the opt-out question did not have a significant impact on the accuracy of reports of the unemployment rate or GDP, but had a marginal significant impact on reports of the CPI.

4 Similar results were obtained when the regressions were estimated separately for each opt-out group and indicated no interactions.

5 It should be emphasized that the digital databases that were searched are likely to contain errors. Every attempt was made to verify that the results of our searches accurately reflected the criteria that the latest official rate of unemployment, change in the CPI, or the change in GDP was mentioned at least once in the four-day window.

6 Total viewership for the nightly newscasts on ABC, CBS, and NBC totaled about 20.2 million in April/May 2007 according to MediaBistro.

7 The newspapers, in order of circulation, were: USA Today, Wall Street Journal, New York Times, Los Angeles Times, Washington Post, Chicago Tribune, New York Daily News, Philadelphia Inquirer, Denver Post, Houston Chronicle , New York Post, Detroit News, Dallas Morning News, Minneapolis Star Tribune, Boston Globe, Newark Star-Ledger, Atlanta Journal Constitution, Arizona Republic, Long Island Newsday, San Francisco Chronicle, Plain Dealer, Seattle Times, St. Louis Post-Dispatch, St. Petersburg Times, San Diego Union-Tribune, Milwaukee Journal Sentinel, Baltimore Sun.

8 Similar comparisons were done for year-ahead forecasts of the national unemployment rate. Curtin (1999, 2003) found that consumers' forecasts of the year-ahead unemployment rate outperformed those of professional forecasters as well as forecasts from two prominent macroeconomic models.

References

Akerloff, George, William Dickens, and George Perry, "Near-Rational Wage and Price Setting and the Optimal Rates of Inflation and Unemployment," Brookings Papers on Economic Activity, vol 1, 2000.

Bacchetta, Philippe and Eric van Wincoop, "Rational Inattention: A Solution to the Forward Discount Puzzle," Working paper 11633, National Bureau of Economic Research, September 2005.

Blendon, Robert J., John Benson, Mollymann Brodie, Richard Morin, Drew E. Altman, Daniel Gitterman, Mario Brossard and Matt James, "Bridging the Gap Between the Public's and Economists' Views of the Economy", Journal of Economic Perspectives, Vol. 11, No. 3. (Summer, 1997), p. 105-118.

Blinder, Alan S., Alan B. Krueger, "What Does the Public Know about Economic Policy, and How does It Know It?", Brookings Papers on Economic Activity, 1:2004.

Branch, William A., "Sticky Information and Model Uncertainty in Survey Data on Inflation Expectations," mimeo, July 2005.

Bryan, Michael and Guhan Venkatu, "The Curiously Different Inflation Perspectives of Men and Women," Federal Reserve Bank of Cleveland, Economic Commentary Series, 2001.

Carroll, Christopher, "Macroeconomic Expectations of Households and Professional Forecasters," Quarterly Journal of Economics, vol. 118 (1), p. 269-298, 2003.

Curtin, Richard, "Unemployment Expectations: The Impact of Private Information on Income Uncertainty," Review of Income and Wealth, vol. 49, no. 4, 2003.

Curtin, Richard, "What Recession? What Recovery? The Arrival of the 21st Century Consumer," Business Economics, vol. 39, no. 2, 2003.

Curtin, Richard, "Price Expectations: Theoretical Models and Empirical Tests," in The Role of Inflation Expectations in Modeling and Monetary Policy, ed. Tomas Lyziak, NBP: Poland, 2006.

Curtin, Richard, "The Outlook for Consumption in 2000," The Economic Outlook for 2000, Ann Arbor, Michigan: University of Michigan, 1999.

Delli Carpini, Michael X. and Scott Keeter, What Americans Know about Politics and Why It Matters. Yale University Press, 1996.

Doms, Mark and Norman Morin, "Consumer Sentiment, the Economy, and the News Media," Working paper 2004-51, Federal Reserve Board, 2004.

Grant, Alan P. and Lloyd B. Thomas, "Inflationary Expectations and Rationality Revisited," Economic Letters, vol. 62 (March), p. 331-338, 1999.

Gramlich, Edward M., "Models of Inflation Expectations Formation," Journal of Money, Credit, and Banking, vol. 15, pp 155-73, 1983.

Hagemann, Robert P., "The variability of Inflation Rates Across Household Types," Journal of Money, Credit, and Banking, Vol 14, no. 4, part 1, Nov 1982, pp 494510.

Hamilton, James, All the News That's Fit to Sell, Princeton University Press, 2004.

Hobijn , Bart and David Lagokos, "Inflation Inequality in the United States," Review of Income and Wealth, vol 51, no. 4, December 2005, p. 581-606.

Khan, H. and Z. Zhu, "Estimates of the Sticky-Information Phillips Curve for the United States," Journal of Money, Credit, and Banking, vol 38, no. 1, 2006.

Kohut, Andrew, Richard Morin and Scott Keeter, "What Americans Know: 1989-2007, Public Knowledge of Current Affairs Little Changed by News and Information Revolutions", Pew Research Center for The People & The Press, 2007

Lucas, Arthur M, "Public knowledge of biology," Journal of Biological Education, vol. 21, no. 1, p. 41-45, 1987.

Lucas, Arthur M, "Public knowledge of elementary physics", Physics Education, vol. 23, p. 10-16, 1988.

Mankiw, N. Gregory, Ricardo Reis, and Justin Wolfers, "Disagreement about Inflation Expectations," Macroeconomics Annual 2003, eds. Mark Gertler and Kenneth Rogoff, NBER: 2004.

Mankiw, N. Gregory and Ricardo Reis, "Sticky Information Versus Sticky Prices: A proposal to Replace the New Keynesian Phillips Curve," Quarterly Journal of Economics, vol. 117 (4), p. 1295-1328, 2002.

Mankiw, N. Gregory and Ricardo Reis, "Sticky Information: A Model of Monetary Non-neutrality and Structural Slumps," in P. Aghion, R. Frydman, J. Stiglitz and M. Woodford, eds., Knowledge, Information and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps, Princeton University Press, 2003.

Mehra, Yash P., "Survey Measures of Expected Inflation: Revisiting the Issues of Predictive Content and Rationality," Federal Reserve Bank of Richmond Economic Quarterly, vol. 88 (3), p. 17-36, 2002.

Michael, Robert T., "Variations Across Households in the Rate of Inflation," Journal of Money, Credit, and Banking, vol 11, p. 32-46, February 1979.

OECD, "PISA Assessing Scientific, Reading, and Mathematical Literacy, OECD, Paris, 2006.

Reis, Ricardo, "Inattentive Consumers," Journal of Monetary Economics, vol. 56, p. 1761-1800, 2007.

Schibeci, R. A., "Public Knowledge and Perceptions of Science and Technology", Bulletin of Science Technology Society, Vol. 10, p. 86-92, 1990.

Sims, Christopher, "Implications of Rational Inattention," Journal of Monetary Economics, vol. 50 (3), p. 665-690., 2003.

Souleles, Nicholas S., "Consumer Sentiment: Its Rationality and Usefulness in Forecasting Expenditures—Evidence from the Michigan Micro Data," Journal of Money, Credit, and Banking, 2001.

Thomas Jr., Lloyd B., "Survey Measures of Expected U.S. Inflation," Journal of Economic Perspectives, vol 13, no 4, p. 125-144, Autumn 1999.

Vanoli, André, A History of National Accounting, Translated from the French original by Marion Pinot Libreros and Gayle H. Partmann, IOS Press, 2005

Chapter 13
Special Eurobarometer
Europeans Knoweledge on Economical Indicators[i]

Antonis Papacostas
Head of Unit Public Opinion and Media Mornitoring, European Commission

Abstract

This paper presents the detailed findings of the first-ever Standard Eurobarometer survey on how the public's perception of three major economic indicators – growth rates, unemployment and inflation rates – compares with the official data for these indicators (Eurostat). The analysis is focused on defining the terms of social dichotomy – perception vs. reality. Respondents were aged 15 years and over and residents of European Union Member States (EU 27), as well as Croatia and Turkey and the Turkish Cypriot community. The first three questions of the survey asked the respondents what the growth rate, unemployment and inflation rates were in their countries and compared those responses to the actual data for these indicators. The next three questions focused on the pertinence of economic statistics: (1) is it necessary to know economic indicators; (2) are political decisions made on the basis of statistical information; and (3) do you trust official statistics?

Introduction

This survey carried out between the 10th of April and the 15th May 2007 was conducted under the framework of the Standard Eurobarometer 67.[1] The timing of this survey is a key element in understanding the results presented, as there have been significant developments across the continent during the past months.

Presentation of the Results

It is the first time that the Eurobarometer proposes questions about the perception and evaluation of economical indicators in correlation with official data (cf. Eurostat) concerning growth rate of every country's economy, unemployment and inflation rates.

i This preliminary analysis is done by Antonis Papacostas and Valentina Pricopie.

The study of the European perception concerning these indicators aims to reveal the need of European citizens to be informed on economical data, on national and European implications of statistics concerning the political decisions and on statistic official data reliability. At the same time, this analysis is also focused on defining the terms of social dichotomy – perception versus reality.

The first three questions evaluate the European perception on growth rate, unemployment and inflation in all European member states. The three questions are open (superior and inferior large limits are indicated for the three indicators).

- QA50a[2] What was the official growth rate of the economy (measured in terms of Gross Domestic Product) in (OUR COUNTRY) in 2006? I can tell you that this figure is between -1% and 15%.

- QA52 What was the official inflation rate, the rate of which consumer prices increased or decreased, in (OUR COUNTRY) in 2006? I can tell you that the exact figure is between -1% and 20%.

- QA54 What was the official unemployment rate, the percentage of active people who do not have a job, in (OUR COUNTRY) in 2006? I can tell you that the exact figure is between 0% and 20%.

From a methodological point of view, "don't knows" are considered incorrect answers, but in the case of these questions, we will present the three variants of answers: correct, incorrect and DK. Any proposition that could be included on the defined interval (established by ponderation) of a correct answer is also considered as a good answer.

Analysing the data collected during the field work, it was obvious that most respondents preferred to give round numbers for all the indicators requested. Also, the fact that the figures which were circulated in various national media were rounded or slightly different from the official ones was taken into account. Bearing in mind that one of the surveys aim was to measure the knowledge/information level, it was decided to consider as correct those answers found within the interval of +/- 20% of the official value.[3]

For example, in the case of Belgium, the official growth rate is 3.2. Every proposition that could be easily included in the interval {2.6 and 3.8} is valid; this interval corresponds to a formula:

*3.2*0.8 = 2.56 and 3.2*1.2 = 3.84.*

First of all, the rate of the "don't know" variant is very high: for instance, for 16 countries in the case of growth rate perception (QA50a) and also for 16 countries in the case of inflation rate, more than 50% of the respondents prefer not to answer. People tend to think that they know the official rate of unemployment in their countries and they give an answer in proportion of 52%, In this case the EU 27 rate of correct answers is indeed higher than for the two other indicators (GDP growth rate and inflation).

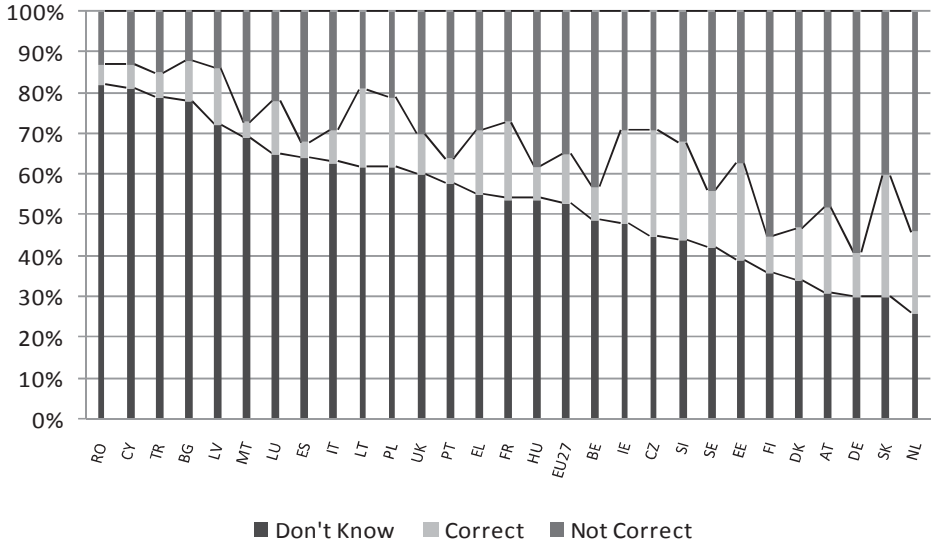
GDP Growth Rate

Concerning the official GDP growth rate, more than 50% of the respondents from 16 countries choose not to give an answer to that question. 82% of the respondents from Romania declare they do not know and prefer not to give an answer; only 5% of

Romanians answered correctly (Figure 13.1).[4] In the Netherlands for example, only 26% of respondents prefer not to give an appreciation concerning the national growth rate and 20% answered correctly. In 12 countries, more than 60% of the respondents prefer not to give their opinion on the official growth rate of the national economy. The countries that register the best scores of correct answers to that question are Slovakia (30%), Czech Republic (26%), Estonia (24%), and Slovenia (24%).

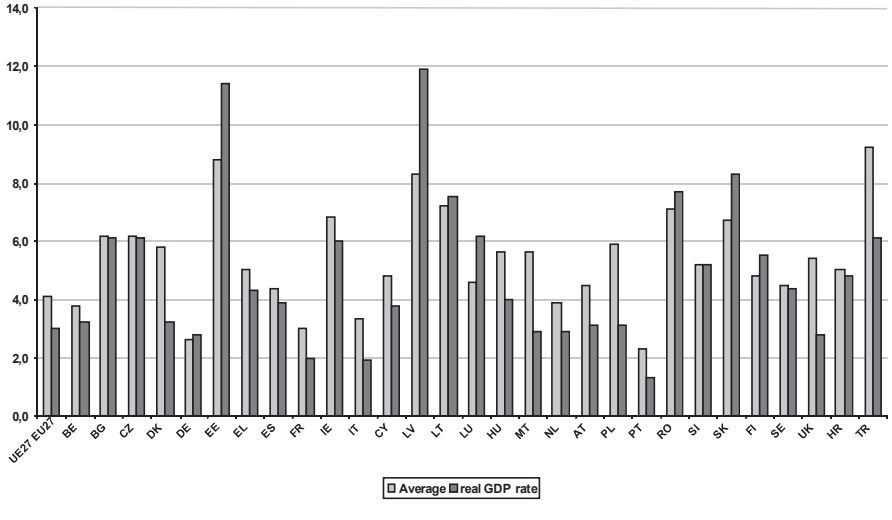
Only 12% of the respondents for EU27 answered correctly to the question.

Figure 13.1 Estimated GDP Growth Rate for 2006



Country Code: Refer to Annex 13.2

Figure 13.2 The Average of Answers for GDP Growth Rate



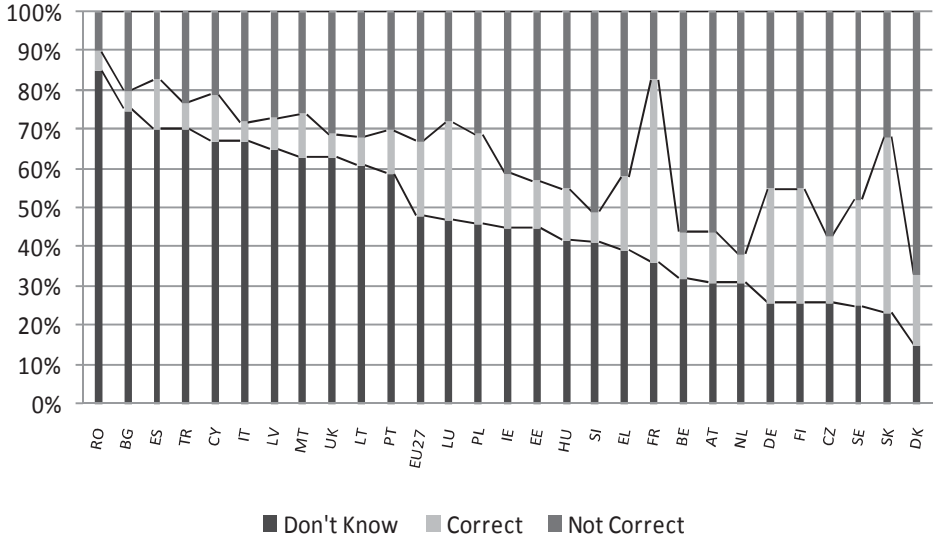
Country Code: Refer to Annex 13.2

Looking per country at the average of answers, out of the three economical indicators that we evaluate in this paper, growth rate seems to be the most well estimated economical issue in Europe: the majority of countries from our panel have a rather correct perception on growth rate, except for Latvia and Estonia where people think that growth rate is smaller in their respective countries, and Turkey, where the average of people answers overrates the actual growth rate. However, the trend on this issue is to overrate growth rate, except for Germany, Greece, Latvia, Lithuania, Luxembourg, Romania, Slovakia and Finland, where people do not have a positive perception of growth rate.

Unemployment Rate

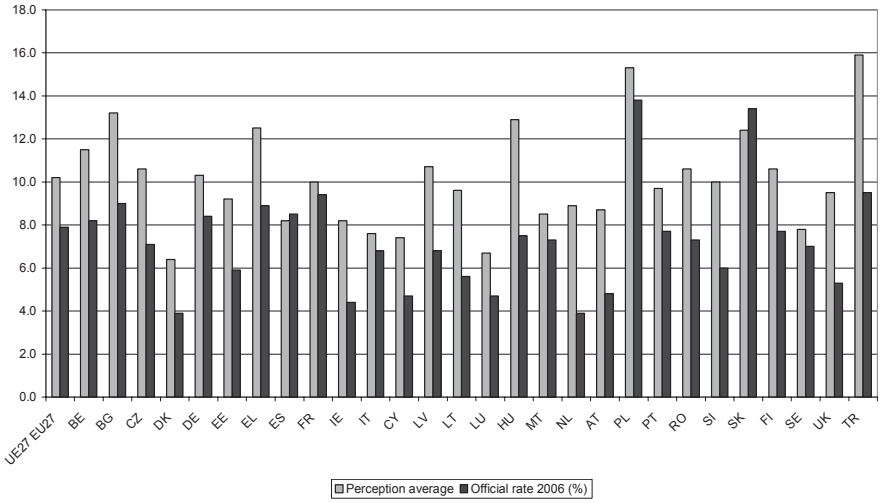
The second indicator would be unemployment rate. The European perception of unemployment rate differs somewhat from the perception of growth rate: 19% of the respondents from EU27 could give the correct official rate of unemployment for 2006 in their respective countries (Figure 13.3). In four countries, (Romania, Bulgaria, Spain and Turkey), more than 70% of the respondents did not answer the question. However, 47% of French respondents could indicate the correct rate of unemployment in their country, as well as 47% of the Slovaks, 29% of the Germans and 29% of Finish respondents.

Figure 13.3 Estimated Unemployment for 2006



Country Code: Refer to Annex 13.2

Figure 13.4 Unemployment: Perception and Official Rate



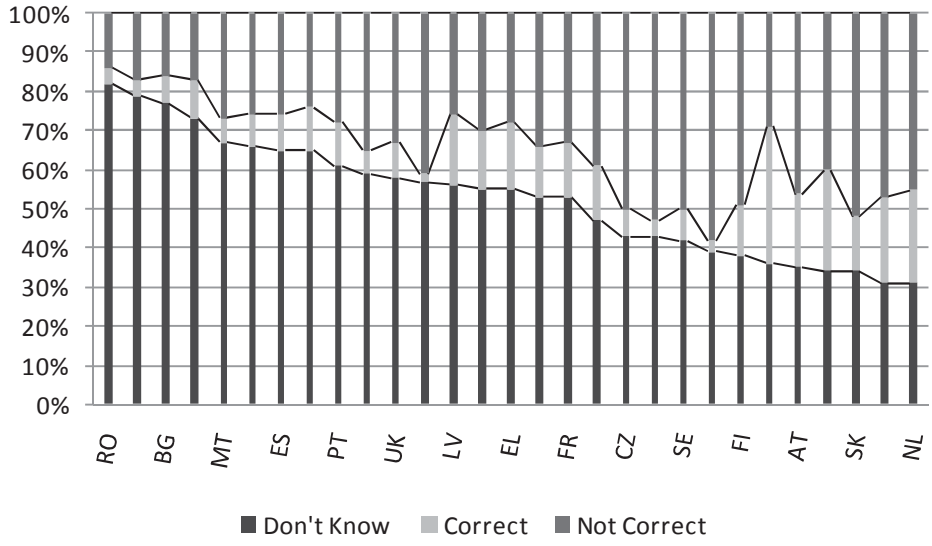
Country Code: Refer to Annex 13.2

Concerning the perception that European citizens have on the unemployment rate in their countries, we can also consider another analytical perspective by comparing these rates with the official data delivered by Eurostat (Figure 13.4). First of all, we have noticed that the average of people's answers is 10% higher than the official rate. This is the case for EU27 for example. There are some countries such as Italy, Sweden and Malta, where the perception average is closer to reality. In some other countries like Turkey, Hungary and Austria, the difference between perception and reality of the unemployment rate is very significant.

Inflation Rate

The perception of the inflation rate is rather similar with the two other economical indicators (Figure 13.5). Less than 50% of the respondents from all countries give an answer, the percent of good answers in EU27 is 13%: with the highest scores in Slovenia (35%), Germany (26%), the Netherlands (24%). Fewest respondents from Poland answered correctly to the question.

Figure 13.5 Estimated Inflation Rate

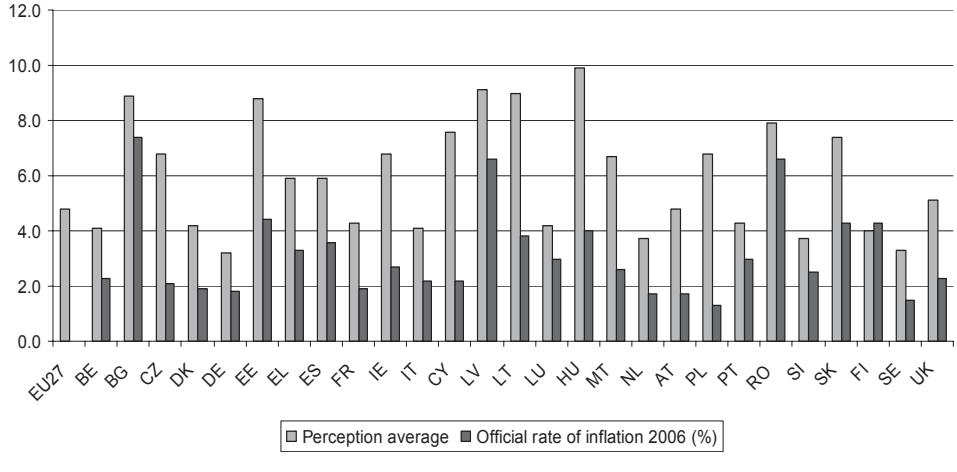


Country Code: Refer to Annex 13.2

The difference between the inflation averages proposed by the respondents from all countries and the real inflation rate is even more significant compared to the difference between the averages of the unemployment rate and the official data on this issue (Figure 13.6).

Countries such as the Netherlands, Luxembourg, Austria, and Spain have a very negative perception concerning the inflation rate. Only Romanian respondents are very close to reality on this issue. However, inflation rate seems to be the most unpopular economical item in Europe.

Figure 13.6 Inflation: Perception and Official Rate



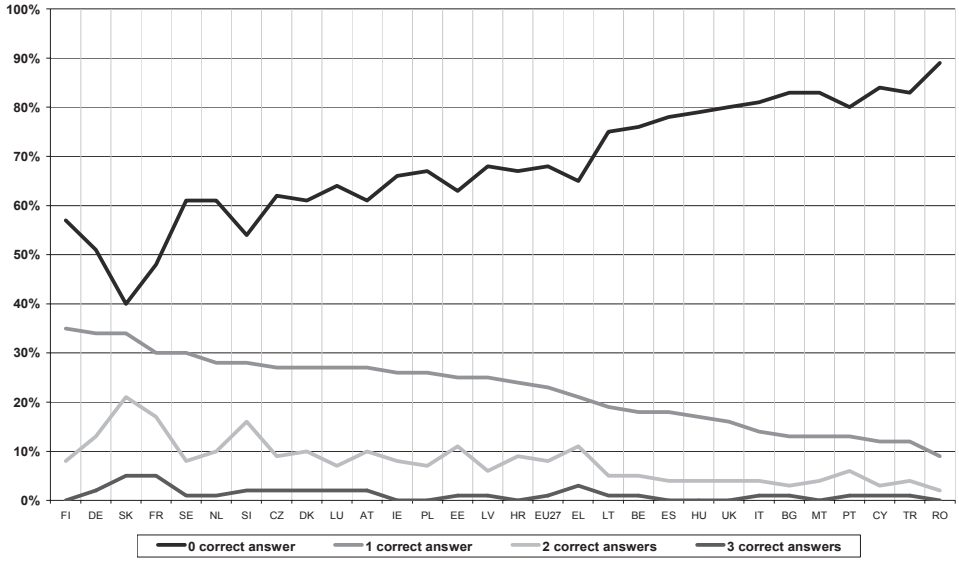
Country Code: Refer to Annex 13.2

Knowledge Score

The correct answers correlation reveals some very important aspects of the European perception of economical indicators (Figure 13.7). We decided to present this correlation graph, in order to identify the rate of correct answers for these three questions that evaluate economics, as a sign of real knowledge, and not as a sign of hazard. The results are significant: more than 60% of the respondents couldn't answer correctly to any of these questions, 23% of respondents from EU27 gave one correct answer, and only 9% from European citizens could respond correctly to two out of three questions.

The next category of questions is focused on the pertinence of economical statistics: It is necessary to know these figures, political decisions are made on the basis of statistical information and finally how much do you trust official statistics?

Figure 13.7 Knowledge Score

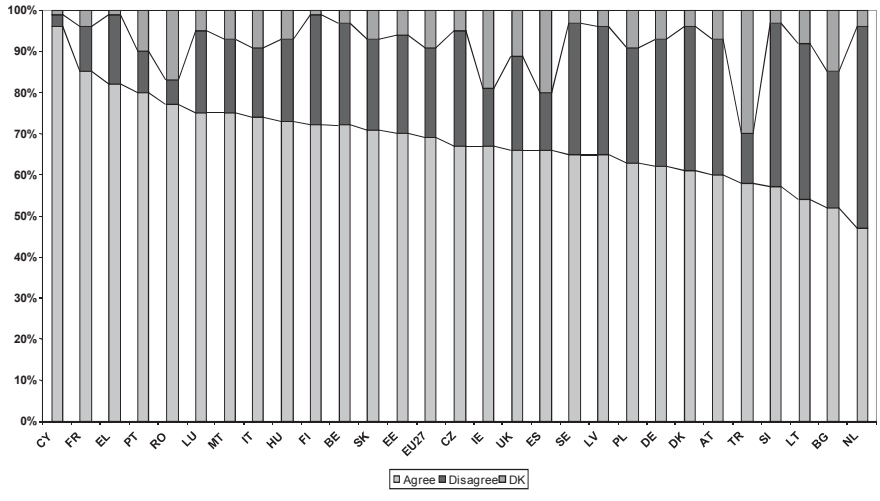


Country Code: refer to Annex 13.2

Is it Necessary to Know Economical Indicators?

For this question, more than 50% of the respondents from all countries we studied (except the Netherlands) consider that it is necessary to know this economical data (Figure 13.8). We can also measure the interest that people in Europe have for official economic data concerning their country; The Netherlands is the only county of our panel where the interest rate for this data is negative. 69% of the respondents from EU27 declare that personally they consider this data important to know, and 77% of the respondents from the Netherlands declare that they do not agree with this statement. On the contrast, more than 90% of the respondents in Cyprus consider that it is necessary to know these figures.

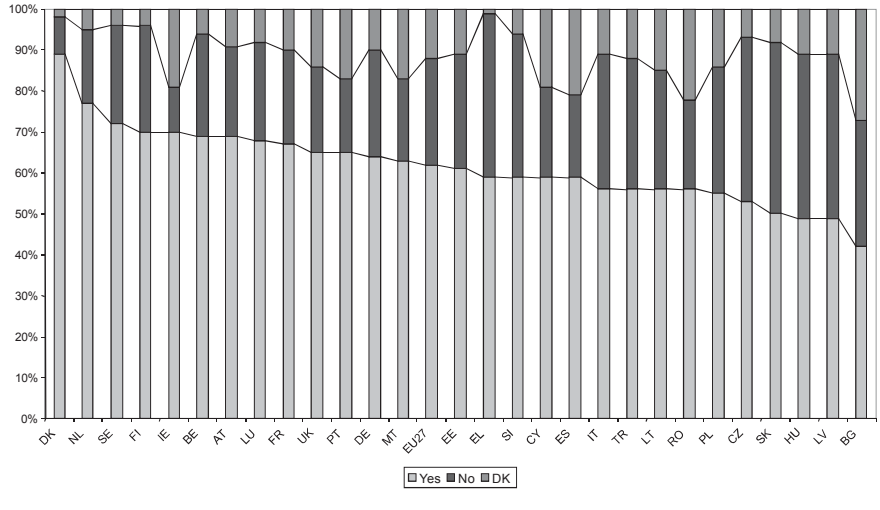
Figure 13.8 Is it Necessary to Know Economical Indicators?



Country Code: refer to Annex 13.2

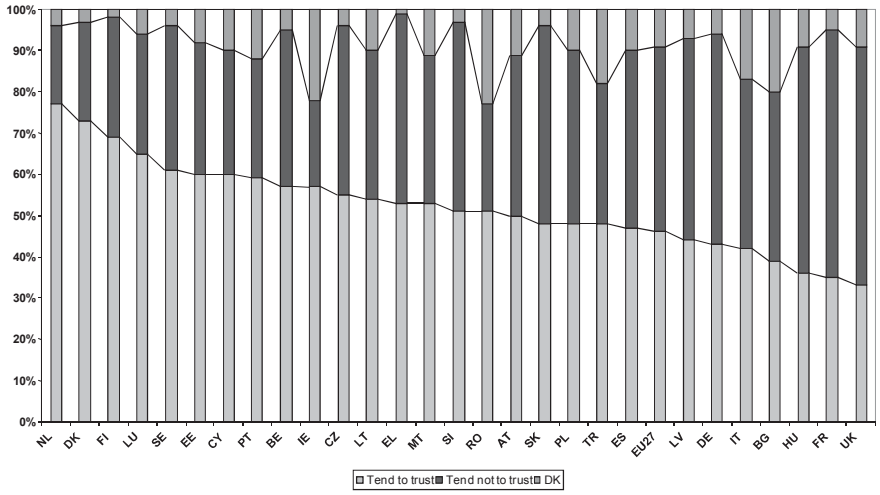
Furthermore, in the case of the second question of this new category we can observe that people's perception on political effects of statistics does not vary a lot from country to country: 62% of the respondents from EU27 consider that in their respective countries political decisions are made on the basis of statistical information (Figure 13.9). 89% of Danish respondents have the same opinion, as well as 77% of respondents from the Netherlands. At the same time, only 42% of Bulgarians share this opinion.

Figure 13.9 Political Decisions are Made on the Basis of Statistical Information?



Country Code: refer to Annex 13.2

Figure 13.10 Do you trust official statistics?



Country Code: refer to Annex 13.2

Finally, the European medium trend is to trust official statistics, but results for EU27 are not too explicit: 45% of European citizens tend not to trust and 46% tend to trust official statistics such as growth rate, inflation rate and unemployment rate (Figure 13.10). In 16 countries from our panel more than 50% of the respondents tend to trust official statistics.

Conclusions

This preliminary evaluation of European knowledge on economical indicators gives us the possibility to understand the interest of European citizens in knowing statistical data in their respective country. We have to specify that this is the first time that the Eurobarometer uses questions that evaluate European perception on economical indicators in correlation with official data (cf. Eurostat).

Obviously, the first conclusion of our study is that people in Europe have a false perception of economical reality in almost every European country. Concerning the three most important economical indicators that we introduced in this survey, the average perception is not equal from one country to another; however, we can notice a major lack of knowledge on all these issues.

The most "popular" item that defines national economies in Europe seams to be the growth rate because of the equivalence that we identify between the average responses and the official rate; contrarily to growth rate, inflation rate becomes the most “unpopular” economical indicator with a significant difference between perception and reality. But, usually, the trend is to overrate unemployment and inflation rates, and to underrate growth rate. It means that people in Europe tend to point out the negative aspects of economy such as inflation and unemployment rates, and to underrate the positive aspects such as growth rate.

However, even if 70% of European respondents consider that it is important to know this economical data, 62% consider that political decisions are made on the basis of

statistical information; that's why the trend to trust official statistics in the EU is not bigger than 46%.

Notes

1 This survey was requested and coordinated by the Directorate General Communication with the assistance of the OECD and conducted by TNS-OPINION & SOCIAL http://ec.europa.eu/public_opinion/index_en.htm

2 This numbering is part of the standard questionnaire.

3 In the paper presented in 2007 at the Second OECD World Forum the correct answers were considered those within +/- 10% of the official values. When drafting the final results it was concluded that for the reasons presented above it was more realistic to consider as correct the answers within +/-20% error margin. Therefore, some of the graphs in this paper are different from the ones actually presented at the forum works.

4 We can notice starting from the next graph that the more significant the rate of DK answers is, the less accurate the results are.

Annex 13.1

Technical specifications

Between the 10th of April and the 15th of May 2007, TNS Opinion & Social, a consortium created between Taylor Nelson Sofres and EOS Gallup Europe, carried out the 67.2 wave of the EUROBAROMETER, on request of the EUROPEAN COMMISSION, Directorate General Communication, "Public Opinion and Media Monitoring".

The STANDARD EUROBAROMETER 67 is part of the 67.2 wave and covers the population of the respective nationalities of the European Union Member States, resident in each of the Member States and aged 15 years and over. The STANDARD EUROBAROMETER 67 has also been conducted in the two candidate countries (Croatia and Turkey) and in the Turkish Cypriot Community. In these countries, the survey covers the national population of citizens and the population of citizens of all the European Union Member States that are residents in these countries and have a sufficient command of the national languages to answer the questionnaire. The basic sample design applied in all states is a multi-stage, random (probability) one. In each country, a number of sampling points was drawn with probability proportional to population size (for a total coverage of the country) and to population density.

In order to do so, the sampling points were drawn systematically from each of the "administrative regional units", after stratification by individual unit and type of area. They thus, represent the whole territory of the countries surveyed according to the EUROSTAT NUTS II (or equivalent) and according to the distribution of the resident population of the respective nationalities in terms of metropolitan, urban and rural areas. In each of the selected sampling points, a starting address was drawn, at random. Further addresses (every Nth address) were selected by standard "random route" procedures, from the initial address. In each household, the respondent was drawn, at random (following the "closest birthday rule"). All interviews were conducted face-to-face in people's homes and in the appropriate national language. As far as the data capture is concerned, CAPI (Computer Assisted Personal Interview) was used in those countries where this technique was available.

For each country a comparison between the sample and the universe was carried out. The Universe description was derived from Eurostat population data or from national statistics offices. For all countries surveyed, a national weighting procedure, using marginal and intercellular weighting, was carried out based on this Universe description. In all countries, gender, age, region and size of locality were introduced in the iteration procedure. For international weighting (*i.e.* EU averages), TNS Opinion & Social applies the official population figures as provided by EUROSTAT or national statistic offices. The total population figures for input in this post-weighting procedure are listed above.

Readers are reminded that survey results are estimations, the accuracy of which, everything being equal, rests upon the sample size and upon the observed percentage.

With samples of about 1 000 interviews, the real percentages vary within the following confidence limits:

ABBREVIATIONS	COUNTRIES	INSTITUTES	N° INTERVIEWS	FIELDWORKDATES		POPULATION 15+
BE	Belgium	TNS Dimarso	1011	11/04/2007	07/05/2007	8.650.994
BG	Bulgaria	TNS BBSS	1039	13/04/2007	26/04/2007	6.671.699
CZ	Czech Rep.	TNS Aisa	1043	13/04/2007	04/05/2007	8.571.710
DK	Denmark	TNS Gallup DK	1002	10/04/2007	14/05/2007	4.411.580
DE	Germany	TNS Infratest	1513	10/04/2007	09/05/2007	64.361.608
EE	Estonia	Emor	1005	16/04/2007	14/05/2007	887.094
EL	Greece	TNS ICAP	1000	10/04/2007	08/05/2007	8.693.566
ES	Spain	TNS Demoscopia	1000	10/04/2007	10/05/2007	37.024.972
FR	France	TNS Sofres	1013	10/04/2007	08/05/2007	44.010.619
IE	Ireland	TNS MRBI	1000	10/04/2007	10/05/2007	3.089.775
IT	Italy	TNS Abacus	1010	10/04/2007	08/05/2007	48.892.559
CY	Rep. of Cyprus	Synovate	502	11/04/2007	07/05/2007	596.752
CY(tcc)	Turkish Cypriot Comm.	KADEM	500	10/04/2007	06/05/2007	157.101
LV	Latvia	TNS Latvia	1013	13/04/2007	15/05/2007	1.418.596
LT	Lithuania	TNS Gallup Lithuania	1018	10/04/2007	05/05/2007	2.803.661
LU	Luxembourg	TNS ILReS	511	10/04/2007	07/05/2007	374.097
HU	Hungary	TNS Hungary	1006	12/04/2007	10/05/2007	8.503.379
MT	Malta	MISCO	500	10/04/2007	30/04/2007	321.114
NL	Netherlands	TNS NIPO	1009	10/04/2007	10/05/2007	13.030.000
AT	Austria	Österreichisches Gallup-Institut	1011	10/04/2007	01/05/2007	6.848.736
PL	Poland	TNS OBOP	1000	10/04/2007	02/05/2007	31.967.880
PT	Portugal	TNS EUROTESTE	1011	14/04/2007	07/05/2007	8.080.915
RO	Romania	TNS CSOP	1019	10/04/2007	08/05/2007	18.173.179
SI	Slovenia	RM PLUS	1013	10/04/2007	09/05/2007	1.720.137
SK	Slovakia	TNS AISA SK	1106	13/04/2007	01/05/2007	4.316.438
FI	Finland	TNS Gallup Oy	1038	16/04/2007	09/05/2007	4.348.676
SE	Sweden	TNS GALLUP	1005	10/04/2007	10/05/2007	7.486.976
UK	United Kingdom	TNS UK	1319	10/04/2007	07/05/2007	47.685.578
HR	Croatia	Puls	1000	10/04/2007	07/05/2007	3.722.800
TR	Turkey	TNS Piar	1005	10/04/2007	10/05/2007	47.583.830
TOTAL			29222	10/04/2007	15/05/2007	444.406.021

For each country a comparison between the sample and the universe was carried out. The Universe description was derived from Eurostat population data or from national statistics offices. For all countries surveyed, a national weighting procedure, using marginal and intercellular weighting, was carried out based on this Universe description. In all countries, gender, age, region and size of locality were introduced in the iteration procedure. For international weighting (*i.e.* EU averages), TNS Opinion & Social applies the official population figures as provided by EUROSTAT or national statistic offices. The total population figures for input in this post-weighting procedure are listed above.

Readers are reminded that survey results are estimations, the accuracy of which, everything being equal, rests upon the sample size and upon the observed percentage. With samples of about 1 000 interviews, the real percentages vary within the following confidence limits:

Observed percentages	10% or 90%	20% or 80%	30% or 70%	40% or 60%	50%
Confidence limits	± 1.9 points	± 2.5 points	± 2.7 points	± 3.0 points	± 3.1 points

In this report, the countries are represented by their official abbreviations. The abbreviations used in this report correspond to:

Abbreviations 13.2

EU27 European Union – 27 Member States
EU15 European Union - 15 Member States before the enlargement of 1st May 2004
NMS New Member States – 12 Member States which joined the EU during the last two enlargements
DK/NA Don't know/No answer
BE Belgium
BG Bulgaria
CZ Czech Republic
DK Denmark
D-E East Germany
DE Germany
D-W West Germany
EE Estonia
EL Greece
ES Spain
FR France
IE Ireland
IT Italy
CY Republic of Cyprus*
LT Lithuania
LV Latvia
LU Luxembourg
HU Hungary
MT Malta
NL The Netherlands
AT Austria
PL Poland
PT Portugal
RO Romania
SI Slovenia
SK Slovakia
FI Finland
SE Sweden
UK The United Kingdom
HR Croatia
TR Turkey

* Cyprus as a whole is one of the 27 European Union Member States. However, the "acquis communautaire" is suspended in the part of the country that is not controlled by the government of the Republic of Cyprus. For practical reasons, only the interviews conducted in the part of the country controlled by the government of the Republic of Cyprus are recorded in the category "CY" and included in the EU27 average.

GDP Growth rate

QA50 What was the official growth rate of the economy (measured in terms of Gross Domestic Product) in (OUR COUNTRY) in 2006? I can tell you that this figure is between -1% and 15%.

Growth rate	EU27	BE	BG	CZ	DK	DE	EE	EL	ES	FR	IE	IT
Average	4.1	3.8	6.2	6.2	5.8	2.6	8.8	5.0	4.4	3.0	6.8	3.3
Official rate	3	3.2	6.1	6.1	3.2	2.8	11.4	4.3	3.9	2	6	1.9

Growth rate	CY	LV	LT	LU	HU	MT	NL	AT	PL	PT	RO
Average	4.8	8.3	7.2	4.6	5.6	5.6	3.9	4.5	5.9	2.3	7.1
Official rate	3.8	11.9	7.5	6.2	4	2.9	2.9	3.1	3.1	1.3	7.7

Growth rate	SI	SK	FI	SE	UK	HR	TR
Average	5.2	6.7	4.8	4.5	5.4	5.0	9.2
Official rate	5.2	8.3	5.5	4.4	2.8	4.8	6.1

	RO	CY	TR	BG	LV	MT	LU	ES	IT	LT	PL	UK
DK	82%	81%	79%	78%	72%	69%	65%	64%	63%	62%	62%	60%
Correct	5%	6%	6%	10%	14%	4%	13%	4%	8%	19%	17%	10%
Not correct	13%	13%	15%	12%	14%	27%	22%	32%	29%	19%	21%	30%

	PT	EL	FR	HU	EU27	BE	IE	CZ	SI	SE	EE
DK	58%	55%	54%	54%	53%	49%	48%	45%	44%	42%	39%
Correct	6%	16%	19%	8%	12%	8%	23%	26%	24%	14%	24%
Not correct	36%	29%	27%	38%	35%	43%	29%	26%	32%	44%	37%

	FI	DK	AT	DE	SK	NL
DK	36%	34%	31%	30%	30%	26%
Correct	9%	13%	21%	11%	30%	20%
Not correct	55%	53%	48%	59%	40%	54%

Inflation rate

QA52 What was the official inflation rate, the rate of which consumer prices increased or decreased, in (OUR COUNTRY) in 2006? I can tell you that the exact figure is between -1% and 20%.

Inflation rate	EU27	BE	BG	CZ	DK	DE	EE	EL	ES	FR	IE	IT
Average	4.8	4.1	8.9	6.8	4.2	3.0	3.2	4.1	8.8	5.9	5.9	4.3
Official rate		2.3	7.4	2.1	1.9	1.8	4.4	3.3	3.6	1.9	2.7	2.2

Inflation rate	CY	LV	LT	LU	HU	MT	NL	AT	PL	PT	RO
Average	6.8	4.1	7.6	9.1	9.0	4.2	9.9	6.7	3.7	4.8	6.8
Official rate	2.2	6.6	3.8	3.0	4.0	2.6	1.7	1.7	1.3	3.0	6.6

Inflation rate	SI	SK	FI	SE	UK	HR	TR
Average	4.3	7.9	3.7	7.4	4.0	3.3	5.1
Official rate	2.5	4.3	1.3	1.5	2.3		9.3

	RO	CY	BG	TR	MT	LU	ES	IT	PT	LT	UK	PL
DK	82%	79%	77%	73%	67%	66%	65%	65%	61%	59%	58%	57%
CORRECT	4%	4%	7%	10%	6%	8%	9%	11%	11%	6%	9%	2%
NOT CORRECT	14%	17%	16%	17%	27%	26%	26%	24%	28%	35%	33%	41%

	LV	BE	EL	EU27	FR	EE	CZ	IE	SE	HU	FI
DK	56%	55%	55%	53%	53%	47%	43%	43%	42%	39%	38%
CORRECT	18%	15%	17%	13%	14%	14%	7%	4%	8%	3%	13%
NOT CORRECT	26%	30%	28%	34%	33%	39%	50%	53%	50%	58%	49%

	SI	AT	DE	SK	DK	NL
DK	36%	35%	34%	34%	31%	31%
CORRECT	35%	19%	26%	14%	22%	24%
NOT CORRECT	29%	46%	40%	52%	47%	45%

Unemployment Rate

QA54 What was the official unemployment rate, the percentage of active people who do not have a job, in (OUR COUNTRY) in 2006? I can tell you that the exact figure is between 0% and 20%.

Unemployment rate	EU27	BE	BG	CZ	DK	DE	EE	EL	ES	FR	IE	IT
Average	10.2	11.5	13.2	10.6	6.4	10.3	9.2	12.5	8.2	10.0	8.2	7.6
Officiel rate 2006 (%)	7.9	8.2	9.0	7.1	3.9	8.4	5.9	8.9	8.5	9.4	4.4	6.8

Unemployment rate	CY	LV	LT	LU	HU	MT	NL	AT	PL	PT	RO
Average	7.4	10.7	9.6	6.7	12.9	8.5	8.9	8.7	15.3	9.7	10.6
Officiel rate 2006 (%)	4.7	6.8	5.6	4.7	7.5	7.3	3.9	4.8	13.8	7.7	7.3

Unemployment rate	SI	SK	FI	SE	UK	TR
Average	10.0	12.4	10.6	7.8	9.5	15.9
Officiel rate 2006 (%)	6.0	13.4	7.7	7.0	5.3	9.5

	RO	BG	ES	TR	CY	IT	LV	MT	UK	LT	PT	EU27
DK	85%	75%	70%	70%	67%	67%	65%	63%	63%	61%	59%	48%
CORRECT	5%	5%	13%	7%	12%	5%	8%	11%	6%	7%	11%	19%
NOT CORRECT	10%	20%	17%	23%	21%	28%	27%	26%	31%	32%	30%	33%

	LU	PL	IE	EE	HU	SI	EL	FR	BE	AT	NL
DK	47%	46%	45%	45%	42%	41%	39%	36%	32%	31%	31%
CORRECT	25%	23%	14%	12%	13%	8%	19%	47%	12%	13%	7%
NOT CORRECT	28%	29%	41%	43%	45%	51%	42%	17%	56%	56%	62%

	DE	FI	CZ	SE	SK	DK
DK	26%	26%	26%	25%	23%	15%
CORRECT	29%	29%	17%	27%	45%	18%
NOT CORRECT	45%	45%	57%	48%	32%	67%

QA50-QA52-QA54 - Number of correct answers

	EU27	BE	BG	CZ	DK	D-W	DE	D-E	EE	EL	ES	FR	IE	IT	CY	LV	LT
TOTAL	26717	1011	1039	1043	1002	1005	1513	508	1005	1000	1000	1013	1000	1010	502	1013	1018
	18048	765	861	652	616	506	772	272	636	648	781	485	664	818	422	690	760
0 correct answer	*68%*	*76%*	*83%*	*62%*	*61%*	*50%*	*51%*	*54%*	*63%*	*65%*	*78%*	*48%*	*66%*	*81%*	*84%*	*68%*	*75%*
	6088	182	140	282	267	348	518	167	246	214	179	299	256	140	60	250	195
1 correct answer	*23%*	*18%*	*13%*	*27%*	*27%*	*35%*	*34%*	*33%*	*25%*	*21%*	*18%*	*30%*	*26%*	*14%*	*12%*	*25%*	*19%*
	2216	56	32	91	102	134	201	67	113	111	36	177	78	45	14	59	49
2 correct answers	*8%*	*5%*	*3%*	*9%*	*10%*	*13%*	*13%*	*13%*	*11%*	*11%*	*4%*	*17%*	*8%*	*4%*	*3%*	*6%*	*5%*
	365	9	6	18	17	17	21	1	11	27	3	53	1	6	6	14	14
3 correct answers	*1%*	*1%*	*1%*	*2%*	*2%*	*2%*	*2%*	-	*1%*	*3%*	-	*5%*	-	*1%*	*1%*	*1%*	*1%*

	LU	HU	MT	NL	AT	PL	PT	RO	SI	SK	FI	SE	UK	HR	TR
TOTAL	511	1006	500	1009	1011	1000	1011	1019	1013	1106	1038	1005	1319	1000	998
	327	798	414	616	612	669	809	906	541	445	592	611	1051	670	827
0 correct answer	*64%*	*79%*	*83%*	*61%*	*61%*	*67%*	*80%*	*89%*	*54%*	*40%*	*57%*	*61%*	*80%*	*67%*	*83%*
	140	164	67	282	279	254	134	89	286	379	365	307	212	236	116
1 correct answer	*27%*	*17%*	*13%*	*28%*	*27%*	*26%*	*13%*	*9%*	*28%*	*34%*	*35%*	*30%*	*16%*	*24%*	*12%*
	36	43	19	100	98	74	59	20	166	228	78	78	49	88	42
2 correct answers	*7%*	*4%*	*4%*	*10%*	*10%*	*7%*	*6%*	*2%*	*16%*	*21%*	*8%*	*8%*	*4%*	*9%*	*4%*
	8	1	0	11	22	3	10	4	21	54	3	8	7	5	13
3 correct answers	*2%*	-	-	*1%*	*2%*	-	*1%*	-	*2%*	*5%*	-	*1%*	-	-	*1%*

QA56 Do you agree or disagree with the following statement concerning economic figures like the ones just discussed: "It is necessary to know these figures"?

	EU27	BE	BG	CZ	DK	D-W	DE	D-E	EE	EL	ES
TOTAL	26717	1011	1039	1043	1002	1005	1513	508	1005	1000	1000
DK	2366 *9%*	32 *3%*	156 *15%*	47 *5%*	42 *4%*	75 *7%*	103 *7%*	22 *4%*	65 *6%*	14 *1%*	195 *20%*
Agree	18405 *69%*	725 *72%*	543 *52%*	708 *67%*	607 *61%*	625 *63%*	949 *62%*	331 *66%*	709 *70%*	817 *82%*	666 *66%*
Disagree	5946 *22%*	254 *25%*	340 *33%*	288 *28%*	353 *35%*	306 *30%*	460 *31%*	155 *30%*	231 *24%*	170 *17%*	140 *14%*

	FR	IE	IT	CY	LV	LT	LU	HU	MT	NL	AT
TOTAL	1013	1000	1010	502	1013	1018	511	1006	500	1009	1011
DK	44 *4%*	186 *19%*	92 *9%*	7 *1%*	36 *4%*	83 *8%*	24 *5%*	66 *7%*	35 *7%*	37 *4%*	66 *7%*
Agree	853 *85%*	678 *67%*	751 *74%*	478 *96%*	660 *65%*	543 *54%*	381 *75%*	736 *73%*	373 *75%*	477 *47%*	611 *60%*
Disagree	117 *11%*	137 *14%*	167 *17%*	17 *3%*	317 *31%*	392 *38%*	107 *20%*	204 *20%*	92 *18%*	495 *49%*	334 *33%*

	PL	PT	RO	SI	SK	FI	SE	UK	HR	TR
TOTAL	1000	1011	1019	1013	1106	1038	1005	1319	1000	998
DK	89 *9%*	102 *10%*	173 *17%*	30 *3%*	81 *7%*	15 *1%*	29 *3%*	141 *11%*	65 *7%*	296 *30%*
Agree	625 *63%*	807 *80%*	790 *77%*	578 *57%*	775 *71%*	744 *72%*	657 *65%*	876 *66%*	625 *62%*	581 *58%*
Disagree	286 *28%*	102 *10%*	56 *6%*	404 *40%*	249 *22%*	279 *27%*	320 *32%*	302 *23%*	310 *31%*	122 *12%*

QA57 Some people say that statistical information play an important role in business, public and political decision making. Personally, do you think that, in (OUR COUNTRY), political decisions are made on the basis of statistical information?

	EU27	BE	BG	CZ	DK	D-W	DE	D-E	EE	EL	ES
TOTAL	26717	1011	1039	1043	1002	1005	1513	508	1005	1000	1000
DK	3288 *12%*	57 *6%*	276 *27%*	76 *7%*	22 *2%*	107 *11%*	154 *10%*	42 *8%*	113 *11%*	15 *1%*	215 *21%*
Yes	16495 *62%*	700 *69%*	431 *42%*	557 *53%*	883 *89%*	652 *65%*	965 *64%*	301 *60%*	611 *61%*	586 *59%*	586 *59%*
No	6934 *26%*	254 *25%*	332 *31%*	410 *40%*	96 *9%*	246 *24%*	395 *26%*	166 *32%*	281 *28%*	399 *40%*	199 *20%*

TOTAL	FR	IE	IT	CY	LV	LT	LU	HU	MT	NL	AT
DK	98 *10%*	190 *19%*	113 *11%*	98 *19%*	105 *11%*	154 *15%*	42 *8%*	108 *11%*	83 *17%*	51 *5%*	90 *9%*
Yes	682 *67%*	703 *70%*	569 *56%*	294 *59%*	497 *49%*	569 *56%*	345 *68%*	493 *49%*	318 *63%*	781 *77%*	703 *69%*
No	232 *23%*	107 *11%*	328 *33%*	110 *22%*	411 *40%*	295 *29%*	124 *24%*	405 *40%*	99 *20%*	178 *18%*	219 *22%*

TOTAL	PL	PT	RO	SI	SK	FI	SE	UK	HR	TR
DK	137 *14%*	174 *17%*	228 *22%*	59 *6%*	85 *8%*	38 *4%*	37 *4%*	184 *14%*	74 *7%*	122 *12%*
Yes	554 *55%*	657 *65%*	566 *56%*	593 *59%*	558 *50%*	731 *70%*	729 *72%*	853 *65%*	505 *51%*	556 *56%*
No	309 *31%*	181 *18%*	225 *22%*	361 *35%*	463 *42%*	269 *26%*	239 *24%*	282 *21%*	421 *42%*	320 *32%*

QA58 Personally, how much trust do you have in the official statistics in (OUR COUNTRY), for example the statistics on unemployment, inflation or economic growth? Would you say that you tend to trust these official statistics or tend not to trust them?

	EU27	BE	BG	CZ	DK	D-W	DE	D-E	EE	EL	ES
TOTAL	26717	1011	1039	1043	1002	1005	1513	508	1005	1000	1000
Tend to trust	12129 *46%*	573 *57%*	408 *39%*	568 *55%*	730 *73%*	435 *43%*	649 *43%*	210 *41%*	601 *60%*	528 *53%*	468 *47%*
Tend not to trust	12077 *45%*	382 *38%*	423 *41%*	431 *41%*	245 *24%*	502 *50%*	773 *51%*	283 *56%*	327 *32%*	461 *46%*	427 *43%*
DK	2511 *9%*	56 *5%*	208 *20%*	44 *4%*	27 *3%*	68 *7%*	91 *6%*	15 *3%*	77 *8%*	11 *1%*	105 *10%*

	FR	IE	IT	CY	LV	LT	LU	HU	MT	NL	AT
TOTAL	1013	1000	1010	502	1013	1018	511	1006	500	1009	1011
Tend to trust	357 *35%*	567 *57%*	422 *42%*	304 *60%*	449 *44%*	547 *54%*	330 *65%*	359 *36%*	263 *53%*	775 *77%*	507 *50%*
Tend not to trust	609 *60%*	214 *21%*	415 *41%*	148 *30%*	492 *49%*	365 *36%*	150 *29%*	554 *55%*	181 *36%*	193 *19%*	389 *39%*
DK	47 *5%*	220 *22%*	173 *17%*	50 *10%*	72 *7%*	106 *10%*	31 *6%*	93 *9%*	55 *11%*	41 *4%*	115 *11%*

	PL	PT	RO	SI	SK	FI	SE	UK	HR	TR
TOTAL	1000	1011	1019	1013	1106	1038	1005	1319	1000	998
Tend to trust	479 *48%*	596 *59%*	519 *51%*	515 *51%*	534 *48%*	714 *69%*	619 *61%*	441 *33%*	428 *43%*	482 *48%*
Tend not to trust	423 *42%*	289 *29%*	266 *26%*	466 *46%*	524 *48%*	303 *29%*	349 *35%*	765 *58%*	507 *51%*	338 *34%*
DK	98 *10%*	126 *12%*	234 *23%*	32 *3%*	48 *4%*	21 *2%*	37 *4%*	113 *9%*	65 *6%*	178 *18%*

Chapter 14
What Do Citizens Know about Statistics?: The Results of an OECD/ ISAE survey on Italian Consumers

Fiona Fullone; Marco Malgarini
Istituto Di Studi E Analisi Economica, Italy

Michela Gamba; Enrico Giovannini
OECD

Abstract

This paper presents the results of a survey by Italy's ISAE (the Institute for the Study and Analysis of Economics) conducted in March-April 2007. The survey ascertains Italians' knowledge of key economic indicators (unemployment, inflation, growth rate) and is a companion survey to those conducted by Eurobarometer (throughout Europe) and the University of Michigan (in the US), as part of the OECD's project on measuring the progress of societies. First, the methodology and questionnaire of the OECD/ISAE survey on statistical knowledge are presented, and the results obtained at the aggregate level are analysed. An econometric model is then presented that is regressed by socio-economic conditions (age, gender, income, etc.), cultural attitudes and the information channels used to access economic data. Commentary about the results obtained concludes the paper

Introduction

The importance of statistical information for democratic processes has been underlined by "public choice" models. The recent literature on the relationships between public opinion, political choices and the functioning of modern democracies argues that there are big differences in what the general public and specialists, such as economists, think about key issues. Increasing attention is given to public opinion, even when it is poorly informed.

Blinder and Krueger (2004) presented some evidence about what U.S. citizens actually know about key economic facts. They found that a significant number of Americans do not know very much about the country's economic situation. They also tested a range of factors that might explain how people's beliefs are shaped. They found that ideology was the most important determinant in shaping the public's opinion, self-interest was the least important, and economic knowledge was in between. Therefore,

their findings seem consistent with an idea from political science: people often use ideology as a short cut for deciding what position to take, especially when properly informing oneself is difficult. They conclude that "there is room for hope that greater knowledge will improve decision making, even though it appears from our survey that efforts in this direction have shown less than impressive results to date".

Recent work relates more directly to statistics and their dissemination. Carroll (2003) tests a model of how empirical expectations are formed. His approach takes the news as the key provider of information on macroeconomic variables. He adds to this, firstly, the idea that people do not update their expectations and personal forecasts continuously but probabilistically. In addition, he looks at the role professional forecasters play in informing the media. Empirical work by Doms and Morin (2004) supplements Carroll's (2003) analysis. These authors elaborate the role of the media. Particularly, they establish three important ways through which the media affects the public's views on the state of the economy: *(i)* by conveying economic data and expert opinions; *(ii)* by sending a signal based on the tone of the economic report; and *(iii)* by the volume of reporting, which influences the likelihood of people updating their expectations (this adds to the signal value of the amount of reporting).

Following the Blinder and Kruger's example, the OECD has promoted the first co-ordinated international survey on what citizens know about key economic statistics. The full results of the survey (carried out, in April 2007, in 29 European countries – by Eurobarometer – and US – by the University of Michigan) will be presented during the OECD World Forum.

This paper presents the results obtained for Italy through a survey carried out by ISAE in March-April 2007 on Italian citizens. In particular, Section "The OECD Producject on 'Measuring the Progress of Societies'" introduces the OECD project on Measuring the progress of societies, presenting the companion surveys performed by Eurobarometer and the University of Michigan on the issues dealt with in this paper. The following section introduces the ISAE survey on Italian consumers and the questionnaire used for the OECD/ISAE survey on statistical knowledge. Section "the OECD/ISAE Survey on Statistical Knowledge" presents the results obtained at an aggregate level and Section "An Econometric Model of Knowledge" introduces an econometric model of knowledge, in which an aggregate measure of statistical knowledge appropriately derived from survey results is regressed on socio-demographic characteristics of the respondents, the desire and willingness of being informed and the sources of information used. Some considerations on the results obtained conclude the paper.

The OECD Project on "Measuring the Progress of Societies"

As cited in Giovannini (2006), initiatives to measure progress at the international, national and local levels are proliferating all over the world in response to a growing demand for more meaningful measures of economic, social and environmental change and for more accountability of public policies and politicians. Meetings organised in every continent in preparation of the second World Forum, with more than 100 countries, as well as the enthusiasm demonstrated for this initiative by organisations and individuals working around the world in the public, private and citizen sectors, show that there is a concrete demand for a global community of practice for those wanting to measure progress.

The Project will achieve its mission through advocating the importance of this work, improving the state of the art on the measurement and dissemination of progress measures and assisting countries to undertake this work. There are four key goals:

- **Foster a global conversation about what progress actually means**. In order to measure and achieve progress, people need to know what "progress" looks like. There can be no single answer, but by bringing together different communities, cultures and interest groups the project will debate and recognise differing views and find common ground. Such a discussion will benefit anyone seeking to measure progress at the sub-national or national level, but it will be important at the global level too.

- **Galvanise people and institutions to action**. By bringing together an engaged global community of practice, the Project will facilitate the collaboration of diverse groups and the sharing of success stories about the development and use of progress indicators, thereby fostering the development of evidence-based public choice and a facts-based civic dialogue, improving the democratic functioning of modern societies.

- **Improve the effectiveness of indicator work and their use for policy making**. By sharing best practices among those working on indicator initiatives, and strengthening international comparisons, the Project will improve the ways in which indicator sets are developed, disseminated and, most importantly, used. In addition to a technical discussion about indicators, an important element of the project will be to foster the debate on the ways in which policies can be improved through the use of indicators.

- **Make a key contribution to the international discussion** in the run up to 2015 when the set of existing Millennium Development Goals and Indicators (mainly designed for developing countries) will be reviewed. The project will integrate the current top-down approach to the development of international indicators with a **bottom-up effort**, to take into account cultural, social and economic differences around the world.

To achieve its goals, the Global Project will carry out activities in the following areas:

- Carry out statistical research on the measurement of societal progress in all its dimensions;

- Design, develop and promote the use of innovative ICT tools to facilitate the transformation of statistics into knowledge;

- Establish a global network to foster the measurement of progress in each and every country;

- Develop a global infrastructure to facilitate the assessment of societal progress at national and global levels to bring about evidence-based policy discussions and decision making.

More information about the Project is made available to Forum participants. In this context, the Project advocates the organisation of an international survey on what the

general population knows about key economic, social and environmental phenomena in their countries.

The ISAE Consumers Opinion Survey: A Methodological Introduction

In the framework of an EU project harmonised by the European Commission,[1] ISAE has been conducting since 1973 a monthly survey on consumers' opinion. The survey consists of qualitative questions on the economic and personal situation of consumers. Questions generally allow five possible answers, ranging from strongly positive to strongly negative; results are usually expressed as weighted balances of positive and negative replies, assigning double weight to extreme (positive and negative) answers.

The survey is carried out via telephone and combined with Computer Assisted Telephone Interviewing (CATI) System; it is based on a monthly sample of 2 000 Italian consumers, changing each month, for a total of 24 000 persons interviewed per year. The sample is extracted from the public telephone book registers and selected on the basis of a two-stage technique: in the first step, it is stratified according to the zone of residence and the size of municipalities (see Table 14.1); the second step is based on the selection of a specific consumer within the household selected in the first step. This selection is based on quota sampling according to gender (48.5% males, 51.5% females).[2]

Table 14.1 The ISAE Sample Size of Municipalities

Geographic zone	Less than 5 000	Between 5 000 and 10 000	Between 10 001 and 20 000	Between 20 001 and 50 000	Between 50 001 and 100 000	Between 100 001 and 500 000	More than 500 000	**Total**
North West	57	23	22	35	19	4	53	213
Centre-North	76	59	53	53	25	15	45	326
North-East	70	73	76	46	28	90	0	383
Centre	44	41	51	76	49	42	88	391
South	78	59	76	97	82	41	32	465
Islands	35	29	28	51	25	32	22	222
Total	360	284	306	358	228	224	240	2 000

For the aggregation of individual replies, ISAE has recently proposed a double-weighting system based on probability and post-stratification weights (see Fullone & Martelli 2006). More specifically, probability weights are the inverse of the selection probability; they are used to correct for possible selection bias deriving from unequal selection probabilities, linked to the nature of the list of reference and the size of the family selected. On the other hand, post stratification weights aim to correct for possible representativeness problems deriving, for instance, from the fact that unemployed and retired people are easier to contact than employee or self-employed. To correct for this kind of bias, ISAE uses an ex-post calibration method based on auxiliary information derived from official structural population statistics: more specifically, we make use of auxiliary information about size of municipality, geographic region, education, type of occupation and age of the respondents.

The OECD/ISAE Survey on Statistical Knowledge

In March and April 2007 surveys ISAE in collaboration with the Statistics Directorate of OECD, has added a number of questions on the knowledge of some key statistical and demographic figures, together with three questions that aim to measure respectively, the willingness/desire to be informed about statistical issues and the media tools that are used to obtain such information.

The Questionnaire

The March questionnaire focussed on the knowledge of official figures for key economic variables such as GDP growth, inflation, unemployment rate, deficit to GDP ratio, and the Euro/ Dollar exchange rate. The April questionnaire, while repeating the questions on GDP growth, Inflation rate, willingness and the media used to acquire information, engaged on the knowledge of key demographic variables such as the size of Italian population and life expectancy at birth. Some other questions on structural aspects of the economy (concerning respectively air pollution, students' literacy and R&D expenditure) were also added to the April questionnaire. Questions on economic figures were generally quantitative, *i.e.* we simply asked about the knowledge on the most recent published value of a given variable, as for the GDP question reported below:.

- ISTAT has recently published the official figures for Italian GDP growth in 2006. Can you tell us the rate of growth of Italian GDP in 2006?
  In the case of the Euro/ Dollar exchange rate we preferred to ask a qualitative question structured on a Likert scale about the increase/ decrease of the exchange rate:

- One year ago 1 Euro was worth 1.20 US Dollars. In other words, with 1 Euro it was possible to buy 1.20 Dollars. Can you tell us if the Euro is now worth more or less than 1.20 dollars; in other words, do you think that now with 1 Euro you can buy more, an equal amount or less than 1.20 Dollars?
  Questions on demographics issues, which are quantitative as those on GDP and respondents were asked to choose among different possible answers, each indicating a possible interval of values. For example:
  - Life expectancy at birth is an important indicator of the welfare of a country; according to official statistics, can you tell us how many years a person born today in Italy may expect to live? Please tick one of the following options:
    - Less than 60 years
    - Between 60 and 65 years
    - Between 66 and 70 years
    - Between 71 and 75 years
    - Between 76 and 80 years
    - Between 81 and 85 years
    - Between 86 and 90 years
    - More than 90 years
    - Don't Know
    - Refuse to answer

In both the March and April's questionnaires, two sets of questions intended to measure, respectively, the importance attributed to economic information (from "extremely important" to "not important"), the desire to be informed (yes/ no) and the main channels used to acquire information (possible answers being TV; newspapers; internet; radio; friends and relatives; political and civic leaders) were added to the existing list.

Participation Rates and Aggregate Results: Quantitative Questions

Questions have been proposed alternatively in March and April, with only those on GDP, inflation and on the desired to be informed/media repeated for both the questionnaires. In order to reach a better assessment on the quality of the survey, for each question, respondents refusing to answer were distinguished from those answering "I don't know" to the given question. Table 14.2 reports the main results for the quantitative questions; response rates are on average quite low, as it is common in this kind of surveys (Blinder & Krueger 2004); it is interesting to note that on average around 30% of the respondents referrers to questions on inflation and on the unemployment rate, whereas only a 13.8 to the deficit to GDP ratio. However, most of the people refusing to answer are claiming that they are not capable of reporting an appropriate value, that is to say that they refuse to answer simply because they do not know the answer to that specific question.

Knowledge about the Phenomena of Interest

After having weighted the results so to consider possible selection bias and representative problems, Italian consumers tend to be slightly "over-optimistic" when reporting about the GDP growth rate and over "pessimistic" for inflation and, especially, for the unemployment rate and the deficit to GDP ratio: a 3.8% average rate is reported for inflation (as opposed to the 1.8-1.7 % official figures for March/ April 2007), whereas an over-estimated 14.5% (official figure 6.8%) and an 8.5% (official figure 4.4%) are the reported figures for unemployment and deficit to GDP ratio.

Table 14.2 Statistical Knowledge - Quantitative Questions

	GDP			Inflation			Unemployment	Deficit/ GDP
%	March	April	*Average*	March	April	*Average*	March	April
Response rate	27	19.2	*23.1*	32.6	23.6	*28.1*	32.1	12.7
Don't know	70.5	72.8	*71.7*	65.1	68.8	*66.9*	65.4	84.5
Refuse to answer	2.5	8	*5.2*	2.3	7.6	*5*	2.5	2.8
Average	2.7	2.1	*2.5*	4.5	2.9	*3.8*	14.5	8.5
Median	2	1.9	*2*	2.4	2.1	*2.2*	10	3.4
P25	1.5	1.3	*1.5*	2	1.9	*2*	7	2
P75	2.4	2.2	*2.2*	3	3	*3*	16	4.4
Std. Dev.	3.7	1.8	*3.1*	8.7	2.9	*6.9*	*13*	*14.5*
True value	1.9*	1.9*	-	1.8**	1.7***	-	6.8****	4.4*

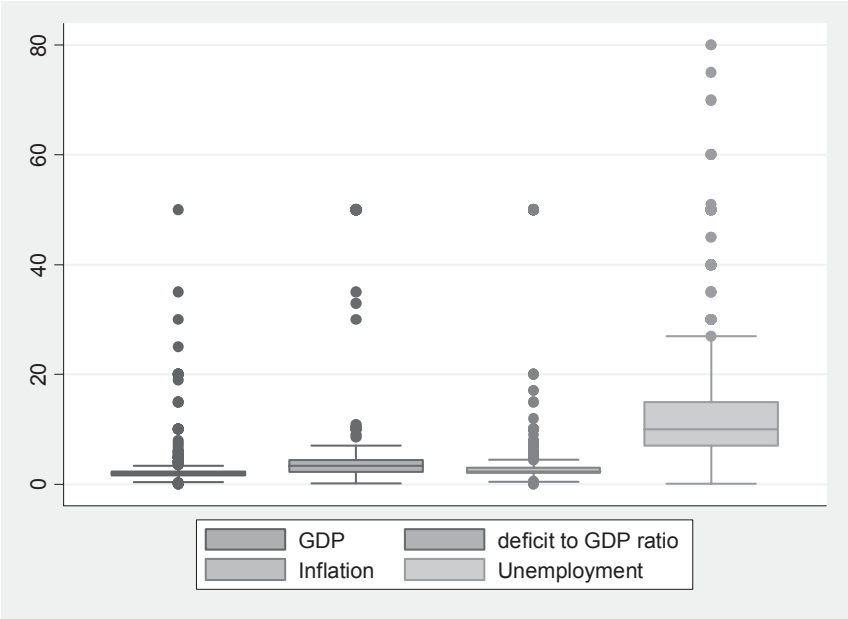
(*) 2006, (**) February, (***) March, (****) III Quarter 2006, seasonally adjusted.

The overall answers are strongly skewed to the right; in such cases, the sample mean is not an accurate estimator of population results, as it is confirmed by the fact that the

mean is always greater than the median of the distribution. In this sense, the median may be considered as a more accurate estimator of sample results: indeed, Italian consumers are shown to have an accurate "median" knowledge of official GDP data since the median is equal to 2 (1.9) in March (April); also median inflation assessments result to be fairly accurate- Italian consumers report a median of 2.2 inflation rate, only slightly overestimating the true value.

Overestimation is much more severe for unemployment, where the median unemployment rate is equal to 10% whereas the true value is 6.8%: a possible interpretation for this is that people are still not fully aware that unemployment in Italy is now well below the – probably psychologically important – 10% threshold, and it has been like this since 2001, after almost 10 years of double-digit unemployment figures (*i.e.*, the 1992-1993 recession). On the other hand the "median" consumers underestimate the official figures for the Maastricht parameter accounting for the state of public finances in Italy in 2006. In this case, it is possible that the underestimation is linked to the fact that besides the 4.4 figure for 2006, ISTAT has also disseminated a lower figure for the same variable –2.4 – corrected for "out of the ordinary" expenditures.

Figure 14.1 Box Plot Distribution
GDP growth, Inflation, Deficit to GDP ratio, Unemployment



Across all four subjects, the geographical distribution of the respondents is quite homogeneous, while a clear gender gap emerges: male response rates is always at least double than female (40% versus 20%), with the exception of the question about deficit/GDP, where female responses drop to 6% (male 21%). People aged 30-64 replied two times more than the older part of the sample. A clear difference emerges with respect to the degree of education: those who hold a university degree have a response rate three times higher than those with a primary education (four times for the question on the deficit/GDP ratio).

Income level also affects response rates. Those who earn more than 3 000 euros have a response rate for GDP and deficit/GDP three times higher than those who earn less than 1 500 euros per month, but smaller differences appear for questions about inflation and unemployment. Those who are self-employed present the highest response rates, up to twice higher than those of people unemployed or out of the labour force.

Socio-economic conditions also affect the precision of answers. For example, it is interesting to note that females indicate higher GDP growth (2.9) than males (2.6), inflation (5.9 and 3.7 respectively), unemployment (18 vs. 12.2) and deficit/GDP ratio (10.9 vs. 7.1). Females also present higher standard deviations for all variables.

Uncertainty

More generally, both standard deviation and the interquartile difference (*i.e.* the difference between the 25th and 75th percentile of the distribution answers, often interpreted as a proxy to the variance of the distribution) indicate that Italian consumers are highly hesitant about the official figures on the deficit to GDP ratio and the unemployment rate; dispersion around the mean is much lower for the knowledge on GDP growth and inflation. As a confirmation of the hypothesis of learning outlined above, the standard deviation is falling steadily between April and March for GDP growth (*i.e.*, one month after data dissemination), possibly also in relation to the observed decline in the response rate. In other words, it is possible that two opposite forces are taking place here: learning may ensure – for instance thru media discussions on the data – a more accurate knowledge after some time that the data are officially available; on the other hand, for the less attentive citizen, the time elapsed from the moment the data have reached the "headlines" implies a "I don't know" reply rather than an inaccurate evaluation on the phenomena of interest.

Participation Rates and Aggregate Results: Other Quantitative and Qualitative Questions

As expected, participation rates are much higher when we consider other quantitative questions (*i.e.* those in which the respondents have to choose among different alternatives) and some qualitative questions. Questions on life expectancy, Co_2 emissions and students' literacy levels are the most "successful" in terms of response rates, with those on population and R&D expenditure receiving the lower level of answers. Also in this case, the proportion of the sample that refused to answer is quite low, confirming on average the quality adequacy of the estimations.

Knowledge on the Exchange Rate and R&D Expenditure

Concerning the more economic-oriented questions – those on the Euro/ Dollar exchange rate and the R&D expenditure in Italy – people show an adequate qualitative knowledge of the underlying phenomena (Table 14.3): over 40% of the sample reports that the exchange rate is increased with respect to the 1.20 figure of twelve months before; similarly, almost half of the sample (48.7%) correctly reports that R&D expenditure of both the public and private Italian sector is lower than the European average. However, almost a 20% of the Italian consumers report that the exchange rate is "stable" and a significant 13% state it is actually decreased with respect to 12 months before. On the other hand, only a small proportion of the population (as low as 5%) believes that Italian firms and the Government spend in R&D more than our European

competitors, albeit more than 17% of them think that expenditures are equal to the European averages.

Knowledge on Socio-Demographic Statistics

People's knowledge on demographic variables results to be quite unsatisfactory: only a very small proportion of the sample (as low as 4.5%) correctly reports that Italian population is between 58 and 59 millions; more than 10% of Italian consumers (the mode of the distribution) affirms that population is "above 60 millions", and another 9.2% estimates it "in-between 59 and 60 millions" (with smaller groups reporting possible values below the "true" one). On the other hand, almost 32% of Italian consumers correctly estimate life expectancy at birth ranging between 76 and 80 years (the mode of the distribution), with almost the same amount of people reporting values (well) below or above the "true" one.

Concerning Co_2 emissions, people correctly reports that emissions have increased in the last five years; however they overestimate the impact of such increase which seems to be particularly "strong" for over 61% of the respondents. Finally, as for students' literacy (OECD 2003), almost 40% of Italian consumers correctly perceive that they are "under-qualified" with respect to their European counterpart. However, 8% of the sample claims that students are "much less qualified" and almost a 27% that they are "equally qualified".

Table 14.3 Response Rate on Qualitative Questions

	Response rate	Don't Know	Refuse to answer	Weighted Survey answers (true values in bold)	
Euro/ Dollar (%)	73.1	25.3	1.7	**Increased**	40.1
				Stable	19.8
				Decreased	13.2
R&D Expenditure(%)	71.4	26.6	2	Higher	5.1
				Equal	17.7
				Lower	48.6
Population (%)	63.4	34.5	2.1	Between 50ml and 58ml	39.5
				Between 58ml and 59ml	4.5
				More than 59ml	19.4
Life expectancy (%)	77.4	20.4	2.2	Less than 76 years	23.5
				Between 76 and 80 years	31.6
				More than 80 years	22.3
CO2 emissions (%)	89.7	8.5	1.9	Increased a lot	61.5
				Increased	16.4
				Stable	8.7
				Decreased	0.4
				Decreased a lot	2.7
Literacy (%)	78.5	19.5	2	Much more qualified	5
				More qualified	6.7
				Equally qualified	26.9
				Less qualified	31.8
				Much less qualified	8.1

How Important is Knowledge for Italian Citizens, and How Do They Know It?

The last part of the questionnaire, common to both March and April surveys, concentrates on the desire of Italian consumers to be informed, and on the main media

used to acquire such information. Table 14.4a reports survey results for the desire and importance of being informed: 46% of Italian consumers think that it is "important" to be informed, but only 8% of them believe that it is "extremely important", a result much lower than that obtained in an analogous survey for the United States[3] where almost 24% of the sample believed that information is "extremely important". Moreover, more than 15% of the sample believes that it is not important to be informed, and over 40% of people are not interested in being more informed on such issues.

Finally, Table 14.4b reports the media used more often to acquire such information; television is by far the most frequently used channel of information, mentioned by over 82% of Italian consumers. Newspapers and periodicals are important for almost the 52% of the population, followed by Internet (23.3%) that – quite surprisingly – precedes the radio (18.6%) as well as conversations with friends and relatives (11.2%) and political leaders (7%). Also in this case, the comparison with the US data is remarkable, showing that Italian consumers pay comparatively much less attention to informal information channels such as political and civic leaders opinions (cited as important sources of information by over 45% of the US sample) and discussions with friends and relatives (cited by over 35% of the sample in the US survey). On the other hand, television strongly dominates all the other media, also in comparison with the US data, according to which only the 61% of the sample use it regularly to acquire such information.

Table 14.4a Importance of Knowledge and Desire to be Informed

How important is to be informed on the issues we have discussed so far?	Average (%)
Extremely important	8.21
Very important	24.43
Important	46.05
Not very important	12.39
Not important	3.2
Don't know	3.96
Refuse to answer	1.77
Would you like to be informed further on these issues?	
Yes	54.84
No	40.44
Don't Know/ Refuse to Answer	4.72

Table 14.4b Information Channels (%)

Information Channels	
Television	82.05
Radio	18.58
Newspapers, periodicals	51.94
Internet	23.27
Political, opinion leaders	6.99
Friends, relatives	11.19
Don't know	2.35
Refuse to answer	1.8

An Econometric Model of Knowledge

Section "The OECD/ISAE Survey on Statistical Knowledge" has introduced a description of survey results, providing some first interesting insights of the quantitative and qualitative level of knowledge of Italian consumers about important socio-economic variables such as GDP, inflation, the Italian population, life expectancy and so forth. This section will then move a step forward trying to assess the determinants of people' knowledge, assuming that this is influenced by socio-demographic characteristics of the respondents, their desire to be informed and the media they use to acquire the relevant information. However in order to derive and estimate such a model, a synthetic measure of knowledge based on the data previously presented needs to be built.

The Knowledge Score

Theory

The intuition behind the construction of our knowledge measure, named "knowledge score" after the work of Blinder and Krueger (2004), is that for each question on knowledge a score is assigned to each respondent based on the accuracy of his/ her answer. As a starting point, we consider only the quantitative questions listed in the March survey, namely those on GDP growth rate, inflation, the deficit to GDP ratio, the unemployment rate and the Euro/ Dollar exchange rate; we therefore limit our sample to the 2 000 people that answered to the March survey. The sample is further reduced by the amount of missing values emerging for both questions.

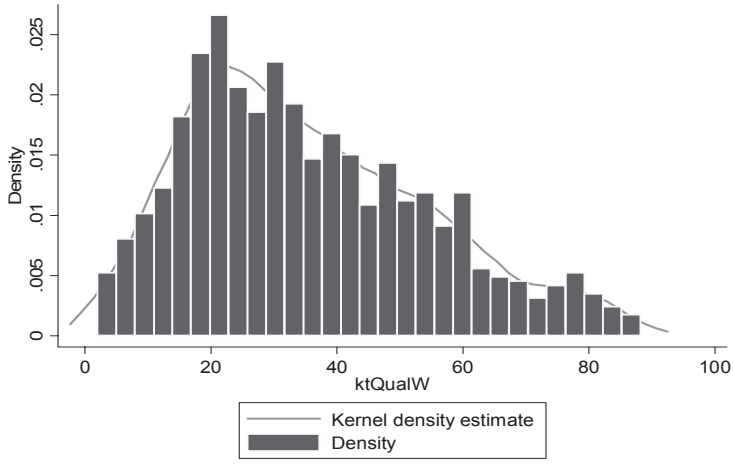
In literature, there are different methods to assign a score to a population sample, based on the results achieved in some kind of test, such as this survey on economic knowledge (Percentile Ranks, Standard Point (z), Normalised Standard Point, T point, Stanine Point). Generally speaking, when the distribution of the results is not normal or particularly skewed, as it was the case for the knowledge questions, the Percentile Rank method is employed.

The Percentile Rank[4] is a simple method that results in an ordinal measure of survey answers. First of all, we calculate the absolute value of the individuals errors (*i.e.*, the difference between the individual answer and the official data), and then individual errors are sorted calculating a "raw score". In the final step we calculate the percentile rank on the basis of the ordinal position of the "raw score", weighting to correct for possible sampling bias[5] and assigning a higher rank the lowest the absolute value of the error. The resulting percentile rank gives an idea of the position of a respondent in the sorted sample. For example, a respondent which has a test result ("raw score") better than 90% of the sample is said to be at the 90th percentile of distribution answers.

Implementing the Knowledge Score to the Survey Data

The Knowledge Score is built on March data, including the quantitative questions on GDP, Inflation, Deficit/ GDP ratio and unemployment rate, together with the qualitative question on the Euro/ Dollar exchange rate. As an example, let us consider the first question about GDP growth; the response rate is roughly equal to 25% and the sample reports a slight overestimation of the true value. As a first step, we compute the absolute value of the error and consider it as a raw score. Secondly, we assign to each answer a percentile rank based on people's accuracy, giving the lowest percentile rank score to the greater error (consequentially, the highest percentile rank is assigned to the lowest error). We then derive the distribution of percentile ranks. In the next step, we assign a "zero point" label to people not answering that particular question. After repeating the experiment for the other questions comprised in the March survey, we can finally evaluate the aggregate Knowledge score as the average of the 4 different scores calculated for each question (Figure 14.2). In this case we find that there are 961 meaningful respondents in our sample, which answered to at least one of the five economic questions. The mean of the aggregate knowledge score distribution is 35.6 and its standard deviation is equal to 19.8.

Figure 14.2 Distribution of Aggregated Knowledge ScoresKernel Density Estimate Superimposed GDP growth, Inflation, Unemployment, GDP ratio, Euro/Dollar: March 2007- 961 observations



The Model and its Econometric Outcomes

Once the appropriate aggregate measure of statistical knowledge is derived, we proceed studying the determinants of knowledge according to the demographic characteristics of the respondents and their desire of being informed. More specifically, the model is:

$$K_i = f(SD_i, D_i, S_i) + u_i \qquad (1)$$

Where is the knowledge score, are the socio-demographic characteristics of the respondent, is the desire of being informed and is the source they use for being informed. The unobserved error term is assumed to be uncorrelated with the covariates; on the other hand we allow residuals to be heteroschedastic and we use robust methods in the OLS estimates.

We consider as control variables information extracted from the ISAE consumers surveys concerning gender (M/F), age (4 classes, from <30 years to 65+), zone of residence (North West, North East, Centre and South), employment status (4 categories, employees and self-employed, unemployed and inactive people), level of education (3 classes, lower, intermediate and University level) and income (divided into quartiles). Desire to be informed and the channels used to acquire information are measured on the basis of the answers provided to the survey questions.

Table 14.5 reports the results obtained using OLS; a total of 961 observations are available for estimation. We normalise with respect to male respondents, being dependent workers, in the first income quartile under 30 years of age, living in the North West of Italy, with the lowest education and having answered that information is extremely important and that they do not want to be informed more about these issues. In this sense the constant term may be interpreted as the average knowledge score for this very specific sub-group of the sample, and the coefficients of the various dummies represent – if significant – the increase/decrease in the knowledge score that may be associated with the possession of the specific characteristic that the dummy itself represents. OLS regressions are able to explain more than 26% of the total variability of the knowledge score.

Table 14.5 OLS Regressions for Economic Knowledge Score (K)

Variables	K-Score (GDP, Inflation, Unemployment, Deficit/GDP) March Survey			
	Coefficient	Std. Dev	t value	Stat sign
Constant term	29.94	3.37	8.88	***
Socio-demographic controls				
Professional category (baseline: employed)				
Self employed	4.1	1.76	2.33	**
Unemployed	1	1.06	0.94	
Out of labour force	1.35	0.38	3.54	***
Age (baseline: up to 30 years)				
30-50 years	4.01	2.1	1.91	*
50-65 years	9.49	2.15	4.41	***
> 65 years	1.88	2.35	0.8	
Gender (baseline: Male)				
Female	-12.08	1.12	-10.75	***
Zone of residence (baseline: North-West)				
North-East	0.01	1.62	0.01	
Centre	-0.94	1.6	-0.59	
South	-5.7	1.46	-3.91	***
Education (baseline: primary school)				
High School or less	5.05	1.27	3.99	***
University degree	8.72	1.89	4.62	***
Income (baseline: 1st quartile)				
Second Quartile	-0.25	1.56	-0.16	
Third Quartile	0.7	1.57	0.44	
Fourth Quartile	2.03	1.51	1.34	
Importance and desire of being informed				
Importance of being informed (baseline: Extremely important)				
Very Important	-1.8	1.99	-0.9	
Important	-7.42	1.86	-3.98	***
Not very important	-11.45	2.47	-4.64	***
Absolutely not important	-19.29	4.15	-4.65	***
Desire to be more informed? (baseline: No)				
Yes	0.71	1.26	0.56	
Information channels				
Television	0.43	1.58	0.27	
Radio	1.39	1.37	1.01	
Newspapers	5.15	1.15	4.46	***
Internet	2.95	1.47	2	**
Political, civic leaders	-2.08	1.73	-1.2	
Friends, relatives	-0.19	1.74	-0.11	

Number of available observations: 961, R^2=0.26
Statistically Significance: *** = 10%, **+ 5%, ***= 1%**

Results are broadly in line with what it may be expected: knowledge of Italian consumers is significantly influenced by professional category, age, gender, zone of residence, education and personal income. It is much lower for women than for men and for Southern residents, being instead higher with higher education, for independent workers and inactive people (among which there are the students) and for people between 30 and 65 years of age. We do not find on the other hand any significant influence of income: in fact, it is possible that people are not particularly willing to provide information about their income, as it is testified by the relatively high non response (24% in March 2007) to this question.

People attributing lower importance to being informed are also those scoring worst results, whilst the declared desire of being informed does not significantly influence the results. As for the main information channels, a "knowledge divide" runs between those that are currently reading newspapers and magazines and those that are not, with the former realising a much higher score than the latter; also the use of the Internet has a positive impact on the knowledge of Italian citizens. All the other channels do not significantly influence the results.

Conclusions

This paper represents a first attempt to evaluate to what extent Italian citizens know official statistics concerning key economic variables. The results indicate a very little capacity of indicating, in quantitative terms, growth rates or levels of variables very often quoted in the public debate and frequently reported by media. Better results are obtained when people are asked to indicate, in qualitative terms, trends or levels of various phenomena. Differences in people's knowledge largely depend on socio-economic conditions, as well as on cultural attitudes and the channels used to access information.

These results should be taken very seriously by official statisticians, media and policy makers. Statisticians should rethink about their communication strategies, as well as the necessary investments to increase the statistical culture of citizens. Media should also think about new ways to make their communication on key economic figures more effective, especially vis-à-vis the more disadvantaged groups. Finally, policy makers should think about new ways to foster civic engagement and improve people's understanding of the main challenges faced by the country, as well as of alternative policy proposals.

Notes

1 For the complete questionnaire, see the DG Ecfin website at: http://europe.eu.int/economy_finance/indicators/business_consumers_surveys/userguide _en.pdf . See also Malgarini and Margani (2007), for a description of the ISAE survey.

2 With quota sampling, response rates are always equal to 100%: non responses lead to replacement in the sample until the quota is achieved. In the case of the ISAE survey, in order to achieve the goal of 24, 000 interviews each year a total of 120, 000 consumers are extracted from the telephone registers according to the stratification outlined above.

3 See Blinder and Krueger, 2004.

4 The mathematical formula for the percentile rank (PCTRank) is the following:
where is the cumulative frequency for all raw scores lower than the raw score of interest, is the frequency of the score of interest, and N is the number of examinees in the sample. See Crocker & Algina (1986).

5 In addition to the weights described in section 3 above, at this stage we also weight according to gender, given the fact that response rates for the knowledge questions were higher for men than for women.

References

Blinder A.S. and Krueger A.B. (2004), "What does the people Know About Economic Policy, and How Does It Know It?" , *NBER Working Paper* n. 10787, September.

Carroll C.D. (2003). "Macroeconomic expectations of households and professional forecasters", Quarterly Journal of Economics, 118 (1), p. 269-298.

Crocker, L. and Algina, J. (1986), "Introduction to classical and modern test theory", Harcourt Brace College Publishers.

Doms M. and N. Morin (2004), "Consumer Sentiment, the Economy, and the News Media", FRBSF Working Paper 2004-09, San Francisco: Federal Reserve Bank of San Francisco.

Fullone, F. and Martelli M. (2006), "Re-Thinking the ISAE Consumer Survey Processing", paper presented at the 28th CIRET Conference, Rome, September.

Giovannini, E. (2006), "Statistics and Politics in a *Knowledge Society*", OECD working paper, Paris.

Malgarini M. and Margani P. (2007), "Psychology, Consumer Sentiment and Household Expenditures: A Disaggregated Analysis", forthcoming in *Applied Economics*.

OECD (2003), "PISA Assessment Framework: Mathematics, Reading, Science and Problem Solving Knowledge and Skills", OECD Paris.

Part Five
Subjective Measures of Well-being

Chapter 15
A Short Introduction to Subjective Well-being: Its Measurement, Correlates and Policy Uses

André van Hoorn[i]
Nijmegen Center for Economics (NiCE), Radboud University Nijmegen, Netherlands

Abstract

Since the emergence of the field over five decades ago, the subjective well-being (SWB; e.g. *self-reported happiness or life satisfaction) literature has progressed rapidly. This overview discusses the state of the art in the measurement of SWB, the understanding of the factors determining SWB, and some uses of SWB in policy. Three broad conclusions emerge. First, SWB measures are meaningful in the sense that they are able to provide valid and reliable information on how well people and societies as a whole are doing. Second, a broad range of factors is found to correlate with or be a causal factor in SWB, both at the level of individuals and that of countries. Finally, SWB data can help shape and appraise policy. In several areas our knowledge is still limited, but the most recent and future research promises to secure further the usefulness of SWB research, not least in policy applications.*

Key words*: Happiness, life satisfaction, welfare, hedonic psychology, Easterlin paradox, cost-benefit analysis*

Introduction

Few people have ever doubted that happiness is very important. In fact, starting at least with the Ancient Greeks, the concept has been subject of unremitting debate, and, surely, this would not have been the case if people generally felt it did not matter.[1]

Since happiness has captured, and continues to capture, the interest of so many people, it should not come as a surprise that philosophers and many others pondering the concept have long yearned for a way to measure it. The breakthrough came in the 1950s. Psychologists – until then mainly interested in negative emotional states such as depression and anxiety – became interested in positive emotions and feelings of well-

i Correspondence: A.vanHoorn@fm.ru.nl. Helpful comments by Anat Itay, Ramzi Mabsout and Robbert Mabsout and Robert Maseland are gratefully acknowledged. Special thanks go to

being (cf. Sirgy *et al*. 2006). Within the discipline a consensus grew that self-reports on how well life is going, can convey important information on underlying emotional states, and so the field pushed ahead with measuring what is best referred to as subjective well-being (commonly abbreviated as SWB). SWB, though not the same as happiness in a strict sense, comprises several notions typically associated with happiness and the good life. It is 'a broad category of phenomena that includes people's emotional responses, domain satisfactions, and global judgements of life satisfaction' (Diener *et al*. 1999, p. 277). Specifically, reported SWB consists of two distinctive components (cf. Diener 1994, p. 106): an affective part, which refers to both the presence of positive affect and the absence of negative affect, and a cognitive part. The affective part is a hedonic evaluation guided by emotions and feelings, while the cognitive part is an information-based appraisal of one's life for which people judge the extent to which their life so far measures up to their expectations and resembles their envisioned 'ideal' life.

Since the emergence of the field over five decades ago, the SWB literature has progressed rapidly. First, as recent surveys show, psychologists and other social scientists have taken huge steps in their understanding of the factors influencing people's SWB. In addition, the methods through which empirical content is given to the concept of SWB have drastically improved, and are expected to continue to improve as increasing use will be made of advances in information- and communication technology. As such, SWB research solicits increasing attention of politicians, government officials and the public alike.

This paper must be seen in light of the increasing attention for SWB and provides a short introduction to the field. It starts with a discussion of the measurement of SWB in Section "Measurement of Subjective Well-being". In Section "Correlates and Determinants of Subjective Well-being" we examine some factors underlying observed differences in SWB, both within and between nations. Section "Some Policy Uses of Subjective Well-being" explores some policy uses of SWB, notably in cost-benefit analyses and in the form of national SWB statistics. We conclude with some summarising remarks in the last section.

Measurement of Subjective Well-being

Concerning the measurement of SWB, two issues are of key importance: the methods and instrument used to measure the SWB construct, and the question whether the measures thus obtained are able to convey meaningful information, *i.e.* whether they are reliable and valid. Both aspects of measuring SWB will be discussed below.

Measurement Scales and Methods

The hallmark of measures of SWB is that they are obtained through self-reports: people are asked to evaluate their lives as a whole or some aspect of it. The questions can be relatively straightforward and a widely used one simply asks: 'Taking all things together, would you say you are …: very happy, quite happy, not very happy or not at all happy'. Measures that are more elaborate use multiple items to target a specific part of SWB and consequently render more reliable results than single-item measures do.

Much-used multi-item scales are the Positive and Negative Affect Schedule (PANAS) scales (Watson *et al.* 1988) and the Satisfaction With Life Scale (SWLS; Diener *et al.* 1985). The former provides a list of feelings and emotions, asking respondents to indicate

the extent to which they felt this way during a given period, for example the past week. The PANAS is very flexible in that it does not specify particular feelings and emotions and many different combinations are possible: interested, distressed, excited, upset, strong, guilty, et cetera. Items from the SWLS elicit responses from individuals about the extent to which they agree or disagree with certain statements and it is designed specifically to capture satisfaction with life.[2]

A critical aspect of measuring SWB is the method through which SWB reports are obtained. Two important measurement approaches are the Experience Sampling Method (ESM) or Ecological Momentary Assessment (EMA) (see Scollon *et al*. 2003 and Stone *et al*. 1999 for overviews) and the Day Reconstruction Method (DRM; Kahneman *et al*. 2004). The hallmark features of ESM/EMA are that they solicit frequent and immediate reports from individuals in their normal surroundings. These qualities, in turn, provide them with important advantages over other methods for soliciting self-appraisals (see Stone *et al*. 1999, p. 27-28). Firstly, because individuals are in their natural settings when they complete the reports, the results are not distorted by unusual circumstances such as being in a laboratory; the method has high 'ecological validity'. Furthermore, ESM and EMA do not have to be limited to self-reports but can also solicit randomly timed reports on the details of people's (momentary) environment. In subsequent analysis, environmental circumstances can therefore easily be related to the subjective emotional assessments. A second advantage is that having subjects complete the self-evaluation on the spot avoids retrospective distortion. A large literature shows that retrospective biases can be substantial and the use of momentary appraisals helps to limit them as much as possible. Finally, high-frequency assessment will increase the reliability and validity of measured SWB and improve its empirical analysis (see the discussion below).

ESM and EMA refer to a general method of measurement and do not specify a particular way of implementing it. Researchers are completely free in choosing the specific scale or scales used to measure SWB, *e.g.* a PANAS scale, a physiological test (for instance blood pressure) or a combination of different measures. In addition, the way in which the actual sampling takes place is not pre-specified, although practical difficulties of course set boundaries on what is actually feasible. Modern information- and communication technologies offer interesting possibilities: when prompted to do so, respondents fill in the questionnaire send to their palmtops, and completed appraisals are included in a central database and analysed shortly after.

DRM is a more recently developed technique that shares most of the advantages of ESM/EMA but appears overall to be a more practical method. The technique asks subjects to keep a diary corresponding to events on or episodes of the day before. For the diary, they report on their experience, breaking it down in specific periods, *e.g.* by hour or by specific activities (commuting, household chores et cetera). In their research, Kahneman *et al.* (2004) simply asked respondents how they felt during each of the identified episodes, but again the method can be applied with any SWB scale.

Reliability and Validity

Reliability

The reliability of an indicator can be defined as its overall quality, *i.e.* its consistency and its ability to give the same results in repeated measurement. The most outstanding feature of reliability is the test-retest correlation of the specific measure under scrutiny. It

is well-known that minor differences in circumstances and technical features of the specific questionnaire used, affect the reported level of SWB (see Schwarz & Strack 1999, p. 62). Correspondingly, the test-retest correlation for most single-item measures is about 0.40, reaching 0.60 when the same question is asked twice during the same one-hour interview (ibidem). In general, the reliability of SWB measures as evidenced by test-retest correlation is substantially lower than that found for common microeconomic variables such as personal income (Krueger & Schkade 2007). It is important to realise, however, that test-retest correlation is not only substantially influenced by the lag between times of asking (*e.g.* one hour, two weeks), but also, and more importantly, by the specific measurement scale used. Not surprisingly, studies show that the more advanced measures, such as multi-item questionnaires, produce more reliable SWB scores (see the overview in Krueger & Schkade 2007).

Validity

Validity, not unlike reliability, comes in many sizes and shapes. In its regular use, validity denotes so-called 'construct validity' and refers to the actual meaning of a term or construct (Larsen & Fredrickson 1999, p. 44). That is, a measure has construct validity if it indeed is able to capture the construct it intends to capture, in this case SWB. Other important dimensions of validity, often contributing to an indicator's overall construct validity, are convergent validity and discriminant validity. The former refers to the correspondence between the indicator of interest and other indicators purporting to measure the same construct. If the measure of interest is to have convergent validity the results obtained should converge on the results obtained using these other measures. Discriminant validity is somewhat the converse of convergent validity: a measure does well on this account if it is unrelated to measures it should not be related with, *i.e.* indicators purporting to measure some opposing construct.

The SWB literature pays a lot of attention to the validity of its measures. The conclusion in most of the reviews is that measures of SWB do quite well on the various dimensions of validity mentioned above (see, for instance, Di Tella & MacCulloch 2006; Diener 1994; Diener *et al*. 1999; Frey & Stutzer 2002; Kahneman & Krueger 2006; Nettle 2005; Layard 2005). This is not to say, however, that SWB does not lack a universal definition. Indeed, as a concept, SWB as of yet does not have a specific theoretical basis. Accordingly, the field appears largely driven by empirical work: SWB is generally conceptualised through the specific indicators used, *i.e.* by the specific questions asked.

The most recognised challenge to measures of SWB and their validity concerns their demonstrated sensitivity to minor life events. Most importantly, whimsical circumstances – a famous example is the weather, whether it is rainy or sunny – have a significant impact on the level of SWB people report (Schwarz & Strack 1999, p. 62). Such effects are obviously an important weakness of SWB measures and, in fact, survey data in general (cf. Bertrand & Mullainathan 2001). Yet, the errors generally seem to be of a random rather than a structural nature, and consequently do not imply that the measurement instrument is systematically biased. Indeed, using large enough samples would go a long way in addressing possible problems introduced by contextual factors influencing the reported level of SWB.[3] This is actually a clear advantage of using ESM, EMA or DRM; in repeated measures of SWB for a single individual, the influence of trivial events is lessened.

At the same time, SWB indicators do face a more substantial challenge to their validity in cross-national comparisons. In particular, there is concern about the possible impact of culture and language on SWB ratings (see Wierzbicka 2004). By the very nature of SWB and the way in which it is measured, *i.e.* through questionnaires, cultural and linguistic factors will introduce biases in country-level SWB scores. The extent to which they indeed do so is largely unexplored: overall, culture is found to matter (*e.g.* Vittersø *et al*. 2005) though it is not necessarily an explanatory factor in observed differences in SWB (*e.g.* Diener *et al*.1995).[4]

Beyond these challenges, increasingly attention is paid to the relation between SWB on the one hand and human physiology and brain activity on the other. This work shows that variations in SWB are associated with concrete objectively observable biological phenomena.[5] Ultimately, work along these lines may lead to an objective anchor for SWB that would improve the comparability of SWB levels across national and linguistic barriers. Obviously many issues – not least the question of causality – need resolving, before a reliable biological basis is given to observed variation in SWB, but a very preliminary step in this direction has already been taken. In a recent study, Blanchflower and Oswald (2007) relate differences in SWB across countries to differences in blood pressure.

Correlates and Determinants of Subjective Well-being[6]

Surveys show that social sciences have taken huge steps in their understanding of the factors underlying differences in SWB ratings (*e.g.* Diener *et al*. 1999).[7] This section briefly discusses some of the uncovered correlates and determinants of SWB, classifying them in six familiar groups: (i) personality factors; (ii) contextual and situational factors; (iii) demographic factors; (iv) institutional factors; (v) environmental factors; and (vi) economic factors.[8]

Psychologists have deeply studied the influence of personality on SWB, and found it to be the strongest and most dependable factor underlying differences in SWB between persons. In a famous study, Tellegen *et al*. (1988) compared levels of SWB for monozygotic and dizygotic twins raised together and raised apart. Their study shows that 40% of the variance in positive emotionality and 55% of the variance in negative emotionality is attributable to genes, whereas shared familial circumstances account for only 22% and 2% of observed variance respectively. Other work has assessed the role of measured personality characteristics and these are consistently found to be highly significant predictors of SWB as well. Neuroticism and extraversion in particular go along way in accounting for differences in levels of SWB (see, for example, Hayes & Joseph 2003).

Although inherent factors play a fundamental role in SWB, individual, contextual and situational factors are also important sources of difference in SWB scores. Notably, a consistent finding across samples of individuals reporting on, amongst others, SWB is that better health is associated with higher SWB, and that married people report higher SWB than others.

The third group of factors associated with SWB concerns demography. Particular attention has been paid to how age matters for SWB and how reported SWB ratings differ between men and women, but overall the relationships are not clear-cut. For example, SWB is said to be U-shaped with age, but the opposite relation, an inverse U-shape is also found (*e.g.* Blanchflower & Oswald 2007; Easterlin 2006). Similarly, sex (and gender)

appears to matter but not in a straightforward matter and observed differences are generally small (cf. Diener *et al*. 1999, p. 292-293).

Institutional conditions constitute a fourth group of factors found to have a systematic relationship with SWB. For instance, the results of Frey and Stutzer (2000) suggest that forms of direct democracy such as referenda increase the level of SWB. At a more abstract level, Radcliff (2001) finds a positive relation between the ideological complexion of governments and levels of SWB. He also reports a positive correlation between qualitative features of the welfare state and SWB. Finally, Veenhoven (2000) finds that political and private freedom add to SWB but only in rich countries.

Environmental conditions are an important factor in observed differences in SWB that operates strictly at a macro level. For example, Rehdanz and Maddison (2005), using data on 67 countries between 1972 and 2000, find that climate variables have a highly significant effect on SWB and that climate changes due to global warming might reduce SWB around the world in the next decades (see also the next section). In much the same fashion, the analysis by Becchetti *et al*. (2007) confirms the link between climate and SWB but indicates that global warming might, in contrast, lead to higher SWB worldwide.

Finally, part of individual- and cross-country differences in well-being is attributable to differences in economic circumstances. The literature has developed a quite clear understanding of the role factors like unemployment and inflation play in SWB (*e.g.* Clark & Oswald 1994; Winkelmann & Winkelmann 1998; Di Tella *et al.* 2001; Becchetti *et al.* 2006). In particular, unemployment is found to affect SWB through two channels: it has a direct negative effect on people who lose their job (keeping income constant), and an indirect negative effect on the entire population (higher risk of losing a job). The most significant of all economic variables seems to elude us, however. It is still not entirely clear what the role of GDP (gross domestic product) and GDP growth is. In this context Abramovitz (1979, p.7) discusses the findings of Easterlin (1974) and first mentions the 'Easterlin paradox'. According to this paradox, within a country richer people on average report higher SWB than poorer people in the same country do, whereas a comparison between countries reveals only a minor relation between income levels and SWB. Moreover, within nations an increase in per-capita income over time is not associated with higher SWB. Much effort has been put in further investigation of Easterlin's paradox. This work has robustly established the importance of relative income, but for aggregate-level income the picture is less clear. In the cross-section, the level of per-capita GDP matters (at least up to a certain point), but time-series evidence for advanced countries points to a small, possibly zero, effect of income on SWB. A particularly interesting finding is by Di Tella *et al.* (2003). They examine time-series date showing that GDP and GDP growth are significantly correlated to SWB, but only if downtrending time dummies are included in the regression equation.[9]

Some PolicyUses of Subjective Well-being

Following the major advances in the field of SWB, some work is currently underway that evaluates the contribution it can make to shaping and appraising policy. More concretely, several ideas exist for the construction of 'a national index of subjective well-being' or a related indicator, which would subsequently serve as a key policy goal. In addition, people are experimenting with the use of SWB data to assess the costs and benefits of policy alternatives. We discuss both policy uses next.[10]

Towards a Comprehensive Picture of Well-being

Frameworks designed to give a comprehensive account of how well a nation and its citizens are doing generally discern three key areas. Measures of economic performance are complemented with indicators on important social and environmental issues. For example, the Belgian Federal Planning Bureau (2005) in its insightful TransGovern model regards society's access to human, environmental and economic resources as determining the evolution of well-being. Correspondingly, it discerns three basic capitals of development:

1. Human capital: comprising the standard of living (material well-being), health (both mental and physical) and knowledge/capacities (what individuals know and are able to do);
2. Environmental capital: including both natural resources (water, air, land and mineral resources) and the biosphere with all its biological diversity;
3. Economic capital: subdivided in physical and technological capital (equipments, buildings, infrastructure, and intangible assets including software and technology patents) and net financial assets.

Typically, well-being frameworks such as this one include only objective indicators; whether it be GDP or related measures of economic activity, water quality or telephone connections, illiteracy rates or patent applications, suicide numbers or victimisation rates, traffic accidents or life expectancy, there is no doubt that from these indicators much can be learned about the quality of life. At the same time, there is growing recognition that, though insightful, there is more to well-being than these measures are able to capture.

The central tenet of SWB research is that the study of well-being cannot get around measuring people's cognitive and affective reactions to life as whole or specific domains thereof (Diener & Suh 1997, p.200).[11] Most importantly, SWB is able to capture's people actual experience in a direct manner, while economic, social and environmental indicators do so only indirectly (p.205). This, in turn, matters because what is experienced does not have to coincide with objective conditions, and indeed a large deviation may be observed. In fact, it is often argued that SWB indicators are useful complements to objective indicators precisely because there is a divergence between what people (reportedly) experience on the one hand and what is captured in the objective indicators on the other (see, in particular, Diener & Seligman 2004, p.2-3).

Building on the above-discussed methods for measuring SWB and the perceived feeling that something important is missing in the current set of well-being indicators, researchers in the field have given the idea of national SWB statistics serious consideration (Diener 2000, 2006; Kahneman *et al*. 2004). To give a taste, a nationally representative sample could be given palmtops and asked to report on their SWB using ESM/EMA or DRM. The cross-section would comprise different age groups, geographical regions, occupational categories and income levels (Diener 2000, p.40). The actual survey items could further target all components of SWB, the affective component (the presence of PA and the absence of NA) and the cognitive component, so as to maximise the usefulness, reliability and validity of SWB thus measured. A 'national index of subjective well-being' or 'national well-being accounts' constructed along these lines promise to be very useful complements to existing objective indicators and to add greatly to a more comprehensive picture of the quality of life.

The Measuring-rod of Happiness

Pigou (1952, p.11) famously thought of money as, amongst others, a 'measuring-rod', defining economic welfare as 'that part of social welfare that can be brought directly or indirectly into relation with the measuring-rod of money'. For much of applied economics, money indeed is the measuring-rod Pigou foisted upon it. In many policy areas, public goods and externalities being the ones most familiar to economists, there is just no way of putting a price tag on the relevant dimensions of the different options available, however. In some of these cases, SWB can provide a way to deal with the problem of missing prices. Below we discuss three examples of how SWB can be used to put a price on certain developments or externalities otherwise escaping (monetary) valuation.[12]

Van Praag and Baarsma (2005) address a concrete valuation problem. They combine a more mainstream approach to valuing an intangible, viz. noise nuisance from an airport, with the use of SWB data. Ordinarily, a hedonic regression analysis of house prices is used to calculate the shadow costs of noise pollution. The hedonic approach to valuing intangibles is prone to leave some residual costs, however, and these can be valued using measured SWB. The method thus developed is particularly useful when for some reason, notably market imperfections, price differences do not fully capture the costs of an externality.

An important area in which SWB data can usefully be applied to circumvent the problem of missing prices is in environmental valuation issues. Notably, Welsch (2006) uses panel data on ten European countries to analyse the effect of air pollution on SWB. Controlling for income, he finds that differences in SWB, both between countries and within countries over time, can partially be attributed to environmental quality. The improvement in air quality over the period 1990-1997 is valued at roughly $750 per capita per year in the case of nitrogen dioxide and about $1 400 per capita per year in the case of lead (averaged over all countries).

Frey *et al.* (2007) notably expand the dimensions of life to which valuation techniques are applied and assess the costs of terrorism using SWB data. The case applies to the Republic of Ireland, during the period 1970-1999. Data is taken from the Eurobarometer survey, and in the analysis differences in life satisfaction are explained from levels of terrorism, household income and other personal characteristics. Their findings reveal that people, on average, would be willing to give up some 41% of their income to have the level of terrorism reduced to the level prevailing in more peaceful parts of the country.

Finally, Oswald and Powdthavee (2007) propose a method to use SWB data to calculate compensatory claims in tort cases, notably bereavement. The idea is that the emotional damage inflicted by the perpetrator to the victim is exactly offset by the positive well-being effect of a monetary settlement. A typical amount for such hedonic compensation is £100 000 or roughly $200 000.

Summarising Remarks and Conclusion

This paper has presented a very brief introduction to the field of SWB or subjective well-being along three lines, its measurement, its correlates and its policy uses. Key points, corresponding to the three aspects of SWB research discussed, are:

1. SWB measures are meaningful in the sense that they are able to provide valid and reliable information on how well people and societies as a whole are doing;
2. a broad range of factors are found to correlate with or be a causal factor in SWB, both at the level of individuals and that of countries;
3. SWB data can be used to shape and appraise policy.

These conclusions notwithstanding, it must be pointed out that important challenges remain. Notably, work on SWB thus far has not been very systematic and in important areas understanding is still limited. The implication is that much work remains to be done before SWB is able to play the role in policy many authors envision for it.[13]

However, as SWB research is a rapidly developing field, many gaps in the subject area are likely to be addressed in the near future. For instance, linking SWB measures to objective, biological phenomena could ultimately help cross-country comparison of SWB scores, which, in turn, could inform public policy (although, of course, without providing a recipe for government action). Similarly, some areas have already been identified in which measures of SWB can complement existing measures and extent the scope of cost-benefit analyses. With more such work underway, SWB indeed appears a very promising area of research, from the perspective of politicians, government officials and the public alike.

Notes

1 King and Napa (1998) actually investigate how much value people attach to happiness by examining its contribution to the desirability of a certain hypothetical life and compare it with the contribution wealth and the presence of meaning make to the desirability of a given life. Their results confirm the folk wisdom developed over the ages: happiness and meaning make for the good life, and their effect on the desirability of a certain life is some five (happiness) to six (meaning) times higher than that of wealth. Relatedly, Diener and Oishi (2004) asked a sample of college students from different countries to rate the importance of happiness and other values on a scale from 1 to 9. Happiness came out first with a score of 8.0, slightly above health and love & affection (7.9), but well above wealth (6.8), amongst others.

2 The statements are: (1) 'In most ways my life is close to my ideal', (2) 'The conditions of my life are excellent', (3) 'I am satisfied with my life', (4) 'So far I have gotten the important things I want in life', and (5) 'If I could live my life over, I would change almost nothing'.

3 Boulding (1972: p. 466) describes this effect of the law of large numbers as follows: 'Even if we cannot get very reliable measures for the individual, by the famous principle which I have sometimes called "Katona's Law", that the summation of ignorance produces knowledge, we may find an operation or instrument with self-cancelling random factors which will give us a much better measure for a hundred or a thousand individuals than we can get for one.'

4 This view on cultural and linguistic factors as a source of bias in cross-country comparisons is not meant to deny that the role of these factors in reported SWB itself is interesting and worthy of attention.

5 For example, using positron emission tomography (PET scans) of regional cerebral blood flow, Lane et al. (1997) found, amongst others, that pleasant and unpleasant emotions differed from neutral ones by statistically highly significant elevated blood flow in the vicinity of the medial prefrontal cortex, thalamus, hypothalamus and midbrain. Levesque et al. (2003) applied functional magnetic resonance imaging (fMRI) and found that sad feelings were associated with significant bilateral activations of the midbrain, the medial prefrontal cortex, the anterior temporal pole and the right ventrolateral prefrontal. Ryff et al. (2004) present some preliminary findings showing that people with higher SWB (specifically more meaning, purposeful engagement et cetera in their lives) have lower

levels of daily salivary cortisol and pro-inflammatory cytokines. In addition, for them the duration of REM sleep is longer than for those with lower levels of reported SWB. Finally, Steptoe et al. (2005) find that positive affect is associated with reduced neuroendocrine, inflammatory and cardiovascular activity. Positive affect was also inversely related to cortisol output during the day (controlling for other factors such as age and gender) and heart rate. During mental stress testing in the laboratory people with higher positive affect had smaller plasma fibrinogen stress responses.

6 This paragraph draws heavily on work by Stefano Castriota.

7 See also Diener et al. (1995), Diener and Seligman (2004), Frey and Stutzer (2002), Lyubomirsky et al. (2005), Nettle (2005) and Layard (2005), amongst others. Earlier accounts of the state of affairs in SWB research are provided by Fellows (1966), Veenhoven (1984) and Wilson (1967).

8 This overview is necessarily brief. Accordingly, important factors such as religious beliefs are not discussed, nor does the section pay attention to the key issue of causality—does high SWB lead to marriage or does marriage lead to high SWB. In general, SWB is found to foster good outcomes in many domains of life, e.g. in work life, relationships and health (Lyubomirsky et al., 2005). In addition, low SWB is associated with general health risk and higher rates of suicide (Koivumaa-Honkanen et al., 2000 and 2001).

9 For further evidence on the importance of both relative and absolute income see, amongst others, Clark and Oswald (1996), Diener et al. (1995), Easterlin (1995) and McBride (2001).

10 Other uses not discussed here are, amongst others, as an alternative to the behaviourist approach to analysing the effects of taxes (e.g. Gruber and Mullainathan, 2005) and in the construction of equivalence scales (e.g. van Praag van der Sar, 1988). Further note that, also not elaborated on here, interpersonal comparability is implicit in virtually all policy uses of indicators, whether objective or subjective.

11 Supporting the field of SWB's approach to well-being is Sumner's (1995) argument that an adequate theory of the nature of welfare or well-being cannot be objective and hence must be subjective.

12 Better known, contingent valuation provides a further approach to solving the problem of missing prices (see the symposium in the Fall 1994 edition of the Journal of Economic Perspectives for a thorough introduction).

13 And even then, it is hard to think of SWB as being the only relevant or all-encompassing measure.

References

Abramovitz, Moses (1979). 'Economic Growth and Its Discontents.' In: Michael Boskin (ed.) *Economics and human welfare: Essays in honor of Tibor Scitovsky*. New York: Academic Press, 3-21.

Becchetti, Leonardo, Stefano Castriota and Osea Giuntella (2006). *The Effects of Age and Job Protection on the Welfare Costs of Inflation and Unemployment: a Source of ECB anti-inflation bias?* Centre for Economic and International Studies (CEIS) Working Paper No. 245.

Becchetti, Leonardo, Stefano Castriota and David Andrés Londoño Bedoya (2007). *Climate, Happiness and the Kyoto protocol: someone does not like it hot*. Centre for Economic and International Studies (CEIS) Working Paper No. 247.

Belgian Federal Planning Bureau (2005). *Transdisciplinarity and the governance of sustainable development*. Belgian Federal Planning Bureau: Task Force Sustainable Development.

Bertrand, Marianne and Sendhil Mullainathan (2001). 'Do People Mean What They Say? Implications for Subjective Survey Data.' *American Economic Review, Papers and Proceedings*, 91 (2), 67-72.

Blanchflower, David G. and Andrew J. Oswald (2007a). *Hypertension and Happiness across Nations*. NBER Working Paper No. 12934. Cambridge MA: National Bureau of Economic Research.

Blanchflower, David G. and Andrew J. Oswald (2007b). *Is Well-being U-Shaped over the Life Cycle?* NBER Working Paper No. 12934. Cambridge MA: National Bureau of Economic Research.

Boulding, Kenneth E. (1972). 'Human Betterment and the Quality of Life.' In: Burkhard Strumpel, James N. Morgan and Ernest Zahn (eds.) *Human Behavior in Economic Affairs. Essays in honor of George Katona*. Amsterdam: Elsevier, 455-470.

Clark, Andrew E. and Andrew J. Oswald (1994). 'Unhappiness and unemployment.' *Economic Journal*, 104 (May), 648-659.

Clark, Andrew E. and Andrew J. Oswald (1996). 'Satisfaction and comparison income.' *Journal of Public Economics*, 61 (3), 359-381.

Di Tella, Rafael, Robert J. MacCulloch and Andrew J. Oswald (2001). 'Preferences over Inflation and Unemployment: Evidence from Surveys of Happiness.' *American Economic Review*, 91 (1), 335-341.

Di Tella, Rafael, Robert J. MacCulloch and Andrew J. Oswald (2003). 'The macroeconomics of happiness.' *Review of Economics and Statistics*, 85 (4), 809-827.

Di Tella, Rafael and Robert MacCulloch (2006). 'Some Uses of Happiness Data.' *Journal of Economic Perspectives*, 20 (1), 25-46.

Diener, Ed (1994). 'Assessing Subjective Well-Being: Progress and Opportunities.' *Social Indicators Research*, 31 (2), 103-157.

Diener, Ed (2000). 'Subjective Well-Being. The Science of Happiness and a Proposal for a National Index.' *American Psychologist*, 55 (1), 34-43.

Diener, Ed (2006). 'Guidelines for national indicators of subjective well-being and ill-being.' *Journal of Happiness Studies*, 7 (4), 397-404.

Diener, Ed, Robert A. Emmons, Randy J. Larsen and Sharon Griffin (1985). 'The satisfaction with life scale.' *Journal of Personality Assessment*, 49 (1), 71-75.

Diener, Ed and Shigehiro Oishi (2004). 'Are Scandinavians happier than Asians? Issues in comparing nations on subjective well-being.' In: Frank Columbus (ed.) *Asian Economic and Political Issues, Volume 10*. Hauppauge, New York: Nova Science, 1-25.

Diener, Ed and Martin E.P. Seligman (2004). 'Beyond Money. Toward an Economy of Well-Being.' *Psychological Science in the Public Interest*, 5 (1), 1-31.

Diener, Ed and Eunkook M. Suh (1997). 'Measuring quality of life: Economic, social and subjective indicators.' *Social Indicators Research*, 40 (1-2), 189-216.

Diener, Ed, Eunkook M. Suh, Richard E. Lucas and Heidi L. Smith (1999). 'Subjective Well-Being: Three Decades of Progress.' *Psychological Bulletin*, 125 (2), 276-302.

Diener, Ed, Eunkook M. Suh, Heidi Smith and Liang Shao (1995). 'National differences in reported subjective well-being: Why do they occur?' *Social Indicators Research*, 34 (1), 7-32.

Easterlin, Richard A. (1974). 'Does Economic Growth Improve the Human Lot? Some Empirical Evidence.' In: Paul A. David and Melvin W. Reder (eds.) *Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz*. New York / London: Academic Press, 89-125.

Easterlin, Richard A. (1995). 'Will raising the incomes of all increase the happiness of all?' *Journal of Economic Behavior & Organization*, 27 (1), 35-47.

Easterlin, Richard A. (2006). 'Life cycle happiness and its sources. Intersections of psychology, economics, and demography.' *Journal of Economic Psychology*, 27 (4), 463-482.

Fellows, Erwin W. (1966). 'Happiness: A survey of research.' *Journal of Humanistic Psychology*, 6 (1), 17-30.

Frey, Bruno S. and Alois Stutzer (2000). 'Happiness, economy and institutions.' *Economic Journal*, 110 (October), 918-938.

Frey, Bruno S. and Alois Stutzer (2002). 'What Can Economists Learn from Happiness Research?' *Journal of Economic Literature*, XL (June), 402-435.

Frey, Bruno S., Simon Luechinger and Alois Stutzer (2007). 'Calculating tragedy: Assessing the costs of terrorism.' *Journal of Economic Surveys*, 21 (1), 1-24.

Gruber, Jonathan H. and Sendhil Mullainathan (2005). 'Do Cigarette Taxes Make Smokers Happier.' *Advances in Economic Analysis & Policy*, 5 (1), Article 4. Available at http://www.bepress.com/bejeap/advances/vol5/iss1/art4.

Hayes, Natalie and Stephen Joseph (2003). 'Big 5 correlates of three measures of subjective well-being.' *Personality and Individual Differences*, 34 (4), 723-727.

Kahneman, Daniel, Alan B. Krueger, David Schkade, Norbert Schwarz and Arthur Stone (2004). 'Toward National Well-Being Accounts.' *American Economic Review, Papers and Proceedings*, 94 (2), 429-434.

Kahneman, Daniel and Alan B. Krueger (2006). 'Developments in the Measurement of Subjective Well-Being.' *Journal of Economic Perspectives*, 20 (1), 3-24.

King, Laura A. and Christie K. Napa (1998). 'What Makes a Life Good?' *Journal of Personality and Social Psycholog*y, 75 (1), 156-165.

Koivumaa-Honkanen, Heli, Risto Honkanen, Heimo Viinamäki, Kauko Heikkilä, Jaakko Kaprio and Markku Koskenvuo (2000). 'Self-reported Life Satisfaction and 20-Year Mortality in Healthy Finnish Adults.' *American Journal of Epidemiology*, 152 (10), 983-991.

Koivumaa-Honkanen, Heli, Risto Honkanen, Heimo Viinamäki, Kauko Heikkilä, Jaakko Kaprio and Markku Koskenvuo (2001). 'Life Satisfaction and Suicide: A 20-Year Follow-Up Study.' *American Journal of Psychiatry*, 158 (3), 433-439.

Krueger, Alan B. and David A. Schkade (2007). *The Reliability of Subjective Well-Being Measures*. NBER Working Paper No. 13027. Cambridge MA: National Bureau of Economic Research.

Lane, Richard D., Eric M. Reiman, Margaret M. Bradley, Peter J. Lang, Geoffrey L. Ahern, Richard J. Davidson and Gary E. Schwartz (1997). 'Neuroanatomical correlates of pleasant and unpleasant emotion.' *Neuropsychologia*, 35 (11), 1437-1444.

Larsen, Randy J. and Barbara L. Fredrickson (1999). 'Measurement Issues in Emotion Research.' In: Daniel Kahneman, Ed Diener and Norbert Schwarz (eds.) *Well-Being: The Foundations of Hedonic Psychology*. New York: Russell Sage Foundation, 40-60.

Layard, Richard (2005). *Happiness: Lessons from a new science*. London: Allen Lane / Penguin Books.

Levesque, J., Y. Joanette, B. Mensour, G. Beaudoin, J.-M. Leroux, P. Bourgouin and M. Beauregard (2003). 'Neural correlates of sad feelings in healthy girls.' *Neuroscience*, 121 (3), 545-551.

Lyubomirsky, Sonja, Laura King and Ed Diener (2005). 'The Benefits of Frequent Positive Affect: Does Happiness Lead to Success?' *Psychological Bulletin*, 131 (6), 803-855.

McBride, Michael (2001). 'Relative-income effects on subjective well-being in the cross-section.' *Journal of Economic Behavior & Organization*, 45 (3), 251-278.

Nettle, Daniel (2005). *Happiness: The science behind your smile*. Oxford / New York: Oxford University Press.

Oswald, Andrew J. and Natavudh Powdthavee (2007). *Death, Happiness, and the Calculation of Compensatory Damages*. IZA Discussion Paper No. 3159. Bonn: Institute for the Study of Labor.

Pigou, Arthur C. (1952). *The economics of welfare*. Fourth edition (1920), reprinted with eight new appendices. London: MacMillan.

van Praag, Bernard M.S. and Barbara Baarsma (2005). 'Using happiness surveys to value intangibles: The case of airport noise.' *Economic Journal*, 115 (January), 224-246.

van Praag, Bernard M.S. and Nico L. van der Sar (1988). 'Empirical Uses of Subjective Measures Of Well-Being.' *Journal of Human Resources*, 23 (2), 193-210.

Radcliff, Benjamin (2001). 'Politics, Markets and Life Satisfaction: The Political Economy of Human Happiness.' *American Political Science Review*, 95 (4), 939-952.

Rehdanz, Katrin and David Maddison (2005). 'Climate and Happiness.' *Ecological Economics*, 52 (1), 111-125.

Ryff, Carol D., Burton H. Singer and Gayle Dienberg Love (2004). 'Positive health: connecting well-being with biology.' *Philosophical Transactions Royal Society London B*, 359, 1383-1394.

Schwarz, Norbert and Fritz Strack (1999). 'Reports of Subjective Well-Being: Judgmental Processes and Their Methodological Implications.' In: Daniel Kahneman, Ed Diener and Norbert Schwarz (eds.) *Well-Being: The Foundations of Hedonic Psychology*. New York: Russell Sage Foundation, 61-84.

Scollon, Christie Napa, Chu Kim-Prieto and Ed Diener (2003). 'Experience sampling: Promises and pitfalls, strengths and weaknesses.' *Journal of Happiness Studies*, 4 (1), 5-34.

Sirgy, M. Joseph, Alex C. Michalos, Abbott L. Ferriss, Richard A. Easterlin, Donald Patrick and William Pavot (2006). 'The Quality-of-Life (QOL) research movement: Past, present, and future.' *Social Indicators Research*, 76 (3), 343-466.

Steptoe, Andrew, Jane Wardle and Michael Marmot (2005). 'Positive affect and health-related neuroendocrine, cardiovascular, and inflammatory processes.' *Proceedings of the National Academy of Sciences*, 102 (18), 6508-6512.

Stone, Arthur A., Saul S. Shiffman and Marten W. Devries (1999). 'Ecological Momentary Assessment.' In: Daniel Kahneman, Ed Diener and Norbert Schwarz (eds.) *Well-Being: The Foundations of Hedonic Psychology*. New York: Russell Sage Foundation, 26-39.

Sumner, L. Wayne (1995). 'The Subjectivity of Welfare.' *Ethics*, 105 (4), 764-790.

Tellegen, Auke, David T. Lykken, Thomas J. Bouchard Jr., Kimerly J. Wilcox, Nancy L. Segal and Stephen Rich (1988). 'Personality Similarity in Twins Reared Apart and Together.' *Journal of Personality and Social Psychology*, 54 (6), 1031-1039.

Veenhoven, Ruut (1984). *Conditions of happiness*. Dordrecht: D. Reidel Publishing Company.

Veenhoven, Ruut (2000). 'Freedom and Happiness: A Comparative Study in Forty-Four Nations in the Early 1990s.' In: Ed Diener and Eunkook M. Suh (eds.) *Culture and Subjective Well-Being*. Cambridge MA / London: The MIT Press, 257-288.

Vittersø, Joar, Robert Biswas-Diener and Ed Diener (2005). 'The divergent meanings of life satisfaction: Item response modeling of the satisfaction with life scale in Greenland and Norway.' *Social Indicators Research*, 74 (2), 327-348.

Watson, David, Lee Anna Clark and Auke Tellegen (1988). 'Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scales.' *Journal of Personality and Social Psychology*, 54 (6), 1063-1070.

Welsch, Heinz (2006). 'Environment and happiness: Valuation of air pollution using life satisfaction data.' *Ecological Economics*, 58 (4), 801-813.

Wierzbicka, Anna (2004). ''Happiness' in cross-linguistic & cross-cultural perspective.' *Daedalus*, Spring, 34-43.

Wilson, Warner (1967). 'Correlates of avowed happiness.' *Psychological Bulletin*, 67 (4), 294-306.

Winkelmann, Liliana and Rainer Winkelmann (1998). 'Why are the Unemployed so Unhappy? Evidence from Panel Data.' *Economica*, 65 (257), 1-15.

Chapter 16
Measures of Gross National Happiness

Ruut Veenhoven[i]
Emeritus-Professor, Erasmus University Rotterdam, Netherlands

Abstract

Happiness is rising on the political agenda and this calls for measures of how well nations perform in creating great happiness for a great number, analogous to measures of success in creating wealth, such as GDP.

Happiness is defined as subjective enjoyment of one's life as-a-whole and this can be measured using self-reports. Question on happiness are currently used in large scale surveys of the general population in nations. As a result we have now comparable data on happiness in 95 contemporary nations and time-series of 25 years and longer on 11 developed nations.

These data can be aggregated in different ways: If the aim is simply greater happiness for a greater number of citizens, Average happiness (AH) is an appropriate measure. If the focus is on enduring happiness, it is better to combine average happiness with longevity in an index of Happy Life Years (HLY). If the aim is to reduce disparity among citizens a relevant indicator is the Inequality of Happiness (IH) in the nations as measured with the standard deviation. Average and dispersion can also be combined in an index of Inequality-Adjusted Happiness (IAH).

Comparison across nations shows sizeable differences on all these measures of gross national happiness and these differences correspond with societal characteristics that can be influenced by policy makers, such as freedom and justice. Comparison over time shows major improvement during the last decade.

Key words: *rule-utilitarianism, policy indicators, happiness, happy life years, inequality of happiness, cross-national, trend/*

Introduction

Attempts to improve the human lot begin typically with treating compelling miseries, such as hunger and epidemics. When these problems are solved, attention shifts to

i E-mail: veenhoven@fsw.eur.nl

broader and more positive goals. For that reason there is now rising interest in 'happiness'. This brings a classic philosophy back on the scene.

Greatest Happiness Principle

Two centuries ago Jeremy Bentham (1789) proposed a new moral principle. He wrote that the goodness of an action should not be judged by the decency of its intentions, but by its consequences on human happiness. He argued that we should aim at the 'greatest happiness for the greatest number'. Bentham defined happiness in terms of subjective feeling, as 'the sum of pleasures and pains'. This philosophy is known as 'Utilitarianism', because of its emphasis on the utility of behavioural consequences. 'Happyism' would have been a better name, since this utility is seen as contribution to happiness.

When applied at the level of individual choice, this theory runs into some difficulties. One problem is that it is often difficult to foresee what the balance of effects on happiness will be. Another problem is that the theory deems well-intended behaviour to be a-moral if it happens to pan out adversely. Imagine the case of a loving mother who saves the life of her sick child, a child that grows up to be a criminal; mothers can seldom foresee a child's future and can hardly be reproached for their unconditional motherly love.

The theory is better suited for judging general rules, such as the rule that mothers should care for their sick children. It is fairly evident that adherence to this rule will add to the happiness of a great number. Following such rules is then morally correct, even if consequences might be negative in a particular case. This variant is known as 'Rule-Utilitarianism'.

Application in Public Policy

Rule-Utilitarianism can be used as a moral guide for legislation and has played a role in discussions about property laws and the death penalty. The principle can also be applied to wider issues in public policy; such as the question of what degree of income-inequality we should accept. This variant of the greatest happiness principle has not been very prominent in public policy as yet, at least not explicit. Currently the principle is gaining ground, especially in the UK (Donovan *et. al.* 2003; Layard 2005).

Happiness can be furthered at two levels; at the level of individuals and at the level of collectivities. At the individual level, happiness can be enhanced by information, training and guidance of individual citizens. This approach is particularly useful in modern nations, where the collective conditions are typically so good that most of the variance in happiness is due to individual differences. At the level of collectivities, happiness can be increased by improving the livability of institutional settings such as schools, work organisations and old age homes. Greater happiness can also be created by improving the livability of society at large, such as by providing a decent standard of living and a climate of trust. Elsewhere I have discussed these options in more detail (Veenhoven 2007)

Need for Measurement

Attempts to create greater happiness for a greater number require measurement of happiness. Measurement is necessary for assessing whether happiness can be realistically raised, for selecting means for furthering happiness and finally for assessing effectiveness.

The measures of happiness used differ somewhat with the level of intervention. At the individual level we need sensitive measures of happiness of persons, but at the collective level we rather need global aggregates.

Plan of this Paper

In this paper I focus on measures of happiness at the societal level, that is, measures of happiness in nations. These measures inform policy makers about the happiness of the great number. The measures of happiness serve to assess how well the country is doing in creating greater happiness and allow comparison across nations and trough time. As such they are comparable to economic indicators, such as GNP per head that indicate how well the country performs in creating wealth. By analogy I call this measures of 'Gross National' happiness.[1]

As a first step I distinguish happiness from other qualities of life and define the concept (Section "What is 'Happiness'"). Next I consider how happiness can be measured and conclude that this can be done using direct questions in survey studies (Section "Can Hapiness be Measured"). I then present four applications of survey data on happiness in nations (Section "Measures of Happiness in Nations"). Next I review the findings yielded with these measures, first findings on the degree of happiness in nations (Section "Degree of Happiness in Nations"), second findings on trends in happiness over time (Section "Trend of Happiness in Nations") and third findings on the relation with societal characteristics (Section "Determinants of Happiness in Nations").

What is "Happiness"?

The word 'happiness' is used in various ways. In the widest sense it is an umbrella term for all that is good. In this meaning it is often used interchangeably with terms like 'well-being' or 'quality of life'. Below I will delineate four qualities of life and show that the concept of happiness fits only one of these.

Four Qualities of Life

Table 16.1 Four Qualities of Life

	Outer qualities	Inner qualities
Life chances	Livability of environment	Life-ability of the person
Life results	Utility of life	Satisfaction with life

Quality-of-life concepts can be sorted using two distinctions, which together provide a fourfold matrix. That classification is discussed in more detail in Veenhoven (2000a). The first distinction is between chances and outcomes, that is, the difference between opportunities for a good life and the good life itself. A second difference is between outer and inner qualities of life, in other words between 'external' and 'internal' features. In the first case the quality is in the environment, in the latter it is in the individual. Lane (1994) made this distinction clear by distinguishing 'quality of society' from 'quality of persons'. The combination of these two dichotomies yields a fourfold matrix. This classification is presented in Table 16.1.

Livability of the Environment

The left top quadrant denotes the meaning of good living conditions, shortly called ‘livability’.

Ecologists see livability in the natural environment and describe it in terms of pollution, global warming and degradation of nature. Currently, they associate livability typically with preservation of the environment. City planners see livability in the built environment and associate it with such things as sewer systems, traffic jams and ghetto formation. Here the good life is seen as a fruit of human intervention.

In the sociological view, society is central. Livability is associated with the quality of society as a whole and also with the position one has in society.

Livability is not what is called happiness here. It is rather a precondition for happiness and not all environmental conditions are equally conducive to happiness.

Life-ability of the Person

The right top quadrant denotes inner life-chances. That is: how well we are equipped to cope with the problems of life. Sen (1992) calls this quality of life variant ‘capability’. I prefer the simple term ‘life-ability’, which contrasts elegantly with ‘livability’.

The most common depiction of this quality of life is absence of functional defects. This is 'health' in the limited sense, sometimes referred to as ‘negative health’. Next to absence of disease, one can consider excellence of function. This is referred to as ‘positive health’ and associated with energy and resilience. A further step is to evaluate capability in a developmental perspective and to include acquisition of new skills for living. This is commonly denoted by the term ‘self-actualisation’. From this point of view a middle-aged man is not ‘well’ if he behaves like an adolescent, even if he functions without problems at this level. Since abilities do not develop along side idleness, this quality of life is close to the ‘activity’ in Aristotle’s concept of eudemonia. Lastly, the term ‘art of living’ denotes special life-abilities; in most contexts this quality is distinguished from mental health and sometimes even attributed to slightly disturbed persons. Art of living is associated with refined tastes, an ability to enjoy life and an original style of life.

Ability to deal with the problems of life will mostly contribute to happiness as defined here, but is not identical. If one is competent in living one has a good chance at happiness, but this endowment does not guarantee an enjoyable outcome.

Utility of Life

The left bottom quadrant represents the notion that a good life must be good for something more than itself. This assumes some higher values. There is no current generic for these external outcomes of life. Gerson (1976, p.795) refers to these effects as ‘transcendental’ conceptions of quality of life. Another appellation is ‘meaning of life’, which then denotes ‘true’ significance instead of mere subjective sense of meaning. I prefer the simpler ‘utility of life’, while admitting that this label may also give rise to misunderstanding.

When evaluating the external effects of a life, one can consider its functionality for the environment. In this context, doctors stress how essential a patient's life is to its intimates. At a higher level, quality of life is seen in contributions to society. Historians

see quality in the addition an individual can make to human culture, and rate for example the lives of great inventors higher than those of anonymous peasants. Moralists see quality in the preservation of the moral order, and would deem the life of a saint to be better than that of a sinner. As an individual's life can have many environmental effects, the number of such utilities is almost infinite.

Apart from its functional utility, life is also judged on its moral or esthetic value. For instance, most of us would attribute more quality to the life of Florence Nightingale than to that of a drunk, even if it appeared in the end that her good works had a negative result in the end. In moral philosophy this is called 'virtuous living', and is often presented as the essence of 'true' happiness.

Core meaning: Subjective Enjoyment of Life

Finally, the bottom right quadrant represents the inner outcomes of life. That is the quality in the eye of the beholder. As we deal with conscious humans, this quality boils down to subjective enjoyment of life. This is commonly referred to by terms such as 'subjective well-being', 'life-satisfaction' and 'happiness' in a limited sense of the word. This is the kind of happiness the utilitarian philosophers had in mind and it is also the kind of happiness addressed here.

Four Kinds of Satisfaction

Even when we focus on subjective satisfaction with life, there are still different meanings associated with the word happiness. These meanings can also be charted in a fourfold matrix. In this case, that classification is based on the following dichotomies: Life-aspects versus life-as-a-whole and passing delight versus enduring satisfaction.

Appraisals of life can concern aspects, such as marriage or work-life, and one's life-as-a-whole. The word 'happiness' is used in both contexts. Obviously, such appraisals are linked. Enjoyment of aspects of life will typically contribute to the satisfaction with life as a whole (so-called bottom-up effect), and enjoyment of one's life-as-a-whole appears to foster the satisfaction with life-aspects (top-down). Still, these are not identical matters. One can have a happy marriage but still be dissatisfied with life-as-a-whole, or be satisfied with life-as-a-whole in spite of an unhappy marriage.

Next, the experience of enjoyment can be short-lived or enduring. Again, the word happiness is used for both phenomena. Sometimes it refers to passing moods and on other occasions to stable satisfaction. Once more, these matters are related but not the same.

When combined, these distinctions produce the fourfold matrix presented in Table 16.2. The distinction between part and whole is presented vertically, and the distinction between passing and enduring enjoyment horizontally.

Table 16.2 Four Kinds of Satisfaction

	Passing	Enduring
Part of life	Pleasure	Domain satisfaction
Life as a whole	Top experience	Life satisfaction (Happiness)

Instant Satisfaction

The top-left quadrant represents passing enjoyments of life-aspects. Examples would be delight in a cup of tea at breakfast, the satisfaction of a chore done or the enjoyment of a piece of art. I refer to this category as 'instant-satisfactions'. Kahneman (2000, p.4) calls it 'instant-utilities'. This quadrant represents hedonistic happiness, especially when the focus is on sensory experience. The concept of happiness used here is broader however. It concerns both overall satisfaction and life-as-a-whole. Though fleeting enjoyment obviously contributes to a positive appreciation of life it is not the whole of it.

Domain Satisfaction

The top right quadrant denotes enduring appreciation of life-aspects, such as marriage satisfaction and job-satisfaction. This is currently referred to as domain-satisfactions. Though domain-satisfactions depend typically on a continuous flow of instant-satisfactions, they have some continuity of their own. For instance, one can remain satisfied with one's marriage even if one has not enjoyed the company of the spouse for quite some time. Domain-satisfactions are often denoted with the term happiness: a happy marriage, happy with one's job, etc. Yet in this chapter the term happiness is used in the broader sense of satisfaction with life-as-a-whole. One would not call a person happy who is satisfied with marriage and job, but still dissatisfied on the whole because his health is failing. It is even possible that someone is satisfied with all the domains one can think of, but nevertheless feels depressed.

Top-Experience

The bottom right quadrant denotes the combination of passing experience and appraisal of life-as-a-whole. That combination occurs typically in top-experiences, which involve short-lived but quite intense feelings and the perception of wholeness. This is the kind of happiness poet's write about. Again this is not the kind of happiness aimed at here. A moment of bliss is not enduring appreciation of life. In fact such top-experiences even seem detrimental to lasting satisfaction, possibly because of their disorientating effects (Diener et. al. 1989).

Core Meaning: Lasting Satisfaction with One's Life-as-a-Whole

Lastly, the bottom-right quadrant represents the combination of enduring satisfaction with life-as-a-whole. This is what I mean with the word happiness. A synonym is 'life-satisfaction'. This is the meaning the utilitarian philosophers had in mind when talking about happiness. When speaking about the 'sum' of pleasures and pains they denoted a balance over time and thus a durable matter.

Definition of Happiness

Overall happiness is the *degree to which an individual judges the overall quality of his/her own life-as-a-whole favorably*. In other words: how much one likes the life one leads. The key terms in this definition may be elucidated as follows.

Degree: The word 'happiness' is not used to denote an optimal appreciation of life. It refers to a degree, like the concepts of 'length' or 'weight' it denotes more or less of

something. When we say a person is happy, we mean that he or she judges his of her life favorably rather than unfavorably.

Individual: The term happiness is used to describe the state of an individual person only; it does not apply to collectivities. Thus, a nation cannot be said to be happy. At best, most of its citizens consider themselves happy.

Judges: The word 'happiness' is used where somebody has made an overall judgment about the quality of his or her life. This implies an intellectual activity. Making an overall judgment implies assessing past experiences and estimating future experiences and estimating average quality of life.

One consequence of this conceptualisation is that the word 'happiness' cannot be used for those who did not make up their mind. One cannot say whether a person is happy or not, if that person is intellectually unable to construct an overall judgment...

Overall: The word 'happiness' refers to a judgment, which integrates all the appreciation criteria used. Thus, the idea that one has all one has ever desired does not necessarily make a person happy. Despite all material endowments such a person may feel pain or be depressed. Similarly, the appraisal that one's life is 'exciting' does not necessarily mark oneself as happy either; life may be too exciting to be enjoyable. A Chinese curse says: "May you have interesting times".

Life-as-a-whole: We do not use the word 'happiness' to characterise satisfaction with specific aspects of life, such as marriage or work. 'Happiness' refers to satisfaction with life in total.

Own life: The term 'happiness' concerns the evaluation of one's own life, not of life in general. A pessimistic 'Weltanschauung' does not necessarily characterise someone as 'unhappy'.

Favorably: Evaluation always embodies appreciation; a conclusion as to whether one likes something or: not. The term 'happiness' refers only to judgments concerning this aspect. Happiness judgments concern the dimension extending from appreciation to depreciation, from like to dislike. All humans are capable of appraisals of this kind, though not all humans can generalise all appraisals into a judgment of life-as-a-whole.

Components of Happiness

Humans are capable of evaluating their life in different ways. We have in common with all higher animals that we can appraise our situation affectively. We feel good or bad about particular things and our mood level signals overall adaptation. As in animals these affective appraisals are automatic, but unlike other animals it is known that humans can reflect on this experience. We have an idea of how we have felt over the last year, while a cat does not. Humans can also judge life cognitively by comparing life as it is with notions of how it should be.

Most human evaluations are based on both sources of information, that is: intuitive affective appraisal and cognitively guided evaluation. The mix depends mainly on the object. Tangible things such as our income are typically evaluated by comparison; intangible matters such as sexual atractiveness are evaluated by how it feels.

In evaluating our life as a whole we can therefore use two sources of information: our affects and our thoughts. We can decide that we feel fine most of the time and we can also judge that life seems to meet our (conscious) demands. These appraisals do not

necessarily coincide. We may feel fine generally, but nevertheless be aware that we failed to realise our aspirations. Or we may have surpassed our aspirations, but nevertheless feel miserable.

Using the word 'happiness' in both these cases would result in three different kinds of happiness, the overall judgment as described above and these two specific appraisals. Therefore the components are referred to as 'hedonic level of affect' and 'contentment'. To mark the difference with the encompassing judgment I will refer to happiness (the core concept) as *overall happiness*.

Hedonic Level of Affect

Hedonic level of affect is the degree to which various affects that someone experiences are pleasant in character and this reflects typically in 'mood'.

A person's average hedonic level of affect can be assessed over different periods of time: an hour, a week, a year, as well as over a lifetime. The focus here is on 'characteristic' hedonic level. That is so to say: the average over a long time-span such as a month or a year. The concept does not presume subjective awareness of that average level.

Contentment

Contentment is the degree to which an individual perceives his/her aspirations are met. The concept presupposes that the individual has developed some conscious wants and has formed an idea about their realisation. The factual correctness of this idea is not at stake. The concept concerns the individual's subjective perception.

Preponderance of Affect

It is commonly assumed that the cognitive component dominates the overall judgment of life, *e.g.* by Easterlin (1984) and Layard (2005). Yet there are strong indications that overall happiness draws on hedonic level of affect in the first place (Veenhoven 2006e).

Synonyms

The above defined concept of 'overall happiness' is denoted with different words. In the 1950s the words *adjustment* and *morale* were sometimes used in this meaning and since the 1960s the term *life-satisfaction* came into use for this purpose. In 1984 Ed Diener introduced the term *subjective well-being*, abbreviated as SWB, and this term is still dominant in psychology.

The term life-satisfaction is mostly used for 'overall happiness', but refers in some cases particularly to its cognitive component and is than synonymous with 'contentment'. In such context, the term happiness is typically used for the affective appraisal of life and then synonymous with 'hedonic level of affect'.

The term subjective well-being is also used in wider meanings than happiness as defined here. Sometimes the term refers to good mental functioning and then denotes the meaning of life-ability in the top right quadrant of Table 16.1. At other occasions the term is used as a generic for all subjective enjoyment and then covers all the quadrants of Table 16.2.

Can Happiness be Measured?

Measurement has long been understood as 'objective' and 'external' assessment, analogous to the measurement of blood-pressure by a doctor. By now we know that happiness cannot be measured that way. Steady physiological correlates have not been discovered, and probably never will be. Nor have any overt behaviours been found to be consistently linked to inner enjoyment of life. Like most attitudinal phenomena, happiness is only partially reflected in behaviour. Suicidal behaviour is an example, though most suicidal persons are unhappy, only a minority of the unhappy behaves suicidal.

By definition, happiness is something we have on our mind and consequently we can measure happiness using questions. That is, simply asking people how much they enjoy their life-as-a-whole. Questions on happiness can be posed in various contexts; clinical interviews, life-review questionnaires and survey interviews. The questions can be posed in different ways; directly or indirectly, and by means of single or multiple items.

Common Questions

Some common survey questions are presented below in Table 16.3.

Validity Doubts

Critics have suggested that responses to questions on happiness actually measure other phenomena. Rather than indicating how much the respondent enjoys life, answers would reflect his normative notions and desires.

No Notion

One of the misgivings is that most people have no opinion at all about their happiness. They would be more aware of how happy they are supposed to be, and report that instead. Though this may happen incidentally, it does not appear to be the rule. Most people know quite well whether or not they enjoy life. Eight out of ten Americans think of it every week. Responses on questions about happiness tend to be prompt. Non-response on these items is low; both absolutely (± 1%) and relatively to other attitudinal questions. `Don't know' responses are infrequent as well.

A related assertion is that respondents mix up how happy they actually are, with how happy other people think they are, given their situation. If so, people considered to be well off would typically report to be very happy, and people regarded as disadvantaged should characterise themselves as unhappy. That pattern is observed sometimes, but it is not general. For instance, in The Netherlands good education is seen as a pre-requisite for a good life, but the highly educated appear slightly less happy in comparison to their less educated counterparts.

Coloured Answers

Another objection concerns the presence of systematic bias in responses. It is assumed that questions on happiness are interpreted correctly, but that responses are often false. People who are actually dissatisfied with their life would tend to answer that they are quite happy. Both ego-defense and social-desirability would cause such distortions.

Table 16.3 Some Currently Used Questions about Happiness

Single questions • Taking all together, how happy would you say you are: very happy, quite happy, not very happy, not at all happy? (standard item in the World Value Studies) • How satisfied are you with the life you lead? Very satisfied, fairly satisfied, not very satisfied, not at all satisfied? (standard item in Euro-barometer surveys) • Here is a picture of a ladder. Suppose the top of the ladder represents the best possible life for you and the bottom of the ladder the worst possible life. Where on the ladder do you feel you personally stand at the present time? (0-10 ladder like rating scale) (Cantril's (1965) present life ladder rating)
Multiple questions (summed) • Same question asked twice: at the beginning and at the end of interview How do you feel about your life-as-a-whole? Delighted, pleased, mostly satisfying, mixed, mostly dissatisfying, unhappy, terrible? (Andrews & Withey's Life 3) • Five questions, rated on a 1-7 scale ranging from strongly agree to strongly disagree. (Diener's 1985 Satisfaction With Life Scale SWLS[2]) – In most ways my life is close to ideal. – The conditions of my life are excellent. – I am satisfied with my life. – So far I have gotten the important things I want in life. – If I could live my life over, I would change almost nothing

This bias is seen to manifest itself in over-report of happiness; most people claim to be happy, and most perceive themselves as happier than average. Another indication of bias is seen in the finding that psycho-somatic complaints are not uncommon among the happy. However, these findings allow other interpretations as well. Firstly, the fact that more people say to be happy than unhappy does not imply over-report of happiness. It is quite possible that most people are truly happy (some reasons will be discussed below). Secondly, there are also good reasons why most people think that they are happier than average. One such reason is that most people are like critical scientists and think that unhappiness is the rule. Thirdly, the occurrence of head-aches and worries among the happy does not prove response distortion. Life can be a sore trial some times, but still be satisfying on a balance.

The proof of the pudding is in demonstrating the response distortion itself. Some clinical studies have tried to do so by comparing responses to single direct questions with ratings based on depth interviews and projective tests. The results are generally not different from responses to single direct questions posed by an anonymous interviewer.

Reliability Doubts

Though single questions on happiness seem to measure what they are supposed to measure, they measure it rather imprecisely.

When the same question is asked twice in an interview, responses are not always identical. Correlations are about +.70. Over a period of a week, test-retest reliability drops to circa +.60. Though responses seldom change from `happy' to `unhappy', switches from `very' to `fairly' are rather common. The difference between response-options is often ambiguous. The respondent's notion about his/her happiness tends to be global. Thus the choice for one answer-category or the next is sometimes haphazard.

Because choice is often arbitrary, subtle differences in interrogation can exert considerable effect. Variations in place where the interview is held, characteristics of the interviewer, sequence of questions and precise wording of the key-item can tip the scale to one response or the other. Such effects can occur in different phases of the response process; in the consideration of the answer as well as in the communication of it.

Bias in Appraisal

Though most people have an idea of how much they enjoy life, responding to questions on this matter involves more than just bringing up an earlier judgment from memory. For the most part, memory only indicates a range of happiness. Typically, the matter is re-assessed in an instant judgment. This re-appraisal may be limited to recent change (are there any reasons to be more or less happy than I used to be?), but it can also involve quick re-evaluation of life (what are my blessings and frustrations?). In making such instant judgments, people use various heuristics. These mental simplifications are attended with specific errors. For instance the ‘availability’ heuristic involves orientation on pieces of information that happen to be readily available. If the interviewer is in a wheelchair, the benefit of good health is salient. Respondents in good health will then rate their happiness somewhat higher and the correlation of happiness-ratings with health variables will be more pronounced. Several of these heuristic effects have been demonstrated by Schwarz and Strack (1991).

Bias in Response

Once a respondent has formed a private judgment, the next step is to communicate it. At this stage reports can be biased in various ways as well. One source of bias is inherent to semantics; respondents interpret words differently and some interpretations may be emphasised by earlier questions. For example, questions on happiness are more likely to be interpreted as referring to `contentment' when preceded by questions on success in work, rather than items on mood. Another source of response-bias is found in considerations of self-presentation and social-desirability. Self-rating of happiness tends to be slightly higher in personal interviews than on anonymous questionnaires. However, direct contact with an interviewer does not always inflate happiness reports. If the interviewer is in a wheel-chair, modest self-presentation is encouraged.

Much of these biases are random, and balance out in large samples. So in large samples, random error does not affect the accuracy of happiness averages. Yet it does affect correlations, random error 'attenuates' correlations. Random error can be estimated by means of multiple-trait-multiple-method (MTMM) studies, and correlations can be corrected (disattenuated) on that basis. A first application on satisfaction measures is reported by Saris *et. al.* (1996).

Some biases may be systematic; especially bias produced by technique of interrogation and sequence of questions. Bias of that kind does affect the reliability of distributional data. In principle it does not affect correlations, unless the measure of the correlate is biased in the same way (correlated error). To some extend, systematic error can also be estimated and corrected. See also Saris *et al*.

Comparability across Nations

Average happiness differs markedly across nations. In Table 16.3 we will see that Russians score currently 4.1 on a 0-10 scale, while in Sweden the average is 7.9. Does that mean that Russians really take less pleasure in life? Several claims to the contrary have been advanced. Elsewhere I have checked these doubts (Ouweneel & Veenhoven 1991; Veenhoven 1993b). The results of that inquiry are summarised below.

The first objection is that differences in *language* hinder comparison. Words like ‘happiness’ and ‘satisfaction’ would not have the same connotations in different tongues. Questions using such terms would therefore measure slightly different matters. I checked that hypothesis by comparing the rank orders produced by three kinds of questions on life-satisfaction: a question about ‘happiness’, a question about ‘satisfaction with life’ and a question that invites to a rating between ‘best- and worst possible life’. The rank orders appeared to be almost identical. I also compared responses on questions on happiness and satisfaction in two bi-lingual countries, and found no evidence for linguistic bias either.

A second objection is that responses are differentially distorted by *desirability-bias*. In countries where happiness ranks high in value, people would be more inclined to overstate their enjoyment of life. I inspected that claim by checking whether reported happiness is indeed higher in countries where hedonic values are most endorsed. This appeared not to be the case. As a second check, I inspected whether reports of general happiness deviate more from feelings in the past few weeks in these countries; the former measure being more vulnerable for desirability distortion than the latter. This appeared not to be the case either.

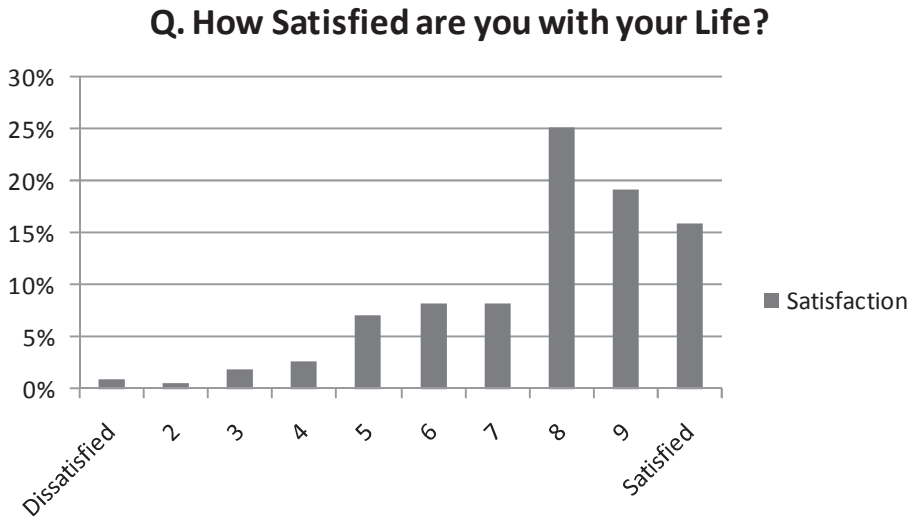
A third claim is that *response-styles* distort the answers dissimilarly in different countries. For instance, collectivistic orientation would discourage ‘very’ happy responses, because modest self-presentation is more appropriate within that cultural context. I tested this hypothesis by comparing happiness in countries differing in value-collectivism, but found no effect in the predicted direction. The hypothesis failed several other tests as well.

A related claim is that happiness is a typical *western concept*. Unfamiliarity with it in non-western nations would lead to lower scores. If so, we can expect more `don't know' and `no answer' responses in non-western nations. However, that appeared not to be the case.

Measures of Happiness in Nations

The above discussed questions on happiness are commonly used in surveys of the general population in nations. A typical pattern of responses is shown in Figure 16.1. Several measures of happiness can be derived from such data.

Figure 16.1 Happiness in the United States 2000



Source: World Value Survey

Average Happiness

Policy makers are typically interested in the *happiness of the greatest number* and this is reflected in the central tendency in responses. This central tendency can be expressed in several summary statistics, such as the percentage above neutral, the modus and the mean. The most commonly used summary statistic is the mean or average. Average happiness is often expressed on range 0 to 10, which invites to an interpretation as a 'school mark'.

Happy Life Years

Policy makers aim typically at *enduring happiness* and this fits the Bentham's recommendation to calculate happiness not only by its intensity, but also by its duration. One of the ways to do that is combining information about length of life, drawn from civil registrations of births and deaths[3], with data on average appreciation of life as assessed in surveys.

Computation

The following simple formula can be applied:

$$\textit{Happy-Life-Years} = \textit{Life-expectancy at birth} \times \textit{0-1 happiness}$$

Suppose that life expectancy in a country is 60 years, and that the average score on a 0 to 10-step happiness scale is 5. Converted to a 0-1 scale, the happiness score is than 0,5. The product of 60 and 0,5 is 30. So the number of happy life years is 30 in that country. If life expectancy is also 60 years but average happiness 8, the number of happy life years is 48 (60 x 0,8).

Advantages

This combination of happiness and longevity has the additional advantage that it provides a more complete view on how well people thrive in a country. In simple animals, good adaptation reflects only in survival, in higher animals, good adaptation also reflects in hedonic experience. Negative affect is indicative of poor adaptation and tends to inhibit the organism, while positive affect is indicative of good adaptation and works as a 'go' signal. So, an animal that does not feel good is probably not doing well. This inner experience is no great issue in biology, because we cannot assess how animals feel. Humans are capable of reflecting on their experiences, and can condense positive and negative affects into an overall appraisal of happiness. They are also capable of communicating that appraisal to investigators. Hence in the case of humans we can use the additional sign of good adaptation and assess how long *and* happy they live.

A related advantage of this combination is that it fits common sense: we prefer a long and happy life above a short but happy life or a long but unhappy life. A further advantage is that this measure links up with another more established goal in public policy and is comparable to the DALY[4] index.

For these reasons HLY scored highest in a scholarly review of indicators of quality of life in nations (Hagerty *et. al.* 2001).

Inequality of Happiness

High average happiness can go together with considerable inequality in happiness, a majority being very happy at the cost of an unhappy minority. Most policy makers will deem that undesirable, since social equality is still high on the political agenda. Hence there is also demand for information about disparity of happiness in nations, comparable to data about inequalities in income and health.

In this case the focus is not on the central tendency in the distribution of responses on questions about happiness as presented in Table 16.3, but on the disparity in the distribution. This can again be expressed in several summary statistics, such as the standard-deviation, the inter-quartile range and the Gini-coefficient. Comparative analysis showed that inequality in happiness can best be quantified using the standard deviation and the Gini-coefficient is not appropriate for this kind of data (Kalmijn & Veenhoven 2005).

Inequality Adjusted Happiness (IAH)

Policy makers will typically aim at both a high level of happiness and a low degree of inequality of happiness. For that reason they need an indicator that marries the utilitarian wish for greater happiness with the egalitarian call for fairness. One such indicator is the index of 'Inequality-Adjusted Happiness' that gives equal weight to both aims (Veenhoven & Kalmijn 2005).

Degree of Happiness in Nations

Survey data on happiness are gathered in the World Database of Happiness (Veenhoven 2006). At the moment this source contains comparable figures for 95 nations on all four these indicators, which are presented in so-called 'finding reports' (Veenhoven 2006a-d). Some illustrative data are presented in Table 16.4.

Table 16.4 Happiness in Nations around 2000

Nation	Average happiness mean on scale 0-10	Happy Life Years life expectancy multiplied by happiness	Inequality standard-deviation on scale 0-10	Inequality adjusted happiness 0-100 index
Switzerland	8.3	62.9	1.9	73
Sweden	7.9	58.9	2.0	69
USA	7.4	56.9	2.1	67
Argentina	7.0	49.6	2.5	60
Germany (W)	6.9	54.8	2.2	64
France	6.7	51.5	2.2	58
Philippines	6.3	43.7	2.7	54
Japan	6.1	49.6	2.1	55
Iran	5.9	41.5	2.7	51
Poland	5.8	42.8	2.8	50
India	4.6	42.8	2.8	48
Russia	4.1	35.7	2.7	35
Zimbabwe	3.3	12.5	3.1	23

Source: World Database of Happiness, Happiness in Nations

Average Happiness (AH)

Table 16.4 presents average happiness in nations on range 0 to 10. The highest score is observed in Switzerland (8.3) and the lowest in Zimbabwe (3.3), so the actual variation on this range is 5 points.

Average happiness tends is above neutral in most countries, which means that great happiness for a great number is apparently possible. Still, there are also nations where the average is below 5. One of these is Russia and this score is typical of most former soviet states, probably as a result of the social transformations in this era. Average happiness is also low in several African countries, but not in all. Possibly, average happiness is also below neutral in some of the countries not investigated, in particular in war stricken countries such as Iraq.

Happy Life Years (HLY)

Theoretically, this indicator has a broad variation, HLY is zero if nobody can live in the country, and will be endless if society is ideal and its inhabitants immortal. The practical range is currently 50 years, the highest number of Happy Life Years is observed in Switzerland (62.9) and the lowest in Zimbabwe (12.5).

The rank-order of HLY is quite similar to the rank order in average happiness, the rank-order correlation being +. 94. The correlation is not perfect however, in some countries people live long, but are not too happy (*e.g.* Japan), while in other countries life

is short but fairly satisfying (*e.g.* Nigeria). The rank-order correlation between life-expectancy and average happiness is +.55.

Inequality of Happiness (IH)

The lowest degree of inequality is observed in the Netherlands (SD = 1,5) and the highest in Zimbabwe (3.1). Again the rank-order is similar to the rank-order of average happiness, rs = -. 65

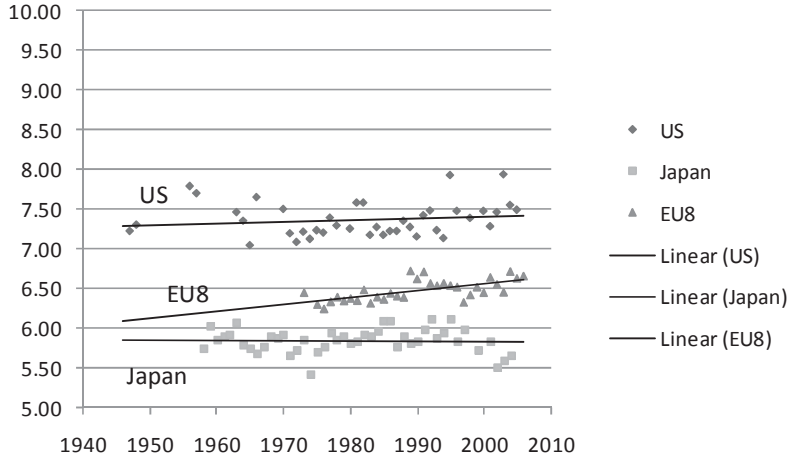
Inequality Adjusted Happiness (IAH)

The variation on this 0 to 100 index is between 73 (Switzerland) and 23 (Zimbabwe). The rank-order of nations is again similar with average happiness, r = +.95

Trend of Happiness in Nations

For only 11 of the 95 nations in the World Database of happiness do we have comparable data on 25 years or longer. These nations are: the United States since 1945, Japan since 1958 and for the first nine member states of the European Union since 1973. These data are available in the data file 'TrendsInNations' that is part of the World Database of happiness.

Figure 16.2 Trend average happiness in the EU8, Japan and United States



Gradual rise of Average Happiness in Most Nations

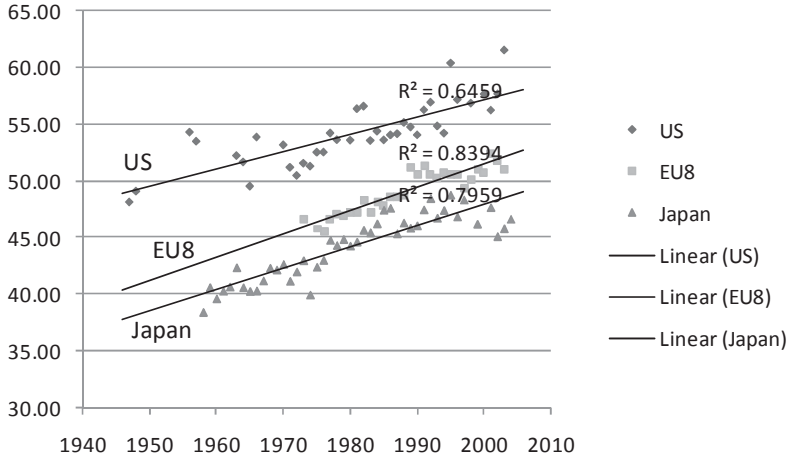
Trend data on average happiness are plotted in Figure 16.2, which shows that happiness has risen somewhat in the US and the EU[5], but stagnated in Japan.

The observed rise in the US is contrary to current opinion, which holds that Americans have become richer, but no happier. This so-called 'Easterlin paradox' pervades the economic discourse on happiness but does not fit the latest data (Veenhoven & Hagerty 2006). This paper presents also indications for rising happiness in several developing nations.

Spectacular Rise of Happy Life Years

Life expectancy has also gone up in the last decade, so it is no surprise that the number of Happy Life Years has also increased, as we can see in Figure 16.3. Still it is surprising to see that the gains are so big, 5.? years in the EU and 4.2 in the United States. More detail on this rise in apparent quality of life is presented elsewhere (Veenhoven 2005a). This growth of Happy Life Years is unprecedented in human history and marks considerable social progress!

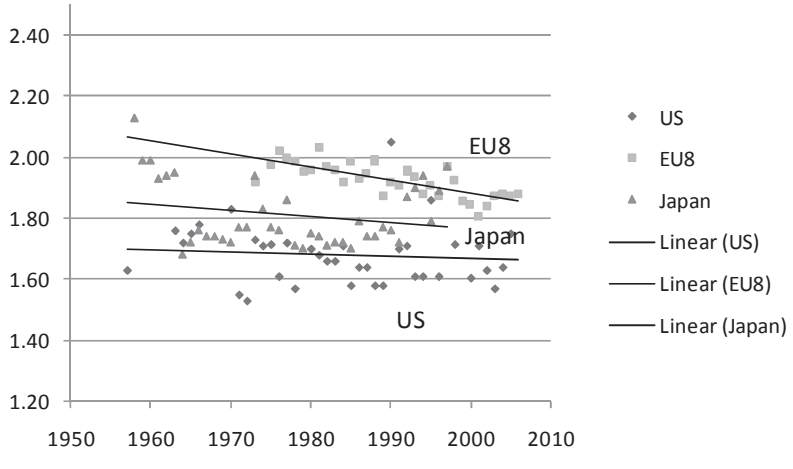
Figure 16.3 Trend HLY in 8 EU-nations 1973-2000



Lessening inequality of happiness in nations

There is much concern about 'new inequalities' emerging in modern nations. Yet data on dispersion of happiness shown a consistent pattern of diminishing inequality, as presented in Figure 16.4. This finding is discussed in more detail elsewhere (Veenhoven 2005b).

Figure 16.4 Trend in Dispersion of Life Satisfaction in the European Union



Rising Inequality Adjusted Happiness

Since average happiness has gone up and inequality of happiness has gone down, it is no surprise that the IAH index has gone up. In the United States this rise was 3.5 points on the 0-100 IAH scale and in the EU8 4.3 points

Determinants of Happiness in Nations

So far the data show that happiness is a realistic goal for public policy. Happiness of a *great number* is apparently possible, since most people rate their happiness above neutral in most nations. *Greater* happiness is also possible in most countries of the world. What is possible in countries like Switzerland should also be possible elsewhere.

Societal Correlates of Happiness

This begs the question what public policies are most conducive to happiness. This requires a view on the determinants of happiness in nations. One way to get determinants in view is to compare nations where people are more and less happy. This is an emerging line of research, the findings of which are also stored in the World Database of happiness, now in the section 'Correlational Findings' under subject category N4 'nation's condition'. An illustrative overview is presented in Table 16.5.

Material Affluence

Table 16.5 Happiness and Society; 90 Nations in the Late 1990s

Condition in nation	Correlation with			
	Average happiness	Inequality of happiness	Inequality Adjusted Happiness	Happy Life Years
Wealth				
Purchasing power per head	+.67	–.64	+.68	+.78
Security				
Lethal accidents	-.51	+.37	–.51	–.50
Social security	+.31	–.51	+.32	+.55
Freedom				
Economic freedom	+.59	–.48	+.61	+.64
Political freedom	+.46	–.34	+.43	+.59
Personal freedom	+.44	–.74	+.51	+.48
Inequality				
Disparity in incomes	+.06	–.33	+.02	–.17
Discrimination of women	–.45	+.38	-.48	–.76
Brotherhood				
Tolerance	+.50	–.33	+.50	+.49
Trust in people	+.37	–.50	+.54	+.39
Voluntary work	+.04	+.22	–.00	–.11
Justice				
Rule of law	+.53	–.57	+.56	+.68
Respect of civil rights	+.56	–.44	+.54	+.61
Corruption	–.60	+.65	–.63	–.74
Explained variance: Adjusted R^2	83%	71%	85%	87%

Data: World Database of Happiness, States of Nations. Indicators described in appendix.

Above in Table 16.5 we see that average happiness is typically higher in the economically most prosperous nations; the richer the country, the happier its inhabitants. The derived indicators of happiness produce similar relations, positive correlations with HLY and IAH and a negative correlation with Inequality of Happiness.

The relationships are curvi-linear; among poor nations the relationship is more pronounced than among affluent countries. When the $20 000 point is passed, the regression line is almost flat, which suggest that the law of diminishing returns applies. A similar pattern has been observed at the individual level: correlations between personal happiness and personal income are strong in poor countries and weak in rich nations. The effect of economic affluence can partly be explained by its effect on prevalence of absolute poverty. Yet, the data show this is not the whole story. Apparently, material welfare provides more gratifications than mere subsistence.

Security

Happiness is also higher in the nations that provide most safety. In Table 16.5, we see strong relationships with incidence of lethal accidents, which is indicative of wider physical safety. The pattern of correlations is again similar across all four indicators of happiness. This relationship appears to be largely independent of economic affluence.

The relationship with state provided social security is less pronounced, and disappears when economic affluence is controlled (Veenhoven 2000b).

Freedom

People are also happier in the nations that allow most autonomy. In Table 16.5 we see strong relationships with indicators of political freedom, which are largely independent of economic affluence. Correlations with indicators of personal freedom are less strong, but all positive. Again all four indicators of happiness behave similarly.

Freedom in society can affect the happiness of citizens in several ways: Political freedom is likely to provide protection against injustice and assault. Personal freedom can make that people choose life-styles that better fit personal needs and capacities. Both effects are likely to result in more rewarding events. Opportunity-to-choose adds to happiness only in publics with a well developed capability-to-choose (Veenhoven 2000c).

Equality

It seems rather evident that people live happier in the most egalitarian societies and that the differences in happiness will be smaller. Yet this appears not to apply to income inequality. Income inequality is essentially unrelated to the average happiness of citizens and only modestly related to dispersion of happiness among them. Possible explanations for this counterintuitive result have been discussed by Ott (2005) and Berg (2007).

The results are more in line with common sense in the case of gender inequality. Happiness is systematically lower in nations where there is discrimination against women, but the correlations are much abated when wealth of the nation is ‘controlled’ (ChinHonFoei 2007).

Social inequality can affect happiness negatively by the frustrations and limitations it involves. Possibly, some kinds of inequality also involve positive effects, which balance the negative ones. This may be the case with income-inequality.

Brotherhood

It is also commonly believed that people live happier in a climate of fraternity and that differences in happiness will be smaller in such conditions. This notion is supported in the correlations with 'tolerance' and with 'trust in people'. These correlations are quite sizeable and appear to be largely independent of the wealth of the nation. Surprisingly, there is no correlation with 'voluntary work'.

Justice

Lastly there are also good reasons to expect that social justice will improve the level of happiness in a nation and will reduce disparities. This expectation is corroborated in strong and consistent correlations with 'rule of law', 'respect for civil rights' and 'absence of corruption. 'Controlling' for wealth of a nation washes away much of this correlation, but we must realise that this does not necessarily mean that the actual effects are negligible.

Explained Variance

Together, these six societal qualities explain 83% of the differences in Average happiness, 71% of the differences in Inequality of happiness and no less than 87% of the differences in Happy Life Years. This suggests a political agenda.

Consistency across Measures of Happiness

The measures of happiness presented in this paper are tailored to somewhat different policy aims. *Average happiness* is an appropriate measure for policies aiming at greater happiness for a greater number, but gains on this measure could be short lived. So when the aim is at enduring happiness the measure of *Happy Life Years* is more apt. The measure of *Inequality of happiness* is designed for policies that aim at fairness in the first place and the index of *Inequality Adjusted Happiness* is tailored to dual purpose policies.

The correlations in Table 16.5 inform us about the compatibility of these policy aims. Divergence in correlation denotes incompatibility, *e.g.* if economic affluence would correlate more strongly with average happiness than with happy life years. Looking at the data from this perspective we can conclude that most aims are served by the same means. For instance, policies that aim at greater physical security seem to add both to average happiness and happy life years, while reducing inequality of happiness.

There is only one case of divergence in Table 16.5. Income inequality in nations is not associated with lover levels of happiness, but does go with greater inequality of happiness.

Conclusion

Happiness is defined as subjective enjoyment of life-as-a-whole. Since this is something people have in mind, it can be measured using direct questions that can be applied in large scale surveys of the general population.

Four measures of happiness in nations can be derived from the responses: 1) Average Happiness, 2) Happy Life Years, 3) Inequality of Happiness and 4) Inequality Adjusted Happiness. All four these measures reveal wide differences across nations and

considerable gains over time. The pattern of correlation with societal characteristics is highly similar.

The findings yielded with these measures show that happiness is a realistic policy aim. Happiness of a great number is apparently possible in modern society and greater happiness as well. It is as yet less clear how that can be achieved; the available data suggest that most gains can be made by policies that focus on freedom and justice. Economic growth is not likely to add much to happiness in affluent nations and neither is reduction of income differences or greater social security.

Notes

1 The term 'Gross National Happiness'(GNH) by the Government of Bhutan, as a name for its political programme that does not focus exclusively on economic development but also aims at preservation of tradition and nature. Happiness in the sense of life-satisfaction is not mentioned as a goal and consequently this matter has not been measured as yet in Bhutan.

2 In my view this last item is not appropriate. One can be quite satisfied with life, but still be open for the opportunity to try something else.

3 Civil registrations provide information on the longevity of citizens who have died. On that basis the life-expectancy of living citizens is estimated. These estimates acknowledge change in living conditions and medical technology.

4 Disability Adjusted Life Years, life-expectancy corrected for years spend in bad health. This measure is used by the World Health Organization.

5 Average of the first 8 member states for which data are available since 1973: Belgium, Denmark, France, Germany (W), Luxembourg, Italy, The Netherlands and the UK.

References

Andrews, F.M. & Withey, S.B. (1976), *Social indicators of well-being,* New York, U.S.A., Plenum Press

Bentham, J. (1789), *An introduction into the principles of morals and legislation*

Berg, M (2007), *Inkomensongelijkheid en geluk in landen (Inequality in income and happiness in nations)* Mens en maatschappij, vol.82, p.28-50

Bleichrodt, N.; Drenth, P.J.; Eds.: "Contemporary Issues in Cross Cultural Psychology, Swets & Zeitlinger, 1990/91, Lisse, Netherlands

Cantril, H. (1965), *The pattern of human concern,* Rutgers University Press

ChinHonFoei, S (2006), *Gender equality and Happiness in nations,* Paper presented at Dag van de Sociologie, Tilburg 8-6-2006

Diener, E., Emmons, R.A., Larsen, R.J. & Griffin, S. (1985), *The satisfaction with life scale,* Journal of Personality Assessment, Vol. 49, 71-75

Donovan, N. Halpern, D. & Sergeant, R. (2003), *Life Satisfaction: The State of Knowledge and Implications for Government.* Discussion Paper, Strategy Unit, Government UK psychology', Russell Sage Foundation, New York, p. 3-25

Hagerty Hagerty, Michael R.; Vogel, Joachim; Møller, Valerie (Eds.) (2002) *Assessing Quality of Life and Living Conditions to Guide National Policy*

The State of the Art, Springer: Social Indicators Research Series , Vol. 11

Kalmijn, W.M. & Veenhoven, R. (2006), *Measuring inequality of happiness in nations: In search for proper statistics.* Journal of Happiness Studies, Special issue on 'Inequality of Happiness in Nations', vol. 6, p. 357--396

Lane, R.E. (1994), *Quality of life and quality of persons: A new role for government?,* Political theory, vol. 22, p. 219-252

Lawton, M.P. (1975), *Philadelphia geriatric center morale scale-revision,* Journal of Gerontology, Vol. 30, 85-89

Ott, J. (2005) , *Level and equality of happiness in nations: Does greater happiness for a greater number imply greater inequality in happiness?* Journal of Happiness Studies, vol. 6, p. 397-420

Ouweneel, P. & Veenhoven, R. (1990) *Cross-National Differences in Happiness: Cultural Bias or Societal Quality?,*

Saris, W.E., Scherpenzeel, A.C. & VanWijk, T., *Validity and Reliability of Subjective Social Indicators* Social Indicators Research, 1998, Vol. 45, 173 – 199

Schwarz, N. & Strack, F. (1991), *Evaluating One's Life: A Judgment Model of Subjective Well-Being* Strack, F.; *et al*; (Eds.): "Subjective Well-Being", Pergamon, Oxford, UK, 27 - 47

Sen, A: (1992), Capability and wellbeing, In: Sen, A. & Nussbaum, M. (Eds.) 'The quality of life', Clarendon Press, Oxford, UK, p. 30-53

Veenhoven, R. (1993), Happiness in nations: Subjective appreciation of life in 56 nations 1946-1992 RISBO, Erasmus University Rotterdam, The Netherlands

Veenhoven, R. (2000a), *The four qualities of life,* Journal of Happiness Studies, Vol. 1, p. 1-39

Veenhoven, R. (2000b), Wellbeing in the welfare state: *Level not higher, distribution not more equitable,* Journal of Comparative Policy Analysis, vol. 2, pp 91-125

Veenhoven, R. (2000c), *Freedom and happiness: a comparative study in 44 nations in the early 1990's,* Published in: Diener, E. & Suh, E.M. (eds) 'Culture and subjective wellbeing', MIT press, Cambridge, MA United States, p.257-288

Veenhoven, R. (2005a), *Is life getting better? How long and happy do people live in modern society?,* Published in: European Psychologist, 2005, special section on 'Human development and Well-being', vol. 10, p.330-343

Veenhoven, R. (2005b), *Return of inequality in modern society? Test by trend in dispersion of life-satisfaction across time and nations,* Journal of Happiness Studies, Special issue on Inequality of Happiness in Nations, vol. 6, 457-487

Veenhoven, R. 2006, *World Database of Happiness, continuous register of scientific research on subjective enjoyment of life,* Erasmus University Rotterdam, Netherlands, Available at: http://worlddatabaseofhappiness.eur.nl

Veenhoven, R. 2006a, *Trend Average Happiness in nations 1946-2004.,* World Database of Happiness, Trend report 2006-1, Available at: http://worlddatabaseofhappiness.eur.nl/hap_nat/findingreports/2006-1.pdf

Veenhoven, R. 2006b, *Trend Happy Life Years in nations 1946-2004.,* World Database of Happiness, Happiness in nations, Trend report 2006-2, Available at: http://worlddatabaseofhappiness.eur.nl/hap_nat/findingreports/2006-.pdf

Veenhoven, R. 2006c, *Trend Inequality of happiness in nations 1946-2004.* World Database of Happiness, Happiness in nations, Trend report 2006-3 Available at: http://worlddatabaseofhappiness.eur.n/hap_nat/findingreports/2006-3.pdf

Veenhoven, R. 2006d, *Trend Inequality-Adjusted Happiness in nations 1946-2004.* World Database of Happiness, Happiness in nations, Trend report 2006-1, Available at: http://worlddatabaseofhappiness.eur.nl/hap_nat/findingreports/2006-4.pdf

Veenhoven, R. (2006e), How do we assess how happy we are? Paper presented at conference on 'New Directions in the Study of Happiness: United States and International Perspectives', University of Notre Dame, United States, October 22-24 2006

Veenhoven, R. (2007), Healthy happiness: *Effects of happiness on physical health and the consequences for preventive health care* Accepted for publication in Journal of Happiness Studies, online since March 2, 2007,

Veenhoven, R. & Hagerty, M (2006), *Rising happiness in nations, 1946-2004. A reply to Easterlin* Social Indicators Research, vol. 79, 421-436, DOI 10.1007/s11205-005-5074-x

World Database of Happiness, See Veenhoven 2006

Part Six
Challenges for National Statistical Offices

Chapter 17
Innovative Statistics to Improve Our Notion of Reality[i]

Henk K. van Tuinen[ii]
Former Deputy General of Statistics, Netherlands

Abstract

This paper serves the preparation of the second World Forum on "Measuring the Progress of Societies", to be held in Istanbul, June 27th-30th 2007. This conference is a key milestone of the OECD's World Forum on Statistics, Knowledge and Policy project. The paper:

- *analyses why current official statistics do not describe modern societies' progress as well as they did fifty years ago;*
- *identifies new information needs stemming from fundamental changes in our societies;*
- *uncovers an inherent tendency to conservatism in both the statistical system and its customers;*
- *concludes that GDP-growth sends misleading signals about the progress of developed economies and that mainstream economic theory and policy turn a blind eye to this problem;*
- *recommends new statistical information needed to describe the realities of the present day;*
- *recommends a more innovative attitude in official statistics;*
- *speculates about a new paradigm to replace mainstream economics.*

i Official statisticians should develop relevant new variables, systems and presentations and must increase their flexibility, creativity, courage and entrepreneurship

ii The author - former Head of National Accounts, Director of Social Statistics and (up to 2003) Deputy Director General of Statistics Netherlands – thanks J.C.J.M. van den Bergh (Free University of Amsterdam), G. Bruinooge and his colleagues, S.B. Algera, B.M. Balk and A.M. van Krimpen (Statistics Netherlands) R. Hueting and S. van Tuinen (University of Gent) for valuable comments. He is much indebted to Jon Hall (OECD) who read two versions of the paper and did numerous suggestions about content as well as language. E-mail: tuinvink@xs4all.nl

Summary for Anybody Needing Statistics

What statistics do we need? To what purpose? In history, statistics were needed for an estimate of how many soldiers could be conscripted into military service, or what amount of taxes could be collected. Later more figures were needed like: the availability of essential commodities, the cost of living of the poor, the labour force, the execution of the penal code, causes of death. Over the past fifty years, the need for statistics has broadened to many aspects of social and economic life including, more recently, the physical environment. The central questions used to be: "how many?" or "how much?" Before asking these questions, there was a *priori* knowledge about what had to be quantified. Decision makers knew what was going on; what they needed was quantification.

Nowadays, our societies have grown more heterogeneous, dynamic and complex. Our a priori knowledge has become fragmented and is outdated more quickly; developments in society have become less predictable. If we ask statisticians only to quantify the phenomena we know, the statistics will describe only fragments of the relevant developments and they will fail to describe unknown new developments. Therefore, in addition to questions of "how many?" or "how much?" we want to ask "what is going on?" What new relevant phenomena are emerging? What interdependencies play a role? Which processes are driving the changes in society?

Is it sensible to ask official statistics these new questions? The answer depends on two issues: (1) are there others who are in a better position to find answers to these new questions and (2) are official statistical institutes capable of finding those answers?

Concerning the first issue, social scientists and economists can be asked. However, why should their *a priori* knowledge not be fragmented or outdated?[1] If they are to observe society open-mindedly in order to find answers, one of their essential vehicles is ... statistics. So, they will turn to official statistics to help answer these questions. We could also ask commercial organisations or non-governmental organisations (NGOs) the new questions. Aren't they together producing an information overload that feeds our wish for authoritative and undisputed data? Moreover, can we expect them to give unbiased answers? It is difficult to find occupations or institutions where open-mindedly looking at what is going on in society is so essential a core value as it is with official statistics. We expect official statisticians to be unbiased, cultivating an undistorted view on reality, not one predetermined by any special interest or by any outdated or disputed theory.

Concerning the second issue, a source of genuinely unbiased information and core values of impartiality and independence are beyond value. Official statisticians have a treasury of basic data that can be explored and their broad experience in estimation and presentation is a powerful tool. Would it not be a waste of publicly-financed knowledge if we did not ask official statistics the new questions?

The paper analyses why contemporary official statistics do not describe the progress of modern societies as well as they did fifty years ago. It identifies new information needs stemming from fundamental changes in our societies. It demonstrates that to meet these new needs, new statistical analyses and more flexibility in innovating statistical work programmes are required as well as the courage to explore daring concepts and to initiate controversial debates on how to give an unbiased statistical description of what is going on in society.

Why did official statistics fail to develop all the new information needed today? In answering this question, the paper analyses two reasons:

- As modern societies have become more complex, relatively simple statistical data (answering the question "how many?") must be supplemented or even replaced by more complex statistical information. This raises budgetary and methodological problems for official statistical institutes. Moreover, the need to present complex statistical information, e.g. to policymakers who need statistics that are easy to communicate, is a challenge to statisticians' creativity. Who calls upon statisticians to produce complex information?

- In the world of official statistics, strong incentives exist to be conservative. This inherent conservatism is not systematically challenged by the clients of official statistics. Indeed, those clients have equally strong incentives to conceptual conservatism. Economic statisticians are lagging behind which is intertwined with the problem that mainstream economists are lagging behind. Who challenges statisticians to produce information that is beyond the mainstream?

The paper recommends strategies for official statistics to develop the innovation, courage and entrepreneurship, needed to keep abreast with the present and the future. Of course, these strategies can only be effective in a society which is aware of the essential value of well developed official statistics. Supporters of official statistics should stimulate statisticians to develop and actively demonstrate vigilance over new developments in reality which require new statistical information or which decrease the relevance of existing information. But the statisticians must lead the way by showing that they are among the first to discover new developments that require new statistics and that they have the knowledge and the creativity to present complex or even controversial information convincingly. Especially, they have to demonstrate that they have become better at producing adequate answers to relevant information needs, rather than on producing accurate answers to outdated questions.

The paper recommends official statistical organisations should develop an outward-looking attitude. An attitude that should result in them becoming pro-active and flexible and so strengthening the Statistics, Knowledge, Policy chain as well as in reaching out to a more heterogeneous clientele. Flexibility and professionalism in the presentation of information (and knowledge) is an essential feature of this.

A major part of the paper is devoted to concrete information problems. Much attention is given to the reasons why GDP growth is an increasingly misleading indicator of economic progress, let alone 'progress' per se. At the same time it is shown that mainstream economics turns a blind eye to essential limitations of GDP as an indicator of economic progress and that economic theory fails to solve the underlying problem.

Attention is also given to the implications of increased heterogeneity and dynamics in modern societies. It is demonstrated how statistical aggregates, answering the question "how many?", fall short in describing essential social, socio-economic and economic developments and how statistics describing processes (in particular longitudinal information) can fill these information gaps. The implications of the increasing complexity of our society are analysed as well. It is shown how the measurement of a common variable like inflation threatens to become a mission impossible and how societal complexity requires information on interdependencies because systems of isolated indicators will fall short.

The paper demonstrates how the virtual reality of aggregates camouflages the fundamental differences between modern and traditional societies. A modern society, described by traditional statistics, will show up as a predominantly traditional society, because most of the heterogeneity and dynamics remain hidden; the timeseries of aggregates suggest relative stability and will thus wrong-foot policymakers and other users of statistics. The statistics may be technically perfect, but they do not portray what is really happening. Therefore, the paper strongly recommends the development of statistical information describing the dynamic processes at the micro-level and effectively present this complex information.

The substantial growth of GDP and consumption in modern economies over the past fifty years did not lead to a significant increase in economic welfare. Mainstream economics is unable to explain this anomaly. Although GDP has become an increasingly misleading indicator of economic progress, economic policy making still focuses on the virtual reality of this aggregate. Official statisticians should stimulate debate on the interpretation of GDP growth, *e.g.* by publishing data on costs (rather than sources of welfare) included in GDP and by publishing data on those sources of economic welfare that are not included in GDP. The paper recommends that official statistics should develop integrated statistical information on happiness and that official statisticians should initiate and get actively involved in the development of estimates of Sustainable National Income.

The recent developments around systematic and valid information on (sources of) happiness will especially open new perspectives for statistical information on the progress of modern societies. The paper sketches contours of an integrated information system which can serve economic and social policy making. It also indicates why this new perspective is distant from mainstream economics. The paper suggests economists should abandon the most unrealistic assumptions of mainstream economics, concerning the autonomous and rational behaviour of consumers, and develop a new paradigm in which economic welfare as a state of mind, economic power and endogenous (instead of completely exogenous) preferences play a realistic role. The paper speculates about this new economic theory and concludes that it will be able to explain why GDP growth did not lead to increasing welfare. And it recommends official statisticians should stimulate the debate on economic theory by publishing data on marketing expenditure as an indication of economic power and the manipulation of consumers' preferences.

The paper recommends that 2% of the total budget of official statistics in developed countries should be earmarked for strategic research projects and that the flexibility of the statistical production process should be increased by creating longitudinal databases of linked microdata, available for the production of unforeseen statistical information.

The Problem

The following section briefly describes the problems raised in the paper. It uses words like "reality", "statistical simulation", "virtuality" and "relevance". These words play an essential role throughout the paper. Therefore, the following sections explain the terminology.

Description of the Problem

Half a century ago, official statistics in developed countries captured their society's "progress" better than today. In most of these countries the statistical output has increased considerably since then and the quality of the output has also improved in many ways. New variables have been added, measurement techniques improved, publications have become more timely and more detailed. Nevertheless, we observe a growing awareness that essential social and economic developments are not really covered. A feeling that the progress of society as measured, *e.g.* in statistics showing continuing economic growth, does not reflect the real progress (or, some might argue, the regress) of our societies. So, the first problem in this paper is to answer this question: Are official statistics monitoring a "reality" which has become too virtual and, if so, why has that happened?

The analysis must be rather comprehensive. For a consideration of the development of official statistics its context must be taken into account. Firstly, today's societies differ in many respects from those of fifty years ago. Therefore, current statistics need to describe new realities. Secondly, perceptions of reality have changed. Both observations pose a challenge to existing statistics which describe reality according to traditional perceptions. And so, the development of economic statistics must be analysed alongside the development of economics and economic policies. But a similar question comes up when one considers economics and economic policies: are they no longer so effective in addressing the essential economic developments? Are they based on a reality which has become too virtual? How does the use of available statistics contribute to that virtualisation?

Another problem raised in this paper is how to create and maintain a process of developing official statistics that adequately monitor reality. What innovations should be made to the statistical work programme in such a way that it continues monitoring the most essential elements in social and economic development? And which elements are most essential today? This paper will illustrate how new developments emerge in such a way that the very content of "relevant reality" changes. The implication is that monitoring reality is a dynamic process that requires statistics to permanently keep up with reality.

If current official statistics do not adequately monitor today's reality, there is urgent work to do. The mission of official statistics requires statisticians to monitor the most relevant developments, including the development of new statistics to monitor new realities. Moreover, if official statistics continues to present information which does inadequately monitor today's reality, this information will, sooner or later, be considered misleading.

What work is most urgent? The paper will present concrete proposals for new statistics or statistical systems. These new statistics are considered essential for monitoring today's reality. In addition, suggestions will be presented concerning statistical techniques to bring more flexibility in innovating statistical work programmes. Finally, a strategy will be sketched for permanently keeping up with reality.

Terminology forAnalysing the Problem

If we accept the idea that our knowledge of the world is incomplete, the implication is that our view of reality is not necessarily coincident with the truth. "Reality" is a representation of the true world, a construction determined by available knowledge and instruments for observation. In this sense, reality is the product of a process of simulation. For instance, what we observe using our eyes is simulated by our brain before it enters

our consciousness and, then, simulated further as far as our interpretation determines our view. Needless to say that the simulation process starts long before our observation when we observe indirectly, *e.g.* through media like television and newspapers, or through statistics.

Statisticians should recognise this concept of reality. While statistical information is used as a description of reality, the producers of the information are aware of the fact that the figures are determined by the concepts adopted and the methods applied. By adopting other concepts, or by applying other methods, the statistician would produce other figures and, consequently, present another simulation of reality.

This concept of reality implies that reality is inevitably "virtual" because it is the product of simulation. Being less strict, one can speak of a reality that is more or less virtual. This will prove to be useful when discussing the development of reality, its perception and its statistical simulation.

Reality may change in two ways. The true world is changing and our simulation processes evolve as well. Good statisticians try to separate the effects of changes in the statistical process from their description of changes in reality, but they do not always succeed. It is very important to ask whether the relevant developments in reality are portrayed by the statistical simulation in a way that the user recognises as the relevant description. That, indeed, is the central problem in the paper. If the relevant developments in reality are decreasingly reflected in the statistical information, we will speak of statistics simulating a reality that is increasingly virtual. We will speak of "too virtual a reality" if the statistical information becomes an unacceptable or misleading simulation of reality.

The paper presents some of the relevant developments in reality which are poorly reflected in official statistics or which are not reflected at all. Which developments are relevant? And how can we identify these developments? The paper will offer concrete and specific answers to these questions.

Although the paper will use the word "progress" at several occasions it will not explicitly define the concept.[2]

Origins of the Problem

The problem that official statistics are now less relevant to measuring a modern society's "progress" than they were fifty years ago has at least five origins. Four arise from basic developments in modern societies during the past fifty years. The fifth concerns the usual limitations of the mission of official statistics.

Heterogeneity

Modern societies are more heterogeneous than they were fifty years ago. Consider a discriminating variable like the level of education. Fifty years ago, the group of people with less than secondary education included roughly 90% of total population; nowadays, the population is widely dispersed over different levels of education. In European societies ethnic heterogeneity has increased considerably. At the same time a strong cultural trend of individualisation became visible, giving rise to more heterogeneous lifestyles.

Increasing heterogeneity became also visible in modern economies. The variety of goods and services produced and available to citizens increased substantially, as did the range of production processes. A comparison of our consumer basket to that of our grandparents may persuade the doubtful.

In traditional statistical presentations the population is aggregated into a limited number of groups, broken down by classifying variables like age, gender, marital status, occupation, etc. The rationale of this procedure is that those groups can be considered relatively homogeneous, which means that 'the average member' of the group has much in common with the vast majority of the real members. Increasing heterogeneity implies that these kinds of statistics will tend to describe a reality that is increasingly virtual: 'the average member' is less and less like the real individuals in the group.

The relevant statistical implication of this is that aggregates become less useful. That means statistical presentations must be more detailed and so their production becomes more expensive (requiring *e.g.* bigger sample surveys) in order to describe the more heterogeneous populations. More fundamentally, social structures are increasingly blurred and the focus of attention shifts from the statistical description of the population to the description of the differences within the population.

Dynamics

New commodities as well as new topics emerge quicker than fifty years ago. In the course of a modern life, people more often change their living circumstance, like marital or cohabitation status, labour-market status or occupation. More years of education, higher incomes, more information and the cultural trend of individualisation enabled and challenged people to live their life in their own way. In many respects modern societies have become increasingly dynamic.

The statistical implication of increasing dynamics is that static or comparatively static descriptions become less meaningful. It not only means that statistical descriptions must become more timely and more frequent. It means that the focus of attention must shift from structures to processes and from cross-sectional to longitudinal information.

Traditional statistical presentations focus on aggregates (totals, averages) or, more seldom, distributions for a reference period. An important special case are timeseries which compare aggregates over time. Timeseries are often interpreted as a description of what happened to the individuals in a population. But, a change (over time) in an aggregate may have two causes: (1) changes (over time) in the elements of the population and (2) changes in the composition of the population (entry and exit of elements). If the changes in the aggregate are dominated by the first cause, they are likely to represent the real changes in the elements of the population. But, if the second cause is dominant, the timeseries does not represent that reality. Therefore, increasing dynamics may lead to timeseries displaying a reality which is increasingly virtual. Chapter "Inadvertent deformation of reality" will present concrete examples.

However, there is a more fundamental implication. When societies become more dynamic, individuals more often change positions, enterprises and other institutions more frequently change business etc., these processes at the micro level become the more dominant force behind societal developments. The implication is that statistical descriptions of developments in reality must include information on those micro processes. Without analysing and summarising longitudinal microdata, official statistics describe an increasingly virtual reality and fail to describe the most relevant realities.

Complexity

To say that modern societies are more complex than fifty years ago seems an understatement. Complexity increased considerably from those same trends that caused increasing dynamics and heterogeneity. But it has also grown because of other trends, like globalisation, a growing awareness of interdependencies and long term bottlenecks like environmental problems, information overload, etc. Our knowledge of society appears to be more fragmented and developments have become increasingly unpredictable.

Describing a complex phenomenon is difficult. There are at least four issues for statistics:

- The methodology of statistical observation and estimation is challenged by the complexity of the phenomena
- The methodology of statistical presentation is challenged, because complex statistical information is not easily understood and, for that reason, can turn out to be used less
- As the general knowledge of society has become more fragmented, which in itself challenges official statistical institutes, these institutes are increasingly vulnerable if they fail to fill the growing gaps in information, because in the long run existing statistics will run higher risks of being considered one-sided or misleading
- As societal developments have become increasingly unpredictable, drawing up long term statistical work programmes has become less adequate, which challenges the strategy of statistical institutions.

These themes will be elaborated in the following chapters. For now, we may conclude that current statistics describe an increasingly virtual reality, by definition, because reality is more complex.

Affluence

Since the origin of statistics on national income and related variables, welfare economics has produced arguments why GDP does not correctly measure economic welfare. A complete list of these arguments is very long. It would include theoretical arguments like *e.g.* the fundamental impossibility of comparing, and thus of aggregating, personal utility, or the ignorance of the consumers' surplus which occurs because of the law of diminishing marginal utility. The list would also sum up a long series of conceptual arguments like *e.g.* the exclusion of major elements of welfare like products of unpaid work or non-produced (but nowadays scarce) goods like a healthy and sustainable physical environment as well as the inclusion of 'costs' like the labour costs on account of employees. Finally the list would include empirical arguments like the fact that satisfaction leads to growing needs (preference drift), one reason why a higher GDP (more goods for satisfaction of needs) leads to a less than proportional increase in welfare, and the fact that in developed economies welfare depends more on relative income than on absolute income (reference drift). The combination of preference drift and reference drift may cause a higher GDP to result in no increase in welfare at all.

Most arguments are as old as GDP itself, but their overwhelming impact emerged gradually from empirical work during the last thirty years. Nowadays, we know that for a

per capita GDP of above $20 000 happiness (and thus economic welfare) does not at all increase with income (Layard 2005, p.34). And we know that middle class Americans during the end of the twentieth century were feeling increasingly poor while their real income and consumption increased unprecedentedly (Schor 1998, p.11-20). Other research concludes that Americans would be happier if they would spend less time on earning money and more on family life and health (Easterlin 2003, p.11 182) which suggests a higher GDP to cause less welfare. A far more complete survey of this evidence in Layard, 2005.

In spite of all this theory and evidence, the volume of GDP growth is the dominant indicator of economic progress worldwide, not least in countries with a high GDP. In Western European countries, policymakers urge citizens to spend more time on earning money in order to speed up or maintain GDP growth. What explains this behaviour?

GDP growth has favourable effects in the short run, especially for policymakers. It reduces unemployment and it solves problems in government budgets. And it is the central variable in the system of National Accounts which authoritatively describes the part of the economic process directly connected with the use of money. And money is an important incentive, also in the process of policy making. But, can we expect these arguments to sufficiently explain policymakers' focus on GDP as the dominant indicator of economic progress?

Perhaps, there is an underlying reason why policymakers and their economic advisors keep focussing on GDP growth. A plausible hypothesis is that these policymakers hold a view on reality which is distorted by existing statistics and by traditional (mainstream) economics. They work with the data they have at their disposal (GDP) and not with concepts that are not yet statistically operational (welfare). The available statistics describe the virtual reality of ongoing growth and that is a welcome message for policymakers and their audience. That message also suggests that mainstream economics is right in predicting welfare maximisation if rational economic subjects are operating in free markets. In this way, GDP growth describes a virtual reality in perfect harmony with the claims of mainstream economic policy.

This hypothesis will be analysed at several places in this paper, including sections "How Mainstream Economics is Lagging Behind" and "GDP Growth and Welfare". For now, we may conclude that the emergence of affluence (a high level of per capita GDP) caused GDP growth to describe an increasingly virtual reality. Unlike fifty years ago, it does not indicate the economic welfare of citizens of developed countries.

The Mission of Official Statistics

The past fifty years saw a lot of new developments in modern societies and official statistics has addressed many of them, often adequately. Instead of summing up these complete or partial successes, the paper will concentrate on the origins of the problem mentioned. Among recent and relevant developments, the concern for environmental problems has a prominent place. Official statistical institutes have created a range of authoritative environmental statistics and systems were designed for analysing the relations between economic and environmental developments.[3] Can this be considered a success? The answer could be "yes" if the ambition of official statistics is limited to producing the basic data, leaving the estimation of broader concepts like welfare or progress to the users of statistics. But if official statistics are to assist users in forming a view of welfare or progress, by striking a balance between GDP growth and

environmental degradation, the answer is "no, not a complete success". If GDP grows and, at the same time, the environment deteriorates, nothing can be concluded about changes in true economic welfare. In order to produce an indication of growth of economic welfare, the statistician must strike a balance between GDP growth and environmental deterioration, for instance by applying a concept like a green GDP.

A similar point can be made about income inequality, unemployment, social protection, etc. Welfare economics has clearly demonstrated that nothing can be concluded concerning aggregate economic welfare if *e.g.* average income increases while inequality increases at the same time. Should official statistics leave 'striking the balance' to the users?

Per capita growth of GDP has remained the most important indicator of welfare in mainstream economic policies. Substantial GDP growth is commonly presented as economic "progress" and even an annual growth of GDP of 1% (or lower) is considered deterioration. About 35 years ago, the growing awareness of a threatening environmental deterioration, including a depletion of natural resources (*e.g.* fossile fuels), gave rise to a wave of criticism around the concept of GDP as well as the emphasis placed on it. Nowadays, a new wave is on its way, rooted in environmental as well as socio-economic arguments. If this wave becomes a tsunami, how safe are official statisticians behind the footnotes in their textbooks which state that GDP is no correct measure of welfare? After decades of presenting these statistics to an audience which interprets them as welfare and "progress", official statistics may run the risk of being considered misleading. If official statisticians leave 'striking the balance' outside the boundary of their mission, they may be considered inadequate.

The paper comes back to this point in the chapter "What economic progress?" and arguments concerning the mission of official statistics will also be considered in the chapter "A dilemma in the mission of official statistics". For now, we may conclude that excluding 'striking the balance' from the mission of official statistics is one of the sources of the problems in monitoring "progress".

Why Statisticians are Lagging Behind

The preceding chapter explained that 'keeping up with reality' has become increasingly difficult. If that were the only reason why statisticians are lagging behind, one should expect at least as much pioneering work going on in statistical institutions as fifty years ago. But, in contrast to the years of the birth of National Accounts and social survey statistics, current statistical research seems less concerned with major conceptual innovation. ICT made statistical production processes more efficient and sometimes more flexible, but statistical work programmes do not show enough flexibility and innovation to keep pace with reality.

The reasons why official statistics lag behind stem from the basics of the statistical system. Not only does the system inherently tend to conservatism but, moreover, the conservatism is contagious.

A Law of Inherent Conservatism in Official Statistics

A 'law of inherent conservatism' operates in advanced statistical work programme design. This section will briefly discuss the mechanisms behind this law. These mechanisms stem from the mission of official statistics, from characteristics of the

statistical production process, from the process of deciding statistical work programmes and from the existence of similar laws in science and in the development of policies.

The mission of official statistics is to provide the society with undisputed information. The statistics must be impartial and authoritative so that all actors, including those with opposite interests, can agree on using them. Statistical institutions have to guard the authority of their statistics. Therefore they will be reluctant to emphasise the shortcomings or to develop competing (conflicting) information. The authority of a set of statistics grows with the duration of its use. This encourages official statistical institutes to maintain existing statistics, and thus to be conservative in developing substitutes.

Producing new statistics of sufficient quality often requires high costs and a considerable period of preparation. Changing the work programme is costly. As most statistics are used in the form of rather long timeseries, the stimulus to be conservative is strong.[4]

With strong 'internal' stimuli to be conservative, external stimuli to be innovative in drawing up work programmes would be a helpful counterbalance. But the dominant external stimuli are conservative too. The 'clients' of statistical institutes are the users of existing statistics and these clients are more involved in deciding statistical work programmes than potential clients of the still non-existent statistics. The demand for more details, improved timeliness or higher frequencies of existing statistics is often more concretely, professionally and authoritatively substantiated than the demand for completely new statistics from which the basic concepts have not yet been made operational.

External stimuli to conservatism become even stronger as soon as statistics are directly applied in administrative processes. In the European Union the system of National Accounts, with all its concepts and definitions, is carried into law (EU 1996) as the figures directly determine huge money flows and major policy issues. It should be clear that conceptual innovation will be hampered by the increased complexity of decision making and the involvement of parties concerned in administrative processes.[5] This illustrates how the success of a set of statistics leads to its inflexibility.

In modern societies, where an important function of official statistics is to reduce uncertainty and to lower transaction costs, the inclination to statistical conservatism seems to be 'natural'. This inclination is intensified by structural tendencies to conservatism in other sectors of modern societies. From epistemology it is well known that the scientific world is conservative. New paradigms are confronted with strong opposition and often face a long struggle, needing completely convincing victories over ruling paradigms before being accepted. This conservatism stems from the scientist's learning process as well as from established interests and institutions (including scientific periodicals and textbooks). The dominant actors in the scientific world have built their reputation by demonstrating their knowledge according to the ruling paradigm. Similar mechanisms are at work in the world of the development of governmental policies.[6] The following section, indicating how mainstream economics and economic policies are lagging behind, may serve as an illustration.

Science and policy are the dominant clients and inspiration of official statistics. Their conservatism contributes to the workings of the law of inherent conservatism in drawing up statistical work programmes.

How Mainstream Economics is Lagging Behind

Welfare economics criticised the use of GDP as an indicator of economic welfare as was illustrated in Section "Affluence". However, welfare economics does not represent mainstream economic theory. Present mainstream economics, like welfare economics, has its roots in neo-classical economic theory. But, unlike welfare economics, it focuses on how markets work. Mainstream economic policies in Western countries and in a growing majority of non-Western countries focus on how to improve the working of free markets. These policies were not dominant during the 1950s-70s when Keynesian policies prevailed. And those policies prevailed too long, leading to the combination of inflation and stagnating growth called 'stagflation'. It is now generally acknowledged that, during the seventies, mainstream (Keynesian) economics clearly lagged behind.

Neo-classical free market oriented policies became the mainstream in the 1980s. The collapse of the centrally planned Eastern European economic systems reinforced the dominance of present mainstream economics. This has dominated policy making, liberating markets, including the labour market, privatising public companies, counteracting dominant monopolies and reducing social security, introducing market-oriented incentives to stimulate participation in the labour force. The resumption and continuation of GDP growth affirms the superiority of today's mainstream economics compared to Keynesian (or Marxist) economics.

The idea behind present mainstream economics is that free markets are the most efficient allocation system. Moreover, this system fits very well in modern democratic societies and in a culture of freedom and individualisation. The core notion is that consumers are free to choose among competing suppliers, thus steering the supply side to produce the basket of consumption goods that maximises consumers' utility. It is postulated that consumers act autonomously and rationally and that they are sufficiently informed about the markets. There is a growing literature,[7] more or less compatible with mainstream economics, on the consequences of incomplete information and there are policies directed to improve consumer information. The idea of autonomous and rational consumers is not seriously challenged within mainstream economics, but it is criticised beyond the mainstream; in behavioural economics as well as in marketing theory. Dominating economic power on the supply side disturbs mainstream economics; the government is assumed to intervene if markets fail.

The efficiency of the free market is today's ruling paradigm.[8] But, is the free market system effective? The fact that the ongoing GDP growth in developed countries during the last decades did not lead to increasing economic welfare, as was pointed out in "Affluence", suggests that the market system was not effective. Although consumers empty a basket of more goods and services than before, their utility does not increase. Does this imply that mainstream economics is seriously lagging behind again? If so, this might have been the case in Western economies for decades; as long as GDP growth did not lead to increasing economic welfare.

Mainstream economics is turning a blind eye to whether growth is effective and prevailing economic policies stay focussed on the efficiency feature. Three possible explanations are relevant. The first is that, as pointed out in Section "A Law of Inherent Conservatism in Official Statistics", the scientific world is generally reluctant to drop ruling paradigms. The second is that in non-Marxist economics the autonomous and rational behaviour (assuming sufficient information) of economic subjects is an indispensable corner-stone in the majority of respected theories. That implies that if the

markets work efficiently, utility *must* be maximised. The third is that most respected theories abstract from dominant economic power.

In economic theory, profits indicate power. In the case of perfect competition profits are zero. Statistics of Western economies show considerable profits and the share of profits in GDP shows a positive trend during recent decades. But this is not the only indication of growing economic power. If power is defined as the ability to manipulate other subjects, marketing expenditure like advertising is another indication of power. If the postulate of rational behaviour is maintained for agents on the supply side of the economy, we may assume these marketing costs to be effective. Therefore we may conclude that consumers are effectively manipulated and thus, that they are not entirely autonomous. It may be that they act rationally in maximising their utility, but that concerns the satisfaction of manipulated needs. Maximising the satisfaction of manipulated needs may result in higher consumption giving a lower increase in welfare than if consumers were maximising the satisfaction of authentic needs. This helps explain the ineffectivity of the free market system. And it supports the hypothesis that mainstream economics is seriously lagging behind. This hypothesis will be discussed further in Sections "Complexity, Consumer Sovereignty and Mainstream Economics" and "Economic Power and Welfare, Including Social Relations, Acknowledgement and Self-expression".

Why Statistical Conservatism is Contagious (and Why Official Statistics Itself Must Produce the Antidote)

The conservatism of science, policy development and official statistics are mutually stimulating. On the one hand the decision-making process of statistical programming, sketched in Section "A Law of Inherent Conservation in Official Statistics", serves as an illustration. On the other hand, the very character of statistical information plays an intriguing role, especially through its concepts.

When a new set of statistics is developed, a lot of conceptual work has to be done. If relevant scientific concepts are available, they will be adopted as a valid point of departure. These scientific concepts, reflecting the ruling paradigm of the age, often are not directly operational in a statistical process because they are too abstract to exactly describe what can be observed or they need translation before they can be applied in the process of statistical observation. Therefore, statistical concepts have to be developed in great detail.

Once the new set of statistics is produced and used, the user explicitly or implicitly accepts the statistical concepts. As long as she uses these statistics as a source of knowledge, her view of reality will be coloured by these concepts. She sees the developments that the statistical concepts permit her to see, but she does not see what these concepts conceal. If the statistics are her only vehicle to monitor the relevant reality, she will not observe those developments in reality that could be described only by applying other concepts.[9] Consequently, her view of reality will give no cause for advocating new statistical concepts. That is why statistical conservatism, especially concerning concepts, can be considered contagious.

Especially in the world of economic and social sciences and policies, statistics are a dominant vehicle to monitor reality. For that reason the contagious nature of statistical conservatism is especially important in these sectors and, thus, in the vast majority of official statistics. Section "Contagious Conservatism in Statistical Concepts: An

Example" presents a concrete illustration focussing on the elementary economic concepts of welfare, production and consumption.

To adequately keep up with reality, therefore, official statistics cannot rely on the stimuli received from its users. The more authoritative a statistic, the less innovative the feed-back. But, who else will stimulate innovative statistical work programmes?

An essential competence of official statisticians is being unbiased, cultivating an undistorted view on reality and having an open eye for all developments which could be relevant. It is difficult to find occupations or institutions where open-mindedly looking to what is going on in society is so important a core competence as it is in official statistics. It enables the statistician to be among the first to discover new developments which ask for new statistical descriptions. A rational society will expect him to do so, because they look to him to produce unbiased descriptions of reality. This logically implies that he is responsible for keeping up with reality. He should be the first person to show the user of his statistics the relevant blind spots, why filling information gaps are important and how this can be done. He has to convince the satisfied user that she should be unsatisfied. He must become a statistical entrepreneur.

Chapters "How to Increase Flexibility?" and "A Dilemma in the Mission of Official Statistics" will suggest strategies to overcome the inherent conservatism and to develop entrepreneurship.

Contagious Conservatism in Statistical Concepts: An Example

Theoretical economics uses abstract concepts. It defines welfare as a state of mind that results from economic actions. More precise, welfare is the satisfaction of needs through the use of scarce means which can be employed alternatively. This concept is far more abstract than the statistical concept of GDP. Other examples of abstract economic concepts in theoretical economics are 'consumption' and 'production'.[10] The theoretical definition of consumption is "an activity which directly leads to satisfaction (of needs)". In the process of production, labour is a sacrifice because it requires time and energy which otherwise could have been devoted to consumption. Like in the case of welfare, the essential criterion is the state of mind of the economic subject. The distinction between consumptive and productive activities is made in the subject's mind: satisfaction or sacrifice?

Applied economics (*e.g.* macro economics, labour economics, monetary economics, and marketing) needs concepts which define phenomena which can be concretely identified. Therefore, welfare has no place in applied economics. Consumption and labour are defined in applied economics as concrete activities, not as states of mind. In macro-economics and labour economics, labour is defined as activities which yield income. In macro-economics and marketing, consumption is defined as the purchase of goods and services by households.

Economic statistics are based on the statistical observation of concrete phenomena. This implies that statistics, like applied economics, does not define its concepts according to the state of mind of economic subjects. Its concepts address observable activities or transactions. In practice, statistical definitions are extremely concrete and detailed; they have to operationalise the concepts of (applied) economics. Thus, economic statistics does two things for applied economics: it defines more precisely and it provides the quantification.

Statistical data are very important to applied economists. The data describe the relevant reality for economists and in order to understand the data economists study the statistical concepts. Many economists learn the basic concepts of the National Accounts during their schooling and empirical economists study these concepts in greater detail. These concepts have become an authoritative foundation of their view on reality. This is why section "Why Statistical Conservatism is Contagious (and Why Social Statistics itself must Produce the Antidote)" of this paper speaks of contagious conservatism. Statistical concepts do not only follow science, they operationalise scientific concepts and then, after the statistics have become indispensable in research, science internalises the statistical concepts.

The implications are very important. Like in theoretical economics, in statistics labour is considered a sacrifice to be compensated by receiving income and consumption is considered a source of welfare. But the agreement is deceiving. If people enjoy their job, their labour is not just a sacrifice but it also yields direct satisfaction of some of their needs (*e.g.* a need for developing their faculties or a need for feeling to have a purpose). Theoretical economics implies that this satisfaction is a component of welfare. If households have to pay for travelling to their workplace, economic statistics treats these payments as consumption. But it is clear that these payments represent a sacrifice: not serving direct satisfaction of needs but serving labour. So, the operationalisation of abstract theoretical concepts into concrete statistical concepts has two consequences. First, essential features of the theoretical concepts get lost. *E.g.* pleasure in work vanishes from sight as a component of welfare; costs of labour (*e.g.* commuting) on account of consumers appear as a source of welfare. Second, as the theoretical concepts are not represented in statistics, users of available statistics interpret the statistical concepts as a representation of the theoretical concepts. GDP and consumption are interpreted as a representation of welfare.

The virtual reality of statistics dominates our view of what is going on. It is relatively easy for official statistics to combat this misinterpretation. At least two things could be done. First, in the main National Accounts aggregates elements of GDP or consumption which represent costs rather than direct sources of welfare should be displayed separately. Inevitable arbitrary judgements can be processed in alternative variants, provoking users to take part in the conceptual discussion. Second, elements of welfare which are not included in consumption (*e.g.* pleasure in work) can be identified, measured (starting with survey measurement of elements of well-being) and presented in a context of economic growth and welfare. In this way, an antidote against contagious conservatism could be developed, stimulating conceptual debate on the basis of statistical data of alternative concepts.

Inadvertent Deformation of Reality

This chapter illustrates how traditional statistics, presenting timeseries of aggregates, simulate a reality that wrong-foots its users.

How Aggregation Distorts Our View of Growth (and of Progress)

Let us take a look at current discussions about the pros and cons of GDP growth in developed economies. Those against growth argue that, on the one hand, growth does not bring benefits because, above a GDP of $20 000 per capita, GDP growth does not increase welfare (see Section "Affluence") and, on the other hand, growth is bad because

it harms the environment. Arguments pro growth include that people's behaviour shows that they want growth (but see Section "Economic Power and Welfare, including Social Relations, Acknowledgement and Self-expression") and that growth generates technological innovation which solves environmental problems. The first pro-argument stems from clear evidence and fits into Western cultural paradigms, which since the Renaissance and Enlightenment have pressed for discovery and progress. It is striking to note that, although the argument addresses human behaviour, its conclusion is formulated at the macro level (GDP). The idea that individual people want growth is projected, apparently unthinkingly, on the aggregate level. This kind of reasoning seems to happen often.

Imagine a society where per capita real GDP (and consumption) growth is zero, the active labour force forms a constant share of the population, people do paid work for 40 years of their lives and tax rates are constant. Are these working people deprived of (real) income growth? The answer, of course, is "no" if they start their career with a lower than average income. As an illustration we assume a simple model. All workers start with 60% of average income and then increase income every year by 2.5%. After 28 years they have doubled income and then keep their income constant during the last 12 years before their retirement. This model fits exactly with the condition of zero GDP growth. There are many alternatives. If you like, you can increase the period of income growth by assuming a lower than 60% of average income starting point or a slower rate of growth than 2.5% a year. *E.g.* starters could earn 70% of average income and see their income growing (by 2% yearly) until it doubles after 35 years. Assuming less homogeneity and more complex career lines, a broad range of outcomes are possible.

The conclusion from this simple model is that zero aggregate growth is compatible with an annual income growth of over 2% for individuals during the major part of their working lives. This illustrates why individual behaviour showing people want growth cannot simply be projected on the aggregate level. But such projections are very common,[11] which illustrates how our thinking is distorted by the virtual reality of aggregates and how we are in need of longitudinal statistics, describing the reality of individual experiences, to restore our view of reality.

It is tempting to explore the problem more broadly. What is growth at the micro level? Is it simply 'more of the same' consumption according to the National Accounts? Let us look at the course of life of western people, ignoring the complicating dynamics of divorce and forming new more person households, childlessness, unemployment, sabbaticals, etc. In the course of their lives, most people leave their parental home to form a one or two person household; they have children and these children become more expensive; after the children have become financially independent more of the parents' budget is available for luxury; after retirement income is lower and, assuming health care insurance, the propensity to consume decreases as people grow old. Many people will leave a substantial inheritance and their heirs will often receive that after their children have become financially independent, when their incomes already allow for more luxury.

This economic sketch of a 'standard' course of life illustrates two things. Firstly, consumption varies substantially over the course of life. This reality stays veiled in aggregate consumption according to the National Accounts. Secondly, even if personal income does not increase after midlife, many people become more affluent when their children become financially independent. This "progress" at the individual level is not reflected in standard aggregates. A disaggregation of the population into generations in a Social Accounting Matrix would be helpful. That will also produce information about

economic problems of population ageing. But only longitudinal statistics, although more complex, will describe the reality of human beings. And only that can prevent us from drawing wrong conclusions from the virtual reality of aggregates.

How Static Statistics Distort Our View

Among social statistics, those on poverty occupy a prominent place. Usually a 'poverty line' (the level of yearly income below which people are considered to be poor) is defined and the number of individuals below that line is counted. This aggregate can be broken down, *e.g.* by age, gender, type of household, type and level of income, immigrants by country of origin. Timeseries show the number of poor people in successive years. What reality do they suggest?

Timeseries of numbers of poor people (as a percentage of the total population) usually show gradual changes, *e.g.* going up during years of economic recession by not more than one percentage point a year, and going down in other years by even smaller percentages. Disaggregations may show some trends, *e.g.* a gradual decrease in the percentage of poor among the elderly or a decreasing proportion of youngsters among the poor. Certainly, this is interesting information that will attract the attention of media and policymakers who will focus on the pro's and cons of policies to influence trends. The data and the interpretation suggest that the overall picture changes slowly and that it can be influenced effectively by policy measures like an extra 0.3% increase of social security benefits for certain groups or a slight increase in income tax thresholds for others. But is that the really relevant reality?

The analysis of a longitudinal database of microdata on household income microdata will lead to statistics describing a far more dynamic reality. It may show that in year t only 10% of the poor (those having an income below the poverty line) were also poor in each of the preceding eight years and that 70% had such an income fewer than half of these years.[12] One might choose to define people as "genuinely poor" in year t if, and only if, they were poor in the years t-3 up to and including t. But in this case only one third of the 'poor' would be defined as 'genuinely poor'. The rationale behind such a definition is that consumption patterns (and thus poverty) depend mainly on long term income, not on yearly income. These longitudinal data suggest a poverty problem of quite another proportion than that which was suggested by the successive cross sections of the preceding paragraph. It shows essential heterogeneity in those aggregates and it reveals considerable poverty dynamics (entry and exit in poverty); both phenomena may suggest that the policy measures mentioned earlier have only limited relevance. It enables one to focus on 'genuinely poor' and on narrowly defined groups of poor people who are most vulnerable because the probability that they stay 'poor' is high.

This example can easily be generalised to income statistics. Successive cross sections of income distributions show relatively stable data. But an analysis of longitudinal microdata reveals that considerable changes in yearly income are widespread and that these changes mainly stem from social dynamics, *e.g.* marriage (cohabitation), divorce, retirement, becoming unemployed or employed, change of job. Only a small proportion of income changes stem from changing wage rates, changing levels of social security benefits or changes in taxation. Again, the essential point is that traditional statistics fail to show the relevant dynamics. They present static snapshots instead of dynamic movies of the processes. The snapshots show relatively stable income distributions. But, these stable distributions are interpreted as a stable reality. And in a stable reality, (relatively small) general changes in wage rates, benefit levels or tax rates are thought to be the main

causes of changes in income. However, in a dynamic reality, where many individuals have considerable income changes, social dynamics and individual careers appear to be the dominant force behind income variability. It is this reality that is experienced at the micro level or, in other words, the reality which is relevant to citizens as well as their policymakers addressing those citizens.

It is reasonable to assume that this can be generalised to many other social statistics, *e.g.* statistics on participation or exclusion, safety or victimisation, labour and use of time, social security and even aspects of health. That leads to the following conclusion.

In traditional societies, social structures were relatively stable, social mobility was relatively low and people's life courses were relatively uniform. This let statisticians aggregate the population into a limited number of fairly homogeneous groups, applying classifying variables like age, gender, marital status, occupation, income. Differences within the groups were relatively small or exceptional, the average group member had much in common with the vast majority of the members of his group. Differences between the groups were much bigger. And so it was possible to describe society adequately using aggregate data on groups. It was reasonable to describe the course of life of individuals by a disaggregation of group data by age.

In modern societies, groups are more heterogeneous. Differences between groups do not overshadow differences within groups. Courses of life are less uniform and show more 'existential changes' than in traditional societies, causing a more dynamic pattern of entry and exit in groups. If a modern society is described by traditional statistics, it will show up as a predominantly traditional society, because most heterogeneity and dynamics stay hidden and the timeseries of aggregates suggest relative stability. This will wrong-foot policymakers and other users of statistics. The statistics may be technically perfect, but they show too virtual a reality in the sense explained in Section "Terminology for analysing the Problem".

Heterogeneity requires statistical analyses of microdata; dynamics need longitudinal data. A relevant statistical description of modern societies requires statistical analyses of longitudinal microdata. This challenges official statistics in two ways:

- They must acquire microdata, develop the skills of linking them at the micro-level (variables from different sources as well as data for different periods, thus producing rich records of longitudinal data at the micro level) and the skills of their analysis and estimation. Section "Flexibility in the Development of New Statistics" elaborates on these points;

- They must develop skills in presenting longitudinal data and other results of these analyses so that they are digestible to users.

The last challenge is easily underestimated. Experience shows that policymakers, who endorse the longitudinal information (on income dynamics) presented to them as an eye-opener, don't go on to use because its presentation is complicated. Statisticians should develop a better combination of creativity and professionalism (statistical as well as subject matter knowledge). Statistical output should be much more than just figures, traditional graphics and explanations of concepts. It should comprise professional presentations of knowledge, illustrated with illuminating moving images. Section "Flexibility in Presentation" elaborates on presentation.

Complex Reality Threatens Simple Measures and Theories

This chapter illustrates how the increasing complexity of modern societies challenges the simplicity of our measures and theories.

Complexity and the Measurement of Inflation

Like statistics on the population and on labour, those on inflation and growth are in the premier league of official statistics. As a rule, the CPI (consumer price index) is interpreted as a cost of living index which implies that price ratios between commodities are considered to reflect marginal utilities, in agreement with current (neo-classical) mainstream economics. This assumption about price ratio's also lays the foundation of the current use of price statistics in the deflation of economic aggregates like GDP and household consumption. The present section goes into the fundamental problem of the measurement of price (and volume) changes in complex and dynamic markets. The next section analyses how the economic theory behind the measurement of inflation and growth is challenged by increasing complexity.

Eleven years ago the Boskin report to the US Senate Finance Committee (Boskin c.s., 1996) shocked the world of official statistics in the US and abroad. The Boskin Commission's report had not been initiated by official statisticians; it had been prompted by the FED and commissioned by the Senate in tough discussions, demonstrating distrust in official statistics. The outcome of the research, a 1.1% upward bias in the US CPI, did not overly surprise policymakers. The Bureau of Labour Statistics (BLS), the producer of the US CPI, accepted important recommendations by the Boskin Commission, criticised a few others and intensified research to remedy the bias. Official statisticians in some other countries reacted defensively, in line with its inherent conservatism as pointed out in section "A Law of Inherent Conservatism in Official Statistics". They suggested that their CPI's suffered less from the problems uncovered by Boskin. But the main problem, that is how to deal with quality changes and new goods, was not easier to solve in other developed economies and its complexity has, if anything, increased.

What is the core problem? If a good in period t+1 is identical to that in t, the price difference between t+1 and t can be considered a pure price change. But if the quality of the good changed between t and t+1, the price difference must be split into two components – one reflecting the change in quality, and he remainder, the pure price change. And if the good no longer exists, because a new good was substituted for it, the price difference between the old and the new good must be split in a similar vein. In both cases, determining the pure price change is difficult. In some cases methods like hedonic estimation can be applied with some success but very often there is no alternative to rough judgements or to ignoring the problem by assuming that the quality differences between t and t+1 are irrelevant. The problem is biggest in dynamic markets, for complex goods and heterogeneous services. To concretely imagine these problems, the reader could reflect a moment on how to measure the price changes of computers (object of very frequent changes in a broad range of technical specifications, including completely new applications), medical care (how much more this year does your general practitioner produce during a consulting-hour than she did ten years ago?) or legal advice (how would you define quantity or volume differences and pure price changes between successive years?). These kinds of problems amounted to a bias of 0.6% a year in the US CPI according to the Boskin report, indicating that statisticians tend to ignore quality improvements or underestimate their magnitude. This amounted to more than half of the

total bias (0.6 out of 1.1%); the rest (0.5%) was a bit easier to combat, *e.g.* by rebasing the CPI more frequently.

Is a CPI bias of 0.6% acceptable? A professional statistician tends to consider any bias unacceptable; in the trade-off between (random) inaccuracy and biasedness she or he accepts the first rather than the second. However, especially in the case of the CPI this professional attitude requires a bit of courage because in many countries external parties with special interests are looking closely to the methods applied. The introduction of a new method which – although better than the old one - is vulnerable to critique, may be resisted by parties who prefer the outcomes of the old method. This strengthens the inherent conservatism of official (price) statistics. On the other hand, the tenths of percentage points of the CPI are very much in the public eye; when inflation rises 0.1% more than expected, share prices react and government budgets suffer substantially. Therefore, it is illogical that a bias of 0.6% would be acceptable. Because public confidence in official price statistics is a top priority, official statistics in many countries should invest in solving this problem before renewed distrust emerges.

The problem increases as markets become more dynamic, goods become more complex and services take a growing share in the basket. It is likely that the bias in our price and volume statistics was much less fifty years ago than ten years ago. The methodology of price and volume measurement has certainly improved during the last fifty years. But, during the first four decades this improvement did not keep pace with the increasing measurement problems. Is today's bias smaller than ten years ago? Problems of complexity and dynamics have increased further, but did official statisticians increase the rate of their improvements? It is recommended that official statistics itself takes the initiative to evaluate the progress of price statistics over the past decade, taking into account (and estimating) the further increased measurement problems. That will be less risky than waiting for another external initiative stemming from mistrust.

Complexity, Consumer Sovereignty and Mainstream Economics

Section "Complexity and the Measurement of Inflation" mentions that the dominant interpretations of price and volume statistics imply that price ratio's are assumed to reflect relative marginal utilities. Focussing on consumption and elaborating on what has been pointed out in Section "How Mainstream Economics is Lagging Behind", the underlying assumptions in mainstream economics are that consumers act autonomously and rationally and have complete information about the market. If markets are free and perfectly competitive, the sovereign consumers steer the supply side to produce the basket of goods that maximises consumers' utility (if we abstract from the complications of aggregation over heterogeneous consumers). If markets are less competitive, consumers will maximise their utility at a lower level. How realistic is this model in a world of complex goods and dynamic markets?

If goods or services are complex, it is unrealistic to assume that consumers have complete information and act autonomously. The consumer of medical or legal services is faced with a 'market' which is far from transparent and he feels highly dependent on the supplier. The case of technologically complex goods is not all that different, especially if technical specifications change rapidly or product-lifecycles are short. It is fair to conclude that for important items in the consumer basket the assumptions of mainstream economics are unrealistic. Have they become increasingly unrealistic? On the one hand, complexity and dynamics increased considerably during the last fifty years, as did the marketing activity and information overload. On the other hand, did market-transparency

increase as well? Are today's consumers more competent in the management of information than fifty years ago?

During the last fifty years market information has been influenced by media and marketing. Media play a more important role in our lives than before television became dominant; news and commercial messages have become more penetrating and reach us in much bigger quantities. Brands are marketed as representations of lifestyles or 'meaning' (Klein 2005). The media introduced the elite in your living-room and you wished to follow the newest trends set by them as shown in marketing messages. If you buy your celebrity's favourite brand, you or your child will become as happy as (you think) the celebrity is. The marketing messages have become effective in stimulating irrationality. They have done less well at creating transparency.

As disposable incomes in modern economies have grown far above subsistence levels and shopping has become a favourite pastime, a growing part of consumers' budgets is spent on impulse purchases. At the same time, spending has become increasingly competitive; 'keeping up with the Joneses next door' has become keeping up with high earners (Schor 1998, p.8 ff.). This suggest that consumer behaviour has become less rational,[13] and so less in agreement with mainstream economics.

How autonomous is today's' consumer? Surrounded by dynamic markets with their complex goods and services, and overloaded with irrational messages and less motivated to behave rationally, she or he is liable to manipulation. And so the supply side invests huge amounts in marketing. The mainstream economic model, assuming autonomous consumers with exogenous preferences, has become too unrealistic. Galbraith put it more radically: "Belief in a market economy in which the consumer is sovereign is one of our most pervasive forms of fraud" (Galbraith 2004, p.26).

A more realistic economic theory must abandon, or at least considerably relax, the assumptions of autonomous and rational consumers with exogenous preferences. It must accept the reality of economic power on the supply side; the power to manipulate consumers' preferences. And it must adopt an economic subject who, rather than an autonomous and rational individual, has the characteristics of a roused herd-animal. Its analysis must be extended to include endogenous preferences (See Bowles 1998).

Hopefully, this more realistic economic theory will produce guidance on the development of economic statistics describing a more complete reality. Unfortunately, there is little reason to be optimistic that this more realistic economic theory will replace present mainstream economics at short notice, because the latter still is deeply rooted in Academia[14] and in current economic policy making. But official statistics should not remain inactive. A relevant contribution could be the publication of statistical data on the costs of marketing, including advertising, modern sales promotion, remodeling,[15] etc. Official statistics should demonstrate the courage to present these data as indicators of the supply side's economic power and of the manipulation of consumers' preferences.

What Economic Progress?

This chapter concentrates on economic progress while the chapter "What information systems?" will broaden the scope to include non-economic elements. Therefore it is about welfare as defined theoretically in Section "Contagious Conservatism in Statistical Concepts: An Example". First it summarises the arguments, mentioned earlier in this paper, about the misleading signals sent by statistics concerning GDP and consumption

and adds recommendations for neutralising these signals. Second it speculates about an economic theory that will be able to explain the anomaly that GDP growth has not led to a significant increase in welfare.

GDP Growth and Welfare

There are good reasons to assume official statistics to underestimate the growth of the volumes of GDP and consumption in the last decades, because of the positive bias in price statistics, mentioned in Section "Complexity and the Measurement of Inflation", leading to a negative bias in deflated values. This underestimation explains at least part of the so called 'productivity slowdown' which created a mystery that mainstream economics failed to unriddle. On the other hand, in developed economies substantial GDP and consumption growth - as measured by official statistics - did not lead to a substantial growth of welfare, as pointed out in Section "Affluence". This anomaly is the more striking if real growth was higher than measured. But most striking is that this anomaly did not enter the consciousness of mainstream economic policymakers.

The core problem, already mentioned in Section "The Mission of Official Statistics", is that official statistics have presented GDP and related statistics to an audience which interprets these data as dominant indicators of economic progress. Half a century ago this was justifiable, although theoretically incorrect, but nowadays this creates too virtual a reality. Of course, official statisticians cannot be held responsible for every improper interpretation of their data. But society may expect official statistics to initiate debate on so prominent a misinterpretation, and to do its utmost to provide society with more adequate measures of economic welfare.[16] With today's knowledge, we can conclude that official statistics did too little too late. That creates an obligation to develop an adequate strategy to neutralise this contagious conservatism.

Let us start by recalling a selection of the suggestions in earlier parts of this paper:

- Display separately those elements, included in consumption, which represent costs rather than direct sources of welfare (*e.g.* costs of commuting); suggested in Section "Contagious Conservatism in Statistical Concepts: An Example";

- Select and measure elements of economic welfare not included in consumption (*e.g.* pleasure in work); suggested in Section "Contagious Conservatism in Statistical Concepts: An Example";

- Develop longitudinal statistics on the dynamics of income changes, as an antidote to the misleading virtual reality of aggregate data; suggested in Sections "How Aggregation Disorts Our View of Growth (and Progress)" and "How Statics Statistics Distort Our View";

- Publish data on marketing expenditure as an indication of the supply side's economic power and the manipulation of consumers' preferences; suggested in Section "Complexity, Consumer Sovereignty and Mainstream Economics".

The first two suggestions should lead to statistical presentations which stimulate the conceptual debate on growth and welfare and the shortcomings of the concepts of the National Accounts. The last two suggestions should expose misleading interpretations of GDP and consumption as indicators of economic progress.

Two additional suggestions are important:

- Publish estimates of the (subjective) satisfaction derived from income.[17] This will also supplement the second suggestion above. The total result should be developed into a picture of satisfaction derived from each of the sources of happiness as sketched in Section "Towards an Integrated System of Indicators of Happiness";

- Initiate, or get involved in studies which aim to "strike the balance" and publish the results, if necessary without taking full responsibility for these results. The rest of this section is devoted to a discussion of this suggestion.

In Section "The Mission of Official Statistics" the problem of 'striking the balance' is illustrated by referring to changes in the environment or in income inequality. A 'green GDP' is an example of striking the balance between production growth as reflected in GDP and changes in the availability of environmental functions. A green GDP requires an economic valuation of these environmental changes; it supplements the produced sources of welfare (GDP) with a category of non-produced sources (environmental functions) which have become scarce and have got an economic value (being zero only in the case of 'free goods' which are abundantly available). A theoretically sound concept of a green GDP, compatible with mainstream economics, has been developed by Hueting (Verbruggen *et al.,* 2001). This 'Sustainable National Income' (SNI) is an estimate of the maximum production level at which, with the technology of the year of reporting, vital environmental functions will remain available for future generations. The changes over time in the gap between GDP and SNI give information about whether society is approaching or drifting further away from environmental sustainability. In the Netherlands the gap increased in the period 1990-2000 by about € 10 billion (MNP 2006), being 2 % of GDP. Estimates of SNI, which arrive at about 50% of GDP, are regularly carried out in cooperation with Statistics Netherlands, which takes no responsibility for the published results. The reasons for not publishing SNI as a regular product of official statistics include:

- The estimation requires very strong assumptions in the application of a general equilibrium model, which are dictated by the requirement of attaining a sustainable burdening of the environment (such as by greenhouse gases);

- The sustainable burden estimates - the so called sustainability standards – are disputable and official statistics obviously cannot take responsibility for them.

But the involvement of official statistics in developing and estimating the SNI is a recommended element in the above mentioned strategy.

In principle, there are as many ways of 'striking the balance' as there are sources of welfare besides GDP. If we could strike all of these balances, a comprehensive indicator of welfare might come in sight. Unfortunately, not all concepts are as sound as SNI. One may hope that the recent wave of research on the sources of happiness (see Section "Towards an Integrated System of Indicators of Happiness") will advance the development of those concepts and their measurement. However, happiness is a broader concept than welfare because it includes non-economic elements. In an integrated system of statistics of happiness and its determinants, the concept of progress will be defined broader than welfare. If that research is successful, 'striking the balance' between GDP and other sources of welfare change may become redundant.

As will be explained in Section "Towards an Integrated System of Indicators of Happiness", we cannot expect that the SNI-project will become redundant; it is difficult to see how happiness research would be able to tackle the multidimensional environmental problem in the same way as it tackles income, unemployment, job insecurity, pleasure in work, health, etc.

Economic Power and Welfare, Including Social Relations, Acknowledgement and Self-expression

In spite of the anomaly that the growth of GDP (and consumption) did not lead to a significant increase in welfare, GDP remains the dominant indicator of economic progress in developed economies. Mainstream economic theory, which still provides the foundation of most economic policy making, postulates that autonomous, rational and well-informed consumers maximise their utility. According to this theory and contrary to the just mentioned anomaly, consumption growth *must* lead to increased satisfaction. Sections "How Mainstream Economics is Lagging Behind" and "Complexity, Consumer Soveriegnty and Mainstream Economics" critically examined this economic theory, leading to the conclusion that mainstream economics is seriously unrealistic and outdated. A more realistic theory would pay attention to irrationality and to the manipulation of consumers' preferences through marketing activities by the supply side. A more realistic theory would be able to explain the anomaly.

Let us speculate a moment about an economic theory which is more realistic than mainstream economics. We assume that consumers have preferences that can be manipulated and that producers have the power to manipulate. We further assume that consumers' preferences are in agreement with the hierarchical model of Maslov (Maslov 1954). This implies that (1) food and drink are the first necessities, then come (2) shelter, safety and work, then (3) love and social relations, then (4) acknowledgement and appreciation and finally (5) creativity and self-expression.[18]

We consider two stages of economic development. At the first stage, incomes from full-time work are just enough for food, drink and shelter; government provides safety; other needs are satisfied outside the market, but only at an elementary level as the need to work full time leaves little room. At the second stage, incomes far exceed subsistence levels. What developments can be expected to occur between the first and the second stage?

At the first stage, marketing budgets are low. Only basic goods and services are produced, markets are transparent and marketing activities are limited to attracting attention and presenting elementary information. Consumers only buy what they need. Their preferences are hardly susceptible to manipulation; they consume food because they are hungry but they do not have the money to spend on more – or more expensive – food.

As incomes grow, consumers will gradually shift attention to the higher needs in Maslov's hierarchy. As love, acknowledgement, self-expression, etc. are not for sale, they try to reduce their working hours in order to spend more time on the higher needs. The supply side will react with at least two strategies:

- Stimulate consumers to buy more of the 'traditional goods' (food, drink, shelter). More luxury variants of these goods are produced and continually renewed, tastes are manipulated, *e.g.* by suggesting that social relations improve by serving the latest luxury drinks, that social acknowledgement benefits from living in a luxury home,

etc. Lifestyle-marketing develops in combination with 'keeping up with the Joneses' (see Section "Complexity, Consumer Soveriegnty and Mainstream Economics"). Among the results are: increased consumption of food, drink and shelter; and a less than proportional increase in welfare because of diminishing marginal utility of food, drink and shelter and especially because the Joneses are as successful as other consumers. And there is more obesity.

- Produce new goods intended to directly satisfy higher needs. At the first stage these goods were unmarketable, because consumers had no money to spend on anything but subsistence. After some time, the new goods are marketed in the same way as the traditional ones which results in a further increased pressure to spend. Although more is invested in the satisfaction of the higher needs, which results in higher welfare, it is doubtful that welfare will increase overall. Negative factors are: the law of diminishing marginal utility, the fact that competitive spending substantially reduces or even undoes the direct positive effects of consumption and the fact that consumption of new goods and services may be an ineffective investment in the satisfaction of higher needs, far less effective than spending time (with one's family for example).

The manipulations of consumers' preferences, resulting in a high pressure to spend, stimulate consumers to go on working full time. Their social relations, self-expression and creativity do not always benefit from the possession of new goods (like the British, Americans with long working weeks spend 3.5 hours a day looking at television (Layard 2005, p.86). They feel under pressure, because the pressure to spend money prevents them from spending time and attention on the higher needs as Maslov's theory indicates. In that way, we arrive at the second stage.

At the second stage, the volume of consumption is much higher, consumers work just about full time, the most competitive consumers work even more than those at the first stage. Productivity and wages increased considerably. The direct opportunity costs of spending time on higher needs increased and the pressures of competitive spending urge workers to work full time. Producers spend huge amounts on marketing. Consumers are not significantly happier (1) because their spending of money and time did not follow the optimal path dictated by their authentic preferences according to Maslov; (2) because of diminishing marginal utility (and the disappointments which follow from the illusions created by lifestyle marketing); (3) because competition in spending reduces increasing consumption to a psychological zero sum game.

Do you recognise in this description of the second stage the anomaly mentioned above? An anomaly inconsistent with mainstream economics. This section suggests that economists should try to develop a more realistic economic theory, assuming rational suppliers who have the power of marketing and less rational consumers whose preferences can be manipulated. And it underlines the recommendation, mentioned at the end of Section "Complexity, Consumer Soveriegnty and Mainstream Economics", that official statistics should produce estimates of marketing expenditure as an indication of economic power and of the manipulation of consumers' preferences.

What Information Systems?

In a strictly economic analysis, it is obvious that the concept of progress boils down to welfare. But why should policymakers[19] confine themselves to economic progress? In

a broader assessment of progress, evidence based policy making requires a combination of economic and non-economic indicators. This chapter explores the requirements concerning the statistical evidence needed.

Indicators or Integrated Systems?

This section concentrates on different kinds of statistics – indicators, frameworks of indicators and integrated systems. An indicator is a timeseries of a variable, usually a target variable for a certain policy. In a framework of indicators a set of variables is selected and relations between these variables are specified conceptually, but not quantified. In an integrated system these relations are quantified.

In a complex society, many target variables exist. In creating indicator systems, it turns out that a limited number of variables has to be selected from a long list. The selection process implies judgements of relative relevance and its result may be disputed earlier or later. As the indicators are timeseries, it is assumed that selected items will keep their prominent place on the societal agenda over time. In an increasingly dynamic and unpredictable society, this is uncertain. When the indicator shows a change for the worse, policymakers may want to take action. But, because the indicator is an isolated variable not connected to other variables, the information it offers will be of limited help in deciding what action to take. The policymaker will need more information about the causes of the undesired development and its relations to variables she or he can directly influence (policy instruments) before deciding what to do. Furthermore, she or he will want to be sure that the action will not cause some unintended consequence elsewhere, which will show up as regress in other indicators (competing policy goals). This information is not available in a set of indicators where mutual relations are not quantified.

This rather pessimistic description of how indicators work in the policy process[20] is not meant to conclude that the indicator approach has to be avoided. Often, there is no good alternative available in the short run. But, for evidence-based decision making in modern societies, the development of systems of isolated indicators can only be considered a first step. The two core problems are (1) that the inevitable selection of a limited set of indicators yields a fragmented picture and (2) that the lack of connection between the indicators reduces the function of the indicator to 'signal only'. The only flexibility in the system is to change the selection of indicators, but that introduces discontinuity.

Frameworks of indicators specify conceptual relations between the indicators. Because these relations are not quantified, they offer little remedy for the 'signal only' functionality. But the framework may stimulate a balanced selection of indicators because it will be based on some coherent vision. A further step, based on such a vision, is the development of composite indicators. These are combinations of separate indicators, each indicating an aspect of some broad concern (*e.g.* social well-being), into a weighted average that indicates some total (*e.g.* total social well-being). The weights result from (disputable) judgement. I am not aware of any serious and systematic use of composite indicators by decision makers (others than those who produce the data).[21]

The selection of indicators as well as the choice of their weights in composite indicators requires judgement and that judgement can be disputed. Official statisticians are not well-placed to make these choices on their own authority. Consequently, the authority of the indicators will depend on the authority and consistency of the parties

involved in the selection process. In modern societies, where many relevant developments emerge unforeseen and politics is less and less synonymous with authority, this may be a disadvantage.

The practice of economic research and policy making has shown that integrated systems, like the National Accounts and like demographic and socio-economic[22] systems, are in many respects superior to indicators. Their main features are that they describe the whole process, not just a fragmented picture, and that they quantify relations between the variables in the system, thus their variables are not 'signal only'. And so, they offer solutions to the two core problems in the indicator approach. And they are more flexible. If the user wants to look at some new variable, it is often possible to rearrange the basic data in the integrated system so that the variable can be derived, even creating a timeseries back in time. The National Accounts have illustrated that such systems can become very authoritative, last for decades, and serve a great variety of policy making and research.

There is an intriguing difference between the relationship indicators and integrated systems have to policy making. Usually, indicators are selected target variables of policy making. They directly address items on the societal (or political) agenda and, thus, they have to follow changes in the agenda. That makes them vulnerable to the vagaries in the process of policy making in modern societies. Integrated systems are designed to describe complete processes. Variables in those systems are defined according to the logic of the system. They can be items on the societal agenda (but other variables can be policy instruments) but if they are deleted from that agenda, they do not loose their function in the system. So, they are less vulnerable. Are they also less relevant? The system of National Accounts has shown that durable relevance is not necessarily connected directly to target variables of special policies (items on the agenda) but to more general information needs (durable demands behind the agenda). Not seldom, new items were placed on the agenda because information from the National Accounts stimulated policy making to do so. The superior statistics are not agenda following, but agenda setting.

What is the main challenge to official statistics? It is to serve decision makers by producing indicators of good quality on (durable) items on their agendas. But even more challenging is to develop systems which produce statistics satisfying durable information needs shaping present and future agenda's. Evidence-based decision making, facing competing policy goals and complex interdependencies, needs flexible integrated statistical information systems.

Two strategies in the future development of integrated statistical systems are especially relevant. The first is the development of integrated systems connected to the National Accounts. In this paper several projects have been mentioned or recommended: NAMEA ("The Mission of Official Statistics", Endnote 3), generational dimension in social accounting matrix ("How Aggregation Distorts our View of Growth (and Progress)"), SNI ("GDP Growth and Welfare"), integrated demographic and socio-economic systems (this section). New variables recommended in the paper, like costs (rather than sources of welfare) included in GDP ("Contagious Conservatism in Statistical Concepts: An Example") and marketing expenditure ("Complexity, Consumer Soveriegnty and Mainstream Economics"), should be estimated as (functional) specifications of National Accounting variables. The main strength of this strategy lies in the authority[23] of the National Accounts and related systems and the richness of the information on interrelations between phenomena. The main weakness is that it is a system of aggregates, presenting static or comparatively static information.[24] Another

weakness is that the system has lost much of its flexibility, as was pointed out in Section "A Law of Inherent Conservatism in Official Statistics". That flexibility should be revived.[25]

The second strategy is the development of integrated systems of dynamic information on the basis of statistical analyses of databases of longitudinal microdata. This strategy is indispensable because official statistics has to describe the relevant realities of heterogeneous, dynamic and complex societies. An extremely promising development is sketched in Section "Towards an Integrated System of Indicators of Happiness".

Towards an Integrated System of Indicators of Happiness

Richard Layard wrote an extremely relevant analysis of recent research about the determinants of human happiness (Layard 2005). He mentions that neuroscience ascertains that subjects are capable of giving valid information about their feelings of happiness, contrary to the postulates of behaviourism which so deeply influenced economics. Mainstream economics, influenced by behaviourism, turns a blind eye to happiness and economic welfare (see Sections "How Mainstream Statistics is Laggin Behind" and "Contagious Conservatism in Statistical Concepts: An Example") and concentrates on observable behaviour; consumers' preferences are supposed to be 'revealed' by their behaviour. That is why mainstream economics cannot accept the fact that consumption has increased without an increase in welfare and sticks to the virtual reality of its propositions. Therefore, the position taken by Layard could be an essential step to a new ruling paradigm in economics.[26] However, his analysis of happiness goes beyond a purely economic analysis of welfare. It confirms that scientific creativity may be stimulated by crossing borderlines.

Layard distinguishes seven main determinants of happiness: family relations, financial situation, work, social environment, health, personal freedom and philosophy of life. He summarises research in developed economies. The financial situation appears to be relatively unimportant; if income decreases by one third, happiness decreases by 0.2 on a ten point scale. More important are family relations (divorce decreases happiness by 0.5), work and health (unemployment decreases happiness by 0.6 as does deteriorating self-reported health by one point on a five point scale). Even job insecurity, statistics showing increasing unemployment, having never been married, the quality of government, or belief in God have more substantial effects on happiness than a one third decrease in income (Layard 2005, p. 64). Of course, Layard's suggestions to economic policymakers differ from those of mainstream economics. In his book, Layard not explicitly presents suggestions for statistical policies. But what recommendations to official statistics could be derived?

Official statistics should produce longitudinal microdatabases including survey data on happiness, pleasure in work etc. The records should contain a good deal of information on the facts in the history of the subjects life, and include subjective data on health, insecurity (*e.g.* about the job, the social environment) etc. Multivariate analyses should investigate what aspects effect happiness and produce information as mentioned in the above paragraph. This, together with the more straightforward indicators of happiness itself, would constitute a framework of subjective indicators on the seven main determinants of happiness. This should be supplemented with a framework of objective indicators on those determinants, *e.g.* the incidence of divorce, work, unemployment etc. Will these frameworks suffer from the disadvantages of indicator frameworks, mentioned in Section "Indicators or Integrated Systems"?

The first of the two core problems of (frameworks of) indicators is the disputable selection of indicators. In the above system, however, this selection can be done empirically: those elements are selected which turn out to substantially influence happiness. The second core problem is that the relations between the indicators and policy instruments are not quantified. But in the above framework, the multivariate analysis provides essential quantification. It provides information on the quantitative impact on happiness from those factors described by the objective indicators. If policymakers want to influence these factors, they have an indication of the impact of each on happiness. In addition, official statistics should apply (or develop) integrated systems in which the relations between the objective indicators are quantified, so that the policymaker can be provided with a more complete picture of the effects of her policies. Imagine how such a system could inform society when policymakers decide to create more flexibility in labour markets. Evidence would be available on the positive effects on happiness caused by the expected decrease in unemployment, as well as on the negative effects on happiness from an increase in job insecurity directly created by those policies.

To predict that this (to a high degree integrated) system of happiness indicators can develop into a system as authoritative as (during the past fifty years) the National Accounts, would be a speculation. But there are good prospects and official statistics should not miss the opportunity.

On one point this system will not succeed. That concerns 'striking the balance' between income (growth) and (deterioration of) the physical environment. It seems highly improbable that survey interviewing can produce the relevant information (with the exception of information on how happiness is effected by publicity about environmental deterioration). The physical environment has many different dimensions; effects of environmental changes may manifest themselves in future, indirectly or without people understanding the cause. It is unlikely that valid survey questions on the effects of environmental issues on happiness can be developed. Moreover, multivariate analysis will fail also because all individuals – or the vast majority – are subject to the same environmental changes at the same moment. For these reasons, it is unlikely that the physical environment will become an operational element which can be added as number eight to the seven determinants of happiness in Layard's system. That is why in the chapter "What economic progress?" the estimation of the SNI is recommended as a sound concept for 'striking the balance' through economic valuation. It is unrealistic to expect direct 'valuation' of environmental functions in terms of happiness to become feasible.

How to Increase Flexibility?

A flexible statistical system is a prerequisite for permanently keeping up with reality. Flexibility is relevant to existing statistics; the development of new statistics; drawing up statistical work programmes; and presenting statistical information.

Flexibility in existing statistics concerns all adaptations which do not change the basic concepts of the output. They include major adaptations of observation techniques, estimation methods, and frequency and minor adaptations of output specifications. These forms of optimisation and adjustment to evolving realities are important types of statistical innovation and their flexibility is enhanced by modern ICT, statistical professionalism and organisational flexibility in statistical institutes. A lot of research in official statistical institutes belongs to this category. If the disruptive effect on timeseries, caused by these changes, is carefully isolated, they will be helpful. Maintaining this type

of flexibility is hardly subject to the law of inherent or contagious conservatism in statistics.

New statistics may be needed when reality changes more fundamentally, *e.g.* in cases of emerging new phenomena, blurring social structures or changing processes in society. In the identification of the need for new statistics, two stages are relevant: (1) the identification of developments in reality, leading to new information needs and (2) the assessment of the feasibility to satisfy these needs by statistical information. These stages are included in the drawing up of statistical work programmes and will be discussed in Section "Flexibility in Innovating Statistical Work Programmes". Here, all elements of the law of inherent conservatism are effective, including the contagious workings. In the development of new statistics, discussed in Section "Flexibility in the Development of New Statistics", only one element of this law is at work.

Flexibility in presentation concerns the heterogeneity of the group of (potential) users of official statistics as well as the increased complexity of (the statistical description of) reality. It refers to the Statistics, Knowledge, Policy chain as well as to the broadening of the clientele of official statistics in modern 'information societies'. This will be briefly discussed in Section "Flexibility in Presentation".

Flexibility in the Development of New Statistics

The traditional process of developing a new large scale survey starts with an often rather detailed specification of the output. In a next stage, focus shifts to the survey process; how to create a process that yields valid data and is efficient for both respondents and statisticians. This leads to the development of completely specified questionnaires, including experiments to validate questionnaires and processes. The first large scale survey follows. If all goes to plan, that first survey leads to a publication of acceptable results within half a year or so.

This first publication comes often years after the development started with the lag depending on the complexity of the statistics, the length of the survey period and the number of parties involved in the development, etc. The production period can be shortened by creating well documented modular systems for computer assisted surveying, especially if the surveys are being held continuously. Introducing new variables in current surveys can be organised more quickly. But even then a substantial preparation period will be required in order to avoid invalid or otherwise 'dirty' information or to prevent disruption to the current survey process, especially in the case of brand new variables.

After the first publication, not all the information needs are satisfied. Many users need timeseries because they are more interested in changes over time than in the levels of variables. The use of timeseries for the estimation of regressions or other models requires rather long series. Hence, the users who want to analyse the relevant changes in society, including their interdependencies, will often have to wait many years for the information they need.

In a society that has grown more complex and more dynamic, a quick response to new information needs has become increasingly important. If new information needs could be satisfied by estimations based on already available raw data, quick responses would be feasible. If the raw data is available for a series of years, users needing new timeseries could be served promptly. Official statistical institutes should develop strategies to shift the focus of the statistical production processes from a unique production line for each category of statistical output to a flexible use of all available raw

data in the production of statistical output, including unforeseen statistics. Important elements of such strategies are:

- Collect as much microdata as you can get (and manage); give priority to register data because that is relatively cheap and because it covers complete populations;
- Store the data centrally in a way that easy access is guaranteed, but keep that simple; standardise the documentation and harmonise the units so that the data can be linked at the micro level;
- Let your organisation develop from a collection of unique production lines to separate units for data capture and for analysis and publication; have all captured data made available in the central storage.

This strategy will lead to an increasing number of occasions where completely new statistical information, including timeseries back in time, can be produced out of existing stocks of raw data. The stock of microdata can be used to quickly develop longitudinal information.

The culture of statistical work shifts from the cultivation of the special characteristics of existing statistics, dedicated to maximal satisfaction of determined information needs, to the quick production of statistical information which can satisfy urgent needs to a reasonable degree, using all available data. The culture of data capture shifts from special data, preferably directly surveyed by the statistical institute, even if in small samples, to available register-data, preferably concerning all units of the population in order to fill the records with a maximum of information through microdata linking. The culture of the estimation process shifts from the analysis of sampling errors to the subject matter oriented analysis of non-sampling errors.

The availability of register data to official statistical institutes differs strongly between countries. But there is a general trend of lowering barriers, stimulated by the developments in modern ICT, improved techniques of data protection and spreading practices of microdata linking for a range of purposes, including the reduction of administrative burdens. Official statistical institutes, being specialists in data protection, must find ways to energetically overcome still existing barriers. International organisations can support them. The strongest argument they have is the fact that statistics, by its very nature, is not about identifying individuals, so that privacy issues are not at stake as in other uses of register data. This argument turned out to be decisive in the eyes of the Dutch data protection authority.

Shifting to quickly produce statistical information which satisfies urgent needs to a reasonable degree, using only available data, implies adopting elements from the culture of statistical work in the compilation of integrated systems like the National Accounts. The strategy, summarised above in the points I – III, will considerably increase the flexibility of those systems. And it enables the quick production of small *ad hoc* systems of integrated information, including longitudinal information, that satisfy new information needs outside the scope of existing statistical systems.

Flexibility in Innovating Statistical Work Programmes

Section "How Mainstream Economics is Lagging Behind" concluded that official statistics must become more entrepreneurial to mitigate conservatism. Statisticians must

widen their outlook from existing clients to potential users and must open-mindedly observe developments in reality which, up to now, are beyond their reach. This implies that they have to become more outward looking and more pro-active. Is this easy?

The workings of the law on inherent and contagious conservatism are very strong. Moreover, much of these workings must not be disrupted, because that could lead to a decrease in the authority of official statistics. A general policy asking the average statistician to become more distant to those existing statistics she or he is responsible for, will be dangerous and, very likely, ineffective. A more effective and less risky strategy is to develop stimuli to innovative programming through separate activities which do not directly interfere with the existing statistics.

The most important activity is strategic research and a statistical institute should permanently devote a not too small part (2%?) of its budget to this. The research could be organised as a bundle of projects, each for a limited period; after that period the budget is moved to new strategic projects. Because the top of the institute is responsible for its strategy, the top decides on which projects are selected (not the authority that decides the statistical work programme). International cooperation in the planning and execution of projects will considerably increase the effectiveness of the strategic research. Cooperation with Academia and other researchers on a national scale may stimulate and accelerate the research and help tackle contagious conservatism.

In the selection of strategic research projects, top priority is given to discovering new concepts, needed to fill major gaps in the description of relevant realities. The research starts with the identification of trends that determine potential information needs and it may end with presenting a prototype of the statistical information that could be developed. This can be illustrated with a few concrete examples, suggested elsewhere in the paper:

- Which dynamic processes are shaping the development of society and what statistical information can describe them? Projects could tackle labour, income and poverty, care, social exclusion, social capital; productivity, globalisation, manipulation of preferences;

- Which developments with (potential) users of statistics (including policy, business, economic and social sciences, citizens) need new statistical information or new tools for presentation;

- Towards comprehensive information about the development of happiness and its relation to essential developments in society;

- Towards the SNI for 'striking the balance' between GDP growth and environmental deterioration, recommended in "GDP Growth and Welfare".

Flexibility in Presentation

In modern societies information plays an essential role for policymakers, other decision makers, scientists and citizens. Many actors produce information leading to information overload and hectic publicity. Official statistics must stand out by its undisputed information available to everybody. This paper's analysis implies two challenges to the presentation of statistical information: (1) how to present statistical simulations of an increasingly complex reality; (2) how to present the information to an

increasingly heterogeneous (potential) clientele. Ignoring other elements of statistical dissemination policies, this section will concentrate on flexibility as an element of a strategy to meet these challenges.

Every description of reality implies a reduction of its complexity and statistical simulations are no exception. As reality grows more complex, more statistical simulations of the same reality may be considered relevant. If, at the same time, the potential clientele of these statistics grows more heterogeneous, it is probable that more different statistical presentations of the same reality will be relevant. Therefore, flexibility in the presentation of statistical information must increase. This concerns the content of the presentation as well as the dissemination policy.

Section "How Statistics Distort Our View" concluded that the challenge of presenting longitudinal information, which is so essential to our view of modern reality, is easily underestimated. And it added that statistical output has to move from a simple use of figures, traditional graphics and explanations of concepts to professional presentations of knowledge, illustrated with moving images. This conclusion had policymakers in mind; they are not only important users of statistical information, but disseminators as well. Statistical information is the more useful for them if they can use it in addressing citizens. One of the elementary implications is that statistical presentations must not be unnecessarily high brow, a lesson successfully learnt during the last decades in the relations between official statistics and the press.

Longitudinal statistics are not unique in being difficult to present. Also other new sets of statistics, recommended in this paper in order to neutralise misleading signals sent by GDP, will require a careful presentation in a context of scientific, subject matter and statistical knowledge. Official statistical institutes should consider to appoint spokespeople for broad area's (socio-demographic, socio-economic, macro-economic, and business) who combine high level knowledge with the gift of convincingly present complex information in simple words and images.

One of the prerequisites of effective presentation is the ability to take the point of view of the audience. Therefore, the management of official statistical institutes should stimulate its professionals to be outward looking and to ask the questions about reality which are relevant to (potential) users. Strengthening the 'Statistics, Knowledge, Policy chain' requires:

- Looking for the relevant blind spots in what potential users know about reality, as recommended earlier in this paper, as well as seeking support for filling these gaps in knowledge. This challenges the communication skills of subject matter specialists as well as spokespeople and the (top) management itself;

- Providing users with tools they can use in presentations to their audience.

An encouraging perspective emerges from developments in statistical websites, where users are able to compile statistical information according to their special needs. As the interactive and flexible tools become more user friendly, the heterogeneous statistical clientele will be more effectively served. Especially the users who are not represented by institutions with a direct influence on the statistical work programme. To improve the statistical service to these less powerful users is an important task for official statistics in modern democracies.

A Dilemma in the Mission of Official Statistics

Official statistical institutes have to provide society with undisputed information. That is why they avoid publishing statistics of low quality. Relatively inaccurate statistical information is only published if consensus exists that the information is by far the best available and it may be accompanied by some sensitivity analysis. In most countries, official statistical institutes do not publish forecasts or statistical analyses based on disputable assumptions. Statistical information that 'strikes the balance' between GDP growth and *e.g.* environmental deterioration is not produced because it requires disputable, or even controversial, assumptions.

Section "A Law of Inherent Conservatism in Official Statistics" argued that guarding the authority of statistics leads to avoiding disputes about existing statistics and to a reluctance to develop competing or conflicting statistics. There is wisdom in the conservatism of the official statistical system. It is often said that 'trust comes on foot and goes on horseback' and guarding the trust in official statistics is a prime task. Confidence in official statistics depends on confidence in the institute that produces the statistics. At the same time, confidence in the institute depends on confidence in its statistics. If the production of official statistics is centralised, the central statistical institute has many opportunities to demonstrate its impartiality, professionalism, relevance and authority all of which could raise confidence. But if one of its statistics runs into problems, the whole system is at risk. Let us call the collection of risks, mentioned in this chapter up to now, risk A.

On the other hand, the preceding chapters have illustrated another risk. The risk that traditional statistics, having been authoritative for decades, are considered misleading because it appears that the reality they correctly describe, is too virtual. This is risk B. In an increasingly complex society, that risk necessarily increases because information on a growing amount of processes is reduced (or neglected) while the traditional statistics are maintained. An effective management of this risk will lead statistical institutes to:

- Introduce new statistics to neutralise the one-sidedness of current statistics by describing other aspects of the same reality;
- Develop statistics presenting a more complete (thus more complex) picture of reality, so that the described reality runs a lower risk of, sooner or later, being considered to be too virtual.

Both strategies may produce information that will be considered competing or even conflicting to traditional statistics. And the second strategy may lead to more disputable statistics, *e.g.* when 'striking the balance' between different phenomena. In other words, both strategies increase risk A. The second strategy abates risk B more fundamentally, but it also may further increase risk A.

How to manoeuvre between Scylla and Charibdis? Mapping the risks will help, permanently looking out is indispensable and a good stratagem (*e.g.* sending a scout *in cognito*) will increase courage. And for the last element, the image of the statistical institute is essential.

If the statistical institute is respected only for its regular production of authoritative statistics, the dilemma is serious. But if the institute is also respected for its scientific research, it has more latitude. The development of alternative statistics can be introduced in a context of experimental scientific research in order to lower risk A. Discussions

about their strengths and weaknesses can be started in an academic context with the same result. This enables statistical institutes to introduce alternative statistics gradually and to gain support for them before putting the authority of existing statistics at risk.

If the statistical institute is also respected for its innovative attitude, it has even more latitude. The alternative statistics can be introduced with more self-confidence. This enables statisticians to act more entrepreneurially and to communicate more convincingly with users about the strengths and weaknesses of existing and alternative statistics. Because the public expects the institute to innovate - and appreciates its ability to discover new developments in reality requiring new statistical descriptions - risk A can be managed more easily. The introduction in the statistical work programme of new information that is more open to dispute than the traditional statistics will not immediately challenge confidence in the institute. It will not, therefore, put the authority of all other statistics at risk.

This leads to the conclusion that official statistics, in order to increase its ability to 'keep up with reality', must develop the scientific attitude, creativity, courage and external communication skills, needed to create a high quality innovative organisation. This includes challenging categories of users entangled in virtual realities by showing them blind spots in their view of reality. It may include initiating controversial debates about the current interpretation of GDP growth and unrealistic elements in mainstream economics. In the long run, these disputes are needed for keeping official statistics undisputed.

Recommendations

This paper's recommendations are summarised under three headings: statistical information, statistical policies, and the context of official statistics. Numbers between brackets refer to places in the paper where the main arguments are mentioned.

Recommendations for Statistical Information

- Because modern societies are heterogeneous and dynamic, develop statistical information describing the dynamic processes at the micro level (Sections "How Aggregation Distorts Our View of Growth (and Progress)", "How Static Statistics Distort Our View"). Develop integrated systems of dynamic information based on the analysis of longitudinal microdata (Sections "Indicators or Integrated Systems?", "Towards an Integrated System of Indicators of Happiness");

- Because GDP is a misleading indicator of economic progress in developed societies, develop integrated statistical information on happiness, including all sources of economic welfare not included in GDP or consumption ("Towards an Integrated System of Indicators of Happiness"); stimulate debate on the interpretation of GDP growth by publishing data on costs (rather than sources of welfare) included in GDP or consumption ("Contagious Conservatism in Statistical Concepts: An Example"); initiate and get actively involved in the development of estimates of Sustainable National Income ("GDP Growth and Welfare");

- Publish data on marketing expenditure as an indication of economic power and the manipulation of consumers' preferences and stimulate debate on the unrealistic

premises of mainstream economics ("Complexity, Consumer Sovereignty and Mainstream Economics", "Economic Power and Welfare, including Social Relations, Acknowledgement and Self-Expression");

- Enrich the National Accounts with integrated socio-demographic and socio-economic information ("Indicators or Integrated Systems?") and with a generational breakdown of the social accounting matrix ("How Aggregations Distorts Our View of Growth (and Progress)").

Recommendations for Statistical Policy

- Create an outward looking culture in statistical institutes so that future statistics answer the relevant questions about reality ("Flexibility in Innovating Statistical Work Programmes", "Flexibility in Presentation");

- Develop the scientific attitudes, creativity, courage and external communication skills, needed to create a high quality innovative organisation; challenge categories of users entangled in virtual realities by showing them blind spots in their view of reality; initiate debate around the current interpretation of GDP growth and unrealistic elements in mainstream economics ("A Dilemma in the Mission of Official Statistics");

- Earmark 2% of the total budget of official statistics for strategic research projects ("Flexibility in Innovating Statistical Work Programmes");

- Increase the flexibility of the statistical production process by creating longitudinal databases of integrated microdata, available for the production of unforeseen statistical information ("Flexibility in the Development of New Statistics");

- Initiate an evaluation of the last decade's progress in price statistics against the background of the further increased measurement problems ("Complexity and the Measurement of Inflation");

- Develop the skills and tools needed to present complex information in a context of subject matter knowledge as well as to flexibly serve a more heterogeneous clientele ("Flexibility in Presentation").

Recommendations for the Context of Official Statistics

- Supporters should stimulate official statistics to develop and actively demonstrate vigilance over new developments in reality which require new statistical information or which increase the virtuality of existing information (Summary, "Why Statistical Conservatism is Contagious (and why official Statistics itself must Produce the Antidote)", "A Dilemma in the Mission of Official Statistics");

- Economists should abandon or considerably relax the most unrealistic assumptions of mainstream economics, concerning the autonomous and rational behaviour of consumers, and develop a new paradigm in which economic welfare as a state of mind, and economic power and endogenous preferences play a realistic role

("Complexity, Consumer Sovereignty and Mainstream Economics", "Economic Power and Welfare, including Social Relations, Acknowledgement and Self-Expression", "Towards an Integrated System of Indicators of Happiness").

Notes

1 It is true that much room for heterodoxy exists in the economic and social sciences. But at the level of individual scientists or scientific organisations orthodoxy dominates in the dedication to a special doctrine. Contemporary mainstream economics, which dominates current policy making, is a clear example.

2 The OECD's global project on Measuring the Progress of Societies is studying the concept of "progress" more deeply and concretely. The reader asking for the meaning of the words "relevance" and "progress" in this paper will find answers in the following: A development in reality may be perceived as relevant. This implies that the development itself is perceived and that it is felt as being relevant. Like all perceptions, this perception is personal. You may call it subjective if you think subjects to be autonomous. But as far as human beings are influenced (manipulated) by media or other simulators of reality, subjectivity will be replaced by collective (manipulated) perception and judgement. From this point we can deduce a meaning of the word "progress".
A subject can perceive a development (simulate its reality) and judge its relevance. He (or she) can also feel that the development can be characterized as "progress". He will consider the development as progress if, in his perception, the development brings reality closer to an ideal (or utopia) and she will consider it as regress if that distance is increased. In a society consisting of autonomous individuals, "progress" can have as many meanings as there are individuals. But if society consists of a limited number of groups of individuals who share common ideals and perceptions, the number of meanings of "progress" is limited. A further limitation may come from the media and other simulators of reality which influence perceptions as well as ideals. This paper will argue that official statistics, being a simulator of reality, must play a role in (initiating) debates about questions like: what (new) developments are "relevant" to our notion of reality? and what developments are relevant to our notion of "progress".

3 This concerns e.g. the NAMEA (National Accounting Matrix including Environmental Accounts) or Hybrid Flow Accounts as they are called in a UN Handbook (UN, 2003, p. 129 ff.)

4 In the section "Flexibility in the Development of New Statistics" this process is more concretely illustrated.

5 The paper assumes official statistics to be independent, and able to develop statistical concepts free from politicisation. But it is clear that it cannot ignore the law or an administrative use of statistics by powerful institutions. Chapter "A Dilemma in the Mission of Official Statistic" suggests a way out.

6 This is sometimes in contrast to the hectic world of day to day politics, where the media increasingly ask for quick responses to hypes. This world may be considered less relevant to statistical work programs because it uses statistics merely ad hoc.

7 An example of special interest is the paper (Giovannini and Uysal, 2006) written for an OECD Workshop in the preparation of the Second World Forum on "Measuring the Progress of Societies".

8 Although Nobel Prize winners stressed market failure (Stiglitz) or presented alternative theories (Kahneman).

9 Examples are given in chapter "Inadvertent deformation of reality"

10 The definitions are based on Metha who presented an elegant elaboration of these concepts (Metha, 1962, p. 69-74).

11 Another example is that it is often argued that growth is a prerequisite for technological progress. It is reasonable to assume e.g. that growth of individual successful firms or industry groups stimulates technological progress. In other words: dynamics is the relevant factor. But this does not imply that growth on the macro level is a prerequisite. This undermines the second pro argument in the second paragraph above.

12 The data, mentioned in this section, are derived from a publication by Dutch official statistics which unfortunately is only available in Dutch (SCP/CBS, 2005)

13 A more elaborate analysis concludes that only in exceptional markets the conditions for rational consumer behaviour are fulfilled. See Poiesz, 2004.

14 It might be useful to take stock of the data needs of behavioural economics and evolutionary economics.

15 Remodelling of durables does two things, especially in a context of competitive spending. It adds to the utility of new buyers who prefer the newest model. But at the same time it decreases the utility of all former buyers. Therefore, frequent remodelling is a marketing strategy. From a viewpoint of welfare, GDP counts the pluses but ignores the minuses.

16 Official statistics should try to do this even if the improper interpretation seems to be based upon a dominant ideology, supported by mainstream economics. This is why Galbraith wrote: "The more than minimal fraud is in measuring social progress all but exclusively by the volume of producer influenced production, the increase in the GDP." (Galbraith, 2004, p. 28) Or why Schor wrote: "One problem with the national discourse is its focus on market exchanges, not quality of life, or social health. Gross domestic product is the god to which we pray." (Schor, 1998, p. 21).

17 Layard mentions the research on the 'preference drift' done by Van Praag and Frijters and on what is sometimes called the 'reference drift' (which quantifies the loss in satisfaction caused by an increase in income of the reference group) by Blanchflower and Oswald, by Clark and Oswald and by Stutzer. (Layard, 2005, p. 46-49)

18 Maslov names his five categories: (1) physiological; (2) safety; (3) love/belonging; (4) esteem; (5) self-actualisation.

19 The word "policymakers" includes agents of government as well as agents of semi-governmental or independent organisations, and ngo's and commercial organisations taking decisions with a significant impact on society.

20 See Van Tuinen (1995) for a more elaborate discussion of indicators versus integrated systems including concrete examples of policy making using indicators.

21 One special category of indicator frameworks are those presentational tools which combine separate indicators to present a birds eye view of a complex development. A successful example is the Business Cycle Tracer. The BCT combines 15 short term economic indicators in a presentation which enables you to assess the stage of the business cycle at a glance (Van Ruth et al., 2005). This kind of indicator frameworks differs from the frameworks mentioned in this paper. It is not a policy-related selection of target variables or a weighted combination which defines the composite variable (e.g. total social well-being). Rather is it an analytical selection of variables which all indicate the same phenomenon (e.g. the business cycle). Could that be a reason why they are more successful than the above mentioned composite indicators?

22 For the strengths of integrated socio-economic systems, see Van Tuinen, Altena and Imbens (1994) concentrating on labour and education, and Van der Laan and Van Tuinen (1997) concentrating on income.

23 It is assumed that the authority of the system, threatened by sending misleading signals about economic progress, will be maintained by applying the recommendations of this paper.

24 A minor weakness, in comparison with (frameworks of) indicators, is that they do not integrate qualitative information like that on consumer or producer confidence.

25 In the take-off of the last revision of the System of National Accounts proposals were made to increase flexibility in the SNA (Van Bochove and Van Tuinen, 1986)

26 The new paradigm should, unlike mainstream economics, explicitly pay attention to the psychological category 'welfare', theoretically defined in "Contagious conservatism in statistical concepts: an example", as well as to the manipulation of preferences as suggested in "Economic power and welfare, including social relations, acknowledgement and self-expression".

References

Boskin, M.J., E.R. Dulberger, R.J. Gordon, Z. Griliches, D.W. Jorgenson (1996) Towards a More Accurate Measure of the Cost of Living, Final Report to the Senate Committee from the Advisory Commission To Study The Consumer Price Index.

Bowles, S. (1998) Endogenous Preferences: The Cultural Consequences of Markets and other Economic Institutions, Journal of Economic Literature, XXXVI, p. 75-111.

Easterlin, R. (2001) Explaining Happiness, Proceedings of the National Academy of Sciences, 100: 19, p. 11176-11183.

EU (1996) Council Regulation 2223/96, June 25th, 1996.

Galbraith, J.K. (2004) The Economics of Innocent Fraud, Penguin, London.

Giovannini, Enrico and Ayhan Uysal (2006) Statistics, Knowledge and Policy: What Do We Know About What People Know? OECD Workshop on Business and Consumer Tendency Surveys.

Klein, Naomi (2005) No Logo, Harper Perennial, London.

Layard, Richard (2005) Happiness, Penguin, London.

Maslov, A.H. (1954) Motivation and personality, Harper and Row, New York.

Metha, J.K. (1962) A philosophical interpretation of economics, London.

MNP (2006), Milieu en Natuur Planbureau, Milieubalans 2006.

Poiezs, Th.B.C. (2004) The Free Market Illusion, Psychological Limitations of Consumer Choice, Tijdschrift voor Economie en Management, Vol. XLIX, 2.

Schor, J.B. (1998) The Overspent American, Why we want what we don't need, Basic Books, New York.

SCP/CBS (2005) Armoedemonitor 2005, Sociaal en Cultureel Planbureau, Centraal Bureau voor de Statistiek, Den Haag.

Van der Laan, P. and H.K. van Tuinen (1997) Increasing the Relevance of Income Statistics; Experiences and Plans in the Netherlands, Papers and Final Report of the First Meeting on Household Income Statistics, Canberra, Australia, 2-4 December 1996, Expert Group on Household Income Statistics, p. 193-220. Belconnen: Australian Bureau of Statistics.

Van Ruth, Floris, Barry Schouten and Roberto Wekker (2005) The Statistics Netherlands' Business Cycle Tracer. Methodological aspects; concept, cycle computation and indicator selection, Statistics Netherlands.

Van Bochove, C.A. van, and H.K. van Tuinen, 1986, Flexibility in the next SNA: the case for an institutional core, Review of Income and Wealth, Series 32, Number 2.

Van Tuinen, Henk K. (1995) Social indicators, social surveys and the integration of social statistics; Strengths, weaknesses and future developments of the main approaches in social statistics, Statistical Journal of the United Nations Economic Commission for Europe, 12 (3/4) p. 379-394.

Van Tuinen, Henk K., Jan Willem Altena and Hanneke C.M. Imbens (1994) Surveys, registers and integration in social statistics, Statistical Journal of the United Nations Economic Commission for Europe, 11 (4) p. 321-356.

Verbruggen, H., R.B. Dellink, R. Gerlagh, M. Hofkes, H.M.A. Jansen (2001) Alternative calculations of a sustainable national income for the Netherlands according to Hueting. In: E.C. van Ierland *et al*. (eds), Economic Growth and Valuation of the Environment, A Debate, Edward Elgar, Cheltenham.

UN (2003) Integrated Environmental and Economic Accounting 2003, Final draft circulated for information prior to official editing; Studies in Methods, Handbook of National Accounting; United Nations, European Commission, International Monetary Fund, Organisation for Economic Cooperation and Development, World Bank.

Chapter 18
Challenges of Data Collection: with Special Regard to Developing Countries

Asad Elahi
Secretary, Statistics Division, Pakistan

Abstract

This paper declares that the quality of statistical data collection in developing countries has to improve in order to aid decision making at the national and international levels, such as tracking progress toward the Millennium Development Goals (MDGs). But in almost all relevant issues of data collection, developing countries face a different set of problems compared with industrialised ones. The author begins by outlining the problems faced by the National Statistical Offices in developing countries, both external (in the society at large) and internal (within the statistical offices). The author then describes the extensive institutional and structural changes the government of Pakistan is undertaking to improve data collection. A list of recommendations, both general and specific in nature, are provided to help developing countries meet the Paris21 Consortium's purpose: greater demand, availability and use of better statistics and statistical analysis in national, international and civil-society decision making.

Data collection here is understood as all activities involving collecting data from respondents or from secondary sources for the purpose of producing official statistics. Data collection may or may not coincide with data entry and data editing and checking. Data collection may be done traditionally "in the field" by enumerators or interviewers or through other media like mail, telephone or internet .Data collection in the developing countries is generally undertaken by the state institutions because of lack of capacity in the private sector. Again in most of these countries the statistical structure is largely federal. In Pakistan, though there are Provincial Bureaus of Statistics but they are weak in terms of infrastructure and capacity to be able to play an effective role in data collection. Therefore, most of the data collection work is handled by the Federal Statistics Division trough its three departments, *i.e.* Federal Bureau of Statistics (FBS), Agricultural Census Organisation (ACO), and Population Census Organisation (PCO). The Central Bank and some Federal Government Ministries too have statistical cells though with limited capacity.

In almost all relevant issues of data collection developing countries face different sets of problems as compared to industrialised countries. These difficulties in most countries are:

- Endogenous lying within official statistics itself on grounds of weaknesses of institutional and organisational set-up, of lack of resources and infrastructure or operational inadequacies;

- Exogenous lying outside of official statistics perceived as inability of the civil society, establishments or even of the administration to part reliable and timely data.

Weaknesses external to the Statistical Offices can be generally attributed to:

1. Level of literacy and lack of appreciation of importance of data collection as a national activity mixed with general cultural ethos of not sharing personal information.
2. Law and order situation in some areas rendering these areas inaccessible for data collection.
3. Cultural constraints in rural/tribal societies in not parting information about gender related issues or sharing information with outsiders.
4. Remote and thinly populated areas pose logistical as well as technical problems for example application of area sampling.
5. Infrastructural deficiencies i.e. insufficient postal services, inadequate transport, road networks, telephone and internet facilities.
6. Large informal economy in subsistence economies make statistical meassurement complicated.
7. Even in formal economy especially in the Small and Medium Enterprises (SMEs), low standard of bookkeeping, inadequate administrative coverage and low standards of social performance. Thus, there are few formalised phenomena statistical observation can refer to.
8. Lack of political will and administrative support discourages development of a vibrant and autonomous statistical structure which tends to support the perception that the production of statistics is a task of minor importance. Lack of credibility and transparency of data create the feeling that the governments tend to utilise statistics for window-dressing of its policy failures. Consequently it leads to resistance to cooperate with the statistical offices
9. Socio-cultural inhibitions towards giving out personal information affect filling of questionnaire by the civil society. Business and industrial class in some economies avoid sharing data for fear of being sued for tax fraud. As a result, there is a low propensity to deliver data for statistics.
10. Developing countries are rapidly changing their structural and administrative set-up. Consequently, scope, items and concepts of statistics and last but not the least their questionnaires have to be adjusted more frequently. This hampers to get civil society acquainted with carrying a certain statistical burden.
11. Quality of enumerators and quality control/vigilance of enumerators need improvement to ensure that the enumerators and interviewers are applying their instructions appropriately
12. For a lot of variables the variance within a developing country is much larger than in an industrialised one. Prices pertaining to agricultural commodities differ a lot between the various provinces and districts due to lack of an organised marketing

chain system and due to the fact that many of these commodities are neither graded nor branded. Their prices are volatile. The baskets of price indices in industrialised countries are much more determined by branded goods of industrial production which makes collection of prices and compiling price indices much easier. The same might be true for the statistical observation of wages and salaries or of private households consumption.

13. The administrative density in industrial countries normally is very high. This enables these countries to collect and compile statistics very efficiently out of secondary sources such as taxation, customs, movements of ships, building licenses and the like. Statistics out of secondary data may replace costly surveys or may at least allow for additional benchmark estimates. In developing countries the option of using secondary sources is limited.

14. There is no control of the National Statistical Office over the statisticians working in statistical cells of other Ministries/ organisations nor is a Statistical service like the one in UK which does not provide a coordinated network.

Weaknesses internal to the statistical offices can be identified as:

1. Statistical offices do not attract quality human resources as in most countries official statistics is perceived as a government task of minor rank. As a consequence the salaries as well as the career planning and motivation of staff are low. Capacity building including technology up-gradation, modern training facilities and linking performance with reward is accorded a low priority. Training is hampered by a lot of fluctuation of staff.

2. Infrastructure for internal information and communication in many countries is poor. The main frame application, for example, is still the prevailing form of data processing. Micro computers and client-server systems have in some countries just started to intrude into the performance of work. Decentralised data entry is not yet the norm. Access to the internet as well as internal networks often have still to be developed.

3. Allocations of financial resources for statistics organisations are inadequate.

4. Information on the institutional setup of the economy lack comprehensive and up-to-date directories. This hampers drawing frames and applying elaborated techniques as, for example, stratified sampling.

5. Statistical organisations lack autonomy while the legal framework is outdated 6to meet requirements of a modern statistical structure.

6. There is absence of Strategic/Corporate and operational planning/management on modern lines.

What can be done to overcome the mentioned problems and weaknesses? Of course, it is easier to amend or mitigate the difficulties we face from internal weaknesses as – in contrast to the external weaknesses – a lot of measures are easier to undertake. In my view the institutional setup of official statistics is a key issue to internal improvements. In Pakistan, the existing three organisations are being merged into an autonomous legal entity .The organisation will be managed on the top by a professional management team led by a Chief Statistician and five members of international repute, steered by a Governing Council comprising of representatives of government and private sectors and supported by a Users' Council. Administrative and a degree of financial autonomy will be

given under the law. All stakeholders will be involved in planning and execution in an organised and formalised way. A new and modern statistical law is being finalised which clearly defines the tenure, the tasks and the role of the Chief Statistician and the five executive Members as well as the duties of the Governing Council, the Supervisory body, and the Users' Council. This will enhance the autonomy of official statistics and its management vis-à-vis the government. Issues like the utilisation of secondary sources or the prerequisites to provide micro data in the form of public and scientific use files are also being tackled within the concept of greater public access to data and building bridges data users. . With restructuring the revamped and modernised official statistics entity of Pakistan will emerge as a reliable, independent and credible institution in its perception amongst media, civil society and other stakeholders.

It has been possible to initiate these changes due to very strong political commitment from both the Government and the opposition parties. Next year when we would have completed our new office building equipped with modern infrastructure in Islamabad we will have all the three statistical organisations *i.e.* FBS, ACO and PCO under one roof, we expect that the performance of the organisation will also benefit from the fact that it is not scattered over various premises any more.

But while at the institutional level we merge and centralise, we go in opposite direction in modernising workflow and organisational performance. With the help of donors we have entered into decentralising data entry in our field organisation. In order to cope with the above mentioned problems of poor infrastructure and illiteracy in remote areas we have strengthened and modernised our regional offices. All of these offices have now been equipped with computers. They now enter the data and simultaneously check them, and they transmit the files via electronic media to our data processing centre. Formerly the questionnaires had been sent as paper sheets. This is a big leap ahead. We started with price statistics and are now expanding decentralised data entry by and by to other surveys, also. This has improved the timeliness and transparency of date. We plan to develop institutional intranet to facilitate real time transmission of data for further processing at our data processing centres. Training in statistical computer techniques is being imparted to all the staff.

All institutional and structural changes and all achievements in the performance of work need human resources which are able to promote this change in a sustainable way. So, capacity building plays a major role in our efforts. In Pakistan with donors support our training institute was rejuvenated and a comprehensive training schedule in modern statistical methodologies and computer techniques has been launched to adequately equip our staff with required qualifications. Training, in my view, is the centre-piece to the multi-dimensional process of modernising and restructuring a statistical organisation. In view of low wages it is difficult to attract top quality youth from the labour market. You normally have to rely on the staff available in your institution and to increase their skills is crucial. We improve and enhance training in two ways. We have established Training Institute as an integral part of the Federal Bureau of Statistics which is providing more and more courses not only for internal staff but for participants of related institutions like statistical offices of the provinces, government departments, State Bank or associations and the like, also. Statisticians are being sent abroad to attend courses which are offered by SIAP of UNESCAP, by the IMF, the Asian Development Bank, the World Bank and many more. A programme has also been launched to send bright youth to Universities abroad for postgraduate education. It is proposed to upgrade our training institute to conduct post-graduate diploma course in Applied Statistics in the near future.

Another object of possible modernisation and improvement is, of course, our statistical programme. As an effective institution we are not only modernising our inputs – this is the traditional way of steering public entities – but also have to look at our outputs or, better to say, our outcome. For a statistical office this means meeting expanding data requirements of a rapidly growing economy which are up-to-date, timely and reliable in a transparent manner according to international standards and methodologies. The period of rebasing of indices and national accounts is also being reduced by undertaking regular surveys in major segments of economy. Regular annual surveys are being conducted on socio-economic sectors for monitoring implementation of MDGs. For professional standardisation and credibility, our survey reports are reviewed and validated by international experts in addition to monitoring and review by technical committees comprising of experts. In the spirit of promoting public-private partnership a joint report on Health Indicators has been compiled in collaboration with an NGO.

Pakistan has a Federal form of Government and our policymakers require data in deep regional structure. But the provincial Statistical Bureaus are not in a position to entertain deep and broad regional statistics. It is the federal statistical organisations who are filling this gap .This has the advantage that the regional statistics are harmonised and reconciled with the national results. However important plans are under implementation to strengthen their capacity. In Pakistan we meanwhile are improving our Mouza (village)-statistics which means that we provide a predetermined set of variables for each village in same content and same format. Thus, statistics supports processes of regional development and planning for equitable distribution of socio-economic benefits and allocating resources for investment in regional infrastructure. It also supports political processes of devolution.

Plans are being prepared to tackle the problems with outdated frames and lack of directories, installing a statistical business register for Pakistan through public-private partnership. This register will be established through coordination of data entries of different state departments and thus try to unify the various directories which are available to us. We put them in a special format and intend to update this register regularly and within short time. This means we exploit all secondary sources and we involve a lot of different parties of the public as well as of the private sector in our attempts to improve things. We will open the results of the register for public use as for as the data of the register are classified as non-confidential. All this will hopefully create some perception of "ownership" with the concerned ministries, semi-public organisations, associations, chambers of commerce and the like. Statistics should more and more be perceived as a joint product of public and private partners. The problems with no or bad response might then be mitigated. But we have to make sure that involving secondary sources, and especially taxation and customs authorities, is a one way exercise not allowing confidential statistical data to be sent to the fiscal bodies. Dealing with secondary fiscal sources, of course, is a sensitive issue. We are also undertaking the exercise of setting up a modern health accounting system.

So far, ladies and gentlemen, I have tackled some Pakistan-related projects and plans. I confess that I am proud of them as I and my team have launched and promoted most of these initiatives. I am deeply convinced that we are marching in the right direction. But I am also convinced that a lot of developing countries need to undertake similar efforts. So, I hope that our way to move forward is of interest for you, also. Let me finally mention some more general aspects, not so much related to my country alone.

Some ten years ago the then new government of Tony Blair in the UK published a so-called White Paper "Statistics: a matter of trust". Trustworthy, credible statistics, ladies and gentlemen, is most of all an issue for the civil society of a country. We should strive for autonomy of official statistics. This would require intensified contacts to the media. This would help to improve the public perception of statistics and its producers as such. This might also contribute to enrich a moderate and facts-oriented public discourse of political issues and to promote modern thinking as, for example, in gender issues. Effects in same direction have been generated in my country by guest lectures by me and other staff of statistics at various forums including universities in order to involve the society at large in the process of data collection on vital issues.

Not only a matter of trust but also a matter of mutual benefit are involving of stakeholders of the public as well as of the private sector into the planning and the performance and the results of official statistics are in my view a must especially for developing countries. In Pakistan we have various technical committees and users councils. These groups incorporate a lot of institutions and persons and may at first glance be deemed oversised. But they are necessary to get the consent and for some crucial decision even the approval of various parts of the government of the provinces, of statistics-related institutes and organisations and – last but not the least – of the civil society as a whole.

Modernisation of statistical offices, of their infrastructure, their performance and their output primarily means to have more funds. Additional funds may accrue from their own economic activities, from government funds or from grants of international development aid. Activities such as selling publications and the like do not have capacity to generate significant resources. I understand that this is also the case for statistical offices in developed countries. And solicitation of public funds for official statistics is difficult in industrialised countries as well as in developing ones. Unfortunately all over the world, official statistics suffer from disregard and from lack of funds.

For the developing countries grants from donor organisations can be a way out. However grants from donors vary based on their own financial and investment projects which are related to their own objectives which may not be according to the specific requirements of an national statistical office Moreover, it is not an easy job to acquire grants from donors and it is even a tougher job to promote coordination between them.. However in my view support to statistics is a good proposal to countries which are willing to enter into development aid or which are willing to extend it. Good official statistics by their very nature have to be neutral vis-à-vis politics. Statistics contribute to transparency, they enable analysis and thus they contribute to good governance.

Last but not the least, I should mention that especially for developing countries it is important to strive to achieve internationally agreed standards concepts, methods and classifications. Even if they did not give many inputs due to scarce resources they, nevertheless, harvest what a lot of countries who are better off have already contributed. But my recommendation for developing countries is to give inputs to international agreements, also. The benefit would be a mutual one as given inputs ensures participation and, most important, learning from others. Pakistan is aware of the importance of international commitments and comparability. It is more and more committed to fulfill or to approach international standards of statistics as, for example, the System of National Accounts. Pakistan will increasingly fill its place on the international landscape of statistics. For any country in the world, the contacts to international bodies and the

participation in their statistical activities is essential for steering the performance, the harmonisation and the modernisation of official statistics.

Ladies and gentlemen, when it comes to international cooperation in statistics, we also have to consider declarations and ideas which have been outlined under the umbrella of the MDG – commitments. The Paris 21 Consortium's goal is to develop a culture of evidence-based policy making and implementation which serves to improve governance and government effectiveness in reducing poverty and achieving the Millennium Development Goals (MDGs). Its purpose is: greater demand, availability, and use of better statistics and statistical analysis in national, international, and civil society decision making. This initiative shows that the international community is aware that we have to improve official statistics in the developing countries and that official statistics is an important factor for achieving the MDGs.

Before I conclude, let me suggest the following for improving the quality of statistical data collection in the developing world:

- Capacity building of statisticians is accorded highest priority by strengthening and coordinating with both national and interregional training centres. Pakistan can share its experiences with the regional countries apart from offering its training institute as a regional training centre. A statistical capacity building fund need to beS created under the auspices of UNSD to fund these projects;

- Under the Marakkish Declaration, countries are required to undertake strategic plans for modernising their National Statistical Offices. However there is a long way to go in this direction. International support is required to provide know-how and guidance in this regard. Paris21, ESCAP and UNSD need to play an active role by undertaking seminars and symposia's at international and regional levels in this regard;

- Developing countries need to be involved actively in the various technical groups on international statistical standards in order to bridge the current gap in implementation of these standards between the developed and the underdeveloped groups of countries;

- International financial institutions should rely on statistical data generated by national statistical offices rather than indulging in self- imputation of data to develop capacities of these institutions. To meet the cherished goals there is a need for appraisal both nationally and internationally. Strong data-base in each country can help in this regard through effective monitoring and appraisal. International community must help in building strong National Statistical Offices (NSOs) to achieve these objectives;

- Analysis of data collected in surveys/census is vital for policymakers and executive arms of the Governments for policy formulation but this is generally not undertaken. Donors need to give due importance to the issue;

- A Statistical Service is created at the national level of statisticians working in all ministries/organisations controlled by the National Statistical Office. This will help in creating better career prospects;

- Due encouragement be given to the private sector in data collection and surveys to expand data base. Standards are specified for surveys/studies to qualify to be graded as national statistics;

- Autonomy be accorded to the NSOs and the Chief Statistician should be provided an assured tenure to discharge his functions independently.

Ladies and Gentleman, we are citizens of a fast shrinking global village, colourful flowers of one beautiful bouquet of human race. Dissemination of knowledge would lead to better understanding and appreciation and thereby create harmonisation in the currently divisive international environment leading to a better world. Statistical facts are the best source of evidence-based truth and knowledge. We must all strive to promote collection of reliable and timely evidence based knowledge for the betterment of humanity.

Part Seven
Measuring Progress: Economic, Social and Environmental Issues

Chapter 19
An Effective Global Statistical System: A Prerequisite for Consistency in Global Measurements

Dennis Trewin
Statistical Consultant, Australia

Abstract

A pre-requisite for consistent global measurement is a strong global statistical system. A lot of improvements have been made in recent years and there are signs of increasing collaboration. However, there is plenty of scope for improvement and the paper make a number of suggestions on to further improve the global statistical system.

Introduction

What is the global statistical system? In this paper, I am referring to the various international organisations that are providers of statistics at both the global and regional levels. This would include those organisations with regional only responsibilities such as Eurostat and the Africa Development Bank. I also include organisations such as OECD whose coverage is global although only for a sub-set of countries. Some have a 'chief statistician', others do not. The global statistical system also includes the various mechanisms for binding the contributing organisations. These include:

- The annual meetings of the UN Statistical Commission, the peak organisation of the global statistical system;
- The meetings of the Committee for the Co-ordination of Statistical Activities (CCSA) which has been meeting two or three times per year;
- International statistical standards; and
- The principles governing international statistical activities.

Why is the global statistical system important? In the next section of the paper, I will explain in more detail why the global statistical system is important and growing in importance. International statistics are increasingly being used in a variety of settings. This World Forum provides a good example. I have not done any analysis but I expect at least half of the presentations in this Forum will use statistics produced through the global

statistical system. To be useful, international statistics must be relevant, of good quality and consistent across countries and across time.

Section "What has been the main advancement in the Global Statistical System in recent years" discussed the many advances in international statistics in recent years. There have certainly been improvements but there is scope to do much more. Section "What is the current state of play with the Global Statistical System" considers the current state of play – both the strengths and weaknesses.

The last section is the crux of the paper. It makes a number of suggestions for a way forward. I have tried to separate by degree of difficulty. The outcome being sought is better international statistics. Better national statistics are a pre-condition for this to occur and the global statistical system also has a responsibility for improving national statistics. And there is a need for improvements in how the global statistical system delivers capacity building effort if it is to deliver on this outcome.

What are my credentials for speaking on such an important topic. My main credential is having been Australian Statistician (Head of Australia's statistical agency) for nearly 7 years and being an active participant in many international statistical activities during this time (eg attendance at many UN Statistical Commission meetings). I have also been President of the International Statistical Institute, Chairman of the ESCAP Committee of Statistics, member of the Bureau of the Conference of European Statisticians, and Chairman of the Global Executive Board of the International Comparison Program). I also spent two months working as an Adviser (on National Statistical Systems) to the UN Statistics Division earlier this year. This provided a different perspective – from inside the global statistical system if you like.

The Importance of Global Statistical System

More and more people are reliant on international statistics. This came through strongly at a High Level Forum on the "(United Nations) Statistical Commission and the Global Statistical System – The Way Forward" held on 26 February 2007. Several of the papers are on the United Nations Statistics Division web site (unstats.un.org).

International agencies are among the key users. Jose Ocampo, the Under-Secretary-General for the Department of Economic and Social Affairs at UN spoke about the importance of statistics as a consequence of the UN Development Agenda. They were needed to target programmes as well as monitoring progress. He spoke about the importance of improving the quality and scope of international statistics. Many developing countries are lacking important indicators. Furthermore some sectors of statistics (eg service industries, informal sector) are not well represented in international statistics.

Francois Bourgignon, Chief Economist of the World Bank, also pleaded for improved coverage of many important statistical indicators. He asked for the delivery of improved and timely statistics that are relevant to policy. He felt that policy development was in some cases restricted by the lack of relevant statistics.

The private sector are also important users of international statistics particularly those that operate globally. Stephen Roach, Chief Economist of Morgan Stanley, spoke at the same Forum. He said that statistics were in some ways the 'glue' that held his organisation together. But he said that in ways the financial system had outstripped the

utility of international and national statistical systems. They were not always relevant to their uses or sufficiently timely.

National governments are also very important users and the level of interest seems to be growing. In particular, comparisons with other countries can be very useful for evaluating the efficiency of existing policy and identifying areas of potential improvement.

Closely related to these uses are the interests of the research community (in academic institutes, NGOs and elsewhere). Many are involved in debate on international and national policies and having access to a good evidence base is fundamental to such research.

And in an increasingly global world, the general community is showing increasing interest in international statistics. The internet has made access to these statistics far easier.

Of course it is not possible to provide high quality, relevant global measurements without strong national statistical systems. For this reason, and others, the global statistical system has a responsibility to develop national statistical systems especially those from developing countries.

What has been the Main Advancement in the Global Statistical System in Recent Years?

The United Nations Statistical Commission (UNSC) is the apex entity of the global statistical system. It has been functioning since 1947 and, in its early years, it was extremely influential on the formation of national statistical systems. There was a period during its life when it was less influential but this has been redressed to a large extent in recent years.

UNSC now meets annually. Meetings every 2 years were not sufficiently timely to influence many international statistical issues. Attendance at UNSC meetings has increased remarkably. This applies to both member countries and international agencies. Furthermore UNSC has had a major influence on a number of important international activities. Examples include the upgrading of the quality of statistics in the Human Development Report; the management arrangements for the 2005 round of the International Comparison Program; reviews of fields of statistics such as service industries, energy and education; and the latest revision of the System of the National Accounts.

But despite these improvements there is scope for further improvements in the operation of UNSC. At present it doesn't have a mandate over many of the international statistical agencies except by co-operation. This includes some of the major players such as the specialised UN agencies (eg ILO, WHO, UNESCO), OECD, Eurostat and the Bretton Woods agencies (ie IMF, World Bank). Most are co-operative but some are not so co-operative. Whilst some organisations will always be outside the remit of UNSC it is not the case for the UN specialised agencies for example, especially if there is a genuine move towards 'one UN'.

The UNSC has been reasonably successful in the development of standards for economic statistics. The System of National Accounts is probably the most notable example but there have also been important developments in trade and industry

classifications for example. Progress on the development of standards for social statistics has been much less advanced.

A major development in 'standards' was the endorsement of the Fundamental Principles of Official Statistics in 1995. These have been very influential particularly among the transition and developing countries. They have provided an excellent point of reference for how National Statistical Offices should operate. This has been invaluable in discussions with Governments. There have been many improvements in national statistics legislation over the last 10 years and the Fundamental Principles have been quite influential in these changes.

There has been a rapid increase in the development of international data bases provided by the international agencies. And they are increasingly being used by both international agencies and countries. The advent of the web has made this data much more accessible. The source data for these data bases mainly comes from national statistical agencies. This can be a considerable burden even though countries are one of the main beneficiaries of these data bases. There was quite a bit of duplication but increased co-operation among the international agencies has largely reduced this duplication.

A particularly pleasing development is the extensive co-operation between the international agencies on the establishment of the indicator data base for monitoring progress against the Millennium Development Goals. As part of this development, meta data standards have also been developed and the system is showing signs of being a world class data reporting system.

The amount of funding for statistical capacity building has increased substantially – so has the in 'kind' effort. This applies to both international and national donors. Whilst this appears to be good news, it is questionable whether the results have justified the increased effort. Too many statistical systems are still underdeveloped. There is a real need to understand how to get most value from the statistical capacity building effort. The recent emphasis on national strategic development plans does seem to be a step in the right direction.

A number of important global statistical projects have been conducted with much co-operation among the various stakeholders. More attention has been given to governance issues to increase the likelihood that more projects will be a success. The 2005 round of the International Comparison Program is one good example. Increased national statistical capacity is one important by-product of this Program. The 2010 Population Census round is a forthcoming opportunity to put a global effort into creating comparable international statistics as well as improving the quality of Census taking in countries.

Finally, the recently agreed principles governing international statistical activities are an important advancement on how international statistical agencies should operate. For legal reasons, some international agencies could not formally endorse the Principles but they should have a major influence on the way they operate even if not formally adopted.

The first principle is that high quality international statistics, accessible for all, are a key element of global information systems. Other principles emphasise the importance of impartiality, professional standards and the transparency of concepts, definitions, classifications sources and methods. There are a number of other important principles and if they were followed by all international agencies we would have a strong global statistical system.

What is the Current State of Play with the Global Statistical System?

I first start with the strengths and there are many. In a paper like this it is easy to dwell on the weaknesses and areas of improvement but the strengths are worth mentioning.

The first strength is that most (but not all) the international agencies collaborate and to a much greater extent than what they did previously. One criticism is that a lot of this collaboration is by way of information sharing but genuinely collaborative programmes are starting to emerge. The ECE region took the lead with its joint work programme, in the early 1990's.

Co-ordination of the work of international agencies is important

- To reduce reporting burden on countries;
- To avoid duplicated and wasted effort particularly in statistical capacity building;
- To ensure consistency in the application of statistical standards and concepts across counties;
- To enable the main gaps in international statistics to be identified; and
- To enable countries to know what is happening so that they can plan accordingly.

Another example is the Committee of Experts formed to assist with the co-ordination of effort on the development and enhancement for the data base on MDG indicators. This appears to be effective and there has been a co-ordinated effort to improve this data base, and make it world class. Countries are now also participating in meetings of the Committee.

Also, as discussed above, UNSC appears to be increasingly influential now that it meets annually. It is widely regarded as the peak international statistical body even though its authority is limited in some areas. These limitations are discussed below.

One of the key outputs of the global statistical system is international statistical standards. Among other things they are invaluable for compiling internationally comparable data. They also provide a base for regions or countries developing their own standards. Most national standards are consistent with international standards at least at the broadest levels. The global statistical system has become more active in promulgating statistical standards for use by countries but more could be done to assist with implementation. This is also discussed below. Eurostat has been especially active for EU countries. Another means of obtaining internationally comparable data is through standard survey instruments. These are increasingly being considered sometimes for a module within a survey of broader scope. The OECD's PISA study is one well-known example. The development of three international standards is discussed in the paper by Wallmann and Evinger being presented at the same meeting.

The UNSC has commenced reviews of particular fields of statistics (eg services, social, energy, industry, education) and these have been helpful in identifying areas where lack of co-ordination is a problem. But it is probably fair to say there has been little effective change put in place since these reviews were considered.

The amount of effort put into statistical capacity building has increased substantially. The challenge is to ensure that this money is well spent and that the enhanced capacity is sustained and that they can make an effective contribution to national and international statistics.

I will now briefly consider the weaknesses of the global statisstical system before making specific suggestions for improvement in the next section.

Although international data bases have become more prevalent, their quality is often of concern. (The OECD quality reviews were an important initiative for improving the quality of data). In 2005, UNSC created a Friends of a Chair group to address concerns with Millennium Development Goal Indicators. I was Chair of that Group. They found some real problems with the quality of the data -

> "As a rough guide, meaningful analysis of an indicator at the global level would require that at least two thirds of countries (where the indicator is relevant) should be able to report at least two measures of it during the time period covered (1990-2005) in order to get an idea of trends. Only 33% of indicators satisfy the criteria".

It is also found that a significant amount of data was imputed resulting in estimates that might be heavily criticised by countries. The assumptions behind the imputation were often unknown. The Friends of the Chair called for greater transparency in imputation methods and the use of country data wherever possible.

UNSC has no influence over the UN specialised agencies, or a number of other important international statistical agencies, except by co-operation. Taking a country perspective, where co-operation is imperfect it creates duplication and makes the global statistical system seem fragmented.

The Regional Commissions, with the exception of ECE, have become much weaker in their statistical efforts. In some Commissions, the Statistical Committees have been abandoned. But there have been steps in ECLAC and ECA to strengthen arrangements but driven by member countries. One outcome of the weakening of the statistical work of the Regional Commissions is that the Regional Development Banks have become more influential. Their increased interest in statistical development should be seen as a positive trend.

There is insufficient lateral mobility of statistical staff between the UN agencies (both global and regional agencies). National Statistical Offices know that lateral mobility is essential for staff development – to learn from different work experiences. It is probably the most important way of learning. But it also important for injecting new ideas into organisations and keeping them fresh. UNSD is best positioned to manage greater flexibility in staffing across the UN agencies but this will only work if it has greater influence than at present over appointment decisions.

UN reform proposals point to problems with duplication and lack of cohesion within the UN system. These criticisms also apply to statistics but, because of the influence of UNSC, possibly not to the same extent as other UN fields of endeavour. Although there has been a considerable reduction in the duplication of data collection, some still exists and should be addressed. These collections impose burden on countries – they are less likely to accept this burden if they are of poor standard and/or countries do not see the results of their efforts. Feedback to countries on the results is important. Also some collections are of poor quality and not developed and tested consistent with good statistical practice.

The Committee for the Co-ordination of Statistical Activities (CCSA) is an attempt to improve co-ordination across the global statistical system. Having not attended any of these meetings, I am not well positioned to comment on their efficiency. But many participants are critical of the usefulness of these meetings. Clearly a body of this type needs to exist. The main challenge is to make it effective.

As previously mentioned earlier, standards for social statistics are not well developed. Yet much of the interest of international agencies is in better understanding trends in social issues.

Although there has been important progress on the development of statistical standards, there is increasing concern that political considerations, not just statistical validity, are playing a larger role in country's views on statistical standards. If this becomes more prevalent, it could put the overall integrity of the global statistical system at risk.

There is another problem with international standards. Insufficient attention is paid to their implementation and as a consequence many countries, particularly developing countries, have not introduced the latest standards. Good examples are the System of National Accounts (SNA) and the International Standard Industry Classification (ISIC) where, for many countries, only versions that are now quite out of date have been implemented. There are many training programmes but 'hands on' assistance with implementation is more limited.

The approval process for international statistical standards is expensive and can take some time. Is there a more efficient way of developing international statistical standards? This is difficult work so the focus should be on those areas of statistics where comparisons are of greatest interest.

Despite the considerable effort put into statistical capacity building, it is still low for many countries. Furthermore there is a real lack of co-ordination within countries. In a Pilot reporting exercise in sub-Sahara Africa, PARIS21 found a number of co-ordination problems within countries. Even the value of technical assistance was not easily identified. Furthermore, the collaboration between bilateral donor agencies and their statistical offices was not well co-ordinated. As a lead agency, Eurostat has been taking the 'co-ordinator' role in many Eastern European countries and believe it has led to significant improvement. Perhaps it is an approach that might be modelled elsewhere. Lack of co-ordination of resources leads to inefficiency in the overall allocation of capacity building funds (eg a single country might be allocated too much or too little of the resources available for development assistance).

A Way Forward

Here I will focus on major initiatives only. However, they should be part of an overall strategic plan for improving the global statistical system a prerequisite for the advancement of international statistics. Isolated, disjointed initiatives are not likely to have much long term impact. A strategic plan should also include the ideal organisational arrangements.

I wonder if statisticians have the full range of expertise required to develop a strategic plan of such complexity. Maybe it should be assigned to a management consultancy group – it is a management problem not a statistical problem that we are trying to solve. But it is important that their work is relevant to the concerns of both the international and

national statistical agencies. One way of doing this might be to establish a consultative group of experts from the countries and the international agencies. UNSC is the body to give endorsement to a strategic plan to improve a global statistical system.

Having said a strategic plan should be developed first, I have chanced my arm at some of the changes that might be made. I would expect that these type of changes would receive consideration in the development of any strategic plan for the global statistical system. Underpinning these proposals is a belief that the fundamental problem is that there needs to be an improvement in co-ordination. This applies at several levels most notably:

- Across the international agencies to provide a coherent global statistical system;
- Across donors and providers of technical assistance within countries.

To my mind these are the main issues that need to be addressed. We have a decentralised global statistical system but that does not mean the stakeholders should not work more closely with each other.

I have separated the recommendations into relatively easy, hard and very hard. Perhaps 'relatively easy' oversimplifies the degree of difficulty.

Relatively Easy

- UNSC should exert its leadership role (through the influence of member countries) in the management and development of international statistical policy and practice; it should rely on its technical strength and expertise and avoid political considerations;
- UNSC should be given final authority for most statistical standards. This does not mean that agencies that have other governance arrangements should not use them as well;
- A federated model should be developed for the dissemination of international statistics. This would mean that individual agencies would retain custodianship and responsibility for their own data bases but they would agree to certain protocols and standards (eg meta data standards for describing the quality and methodology of collections) so that users can search and download data across data bases. It does not mean one physical data base although this may be appropriate for some data bases such as the Millenium Development Goal indicators. This should result in increased user interest, awareness in international data bases as well as improved accessibility;
- Consistent with the previous proposal there should be agreed information standards for the dissemination of international statistics;
- The recent SNA Revision process should be reviewed to identify strengths and weaknesses with the view of developing a model of the process for the development of international statistical standards.

Hard

- If these types of changes to the global statistical system are to be successful, there needs to be a real commitment from all the international agencies to make real change to the system with pressure from countries on those agencies who do not comply;

- The CCSA should be restructured, or its processes revised, to improve its effectiveness;

- There should be a protocol for improved co-ordination among all donors of technical assistance effort within developing countries to ensure greater alignment with national priorities, and that statistical development is sustainable. A key element of the protocol will be to appoint a 'lead agency' to facilitate the co-ordination. The lead agency will vary by country and could be one of the bilateral donors. There continues to be a real need to improve statistical capacity in developing countries; the lack of statistics is often constraining policy development;

- The UN Statistics Division and regional bodies should work together cohesively on a mutually agreed agenda with respect to global statistical issues; unless satisfactory governance arrangements are in place for the regional bodies (such as for ECE) make them regional units of the UN Statistics Division;

- A project management approach along the lines of the International Comparison Program (ICP) should be adopted for the implementation of new/revised statistical standards and other international statistical programmes. A Friends of the Chair Group has been established to review the 2005 round of the ICP. They will no doubt make conclusions about the governance and project management arrangements. In my view they are basically sound. Any problems were due to the way they were implemented rather than the basic model;

- It is necessary to develop ways of engaging more high quality staff within the international statistical system even if only on a temporary basis. This may mean eliminating some of the current bureaucratic restrictions. Consideration should be given to establishing a single recruitment arrangement for professional statistical posts within the UN system. This would most sensibly be managed by UNSD;

- International statistics are becoming increasingly important for both national and international users. The scope and coverage of international statistics should be continually reviewed to ensure they are relevant and timely. This will involve new collections from time to time. Before embarking on such collections, the costs and benefits should be carefully assessed. Both international and national uses and costs should be part of that assessment. As for national statistical offices, lower priority collections should be curtailed from time to time.

Very Hard

Even if only some of the previous proposals are implemented, strong leadership and co-ordination will be necessary. For this purpose a global statistical office should be created (or assigned to a restructured UN Statistics Division) to provide leadership to the

Global Statistical System; a Chief Statistician should be appointed to head the Office. A strengthened Bureau of the UNSC would in effect act as a Governing Board for the global statistical office.

Concluding Remarks

This is quite a long list but the effort will be worthwhile. International statistics will increase in importance – many participants at this Forum will vouch for that. Also, with increasing emphasis on managing for development results, national policy makers and the international development community have become increasingly aware of the essentiality of statistics in informing and designing development interventions. If the international agencies do not take the necessary steps to improve the situation, it may eventually be taken out of their hands.

Improvements also require member countries, individually and through UNSC, to play a leadership role in management and development of international statistical policy and practice.

Notice of correction

A corrigendum has been issued for this chapter.
The entire new chapter (Corrigendum 1) may be found at the end of this file.

Chapter 20
Policy and Statistical Issues Underpinning Financial Stability: The IMF Perspective

Robert W. Edwards[i]
Director, Statistics Department, International Monetary Fund

See corrigendum 1

Abstract

In the past decade, financial globalisation has led to an unprecedented deepening in financial markets and internationalisation in capital allocation. While it has brought benefits, financial globalisation has also brought new risks and challenges for policy makers, particularly through its linking of national economies into a vast network of closely interconnected balance sheets, and consequent potential for severe cross-border spillover effects. In this environment, financial instability in one country may be transmitted to other countries or affect developments in regional or global markets. To meet the new challenges, the IMF has been sharpening the focus of its financial surveillance. While continuing to emphasise its financial sector assessment programme (a joint effort with the World Bank), the IMF has been further integrating finance into its Article IV surveillance; strengthening its global surveillance through multilateral discussions on common issues of systemic importance; and expanding its framework for vulnerability assessments to include the analysis of balance sheet vulnerabilities and use of new data tools such as financial soundness indicators as well as market-based, forward-looking indicators. In the statistical area, the IMF has launched several initiatives in order to gather timely, comprehensive, and internationally comparable data to help meet the new data needs of financial surveillance, better track cross-border flows and positions, and support rigorous, diagnostic, analytical work. In meeting its challenges, the IMF will continue to leverage the use of its comparative advantage - its universal reach and expertise on macroeconomic and financial issues and on statistics.

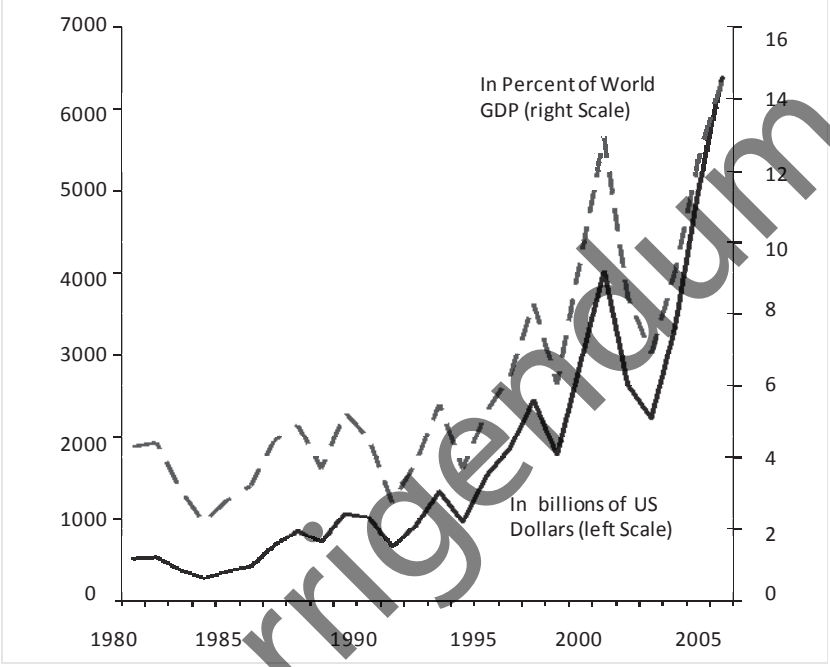
Financial stability in a Globalised World

The last decade has witnessed an unprecedented deepening in financial markets and internationalisation in capital allocation. The rise to $6.4 trillion in total cross-border

i The author wishes to thank the following for their inputs into the drafting of this paper: Marina Moretti and Andrea Maechler of the IMF's Monetary and Capital Markets Department, and Armida San Jose and Ewe-Ghee Lim of the IMF's Statistics Department. The views expressed herein are those of the author and should not be attributed to the IMF, its Executive Board, or its management.

inflows has been impressive – it now represents over 14% of world GDP (Figure 1). These flows have also been associated with the globalisation of financial institutions. Cross-border financial sector mergers and acquisitions, mostly concentrated in the banking sector, are now around 40% of all financial sector acquisitions, compared to less than one% a decade ago. The increase in foreign bank ownership has been particularly rapid in Eastern Europe and Latin America

Figure 20.1 Total Cross-Border flows, 1980-2005
(in % of world GDP and in billions of U.S. dollars)



Source: IMF staff calculations based on International Financial Statistics and World Economic Outlook.

See corrigendum 1

The globalisation of financial institutions has complex implications for financial stability.[1] From the perspective of individual institutions, globalisation helps diversify risks and may well have improved financial stability, particularly in the face of relatively small shocks. But as national economies become part of a vast network of balance sheets, often closely interconnected through the financial sector, severe crises in the future may be more easily transmitted across borders and therefore may have become more broad-ranging and difficult to deal with. For instance, financial systems with substantial foreign bank presence may be more resilient to traditional domestic banking crises, but more vulnerable to foreign shocks that affect parent banks. More broadly, financial instability in one country may easily spill over to other countries or affect developments in regional or global markets, as was the case, for example, during the Asian and Russian crises in the late 1990s.

The global community is thus presented with new challenges. To maximise the benefits of globalisation while containing the potential risks, efforts need to continue to reduce the likelihood of systemic crises and limit their severity if they occur. This requires measures conducive to deep, broad, and resilient financial systems, which can

support an optimal allocation of resources, an effective monetary policy, and a sound fiscal stance. In addition, policy makers need to ensure effectively coordinated oversight of internationalised financial institutions and to put in place cross-border crisis management and resolution arrangements that are sufficiently robust to handle a severe shock and minimise spillovers. Progress is being made in these areas, but more needs to be done, in a wider range of countries.

The International Monetary Fund (IMF) has an important role to play in helping countries meet the challenges of financial globalisation. Interactions between economies and financial sectors are becoming increasingly important, and addressing financial sector vulnerabilities and risks increasingly needs a global approach. Through the continued use of its comparative advantage – that is, its universal reach and expertise on macroeconomic and financial issues, the Fund can further enhance its monitoring of the global economy and help individual countries strengthen the resilience of their economies. In addition, it can use the same comparative advantage in the statistical area to help provide timely, comprehensive, and internationally comparable data to meet the new data needs of financial surveillance, better track cross-border flows and positions, and support rigorous, diagnostic, and analytical work.

The IMF Financial Surveillance Framework

The IMF is sharpening the focus of its financial surveillance to remain in step with a rapidly changing world. Financial surveillance is ultimately rooted in concerns about financial crises that are changing in nature. The late 1990s crises in emerging markets were largely financial account crises caused by abrupt shifts in global asset allocations (that is, crises that affect stocks or balance-sheets) rather than budgetary or current account crises caused by fiscal or terms-of-trade shocks (that is, crises affecting flows). What distinguishes these events from past episodes is the speed and amplitude at which they propagate, both domestically and cross-border. Efforts at preventing these phenomena require new directions in financial surveillance that recognise the complex networks of interconnected balance sheets and emphasise the spillovers of financial market disturbances, both between countries and between real and financial sectors.

New Directions in Financial Surveillance

Given the complexity of balance sheet linkages, there is no single, widely-accepted methodology for assessing financial sector stability. Financial stability assessments need to cover a potentially wide range of topics – depending on country circumstances – and to take a holistic view of the financial system. For this, financial surveillance needs to address two key questions:

- What are the critical channels of interaction between the macroeconomy, financial markets, and the financial institutions?

- What is the role of the financial sector per se in initiating, amplifying, or muting disturbances in the economy and in transmitting such effects internationally?

In addressing the first question, it is important to highlight that the linkage between financial institutions and markets, on the one hand, and macroeconomic performance and policies, on the other, runs in both directions. Financial soundness affects macroeconomic performance and policies, and macroeconomic performance and policies have

See corrigendum 1

consequences for the financial sector. In order to understand this two-way relationship, better use must be made of a broader set of indicators, including data on financial flows, the balance sheets of the financial and other sectors, and market-based price data. A great deal of progress has been made in this area in the last decade but more needs to be done.

In addressing the second question, it is important to understand the role of the financial sector as a risk transmission mechanism. Potential **risks** to the financial system – that is, events that might trigger a crisis – must be distinguished from the underlying **vulnerabilities** that will determine the impact of such events should they occur. Vulnerabilities in the financial sector are not merely exposures to risk, including balance sheet vulnerabilities across the various sectors, but also the system's ability to absorb or withstand a shock (*i.e.* its resilience). To assess those risks and vulnerabilities, possible channels of contagion also need to be analysed explicitly, across borders, and in both directions.

Integrating Finance into the IMF's Work

Continued Emphasis on the FSAP

The Financial Sector Assessment Program (FSAP), a joint effort by the IMF and the World Bank launched in 1999, is a key element of the Fund's engagement on financial issues with its member countries. Supported by experts from a range of national agencies and standard-setting bodies, work under the programme seeks to identify the strengths and vulnerabilities of a country's financial system; to determine how key sources of risk are being managed; to ascertain the sector's developmental and technical assistance needs; and to help prioritise policy responses. The programme is designed to assess the stability of financial systems as a whole, rather than individual institutions, and to emphasise prevention and mitigation rather than crisis resolution. As part of the process, the FSAP provides assessments of observance of various internationally-accepted financial sector standards and codes, set within the broader institutional and macroprudential context.

With a total of 120 initial assessments under the FSAP now completed or underway, and an additional 20 or so in the pipeline, the programme is increasingly moving into a phase of FSAP updates (27 updates have been completed or are underway, and 30 or so more are being planned). FSAP updates provide an opportunity to refresh the initial assessment, albeit with possible differences in scope. With the initial assessment as a reference point, updates can choose focus over comprehensiveness and concentrate on key stability and development issues.

Coverage of Financial Issues in Article IV Surveillance

Annual consultations with member countries under Article IV of the IMF's Articles of Agreement provide another vehicle for financial surveillance. The coverage of financial sector issues and balance sheet vulnerabilities in Article IV reports is being enhanced with a view to better reflect the immediacy and extent to which financial market developments can affect people's lives. This would strengthen the overall effectiveness of Article IV surveillance, in addition to ensuring continuity in the coverage of financial sector issues in-between FSAP assessments, which occur at relatively low periodicity (on average, every five to six years).

See corrigendum 1

Efforts are also underway to integrate more fully financial analysis into the traditional macroeconomic framework examined as part of Article IV work. This requires more attention to macro-financial links, better use of available data and information – both quantitative and qualitative – and a strengthened focus on spillovers and cross-country links. A key challenge in this area is the absence of a clear intellectual framework for integrating financial sector analysis, much of which is microeconomic, with macroeconomic analysis. Incorporating these two different dimensions of surveillance is one of the most compelling aspects of the Fund's work.

A More Global Perspective

Integrating multilateral and regional aspects in financial surveillance represents a further challenge. Strengthening global surveillance can be achieved by facilitating discussions within groups of countries on common issues of systemic importance and strengthening the cross-country analysis of macroeconomic and financial risks and their interactions, as well as formulating regional work plans that focus on particular policy issues facing a region.

Some work on cross-border financial sector issues has already been undertaken and more is in train. Regional FSAPs are being conducted for currency unions, and are particularly appropriate where significant regulatory and supervisory structures are at the regional level. The IMF is also engaging in financial sector regional projects that examine special issues relevant to a particular region or group of countries – recent examples include studies on Central America, the Maghreb, and the Nordic-Baltic region. Efforts are equally underway to incorporate in the FSAP elements that transcend the borders of the country under consideration – such as heavy dependence on cross-border financing or the challenges of home-host supervision.

Analytical Tools and Indicators for Financial Surveillance

Surveillance must strive to extract diagnostically useful information about risks and vulnerabilities from all of the available financial data. This requires an in-depth understanding of existing data sources and their appropriate use through a wide range of financial analysis tools.

Use of Indicators in Surveillance

The ability to conduct financial surveillance presupposes the existence of indicators that can be used as a basis for analysing the health and stability of the financial system. These indicators comprise financial soundness indicators (FSIs) – that is, aggregated data on individual banking institutions and their non-bank clients, and indicators that are representative of the markets in which these institutions operate – as well as market-based data drawn from price and volatility measures of various capital market instruments.

The monitoring of FSIs is a key element of surveillance. They provide a measure of banks and non-banks' exposure to different types of risk and their capacity to handle shocks (their resilience) that affect solvency or liquidity. FSIs consist mainly of aggregate balance sheet or income statement measures and their concept originates from prudential and commercial measurement frameworks that were developed to monitor individual entities. This type of aggregation of individual institution-level indicators (microprudential indicators) into financial soundness indicators (macroprudential

See corrigendum 1

indicators) necessarily involves a loss of information because the distribution of prudential indicators of individual institutions is also a crucial dimension of financial stability. Although aggregation is required for facilitating macroprudential analysis and international comparison, the stability assessments are strengthened by allowing some disaggregation through peer groups or the monitoring of the distributional characteristics of various indicators. In addition, FSIs themselves are concurrent indicators of financial soundness, available often with a lag. Therefore, proper interpretation of FSIs requires a range of forward-looking analytical tools, notably stress testing of individual institutions as well as analyses of the determinants of FSIs and forecasts of their future course.

FSIs can be complemented by market-based indicators, which are forward-looking indicators of soundness and are available with higher frequency. Asset prices, for example, reflect market perceptions and can influence economic developments. In a world of highly integrated financial markets, expectations themselves are subject to a slew of heterogeneous influences. Information from high-frequency financial data (about expectations, arbitrage incentives volatility, risk premia, and the like) is useful to identify policy inconsistencies and warning signals. FSIs are part of the broader framework for vulnerability assessment, which includes a wide range of analytical tools, including scenario analysis, different methods of debt sustainability analysis, and a variety of balance-sheet-type methods of risk analysis.

The Balance Sheet Approach (BSA)

The analysis of vulnerabilities across various balance sheets and sectors is an increasingly important part of a country's risk assessment. Balance sheet weaknesses in the public and private non-financial sectors are now recognised as a key factor that exacerbated the financial account crises of the last dozen years.

The Balance Sheet Approach (BSA) is based on stock variables in countries' sectoral balance sheets (assets and liabilities of financial firms, nonfinancial firms, households, government, and their sub-sectors, as appropriate) and the consolidated aggregate balance sheet (for the country). The balance sheet analysis focuses on the determinants and evolution of stocks of assets and liabilities, and the likely shocks to the stock variables, both of which can trigger large adjustments in flows (including cross-border capital flows, shifts in holdings of domestic or foreign currency assets, etc.). An approach of this type can, therefore, be a useful complement to the traditional flow analysis that is based on data related to fiscal, balance of payments, and financial programming.

Where markets are deep enough to provide relatively reliable information, the Balance-Sheet Risk Approach can also be used to derive a set of risk indicators that can serve as barometers of financial sector vulnerability. This approach combines balance sheet data with high frequency financial market prices to impute the market value and volatility of assets, both of which are needed to understand changes in the overall level of risk facing each sector but are not directly observable. An important advantage is the ability to translate continuously adjusting financial market price information into current market value estimates of asset values, which is particularly important given the speed with which economic conditions change relative to the time span between releases of consolidated accounting balance-sheet information. This framework allows to compute a forward-looking measure of credit quality (probability of default) and could be used as a metric for comparing the risk profiles of a sector/country under different policy assumptions.

See corrigendum 1

Statistical Initiatives in Support of Financial Surveillance

Collecting and presenting high quality statistical data to support the analytical work needed for financial surveillance is a complementary challenge for the IMF. Indeed, it has been noted that the surprise element at the onset of a financial crisis is often the lack of high quality, comprehensive data. The Asian crisis, for example, revealed major gaps in statistical coverage of the domestic financial sector and the external sector that permitted serious vulnerabilities to remain undetected.

In general, two types of data are needed for effective financial surveillance. One, as already noted, is high frequency market-based price data, which are forward-looking and needed to assess changes in market views and expectations, and impending changes in vulnerabilities. The other is timely and internationally comparable economic and financial data. While market-based data are available from commercial vendors, the IMF has in recent years undertaken several initiatives aimed at meeting the economic and financial data needs stemming from these various new features of the financially globalised world:

- The intensification of cross-border financial flows and positions, greater potential for spillover effects, and need for a global, multilateral perspective on surveillance;
- The lack of good quality cross-country comparable financial soundness indicators; and
- The balance sheet nature of the major crises of the past dozen years.

Data on Cross-Border Financial Flows and Positions

To meet the data needs of global, multilateral surveillance, the IMF in recent years launched five major statistical initiatives to improve as well as increase the available data on the cross-border linkages among economies: the Coordinated Portfolio Investment Survey; data for the International Investment Position; the Data Template on International Reserves and Foreign Currency Liquidity; dissemination of quarterly data on the currency composition of official foreign exchange reserves; and the Joint External Debt Hub. In addition, it plans to launch a Coordinated Direct Investment Survey (CDIS).

The first four initiatives and the planned CDIS all involve external sector data which were deemed by the IMF's Executive Board in 2002 as important data for the assessment of an economy's vulnerabilities. The fifth initiative involves creating a platform that brings different databases together for worldwide dissemination.

External Sector Data – Portfolio Investment, IIP, and Direct Investment

In 1997, the IMF launched the first **Coordinated Portfolio Investment Survey (CPIS)**, in response to global asymmetries in reported balance of payments data, particularly those in portfolio investment flows of equities and debt securities. Starting with 29 economies, the CPIS's distinguishing feature is that it provides data by partner countries by requiring all participants to provide a breakdown of their stock of portfolio investment assets by the country of residency of the nonresident issuer. This feature allows the derivation of a country's foreign portfolio investment liabilities from creditor sources, facilitating cross-checking of data and improving data quality. From 2001, the

See corrigendum 1

IMF began to undertake the survey on an annual basis and presently, about 70 countries participate in the survey.

After 1998, the IMF placed a greater emphasis on collecting data on countries' **International Investment Position (IIP)**. The IIP is important for vulnerability assessment because it provides a balance sheet snapshot of the levels, sectoral distribution, and maturity of a country's external liabilities (*e.g.*, external debt), and the size and composition of its external claims (*e.g.*, banks' foreign claims), that may be available to meet its external obligations. The IIP also complements the IMF's existing collection of balance of payments data. In 2001, the IMF Board further promoted IIP data by including it as a prescribed category of the IMF's Special Data Dissemination Standard (SDDS). As a result, the number of countries reporting IIP data increased from 37 in 1998 to around 110 countries presently, albeit with varying degrees of component detail.

Plans are underway to launch a **Coordinated Direct Investment Survey (CDIS)** in order to improve understanding of the pattern and direction of foreign direct investments worldwide. The CDIS, which will be modeled after the CPIS in terms of collecting partner data, has a targeted reference date of end-2009. The survey will be a collaborative effort with several international partners, namely, the European Central Bank, Eurostat, the Bank for International Settlements (BIS), the OECD, UNCTAD, and the World Bank.

External Sector Data – International Reserves

Two of the initiatives involved data on international reserves, another critical input for vulnerability assessments. In 1999, the IMF launched the **Data Template on International Reserves and Foreign Currency Liquidity (Reserve Template**) with the objective of improving the assessment of a country's official foreign currency liquidity position. The Reserve Template disseminates data on international reserve assets and potential short-term foreign currency obligations (and claims) including on off-balance sheet activities (such as those arising from derivative operations). This initiative inter alia helps to rectify the problem indicated earlier regarding the lack of data on off-balance sheet derivative activities during the Asian crisis. All SDDS subscribers are required to provide data for the Reserve Template.

In December 2005, the IMF published for the first time quarterly data on the **currency composition of official foreign exchange reserves (COFER)** on its website. COFER data, which distinguish reserves denominated in U.S. dollars, euros, pounds sterling, Japanese yen, Swiss francs, and other currencies had previously been published only on an annual basis in the IMF Annual Report. In response to heightened public interest, the IMF decided to make the data publicly available on a quarterly basis. Presently, COFER data are reported on a voluntary basis by 119 member countries of the IMF, comprising all 24 industrial countries and 95 out of the 160 developing countries.

External Sector Data – common Platform for Dissemination

The fifth external sector data initiative involved the creation of a platform for worldwide dissemination of different databases on external debt. In May 2006, the IMF, BIS, OECD, and the World Bank jointly launched the Joint External Debt Hub (JEDH) to bring together the data that they each compile on the external debt of countries. This database complements external debt statistics based on national sources, filling important coverage gaps, particularly in the area of private sector external liabilities. It disseminates

See corrigendum 1

data provided by over 60 SDDS subscribers, for whom it is a reporting requirement; data from creditor and market sources; and comprehensive metadata. Altogether, the database facilitates cross-country comparisons of external debt flows and positions for 214 economies.

Table 20.1 Financial Soundness Indicators: The Core and Encouraged Sets

Core Set	
Deposit takers	
Capital adequacy	Regulatory capital to risk-weighted assets Regulatory Tier 1 capital to risk-weighted assets Nonperforming loans net of provisions to capital
Asset quality	Nonperforming loans to total gross loans Sectoral distribution of loans to total loans
Earnings and profitability	Return on assets Return on equity Interest margin to gross income Noninterest expenses to gross income
Liquidity	Liquid assets to total assets (liquid asset ratio) Liquid assets to short-term liabilities
Sensitivity to market risk	Net open position in foreign exchange to capital
Encouraged Set	
Deposit takers	Capital to assets Large exposures to capital Geographical distribution of loans to total loans Gross asset position in financial derivatives to capital Gross liability position in financial derivatives to capital Trading income to total income Personnel expenses to noninterest expenses Spread between reference lending and deposit rates Spread between highest and lowest interbank rate Customer deposits to total (noninterbank) loans Foreign-currency-denominated loans to total loans Foreign-currency-denominated liabilities to total liabilities Net open position in equities to capital
Other financial corporations	Assets to total financial system assets Assets to gross domestic product (GDP)
Nonfinancial corporations sector	Total debt to equity Return on equity Earnings to interest and principal expenses Net foreign exchange exposure to equity Number of applications for protection from creditors
Households	Household debt to GDP Household debt service and principal payments to income
Market liquidity	Average bid-ask spread in the securities market 1/ Average daily turnover ratio in the securities market 1/
Real estate markets	Real estate prices Residential real estate loans to total loans Commercial real estate loans to total loans

See corrigendum 1

Financial Soundness Indicators

To encourage compilation of internationally comparable FSIs for financial sector surveillance, the IMF undertook the following initiatives.

After several rounds of extensive consultation with experts from international agencies, standard setting bodies, and member countries, the IMF completed and published the Compilation Guide: Financial Soundness Indicators (Guide),[2] which provides guidance on the concepts, definitions, sources, and techniques used in compiling FSIs so as to encourage compilation of internationally comparable FSIs. The Guide provides guidance for compiling the 12 core and 27 encouraged FSIs (Table 20.1). To pilot the implementation of the Guide, the IMF in 2004 launched a Coordinated Compilation Exercise (CCE). In particular, the objective of the CCE is to develop member countries' capacities to compile FSIs; promote cross-country comparability of FSIs; and disseminate methodologically sound FSIs. Sixty two countries participated in the CCE, whose terms of reference required the compilation of the 12 core FSIs while encouraging compilation of as many of the encouraged FSIs as possible. As a result of the CCE, metadata and end-2005 FSI data for 57 of the participating countries have been posted on the IMF's website. Posting for the remaining countries is expected in the coming months.

Notes

1 For a detailed discussion on financial stability implications of globalisation of financial institutions, see Chapter 3 of the 2007 Global Financial Stability Report (Washington, International Monetary Fund).

2 The Guide was released in electronic format in 2004 and issued as an official IMF publication in 2006.

See corrigendum!

Notice of correction

A corrigendum has been issued for this chapter.
The entire new chapter (Corrigendum 2) may be found at the end of this file.

Chapter 21
International Integration and Societal Progress: A Critical Review of Globalisation Indicators

Philippe de Lombaerde[i]
Research Fellow, United Nations University (UNU-CRIS), Belgium

P. Lelio Iapadre[ii]
Associate professor of international economics, University of L'Aquila, Italy
Professorial lecturer of international economics, Johns Hopkins University, SAIS Bologna Center, Italy
Associate Research Fellow, United Nations University (UNU-CRIS), Belgium

See corrigendum 2

Abstract

This paper presents a critical survey of different proposals to measure globalisation, from the perspective of a more general view of the relationship between international integration and societal progress.

A higher degree of international integration can be seen in itself as an indicator of societal progress, inasmuch as it reveals that human societies more and more acknowledge their common destiny. In addition, international integration fosters the provision of essential ingredients of societal progress, such as trans-national public goods and economic growth.

The available composite indicators of globalisation, although going beyond the limits of a purely economic definition of international integration, fail to perform adequately their task for a variety of conceptual and methodological reasons.

A promising alternative is based on the recognition that the scope of international integration is not necessarily global, as cross-border interactions among human societies are often limited in their geographic reach. A new generation of statistical indicators is therefore being developed, in order to clearly distinguish between regional and global integration.

i E-mail: pdelombaerde@cris.unu.edu

ii E-mail: iapadre@ec.univaq.it

Introduction

The quality of social life depends on a wide range of environmental, cultural, economic and political factors, whose measurement represents a fundamental challenge for understanding their role and for devising proper policies. The task is made even more complex by the process of international integration, which raises the degree of interdependence among human societies across the globe, so that it is increasingly meaningless to approach the study of any community without explicitly considering its linkages with the rest of the world. 'Globalisation' is the word most widely used to describe this process, even if the scope of international integration is not necessarily global.

The purpose of this short paper is to critically review different proposals to construct globalisation indicators, from the perspective of a more general view of the relationship between international integration and societal progress. Special attention will be given to the underlying conceptual frameworks and to the question whether indicators are built on broader or narrower concepts of international integration (society versus economy, regional versus global integration, etc.), and to the coherence between the measurement needs, on the one hand, and the selection of the variables and indicators, on the other. A thorough discussion of the different technical solutions adopted in the construction of the different composite globalisation indicators is outside the scope of this paper.[1]

After critically revising the existing supply of globalisation indicators, we address the nexus between international integration and societal progress in last two sections concludes.

Genealogy of Globalisation Indicators

The Kearney/FP globalisation index (Kearney/FP-GI) is generally considered as the first proposal to construct a composite multi-dimensional globalisation index, supported by a database (Kearney 2001-2006). The index covers the economic, technological, political and personal aspects of globalisation, taking inspiration from the approach used to build the Human Development Index (UNDP 1998).

The economic dimension of this and other globalisation indicators benefited from previous work on international openness and competitiveness, including: the World Economic Forum's indicator of competitiveness, since 1979 (Lopes-Claros *et al.* 2006), Gwartney and Lawson's work on economic freedom, since 1996 (1996, 2006), and the World Market Research Center globalisation index (G-index) (Randolph 2001). More recently the OECD has taken the lead as a facilitator of new work on economic globalisation indicators, which has materialised in a Handbook (OECD 2005a) and a set of indicators (OECD 2005b), but has not involved the construction of a composite index of globalisation.[2]

Several proposals followed the Kearney/FP-GI, all trying to improve it on some aspect(s). Lockwood and Redoano (2005), consistent with Lockwood's critique of the Kearney/FP-GI (Lockwood 2001, 2004), designed the CSGR globalisation index (CSGR-GI). Whereas they partly present a different set of variables, the index mainly differs from the Kearney/FP-GI on the operational aspects (adjustment, normalisation, and weighting of specific sub-indicators). Martens and Zywietz (2004, 2006), based on Zywietz (2003), proposed a Modified Globalisation Index (MGI). The authors take the Kearney/FP-GI also as their point of reference but start from a broader definition of globalisation,

See corrigendum 2

including environmental and military dimensions, and consequently, reduce the weight of the economic dimension. In addition, some technical improvements to the construction of the indicator are introduced. Heshmati's indicator (Kearney/FP/H) (Heshmati 2006), does not alter anything to the choice of variables and structure of the Kearney/FP-GI but a sophisticated statistical weighting procedure is added.

As far as the different dimensions of international integration are concerned, Dreher's globalisation index (DGI) is a more significant departure from the Kearney/FP-GI than the previous ones (Dreher 2005). Dreher expands the variables concerning personal contact and information flows, includes a cultural convergence variable, and re-introduces economic policy measures, which had been used before the Kearney/FP-GI to assess the degree of international economic integration.

Finally, an alternative for the traditional approaches to the measurement of the economic dimension of globalisation is represented by the recent work of Riezman, Whalley and Zhang (2004), who construct different measures of globalisation by comparing actual data to a counterfactual full integration equilibrium. Although pointing to a promising new strategy of research, they admittedly fail to obtain robust and reliable ordinal measures of globalisation, even if their indicators offer some information about the relative ranking of countries. A further problem of their approach is their reliance upon simple general equilibrium models based on the assumption of perfect competition, which do not appear apt to represent the actual features of global markets, characterised by various degrees of monopoly power.

Conceptual Frameworks

As there is no unique definition for globalisation the conceptual frameworks behind the globalisation indicators are diverse.[3] Until the late 1990s globalisation was still often considered as a synonym of 'global economic integration'. In his review article, for example, Brahmbhatt (1998, p.2) proposes as a definition of globalisation: "the increasing freedom and ability of individuals and firms to undertake voluntary economic transactions with residents of other countries, a process entailing a growing contestability of national markets by foreign suppliers". The definition used by the World Markets Research Center, developers of the G-index refers to "the ever closer knitting together of a one-world economy" (Randolph 2001, p.5). More recently, the OECD in its Handbook also still affirmed that "[g]lobalization refers above all to a dynamic and multidimensional process of economic integration whereby national resources become more and more internationally mobile while national economies become increasingly interdependent" (OECD 2005a, p.11).

However, inspired by the conceptual work of Held *et al*. (1999), Scholte (2000) and several others, a shift towards a multi-dimensional concept of globalisation has been observed. The Kearney/FP index, for example, measures the globalisation of a country in four dimensions (components): (i) the degree of integration of its economy in the world economy, (ii) the internationalisation of the personal contacts of its citizens, (iii) the use of internet technology, and (iv) the extent of its international political engagement. Martens and Zywietz (2004, 2006) add two additional dimensions in their MGI: (i) the involvement of a country's military-industrial complex with the rest of the world, and (ii) the intensity of globalisation in the ecological domain. The spread of ideas, information, images, and people have been grouped together by Keohane and Nye (2000, p.4) and labeled 'social globalisation', as contrasted with 'economic' and 'political' globalisation.

See Corrigendum 2

Martens and Zywietz define globalisation then as: "the intensification of cross-national cultural, economic, political, social and technological interactions that lead to the establishment of transnational structures and the global integration of cultural, economic, environmental, political and social processes on global, supranational, national, regional and local levels" (Rennen & Martens 2003).

Dreher (2005) builds on the definitions proposed by Clark (2000, p.86) and Norris (2000, p.155) and refers to a process of "creating networks of connections among actors at multi-continental distances, mediated through a variety of flows including people, information and ideas, capital, and goods", a process "that erodes national boundaries, integrates national economies, cultures, technologies and governance, and produces complex relations of mutual interdependence".

Table 21.1 shows which dimensions have been used in the construction of the different globalisation indicators, and how the variables are distributed across dimensions.

Table 21.1 Globalisation Indicators - Number of Variables and Dimensions

Globalisation index	Number of variables	Number of categories	Dimensions
G-Index (Randolph, 2001)	6	2	'old' economy (3 variables), 'new' economy (3)
Kearney/FP-GI (2001)	11	4	globalisation in goods and services (2), financial globalisation (3), globalisation of personal contact (3), internet connectivity (3)
Kearney/FP-GI (2003)	13	4	economic integration (4), personal contact (3), technology (3), political engagement (3)
Kearney/FP-GI (2004)	14	4	economic integration (4), personal contact (3), technological connectivity (3), political engagement (4)
Kearney/FP-GI (2005, 2006)	12	4	economic integration (2), personal contact (3), technological connectivity (3), political engagement (4)
CSGR-GI (Lockwood and Redoano, 2005)	16	3	economic globalisation (4), social globalisation (9), political globalisation (3)
MGI (Martens and Zywietz, 2004, 2006)	11	7	global trade (1), global finance (2), global politics (2), organised violence (1), people on the move (2), technology (2), environment (1)
DGI (Dreher, 2005)	23	3	economic integration (8), political engagement (3), social globalisation (12)
Kearney/FP/H (Heshmati, 2006)	13	4	economic integration (4), personal contacts (3), technology (3), political engagement (3)

See corrigendum 2

Notes

1 See, for example, Nardo et al. (2005) for an excellent general treatment of these operational issues.

2 The OECD indicators aim at measuring the magnitude and intensity of economic globalisation, in four areas: international trade, foreign direct investment (FDI), the activity of multinational firms, and the production and international diffusion of technology.

3 There is now already a rich literature on globalisation. It is beyond the scope of this article to review the evolution of the concept of globalisation in depth. We refer to Scholte (2002) for an excellent overview. He convincingly argues for an understanding of globalisation "as the spread of transplanetary – and in recent times more particularly supraterritorial – connections between people […] globalisation involves reductions in barriers to transworld contacts. People become more able –physically, legally, culturally, and psychologically – to engage with each other in 'one world' […] globalisation refers to a shift in the nature of social space". The author further questions the 'methodological territorialism' which is still dominating the social sciences (and hence the construction of globalisation indicators). See also, Caselli (2006).

See corrigendum 2

Chapter 22
Understanding Retirement Savings and Pensions[i]

Len Cook
Former Government Statistician, New Zealand,
Former National Statistician of the United Kingdom, New Zealand

Abstract

This paper surveys the range of considerations which influence how we as citizens provide for our future, during the years of life generally considered as retirement. The notion of retirement is changing, and this has reduced the adequacy of analysis based on demographic measures alone. There are a great variety of forces and influences over coming decades which, if not considered together, could lead to continued policy volatility as we react separately to fragments of information or events. Uncertainties which influence policy volatility can be managed, and we can anticipate where uncertainty could reduce the quality of policy advice in the long term, and reduce the well-being, of citizens of all ages. The paper draws on comparisons between New Zealand and the United Kingdom.

A New Zealand Perspective

Despite many of changes over the last three decades, retirement provision in New Zealand has retained a simplicity, economy and relevance not seen elsewhere. Given the impacts of globalisation and population change expected over the next three to five decades, continued iteration in policy without a commonly accepted focus for change could influence this. New Zealand is in a good position in terms of fiscal risks through its flat rate taxable pension, and the incentives for continued employment that this provides. The absence of an earnings related retirement pensions underwritten by the state leads to less fiscal uncertainty resulting from increasing longevity, compared to the situation in those OECD countries that have such schemes. There is a pragmatic approach taken to maintaining the purchasing power of retirement pensions in New Zealand. Sustaining near-replacement fertility throughout the next four decades would give New Zealand a significant advantage over most OECD countries, as the extent of ageing will be lessened, and the impact of ageing will occur over a longer period, easing whatever adjustments are

i This paper draws on a report published by the New Zealand Retirement Commission http://www.retirement.org.nz/files/Revised%20Retirement%20Income%20in%20NZ%202004%20Redraft.pdf

needed. There is a risk that this advantage could be dissipated through the consequences of child poverty, inadequate educational achievement and high levels of emigration of skilled New Zealanders.

Markets have a critical influence on the final impact of public policy on retirement. The constraints on choice because of the inadequacies of New Zealand's financial markets are critical, and there are many influences on the shifting importance of the labour market as the baby boom cohorts reach the age of eligibility for retirement pensions. The influence of the consumption of the retired is recognised. New Zealanders invest most in housing, but we do not know how vulnerable this leaves the household sector. Alongside a chronic balance of payments deficit, there is an uncertain degree of fragility about the seeming confidence in the accumulation of wealth of individual households and the poor savings of New Zealand as a whole.

Figure 22.1 The Volatility in Key Elements of Retirement Provision in New Zealand 1938-2008

	Key Elements of Retirement Provision					
	Income Tested Benefit	Universal Flat rate Pension (Taxed)	Subsidise Job Related Saving	Affordable Housing	Public Investment	
2007	X	XXXX	XXX	XX	XXX	Stage regulated, subsidised personal savings, externally invested, no pension
1996-2006 (2003)	X	XXXX	X	X	X	No surcharge on public pension, pragmatic wage price adjustment. Fund to invest budget surplus in equities until baby boom demands reach its peak. Continuity in mix.
1987-96 (1991)	XXXX		X	X	X	Tax subsidy removed from occupational pensions, Tax surcharge on NZ Super Age of eligibility rises from 60 to 65 years 1991-2001 Continuity in mix
1977-1987	X	XXXX	XX	X	XXX	New Zealand Superannuation, to all from PAYE, taxed only, linked to wages
1976-1977	XXXX	XX	XX	XX	XX	Return to past system as transition to NZS
1974-1976	XXXX	XX	XXXX	XXX	XXXX	Compulsory contribution to retirement scheme
1938-1974	XXXX	XX	XX	XXX	XXXX	Age benefit consistent for first 26 years

Current State Provision for Retirement Income

New Zealand introduced public retirement pensions in 1898. New Zealand Superannuation, as it is now known, has provided a universal, flat-rate, individual pension since 1977, generally after confirmation of eligibility, determined by age (65) and residence. It is a taxable benefit and for most there is no other abatement or income test. The entitlement is independent of other income, subject only to the progressive elements of the income tax. The basic level of payment is set to lie within a broad range determined by the current income of working New Zealanders, as measured by average weekly earnings; currently, the combined pension rate for a married couple cannot fall below 65% or rise above 75% of average weekly earnings. Within these limits, the level of benefit is adjusted every six months in line with the changes in the all groups consumer price index. This broad formula involves a pragmatic balance of both price and wage adjustment. It relates the value of benefits to the working life earnings of the population, and also guarantees to preserve basic living standards. Those over 65 receive an income

which grows as real national income grows, with a link which is highly transparent, but can be pragmatically managed.

Sustainability of the Current System of Retirement

We expect retirement provision to remain robust in the face of a great mass of possible change, national and global. Its sustainability rests on its continued relevance despite a great variety of possible circumstances, rather than on the chance that government might predict the economic, social and environmental change and innovation that we might experience. What seem to be among the most important factors in ensuring robustness, and therefore sustainability, are the overall breadth and coherence of provision that people and the state make for the future, the consistency, simplicity and certainty of state involvement, the inclusiveness and equity of these arrangements, given the huge variety of lifetime circumstances faced by New Zealanders, the fit with other things we do to strengthen our chances of future prosperity, and the continuity across cohorts and generations of the key elements. We also need a good fit across the elements of the system of retirement provision, and limits to the extent to which projected changes in the age and gender composition of the older population influence the overall living conditions of the group as a whole. Chart 1 shows how the stable foundation from these elements has not prevented a high degree of policy volatility and the consequent uncertainty that the public experience.

An Appraisal of Current Arrangements

New Zealand's first ever survey of household wealth shows that for nearly 60% of the near retired, they have almost no financial assets except for that provided by New Zealand Superannuation. There are huge differences in circumstances and lifestyle among this 60% of the near retired. Measuring their financial assets alone does not tell us enough about the range of ways that they have provided for their future. Capital built up outside the financial system would include housing and stocks of durable and semi-durable goods, as well as education, health, the capacity to drive, repair, cook, communicate, cultural participation, and family and community goodwill. The New Zealand system focuses, first and foremost, on ensuring and maintaining a basic pension level for all, regardless of either pre-retirement income levels or marital status. Because the New Zealand arrangements have always been simple, there has been have no complex legacy to untangle before adapting the system of retirement provision to be relevant for the lives people now lead or will lead in the future.

Other countries such as the United Kingdom (UK) are now iterating towards the New Zealand approach, but they have inherited a complex set of obligations that are difficult to leave behind. Since the changes made in the 1980s, retirement provision in the UK has essentially combined a flat rate income tested benefit scheme with an earnings related top-up (SERPS). This has created a remarkably complicated scheme. The present UK arrangements involve annuity based schemes to provide some retirement income for most wage earners. Those who are substantially reliant on a basic state pension depend on an income supplement scheme, where many retired pensioners receive a means tested benefit. For the ordinary person of average means, the complex mix of benefit components in the UK has made understanding retirement provision something of a nightmare. There seems to be little capacity at a political level to simplify the system without grandfathering a raft of historical and partly inconsistent obligations.

The UK system demonstrates what can happen when maintaining a basic pension level for all becomes a secondary consequence of other goals, rather than being the first target of policy. The simplicity and clarity of New Zealand Superannuation has allowed New Zealand to avoid getting into this situation. The future pension obligations of many countries look less sustainable than those of New Zealand. This is more a result of the nature of the arrangements than of New Zealand's economic capacity. The New Zealand scheme faces significantly lower risks from increases in longevity and declines in fertility than other OECD countries, because the obligations that future New Zealand governments and taxpayers will incur for New Zealand Superannuation are considerably lower. Where public pension schemes guarantee the continuation over the full lifetime of the level of income related to that experienced during working life, as they do in several EU countries, the state assumes the responsibility of managing risks that are beyond those associated with ensuring adequacy of income in old age. New Zealand does not face this risk.

Although the labour force of New Zealand will look older in future, it may not be significantly less active in paid employment than previously. A growing proportion of people now seem likely to prove able to choose to undertake some form of paid employment until their early seventies. The structure of pensions systems in some countries is also encouraging earlier retirement. The payment does not necessitate additional benefits and tests for those whose working life did not involve full time paid employment. Women as a group have quite a different mix of labour market experiences than men. New Zealand Superannuation provides a simple way of ensuring that differences in earnings are not reflected in a different standard of living in retirement.

Sustainability and Demographic Influences

Throughout New Zealand's recent history, the relatively high fertility and low mortality of its young population has meant that natural increase has almost always underpinned strong growth in the total population, regardless of the level of immigration. Over the past 50 years, natural increase has contributed some 85% of population growth. The fertility rate has stayed near to 2.0 births per woman, just under the long-term level (2.1) required for the steady state replacement of the population. Continued near replacement level fertility is a significant competitive advantage for New Zealand, as it has lessened the shock effects of rapid demographic change by extending the period over which change that will significantly influence policy and individuals will occur.

For New Zealand in total, the large gross flows every year from births and deaths, and the smaller flows from immigration and emigration, increasingly offset each other. Projections show that little change to the total population is likely over the next 40 years, as it rises by one million, or less than 20%, then it may start to fall. The total population will grow at a slower rate than it did over each of the past two 30-year periods. The population projections of Statistics New Zealand show that the total population could grow by some 0.6% per year on average between 2004 and 2040. For the first time in New Zealand, from the early 2040s the population could then decline. At that point, it is most likely that births will still number some 50 000 per year, just 9 000 per year fewer than now. By then, annual deaths will have risen considerably to some 59 000, compared with the current annual average level of somewhat under 30 000 deaths.

Life Expectancy Experiences and Uncertainties

We would expect that living longer and better would be a desirable outcome for all in any society. The increase in life expectancy of men and women having reached age 65 that we now expect to experience every decade is greater than that experienced during the whole first half of the last century. 17% of men and 31% of women now alive at age 65 years will reach 90 years of age. The rapid transition to an ageing population has led, in many countries, to a reduction in confidence in the sustainability of their intergenerational arrangements for retirement provision, and for funding health care and other essential services. Country experiences and prospects for population change, as measured by the rate of change in dependency ratios, and other crude measures of population structure, vary considerably. Very low levels of fertility are making many countries dependent for the first time on levels of migration that have always been part of our way of life in New Zealand. There are examples of many possible types of policy responses in the area of retirement provision.

Increasing the age of eligibility for public retirement pensions has occurred in the United States, and will increase in the UK over several decades after 2020. There have been both direct and indirect approaches to reduce the level of provision that the retired can expect to receive beyond that from their own resources, although in the UK commitments were made in 2005 to increase the state retirement pension. The UK government will no longer continue to adjust the basic state retirement pension by changes in consumer prices, from 2020. The UK government is now planning to link public pension increases to that of average earnings, having gone through a long period of price adjustment only. In Australia, eligibility for the state pension has been tightened. The government-enforced compulsory employment-based scheme for saving still has no requirement to purchase an annuity on retirement. In New Zealand, the age of eligibility for a public pension stayed at 60 years until 1992, and it has been 65 years since 2001.

Challenges to Continuity

Life expectancy measures give a simple but imprecise summary of the well-being of a population. However, life expectancy is based on a record at death only, and has several limitations. Life expectancy is intrinsically a forecast projecting out from past experiences, yet those who continue living are not a representative sample of those born. Across many countries, improvements to life expectancy have been consistently underestimated, most especially over the past two decades. This has had implications for confidence in the capacity to plan for retirement provision. While the majority of a cohort is still living, health and other indicators might be able to signal emerging trends. Olshansky *et al*. note that while "the world record for life expectancy at birth in developed nations has been increasing by three months per year since 1850", this steady rise "may soon come to an end. They note a life-shortening effect of 5 to 20 years from severe obesity.

Living longer and better builds more opportunity for the state and individuals to make provision for the future. It also challenges behaviours, processes, expectations, entitlements and laws that fitted how we were. Because each generation of older people has so far generally seemed to be healthier, better educated and more active than those they succeed, as well as longer lived, we cannot easily extrapolate from the experiences of earlier generations what will be the consequences when we have a substantial increase in the number and diversity of circumstances of those aged 65 and over.

An explicit retirement age may well have been what the early architects of welfare states sought to achieve, so that job opportunity for the younger cohorts would be created through obligatory retirement, even when economic circumstances slowed growth overall. Today the boundary between paid employment and retirement from paid employment is becoming less distinct in New Zealand than in many other OECD countries where full retirement from the labour force continues at the same high rate as a decade ago. In New Zealand, since 1991 there has been a rise in the labour force participation of both men and women aged between 55 and 60. There was also a significant increase in labour force engagement of people aged between 60 and 64 years between 1991 and 2001, compared with the previous decade, following the 1991 change in the age of eligibility for New Zealand Superannuation. For those aged between 60 and 64 years, women have increased at a faster rate than men, although they participate less. As a result of the baby boomers' passage through life, the median age of the labour force is rising. The baby boom generation is now concentrated in the older half of the labour force aged 40 and above of the labour force. The median age of the labour force was 36 years in 1991 and it will reach 42 years in 2011, and stay near that until 2030, possibly rising slightly after that, depending on the increase in labour force participation among people aged 65 years and over. This passage of the baby boom generation through the workforce has led to a long-term shift in its age structure. In 2001, there were 9% more people working in New Zealand who are aged below 45 years than before 1991, and 50% more are now aged over 45 years than in 1991. As the first in the large baby boom cohorts reach their seventh decade, there will be a sharp rise in the proportion of those currently employed who will soon be making a decision about retirement. There are many occupations that now face the loss of a large share of their human capital if, over the next five to ten years, all who can take retirement at 65 years (or even younger) do so. The potential impact of this effect in the US is described by Dohm who also notes a considerable dispersion in the mix of occupations that are experiencing the most significant ageing.

In the future we are likely to see a greater variety of decision points in the life course for all, resembling the variable working life experience with which many women are already familiar, as active engagement in paid employment evolves away from a set end point. Less explicit transitions are those between full-time and part-time paid employment, and the various and evolving stages of active engagement in both paid and unpaid work. There is a need to quantify the influences that determine when individuals pass through these transitions, including the effects of changing fitness and health status, unpaid work demands, including the demand for care of the "oldest old" by the "younger old", and the demand for care by grand-parents of grand-children, the propensity to need functional support, and its evolving capability, employment experiences and wealth accumulation, lifestyle, the lifetime nature of education and training, and occupational opportunity, household arrangements, perceptions, attitudes and expectations and debt.

We may find it more relevant to think that the working-life cycle of the future includes a prolonged retirement period which would involve a long period of continually shifting labour market engagement from the ages of 50 to over 70 years. There may be no universal concept of a retirement age. Older people's distinct consumption patterns, which include care, health and other support services, will increasingly influence the economy as a whole, and impact significantly on some industries. We expect an increased demand for leisure and personal services.

Financial Markets and Retirement Provision

There is little evidence that savings incentives increase saving overall, although they do influence the mix of savings. Most incentives have the effect of raising or lowering the capital value of assets in some way, often perversely, so that the cumulative effect of incentives to save may simply result in a shift in where and how saving occurs, rather than in more saving overall. The state in New Zealand does not manage retirement benefits from occupational schemes. Introduced in 2007, Kiwi Saver is a state subsidised saving scheme for the employed in New Zealand, akin to the Australian state savings scheme. Both schemes need a tax subsidy for people that save during their working life. Providing a subsidy through taxation, or some other form of protection for financial savings and occupational pensions, benefits most those who are in the upper 40% of income distribution. Moreover, such subsidies to the more affluent usually reduce the overall level of benefits, unless provision is able to be made elsewhere for their cost.

The arguments favouring this sort of scheme depend on judgments about the need to increase private financial savings, based on two assumptions: that compulsory saving will provide a higher level of individual income in retirement, to "top up" the basic state pension; and that it will also raise levels of saving overall. It can also be argued that when young, people are myopic and do not make the personal provision for their own old age that they would make if they were fully rational and fully informed, and that some form of compulsory financial saving counters this myopia.

Financial markets and services are not yet adequate for meeting the expectations and needs that New Zealanders now have, both for accumulation and for decumulation of savings for retirement. The small size and isolation seems to lead to less efficient markets for savings, annuities, home equity release schemes and other financial services in New Zealand. Globalisation has yet to bring to individual New Zealanders forms of equity investment options and other alternatives which have the credibility for investors that larger countries experience. Future equity returns are uncertain, and the recent past is no guide to the next 30 years. A recent study for the UK Pensions Commission of returns on equities and nominal bonds for ten-year periods during the 20th century highlighted the huge dispersion of returns over time. The UK Pensions Commission (195) gives a stark example: an equity investor who had delayed their retirement from March 2000 to March 2002 would have seen a fall of one third in the value of the pension they could afford.

Fund management costs absorb a significant amount of retirement funds. The UK Pensions Commission has identified a strong relationship between annual management charges and the form and scale of pension provision. The UK Department of Work and Pensions has placed considerable emphasis on annual management charges for personal savings accounts being kept as low as possible. They note that "under a 1.5 per cent management charge, an individual saving for 40 years will loose around 20 per cent of their pension, compared with a charge of 0.5 per cent." Annual management charges within the Australian compulsory savings system are between 0.4 and 1.8% for individual accounts, and 0.4 and 0.8% for industry accounts.

A considerable share of saving involves investing in assets which will be sold to the working population of the future by retired generations. The retired sell assets to those who are now savers, in order to sustain current consumption. In all developed countries, reduced fertility means that a smaller number of non-pensioners will probably be willing or able to buy those assets. Where later generations are smaller in size, asset prices must

be expected to fall. Moreover, the willingness of later cohorts to purchase assets at the price sought by the larger cohorts as they shortly move into retirement will very much depend on their comparative wealth. It has not been the past experience anywhere to transfer the assets of one generation to a later generation that is smaller in size, and perhaps wealth, so there may be more uncertainty in future asset values over the next three decades. Johnston notes that, given the size of the baby boom generations, even switching from equity to fixed-income assets may have difficult to foresee consequences for the functioning of financial markets. It is not clear how far these problems could be avoided by investing overseas, given demographic shifts in other countries. A comparative analysis of demographic transitions among countries of significance to New Zealand would increase understanding of this. This is important, given the shift in pre-funding that the New Zealand Superannuation Fund and Kiwi Saver have created.

Differences in Cohort Experiences

The baby boom cohorts (born between 1946 and 1966) and their predecessors have generally benefited at all levels of income from economic circumstance, public policy and improvements to the standard of living. Later cohorts born from the mid 1960s have seen greater disparity in the impact at different levels of income of social change, labour market opportunity, health and public policy. The consistency and breadth of influences across economic, demographic and social aspects of life have had a severe impact on cohort differences. We need to better understand their effect on the many longstanding trends whose continuity is implicit in demographic and economic projections. Disparities in life expectancy are increasing, as longevity at older ages is rising at rates we have never before experienced. We do not know for long this will continue.

When we examine the evidence we can glean across demographic, social and economic fields of statistics, we can identify areas where the limitations of our information base can limit the quantification of emerging trends. We do not know what the total wealth of each current cohort will be on retirement, but a considerable difference has emerged in the distribution of wealth among those at different income levels, in the cohorts still alive and born during and before the baby boom, and those afterwards. (after the mid 1960's). Public policy over the first three decades of the post war period involved stimulating home ownership, and occurred alongside huge gains in educational levels, improvements in health, and employment opportunity. These major influences on increasing personal wealth were spread across the economic spectrum. The later benefits from housing inflation have also since been spread across the economic spectrum for these cohorts, albeit unevenly. These cohorts will benefit from labour market flexibility with less of the disadvantage in reduced job security that later cohorts have experienced. Whether cohorts born since the mid 1960's will build more or less wealth than earlier cohorts is not yet certain, but the distribution of wealth will be significantly different. Many of the influences on increasing personal wealth overall have acted to contain the economic position and opportunity of those in the lowest economic circumstances. The evidence for some is not robust, although studies elsewhere may have more effectively quantified the shift.

A research agenda that could reduce uncertainties that seem likely to influence future policy on retirement provision. The need to engage in international comparative studies and cohort analyses is emphasised, while the contribution of longitudinal studies reinforces the need to more effectively integrate epidemiological research into long term planning for retirement provision.

Information and Education

In New Zealand, the requirement to regularly review and reflect on arrangements for retirement provision is a statutory obligation of the Retirement Commissioner. Provision for retirement involves lifetime investments, and they take many forms, often unmeasured. Retirement saving is influenced by taxation policy, labour market, housing, health, education, as well as government approaches to investment of transport and energy infrastructures, and the involvement of the state in enterprise, past, present and future. What we do in the future depends on the importance that a continuing mix of New Zealand governments have placed on this longstanding national commitment to the welfare and well-being of those who reach retirement age. We can rarely predict the way any public policy will ultimately have an impact on the public or private capital base of New Zealand, or of individual households. Periodically, we need to reassess how well we are encouraging and sustaining confidence in these lifetime investments of New Zealanders and their government.

Information and education has the potential to increase understanding among the public of their life course and how people can provide for their well being, in partnership with the state and employers. Encouragement through increasing understanding has so far been more visible in New Zealand than elsewhere. People do not always understand the meaning and quality of the information that they do use; they often do not even use what they have. In the areas of retirement provision and savings, there is some evidence that these information effects can be quite substantial. The UK Pensions Commission produced estimates in the United Kingdom that men underestimate their longevity by 4.2 years, and women by 6.9 years. Those in poorer economic situations usually have a lower life expectancy than average, but even then, people in this group may still underestimate their true life expectancy, and hence undervalue the benefits from personal saving. Similarly, in New Zealand, expectations of inflation exceeded experience by a considerable margin for many years after price stability became well established, with consequent effects on the propensity to save rather than consume.

Giving people better access to critical information may increase the extent to which individuals make sufficient provision for the future. UK analysis by Banks and Oldfield has shown that where basic numeracy skills were missing, people were less able to make informed choices among the many options they had. Recently, the Retirement Commission and a major bank surveyed the understanding New Zealanders had about retirement provision. The result reinforced the concerns covered in this section, that much of what is important to working life choices made to increase living standards after retirement is not universally known. Some people do not have enough understanding about how they can influence their personal situation.

Inflation and economic difficulties have been major destroyers of private saving and personal provision for retirement in the past. The breadth and resilience of infrastructure affects the investment risk, and is part of the comprehensiveness and scale of retirement provision policy by the state overall. How we decide on an appropriate balance between placing national savings in the equity markets of other countries, and investing in the education, health and infrastructure of New Zealand, is quite impossible to assess exactly, but some broad understanding of how these alternatives are balanced will become important as the New Zealand Superannuation Fund and the Kiwi Saver Fund expand. Barr and Diamond note that there are four channels through which savings are allocated to investment. They identify these as market transactions – purchasing newly issued bonds and stocks (as opposed to trading existing ones); intermediation – deposits in banks

and other intermediaries, which are then lent to investors; direct pair wise loans: to people starting small businesses, from friends and families; to people one transacts with through trade credit; or seller provided credit more generally; and saving to finance one's own investment.

Managing equity investment brings a mix of risks. A government is not always well equipped to identify and manage commercial risks as they develop, partly because of the way that commercial risks and political risks interact. Large scale commercial investment by governments brings untested risk, and the maintenance of a commercial investment programme may reduce the availability of funding for the collective investments in infrastructure, education and health that are usually led by government. Significant initiatives where the government invested in equities in New Zealand and elsewhere, or engaged in commercial lending, have generally failed.

The commonly cited dependency ratio – that is, the ratio of people who are over and under "working age" to those who are "of working age" – is a very poor analytical and presentation tool for understanding population change. It provides a gross and misleading oversimplification of the nature of demographic change in New Zealand. To ensure effective policies, and also to convince the public of their sustainability, we need more realistic forms of measurement. We need measures which can show changes in generational equity, and shifts in the share of active life that is spent in both paid and unpaid work. This would include monitoring how reward for effort is sustained for all whenever the capacity of society to sustain itself collectively and individually is protected.

Conclusion

To maintain the robustness of the means of provision for the future, there are three major pathways ahead. The first is to increase the capacity to generate wealth in New Zealand; second to change the level of the basic public pension entitlement, or third to change how the level of entitlement is assessed, for example the age of eligibility for a public pension can change. A general tendency for living longer associated with more active lifestyles in later years, would be consistent with changing the age of eligibility for New Zealand Superannuation, assuming that all share in these improvements.

Chapter 23
Key Role of Cultural and Creative Industries in the Economy

Hendrik van der Pol
Director, UNESCO Institute for Statistics, Canada

Abstract

As culture and creativity become an increasing driving force in the international marketplace, it is essential to measure their impact not only on the economy but also on society at large. UNESCO has always been at the forefront of addressing the dual cultural and economic nature of cultural goods and services, working on both the theory and the practice. The Organisation's Institute of Statistics (UIS) is currently seeking to redefine international data collection standards for the culture and creative sector so that they take into account the needs and specificities of the developing world.

Providing an overview of the main approaches to assessing the economic and social importance of culture, this paper highlights the numerous limitations in current statistical information. It also includes a series of proposals based on the pragmatic approach currently being developed at UIS.

Introduction: Increasing Role of Cultural and Creative Industries

In this era of extraordinary change and globalisation, many acknowledge that creativity and innovation are now driving the new economy. OrganiSations and even economic regions that embrace creativity generate significantly higher revenue and provide greater stability into the future.

Based on ideas rather than physical capital, the creative economy straddles economic, political, social, cultural and technological issues and is at the crossroads of the arts, business and technology. It is unique in that it relies on an unlimited global resource: human creativity. Growth strategies in the creative economy therefore focus on harnessing the development potential of an unlimited resource and not on optimising limited resources (as in traditional manufacturing industries).

Many stakeholders are involved in this process: the public sector which includes cultural institutions, *e.g.* museums, public service broadcasting organisations, etc.; the private sector which covers a wide range of commercial operations in all fields of cultural production and distribution; the non-profit sector including many theatre and dance

companies, festivals, orchestras, which may receive government subsidies; and non-governmental organisations such as advocacy agencies, actors and musicians' unions.

Culture and creative industries have been increasingly integrated into the policy agenda of both developed and developing countries. In 2005, the United Kingdom's Commission for Africa reported that there was a 'real danger that a lack of attention to culture in policy making […] will overwhelm many of the collective mechanisms of survival which are part of Africa's cultures' (Commission for Africa 2005, p.130). In early May 2007, the European Commission announced its decision to adopt a strategy on the contribution of culture to economic growth and intercultural dialogue (European Commission 2007).

Culture is increasingly finding a route to the market, which is leading to radical transformations in the way people create, consume and enjoy cultural products. Globalisation and the convergence of multimedia and telecommunications technologies has transformed consumers from passive recipients of cultural messages into active co-creators of creative content. Digital distribution in industries such as design and music has transformed global markets and allowed new industries and consumers to emerge in developing regions such as Africa and Asia (OECD-2 2005). It is estimated that licensed digital distribution of recorded music will rise from $653 million in 2005 to $4.9 billion in 2010, which represented a 49.5% compound annual increase (PWC 2007).

The digital distribution of music is but one aspect of much larger economic phenomena, as will be discussed in this paper. Yet it is important to note that culture and creativity also have a tremendous impact on social cohesion and development. In Europe, the role of culture in development shows that 'the arts enrich the social environment with stimulating or pleasing public amenities.... [and] artistic activity, by stimulating creativity… [and enhancing] innovation. Works of art and cultural products are a collective "memory" for a community, and serve as a reservoir of creative and intellectual ideas for future generations. Arts and cultural institutions improve the quality of life' (Council of Europe 1997). Likewise, Australia has underscored the fact that 'the culture and leisure sector contributes to economic development through facilitating creativity innovation and self reflection' and, as such, recognises culture as a key component of society's well-being (ABS 2001). Culture should not only be considered as a means (or a barrier) to achieve economic growth but also as a factor of social cohesion and human development.

Before exploring the social and economic importance of culture any further, certain conceptual differences should be discussed. One choice of orientation, already highlighted in the very title of this session, is to differentiate between 'cultural' industries and 'creative' industries.[1] (Another approach, adopting the term 'copyright' industries, is considered briefly below while other categorisations, such as design industries, lie beyond the scope of this paper). Cultural industries relate to the creation, production and commercialisation of the products of human creativity, which are copied and reproduced by industrial processes and worldwide mass distribution. They are often protected by national and international copyright laws. They usually cover printing, publishing and multimedia, audiovisual, phonographic and cinematographic productions, crafts and design. Creative industries encompass a broader range of activities than cultural industries including architecture, advertising, visual and performing arts. The United Kingdom's Department of Culture, Media and Sport (DCMS) placed its definition of creative industries at the heart of its policy making agenda in the late 1990s and defined creative industries as those requiring creativity, skill and talent, with the potential for

wealth and job creation through exploitation of their intellectual property. The principal practical difference between these approaches has been the definition of the sectors and occupations to be included in the statistics.

Diverging conceptual terms is not, however, the only difficulty when trying to measure the impact of culture and creativity. For an organisation like UNESCO, which is primarily concerned with developing countries, a lack of key data poses a major problem. Since most data on culture are from the developed world, such as OECD and the European Union (EU) countries, the most significant challenge facing UNESCO's Institute for Statistics (UIS) is to develop cultural indicators which are relevant to the developing world, especially Africa. A different approach to measuring culture must therefore be advocated but what form should it take? A certain number of proposals are presented in the latter half of this paper but first it is important to understand fully the economic and social role of culture.

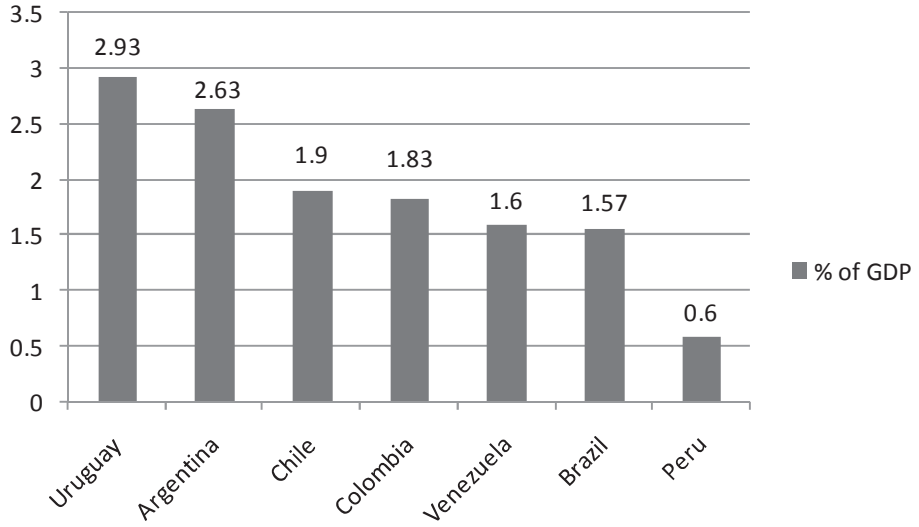
Key Statistics on the Economic and Social Role of Culture

The entertainment and media industry are forecast to grow from $1.3 trillion in 2005 to reach 1.8 trillion by 2010 (PWC 2007).[2] Asia is expected to record the highest growth rate of all regions in the entertainment and media industry, increasing from $274 billion to $425 billion (with a 9.2% compound annual growth rate (CAGR)) and China will have the fastest growing industry in the world, with a 26% CAGR. In 1990s, the creative economy in OECD countries grew at an annual rate twice that of service industries and four times that of manufacturing (Howkins 2001).

The growth of the cultural and creative sector in the European Union from 1999 to 2003 was 12.3% higher than the growth of the overall economy (European Commission 2006). Turnover of the culture and creative sector in the EU, which comprises television, cinema, music, performing arts, and entertainment, generated €654 billion and contributed to 2.6% of the European Union's GDP in 2003. The culture sector employed at least 5.8 million people in Europe in 2004, which is more than the total working population of Greece and Ireland put together. Furthermore, it is often noted that the quality of jobs generated in the creative industries may provide higher levels of job satisfaction, given a strong sense of commitment to the sector and involvement in cultural life.

While these statistics are vital to our understanding of the increasing economic contribution of culture, it should be noted that there are several other ways to measure the importance of cultural industries. One such alternative source of data can be found in national satellite accounts, which are being adopted by several MERCOSUR countries. Initial results are provided in Figure 23.1, where it is shown that the contribution of culture (here understood as publishing, leisure, cultural services and sports) to the GDP for the MERCOSUR countries was less than 3% in 2003. While this figure seems to be lower than for European countries, varying definitions may explain these differences; for example, the European study incorporates gambling and casinos and internet transactions, which is not the case for MERCOSUR.

Figure 23.1 Contribution of the Value Added of Selected Cultural Industries to the GDP for MERCOSUR Countries in 2003 (in percentage)



Source: Cuenta satélite de cultura, Secretaria de cultura, Argentina, 2006.

Another way of assessing the creative economy is by examining what is referred to as the copyright industries. The World Intellectual Property Organisation (WIPO 2003) developed a framework that enables countries to estimate the size of their creative and information sector. The guide separates out this sector into four categories of industries, which are the Core Copyright Industries, the Interdependent Copyright Industries, the Partial Copyright Industries, and the Non-Dedicated Support Industries. These industries differ by their level of involvement in creation, production and manufacturing in the literary, scientific and artistic domain. The core copyright industries[3] are usually those characterized as typical cultural industries.

Table 23.1 Contribution of Copyright Indistries to the GDP and Employment in % for Several Countries in 2001

		Contribution of copyright industries % to	
Countries	**Industries**	**GDP**	**Workforce**
Singapore	Copyright industries	5.7	5.8
	Core copyright industries	2.9	3.6
Canada	Copyright industries	5.3	7.0
	Core copyright industries	3.9	
United States	Copyright industries	12.0	8.4
	Core copyright industries	7.8	4.0
Hungary	Copyright industries	6.8	7.1
	Core copyright industries	4.0	4.2
Latvia	Core + interdependent	4.0	4.4

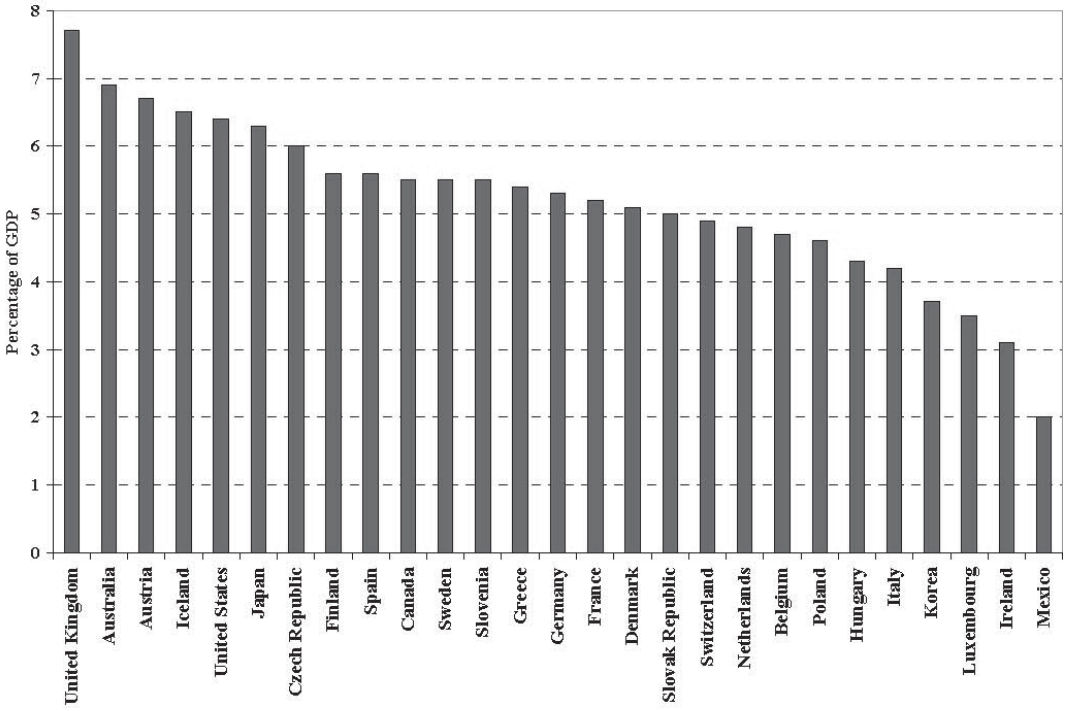
Source: WIPO, National Studies on Assessing the Economic Contribution of the Copyright-Based Industries, 2006.

Table 23.1 shows the contribution to GDP and employment of the Core Copyright Industries and Copyright Industries as a whole (including the Interdependent and Partial Copyright sectors). When all copyright industries are factored into the economic model,

the total value of the contribution to GDP or the workforce can vary considerably depending on how they are defined.

A final key source of data on the creative economy is information related to the consumption of cultural activities or products, which can be captured by statistics on household spending on recreation and culture. The inclusion of recreation covers domains beyond the common definition of creative industries, such as the purchase of leisure equipment for camping. The percentage of GDP spent on household expenditure on recreation and culture for most OECD countries shows a positive correlation with per capita income (OECD 2007). The richer a country is, the more chance there is that the population will spend a higher percentage of their income on culture and leisure. However, there are some anomalies: in Ireland, considered a rich country, the population spends relatively little on recreation and culture, while the Czech Republic, considered a poorer country, spends a rather high share. Figure 23.2 shows that in OECD countries, the percentage of household expenditures on recreation and culture is included in a range between 2% for Mexico and 8% of the GDP for the United Kingdom. The trend has remained fairly stable at around 5% of GDP over the last decade.

Figure 23.2 Household Expenditure on Recreation and Culture as Percentage of GDP in 2005 or Latest Year Available



Source: OECD, OECD Factbook 2007 - Economic, Environmental and Social Statistics, 2007.

Measuring Progress: the Limits of Current Cultural Statistics

Despite these different methodologies to measure the creative economy, there are a number of limitations to such data collection. Not only do varying definitions and categorisations often make information incomparable, as we have already seen, but data is also frequently scarce or at best incomplete. Furthermore, there is a widespread lack of

resources and expertise to ensure high quality statistical work, especially in the developing world, as data collection on the creative sector remains a low priority area for many countries.

Many different initiatives to measure culture and map out cultural and copyright industries have been developed in Latin America in the past ten years, not only at the local and national level but also sub-regionally (*e.g.* Mercosur, Grupo Andino and CONACULTA). Statistical data and indicators have reached an advanced level of detail and sophistication in some cases (such as Colombia and Chile), but are non-existent in other countries of the region, particularly those in Central America. The Organisation of American States and the Organisation of Ibero-American States are currently striving to establish cultural information systems (CIS) through knowledge-sharing activities among countries in the region. However, these tend to be highly descriptive and a common set of core indicators has yet to be identified.

As for Africa, the role of cultural industries in the continent's development was acknowledged in the 1992 Dakar Plan of Action (UNESCO 1992). The document serves as a reference point for current strategies in this area, as described in the New Partnership for Africa's Development (NEPAD) plan for culture and development (UNESCO 2003). Having said this, few culture statistics are currently collected on a regular basis, despite increasing policy needs to assess the importance of revenue generated from cultural industries which thrive on the continent, such as music and crafts. This requires a better characterisation within international classification systems in order to provide clear guidance to National Statistics Offices.

On a more global scale, one of the major failings of current statistical data is that it does not accurately capture copyright flows and other intangible assets. World trade in 'core' cultural goods – taken as recorded media (music, CDs, etc) and printed media (books, newspapers, periodicals, etc.) – amounted to almost $60 billion in 2002 (UIS 2005). However, these figures do not capture the value of copyright cultural products traded worldwide. Trade statistics are captured through customs data and therefore relate only to the physical characteristics of products. What they fail to take into account is the value of ideas, creativity and innovation which, in most cases, are transformed into productive capacity requiring intellectual property protection. As mentioned in the introduction, the cultural sector exploits an infinite raw material – creativity – which proves difficult to trace in physical form.

Data on imports and exports of films can be used to indicate diversity in production and exchange by indicating diversity in the origin of films entering a country, but it presents the severe limitations of customs and balance of payments data (UIS 2005). Typically, films are exported to the destination market, then copied and distributed locally. As a result, the level of exports may bear little relation to the volume distributed in the recipient country. While an exported film has an almost negligible value at customs, the bulk of international exchanges relating to its export are compiled in data from balance of payments, in the form of receipts for royalties and licences through copies, exhibition rights and reproduction licence fees.

Another major challenge to data collection is the question of how to measure cultural employment. In order to understand cultural employment within a country, occupations within cultural industries need to be supplemented with cultural occupations in non-cultural industries. These could include, for example, design activities in manufacturing and other sectors. Furthermore, cultural occupations in developing countries are often a

secondary occupation for agricultural labourers or other workers and, as such, are often not declared or captured in censuses and labour force surveys.

The International Standard Classification by Occupations (ISCO) currently does not provide the level of detail required to identify cultural occupations in a truly comprehensive manner. In some cases, it is necessary to link employment data with industry data to calculate total cultural employment. These hidden or 'embedded' cultural occupations may not include a large enough number of practitioners to be accurately measured in sample surveys. Moreover, self-employed or informal work, and even small companies of less than ten people are not captured in surveys. In this respect, even European statistics may well underestimate cultural employment.

UIS's role in Defining Standards

In 2006, the UIS began to revise the 1986 UNESCO Statistical Framework for Culture. Preliminary proposals have been well received by countries and present a further opportunity to rally countries around a common standard. It is recognised that, in the past, such initiatives have often failed as few countries have sufficient resources for dedicated surveys of culture, and that earlier standards have seen to be centred on OECD or EU perspectives rather than those of developing countries. Instead, UIS will concentrate on a pragmatic approach building on a regular collection of statistical instruments, such as labour force surveys and population censuses. Proposals will focus on standard-setting, which involves marginal adaptations to these existing data collection instruments, thus minimising costs and maximising the potential for regular collection of cultural data once the project has been completed. They will present a comprehensive approach to measuring culture in both social and economic spheres while allowing countries the flexibility to identify different national priorities.[4]

UIS has produced two reports on the international trade in cultural goods, but the Institute must now also consider the question of the consumption of cultural goods. 'Consumption' is seen by some as an economic approach whereas culture demands a wider view in which 'enjoyment' has more than an economic value. In addition, specialists in the culture field are quick to point out that 'participation' in cultural activities has strong social benefits, such as confirming identity with a social group, increasing social cohesion, or, through an encounter with a different culture, allowing people to question assumptions about how they act or think about their own emotional and moral values.

UIS is therefore interested in developing instruments for national and international studies, which will evaluate participation in a wide range of cultural activities. The Institute's approach is a fundamentally pragmatic one since it realises that estimating the full 'value' of cultural goods would be too costly. Developing countries in particular do not have the financial resources to carry out extensive surveys to assess cultural value, and expertise on the special methodologies involved is scarce. Instead, UIS recommends that countries take a broad-based approach, while understanding the inevitable imperfections of the results produced. A measurement strategy, which draws on common existing statistical instruments and standards, is required. As such, UIS advocates the use of the International Standard Industrial Classification (ISIC) and Classification of Occupations (ISCO).

UIS is currently working with the International Labour Organisation to ensure that they represent the full range of cultural employment. At the same time, UIS will advise

national statistical offices that, while adhering to international standards, they should aim to identify the informal sector in any cultural industries that are a national priority. For example, all national statistical offices conduct a labour force survey and, if such surveys asked about a *second* occupation, it is likely that many craftsmen would be identified.

In addition to making small adjustments to existing surveys, it is always possible for countries to undertake dedicated surveys for priority issues. In the realm of culture, dedicated surveys could be used to 'drill down in depth' in a particular sector, to consider the implications of a particular event, or to measure the overall participation of a certain population in cultural activities. In the case of a sector study, considerable care should be taken with sample design to ensure that informal, artisan, and amateur craft production is captured.

Another common form of survey covers a particular event or place, usually registering the number of visitors to a site, event or festival. While in OECD countries it is easy to separate tourists and audiences from locals and performers respectively, this is more difficult in other countries in the world. Many sites, whether natural or architectural heritage, play a central role in the lives of local communities. This means that local participation and use of these sites should be surveyed alongside those of visitors from further a field. Similarly, in many festivals in non-OECD countries audience and performers cannot be separated as people move between roles at different times. This is clearest in music, dance and drama, and artistic production where local shows are often an opportunity for artists to exchange ideas and inspirations as much as they are about displaying works to the general public.

Recognising the many limitations of current cultural statistics, UIS proposes three main aspects in which data collection on a particular cultural industry may require data of other forms: education, traditional knowledge, and archiving and preserving. Education teaches people how to value culture and is the principal means of transmitting culture from one generation to the next. Enrolment rates in cultural education programmes are therefore useful indicators on how different aspects of culture are covered in school and higher education. Traditional knowledge represents cultural production that is not included in the conventional market because it is subject to informal exchange, non-monetary exchange (*e.g.* exchange of gifts) or amateur/informal production. Examples might include the practice of traditional medicine, rural textile production, production of traditional musical instruments or dress. While these articles may not enter the market, their production can be a very significant share of cultural production in a particular sector and many countries will want more accurate data on the scale of this production. Archiving and preserving is not merely the way in which significant works of art are conserved and safeguarded for posterity. Important works of art or traditions can be powerful forces for acculturation and inspiration for future artists. Works of art, heritage sites and intangible heritage are therefore assets, which, if invested in, can provide returns in terms of new cultural products, and visitor revenues. Table 23.3 shows an initial view of some sectors showing how the three suggested transverse dimensions might be measured.

UNESCO's Approach to Measuring Creativity and Culture

UNESCO has always been at the forefront of addressing the dual cultural and economic nature of cultural goods and services. The Organisation has worked extensively on this interface, exploring both the theory and the practice.[5] In May 2008, it is set to

publish its next World Report on the theme of cultural diversity. Building on the considerable amount of critical reflection that lies behind the Universal Declaration on Cultural Diversity (2001) and its Action Plan, the report will aim to identify key policies and policy making instruments to ensure that cultural diversity, sometimes perceived as a threat or a source of insecurity, is truly fruitful and can flourish for the benefit of all. This report is to include a statistical annex provided by UIS.

Table 23.3 Examples of Data and Indicators for four Sectors and the Three Transverse Dimensions

	Heritage Natural Cultural Tangible and Intangible	**Arts** Visual, fine, performing arts	**Books & Press**	**Audio-visual** Film, video, new media
Education	Enrolment Attendance (visitors & locals)	Enrolment Performances in/by schools	Number of school textbooks	ICTs in education
Traditional knowledge	Intangible heritage (no of themes) Biodiversity , Languages spoken Traditional and complementary medicines	Craft artisan Story tellers Festivals ($, attendance)	Languages in print Community newspapers titles & circulation	Traditional knowledge on the Internet Audio-visual documentaries
Archiving & Preserving	Conservation (employment, $ earnings)	Documentation centres - Galleries (number of items, employment)	Libraries (volumes, transactions)	Film archives (volumes)

An example of UNESCO's practical involvement in the field, the Global Alliance for Culture Diversity is a pilot project that forges public-private partnerships to strengthen cultural industries and enterprises in developing countries. In terms of normative action, UNESCO adopted the Convention on the Protection and Promotion of the Diversity of Cultural Expressions in 2005. Following an unprecedented ratification rate, the Convention entered into force on 18 March 2007. This binding legal instrument promotes cultural industries as key to sustainable development and poverty reduction and seeks to strengthen international cooperation in the field of cultural industries. One of the aims of the Convention is to gain a greater understanding of how to measure the diversity of cultural expressions. To that end, UNESCO's international standards in the field of culture will be reviewed in close collaboration with UIS, in particular those related to creative industries and their impact on economic development.

Conclusion

A broader assessment of the economic impact of cultural sectors and products is key to providing a fuller picture of the real impact of culture. As a final example, we should consider the impact of the Great Wall of China. It attracts huge tourist income, investment by national authorities and local people. Yet a quick search finds the Great Wall also used as a symbol to sell a wide range of products, including a brand of car, a packaging machine, the Olympics, Chinese language courses, software, music, hotels, cigarettes, medicine, fireworks, food, hydraulic fluid and so on. The Wall contributes to the sale of all these products, extending its economic impact far beyond the heritage sector. In theory, an assessment of the economic impact of the Great Wall should take into account its value as a brand for all these products, its value to the people of China as a symbol of national identity as well as its value to the people of the world as an

outstanding achievement of humanity. It is of course this latter aspect which has led to its inclusion on the UNESCO World Heritage lists.

The example of the Great Wall illustrates how far we have to go to measure the full impact of culture in the economy, as well as the extensive contribution that culture makes to both economy and society as a whole. Having said this, a sense of optimism should prevail since the studies cited at the beginning of this article indicate that an initial assessment of cultural industries using existing data can bring great benefits. All countries that have carried out such an assessment have been surprised by the extent to which such industries contribute to national economic performance. Such an initial study can help identify which industries are the most important and where the evidence requires further work to present a full picture.

Often the impact of an initial study into such cultural industries creates momentum for statistical agencies to examine the issue in greater depth. The more that the impact of cultural industries is studied using the approaches suggested in this article, the more officials and members of the public will be able to appreciate the breadth and richness of what culture has to offer a country. In addition, the more that the economic value of culture is appreciated, the more people are willing to invest in culture, whether they be public agencies or private companies who recognise the real returns that culture provides.

All these issues can only be addressed through close international co-operation. This must include partnerships between different international agencies (such as the OECD, EUROSTAT and UNESCO), professional associations (such as the International Federation of Library Associations) and national authorities for whom the resulting statistics should be relevant, timely and of high quality. Only then will we be able to measure the full impact on culture and creative on our economies and our societies.

Notes

1 The development of notions of 'creative industries' in the last ten years stems from the UK Dept of Culture Media and Sport Creative Industries Mapping Document (London 1998), with a revised version in 2001.

2 The media and entertainment industry includes filmed entertainment; TV; recorded music; radio and out-of-home advertising; internet advertising and access spending; video games; business information; magazine publishing; newspaper publishing; book publishing; theme parks and amusement parks; casino and other regulated gaming; sports.

3 The core copyright industries are industries that are wholly engaged in creation, production and manufacturing, performance, broadcast, communication ad exhibition, or distribution and sales of works and other protected subject matter.

4 It is suggested that the Framework will allow countries to choose the sectors which are most appropriate for their national policy, but that the sector definitions will be based on international standards. As such, while countries may choose different sectors, they will be able to compare themselves with other countries that choose the same sector or sectors.

5 UNESCO has published a number of works on the subject matter: Our Creative Diversity, report by World Commission on Culture and Development, 1995; Culture, Creativity and Markets, World Culture Report, 1998; Cultural Diversity, Conflict and Pluralism, World Culture Report, 2000 and International Flows of Selected Cultural Goods and Services, 1994-2003, UIS, 2005.

References

Australian Bureau of Statistics, Measuring well being, 2001.

Council of Europe, From The Margins: A contribution to the debate on Culture and Development in Europe, European Task Force on Culture and Development, 1997.

Department for Culture Media and Sport, United Kingdom (DCMS), Mapping Document, 1998.

Department for Culture Media and Sport, United Kingdom (DCMS), Creative Industries Economic Estimates Statistical Bulletin, 2006.

European Commission, Communication for a European agenda for culture in a globalizing world, 2007.

John Howkins, The Creative Economy: How People Make Money from Ideas, 2001.

OECD-1, Culture and Local Development, 2005.

OECD-2, Digital broadband content: mobile content, new content for new platforms, Working party on the information economy, 2005.

OECD, OECD Factbook 2007 - Economic, Environmental and Social Statistics, 2007.

Price WaterHouse Coopers (PWC), Global Entertainment and Media Outlook: 2006-2010, 2007.

Presidencia de la nación, Secretaria de cultura, Argentina, Cuenta satélite de cultura, 2006.

UIS, UNESCO, International Flows of Selected Cultural Goods and Services, 1994-2002, 2005.

UNESCO, Culture industries for development in Africa: Dakar plan of action, Organisation of African Unity, 1992.

UNESCO, Atelier sur la culture et le développement dans le progamme d'action du NEPAD, 2003.

World Commission on culture and development, Our creative diversity, 1995.

World Intellectual Property Organisation (WIPO), Guide on surveying the economic contribution of the copyright-based industries, 2003.

WIPO, National studies on Assessing the Economic Contribution of the Copyright-Based Industries, 2006.

Chapter 24
Measuring and Improving Government Performance

Joe Grice
Chief Economist and Executive Director, Social and Public Service Analysis and Reporting, Office for National Statistics, UK

Abstract

Given the important role the public sector plays in economies, this paper addresses the question of how government productivity might best be measured. The author argues that it is reasonable to expect government spending to become more efficient and productive over time, just as we expect rising productivity in the rest of the economy. Also, government spending should be subject to accountability. But how to do we accurately measure the value of public-sector output? The author explores applying the national accounting method of differentiation to quantitatively measure government output, and then adding to that quality adjustments that reference outcomes. The author then outlines two major levels of inquiry that must be addressed in developing an effective gauge of government output and performance, one at the theoretical level and the other on the practical level.

The last few years have seen a sharp expansion in the amount of activity devoted to measuring government output and indeed government performance more widely. In the European Union, this has become one of the major political and economic issues of the day. In the last couple of years, for example, the Council of EU Economic and Finance Ministers (ECOFIN) has repeatedly called for intensification of the effort to measure government efficiency and to learn from best practice ways to promote it. But in other parts of the world, too, the same issues are receiving enhanced attention.

Underlying this intense political interest can be discerned two sets of considerations:

- One relates to the contribution of the government and public sector to the economy overall;
- The other is more specifically related to the public sector performance per se.

These are discussed in turn.

The contribution to total productivity. In many of our economies, the public sector accounts for around 20% or so of total GDP: in some cases rather more. For the UK and for most other countries, this is more than the manufacturing sector in total. So anyone

who is interested in the overall productivity and growth of our economies, as for example, in the EU's Lisbon Agenda or in the OECD-wide Going for Growth project, has to be at least as interested in public sector productivity as in that of manufacturing. Moreover, there is a further reason why governments should be particularly interested in public sector productivity. With the private sector, they can set policy frameworks, influence incentives, persuade, cajole, educate and so on, to try to improve productivity performance. But so far as the public sector is concerned, governments and public authorities have direct managerial leverage and corresponding direct responsibility. Furthermore, many of the services that governments provide – education and training, legal and other commercial framework systems, health services and so on – directly condition the productivity performance of the rest of the economy.

Public sector performance. The key underlying issue here turns on a simple but powerful piece of political economy. In all of our societies, citizens come to expect ever increasing standards of service from government and from the public services more widely, just as they expect improving services provided by the private sector. But, at the same time, the desire for increasing taxation to pay for these services not present to the same degree. The only way to square this circle is to ensure that whatever governments can afford to spend produces the highest possible outputs, so meeting our citizens' aspirations. In other words the efficiency or productivity of government spending has to increase over time. It is not unreasonable, of course, to expect that it should do so, just as we regard rising productivity in the rest of the economy to be the norm.

Allied to this desire to meet society's expectations, governments are subject to rising demands for accountability. On the whole, governments in developed economies are reasonably good at accounting for the money that they spend, and have been for a long time. But at the same time citizens, taxpayers and users of government services – substantially overlapping groups - have a legitimate demand for accountability as to the use made of the tax revenues that governments spend on their citizens' behalf. Has the money been spent to good effect, achieving the outcomes that society collectively wants, or, at the other end of the scale, has the money been wasted?

Sharpening the urgency of both of these considerations is that fact that virtually all of our economies are ageing, and at an accelerating pace. In some cases, this is already underway. This means that age related expenditures such as on health and social care services will be subject to upward pressures that successive finance ministries will find difficult to resist. At the same time, the proportion of the population of working age - who form the main part of the tax base - will shrink significantly over coming decades. In the limit, without rising productivity of government spending, the sustainability of the public finances would be called into question.

Note also that for these public finance purposes, the scope of the issue becomes very much bigger. While the public sector typically accounts for around a fifth of GDP, governments also engage in large scale transfers through the social security system of roughly the same magnitude. Notwithstanding that they are transfers from one part of the economy to another, they all need to be financed from government revenues. Taxpayers and finance ministers will again want to know whether these expenditures are worthwhile and achieving their intended purpose.

Public Sector Productivity as Part of Overall Productivity

In this context, we are concerned here with the contribution of public sector output to overall GDP. So it is natural and unavoidable that measurement of public output and productivity, for this purpose, should take place within the well developed framework of national accounting. Indeed, it was the decision of the national accounting community in the 1980s to move away from the traditional <output = inputs> convention for measuring public sector output in the national accounts that first sparked off the current major work programmes in this area around the world. Their decision, embodied in the 1993 System of National Accounts convention, recognised that an assumption of perpetual zero productivity change for the public sector was so far from reality as to be untenable.

But what to do in the absence of the convenient "output is whatever you spend" convention? One way to think of gross domestic product is as cumulative added value by the various successive economic processes that make up national production. In the case of the private sector, there is normally a market and a market price so that it is possible to determine value and value added from direct observation. In the case of the public sector, there is usually no market to serve this function. So value has to be inferred indirectly. For typical public services such as health and education, this will not be straightforward, though not impossible.

In his *Review of the Measurement of Government Output and Productivity in the National Accounts* (2005), Sir Tony Atkinson argued first and foremost for a principled approach to this measurement task. This, he believed, was necessary, first, to ensure the intellectual integrity of public output measurement within the national accounts and to ensure consistency with the rest of the accounts, and, secondly, to guard against the problems which had arisen in Britain and elsewhere from an excessively opportunistic and incremental approach. He accepted the argument, and made it his first principle, that public sector output should be measured by reference to the value it generated.

This is a good starting point. But value normally has both a quantity and a quality dimension. How many diamonds do I have and what is the quality of each of them? In the main public services, we normally have reasonably good information on the quantity side. We know how many patients the health service has treated, how many pupils have been taught in schools and so on. But it is the quality dimension which presents the greater challenge.

Some progress can be made by use of the traditional and powerful national accounting method of differentiation. The aim is to differentiate the aggregate into groups sufficiently small that they can be considered homogeneous in all relevant aspects, including quality. To take the health service as an example, in the UK the measurement of outputs distinguishes between some 2000 or so separate treatment categories based on so-called Health Resource Groups which are clinically similar and with similar resource intensity. This is an improvement over an aggregate measurement which would implicitly treat a full heart and lung transplant, for instance, as the same in output terms as a simple in-growing toenail correction.

But this takes us only so far. It seems obvious that there are aspects of health care quality with which differentiation will not deal. One is whether the treatment leads to health gain. A medical intervention which fails due to the poor quality of the clinical input is obviously of less value to the patient than one that successfully treats him or her. This suggests that separate quality adjustments based on reference to the outcome achieved are an additional necessary step. In the field of education, the equivalent

relevant outcome might be the degree of educational attainment achieved. It would be one thing for pupils to be taught in a class room. But if the outcome was little or no educational attainment, due to the poor quality of the teaching, one would want to score the output as lower than if the same process with better quality teaching had led to increased educational attainment.

The process of considering and making quality adjustments based on outcomes forces attention as to what the desired outcomes are. What is from public services that society actually wants? In the case of schools, is it increased educational attainment that is the objective? Or is it to produce well balanced, cultured citizens that can contribute to and benefit from society as a whole? Or is it, at least in part, to provide reasonable quality childcare? Presumably, labour market participation rates, especially of women, would fall sharply if schools were not looking after children for a large part of the working day.

The Value for Money of Public Spending.

Any realistic study of the efficiency of public spending is likely to go wider than just looking at its contribution to GDP. There are good reasons why it should:

- The overall output of public spending may be of interest in itself. But citizens and policymakers are also likely to be interested in particular aspects of spending efficiency and effectiveness. Even if the aggregate output and productivity of a public service looks satisfactory, that may hide a multitude of problems in particular areas or in the delivery of specific aspects of the service;

- The national accounts are judgment free about distributional issues. GDP is concerned only with aggregate value, regardless of who benefits from it and, for that matter how widely or evenly those benefits are enjoyed. In practice, governments and societies are likely to have objectives for such considerations. Access and fairness are therefore additional considerations that need to be taken into account;

- More generally, public services and other areas where governments spend money are large and complex. It is not likely or realistic to suppose that the performance of such services could be adequately described by a single measure. A range of information will be needed.

Various bodies across OECD countries have a natural role in assembling such information, though exact responsibilities vary from country to country:

- Finance ministries have an obvious interest in such information and in some countries play a role in measuring and improving performance. In other countries, separate departments within government have principal responsibility for this;

- Parliaments, often working with and through audit bodies, have an equally key role;

- Inspection bodies, charged with spreading best practice and raising standards;

- National Statistical Institutes, not only by virtue of their responsibility for the national accounts but also via a wider responsibility to provide information and analysis on the state of society and the performance of government;

- Perhaps in some ways curiously late, but certainly energetically over the recent past the academic and research communities have become active in this field.

But having noted earlier that no single indicator or statistic can describe realistically the performance of a complex public service, it does not follow that an ever-growing wide range of disparate performance indicators, furnished by some or all of the above, serves citizens well either. Duplicating, overlapping and potentially intellectually inconsistent measures are more likely to confuse than to shed light.

In one sense, the issue here is the obverse of that for the national accounts. GDP is measured within a rigorous intellectual framework, with a high degree of comparability between countries. However, national accounting's rigour and deliberate restrictiveness mean that it can not do justice to the comprehensive measurement of public sector performance. On the other hand, the craft of measuring the value for money has developed in a very free-hand kind of way. This has yielded no problems therefore of excessive restrictiveness. But if an audit office, NSI or a finance ministry produces a set of performance indicators, benchmarks or targets for a particularly public service or area of government spending, it is not clear by what criteria they should be judged as adequate or otherwise, even in the context of the particular country concerned. Making cross-country comparisons is therefore difficult, and the criteria for constructing indicator and information sets that would allow such comparisons, are even less clear.

Two Key Tasks Ahead

The performance of government, and of the services that it finances, are a major contributors to societal well-being. Their importance is not in doubt and it is entirely right that the topic should figure prominently in a conference of this kind. As noted above, for the same reasons, there has been an explosion of interest in the subject in recent years. What now needs to be done to take this agenda forward? At least two tasks face us.

At a theoretical and conceptual level, progress needs to be made in defining and agreeing a consistent basis for measuring government performance. In some ways the issue is the same as that for the measurement of societal well-being overall. We can probably all agree that GDP is not a good or complete measure of social welfare (although, by default, it is often used for this purpose), just as measuring public sector output within GDP is not a good or complete measure of government performance. But what is the correspondingly rigorous framework going beyond the national accounts and GDP that could take us towards greater completeness?

The task is certainly a daunting one but probably not impossible. It is certainly possible to speculate on the principles that might underpin such a framework. Measures and indicators should, for example, presumably be comprehensive, covering all of the aspects of government performance in which citizens, taxpayers and service users have an interest. They should not duplicate or overlap and they should have close correspondence with what it is in each of these aspects that society actually wants to achieve or obtain.

At a more mechanical level, devices such as social accounting matrices, ands satellite accounts more generally, offer possibilities for relaxing the restrictiveness of national accounts and extending the scope while maintaining their systematic rigour. While such tools have been available for many years, their potential - certainly in the area covered by this paper - remains largely unfulfilled.

The other agenda to be taken forward is a more practical one. For the reasons noted earlier, there is unmet demand from both policymakers and citizens for better information on public sector performance and efficiency. Being able to measure public sector output and efficiency is indispensable in any programme of action to improve performance. At the same time, as also noted earlier, accelerating activity is apparent which could, if well directed, fulfil that demand:

- Many NSIs are actively working in this field. In the European Union, all Member States (with one exception that has a derogation till 2012) have been required to produce measures of public sector output for GDP purposes by end 2006, based on direct indicators;

- In a number of countries, particularly those with a tradition of an outputs approach to public spending, work has proceeded in finance and or modernisation ministries to improve output and performance information. Australia, the Netherlands, the United Kingdom, Sweden, Norway and new Zealand are all cases in point;

- The research community has shown increasing interest. For the two largest public services, health and education, the measurement of outputs and outcomes is closely allied to traditional health and education economics agendas;

- The supranational organisations have, and are playing, a crucial role. The OECD, in particular has expertise in economics, statistics and national accounting, health and social care, and education which are all key ingredients in taking the agenda forward successfully. So, too, is its cross-country comparative standpoint.

If ways can be found to take forward and coordinate these currently somewhat separate work streams, the prospect for a major advance in analysis aiding policy and policy delivery is there for the taking. Perhaps this conference will serve as a catalyst to this being realised.

Chapter 25
Overcoming Indigenous Disadvantage in Australia

Gary Banks[i]
Chairman, Productivity Commission, Australia

Abstract

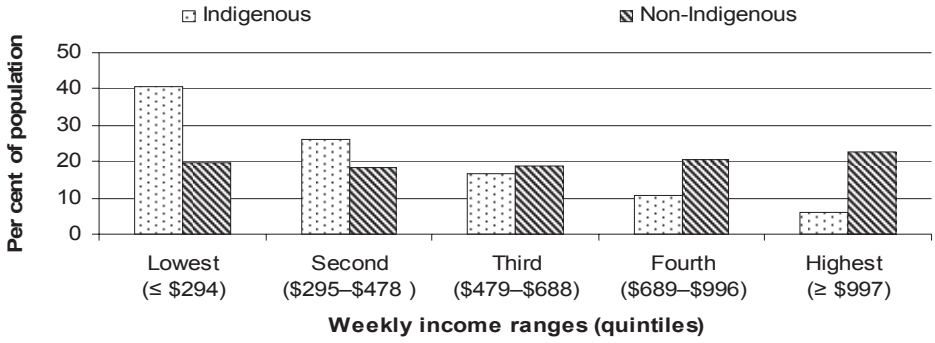
Despite ongoing policy attention over several decades, many of Australia's Indigenous people continue to experience significant disadvantage. As part of a renewed commitment by all governments within Australia, a two-yearly report is produced to map progress. The report, Overcoming Indigenous Disadvantage: Key Indicators, utilises a strategic framework, based on a preventive model covering key drivers of disadvantage, which has parallels with a 'social exclusion' approach. The latest report, just released, indicates that, while progress has been made in some key areas, particularly economic participation, other areas have shown no improvement.

Introduction

As you may know, Australia is a relatively affluent country, even by OECD standards. It has also maintained a strong egalitarian tradition. After two decades of market-based reforms to promote competition and efficiency, average incomes in Australia have grown rapidly, but their distribution remains relatively even – indeed (slightly) more so than the OECD average. Notwithstanding this overall economic success, a segment of Australian society experiences significant poverty, and this poverty is disproportionately concentrated among our Indigenous population (Figure 25.1). As the title for today's session suggests, however, poverty is only one facet of disadvantage, and the figure itself provides only a static picture of incomes. For many Australians, low incomes are a transitory or temporary phenomenon – when they first enter the workforce or choose to work part-time at certain stages of the lifecycle – but for many Indigenous people, low incomes and poverty are intergenerational. Moreover, they reflect wider disadvantage and social exclusion.

i Gary Banks is also Chairman of the inter-governmental Steering Committee for the Review of Government Service Provision in Australia, which is responsible for the Overcoming Indigenous Disadvantage Report. This paper draws heavily on the latest report, but the views expressed in it are not necessarily shared by the Steering Committee.

Figure 25.1 Income Distribution is Skewed Downwards for Indigenous Australians



Distribution of gross weekly equivalised household incomes, people aged 18 years and over, 2004-05
Data Source: SCRGSP 2007b, table 3A.6.3.

The Report I am going to talk about today, Overcoming Indigenous Disadvantage: Key Indicators, is built around a strategic framework of indicators that map thedrivers of Indigenous disadvantage. While not explicitly based on 'social exclusion', the framework has many characteristics in common with that perspective. The UK Cabinet Office has defined social exclusion as:

> *A shorthand label for what can happen when individuals or areas suffer from a combination of linked problems such as unemployment, poor skills, low incomes, poor housing, high crime environments, bad health and family breakdown. (Social Exclusion Taskforce 2007)*

Regrettably, notwithstanding the diversity of Australia's Aborigines and Torres Strait Islanders, this 'label' describes the circumstances of too many Indigenous people, as reflected in this Report.

At the same time, there is a new determination in Australia to address Indigenous disadvantage and, as I will show, this Report has been assigned a central role in that effort.

Some 'Headline' Statistics

First, a few statistics from the latest report, released earlier this month, will convey to you the extent of the challenge. Perhaps the most telling of these – and the culmination of many aspects of disadvantage – is the seventeen year gap in life expectancy between Indigenous and other Australians (Figure 25.2).

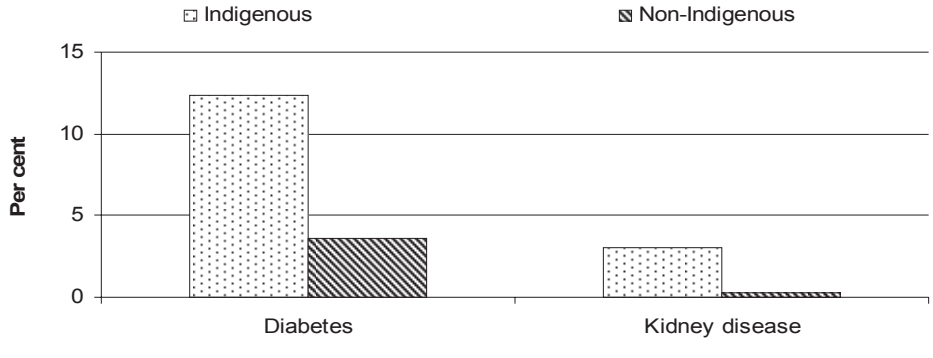
While international comparisons in this area are fraught – an opportunity, perhaps, for the OECD? – this appears to be at least double the life expectancy gaps in three comparable OECD countries with Indigenous populations (New Zealand, Canada and the United States). In those countries, the gap also appears to have narrowed significantly over time, which cannot be said for Australia.

Figure 25.2 A 17 Year Gap in Life Expectancy

Life expectancy at birth. Indigenous for 1996–2001, total population for 1998–2000.
Data source: SCRGSP 2007b, figure 3.1.1.

One significant contributor to life expectancy is infant mortality. Despite some improvement, mortality rates for Indigenous babies remain three times greater than for all Australian infants (which is about twice the difference observed in the other countries). There are also major gaps in virtually all children's health-related indicators, with the death rate for Indigenous children from 'external causes and preventable disease', for example, being five times as high as for other Australian children. Indigenous health outcomes progressively worsen into adulthood, with much higher rates of debilitating chronic diseases. For example, in 2004-05, the incidence of kidney disease for Indigenous people was 10 times higher, and diabetes three times higher, than for other Australians (Figure 25.3). The biggest difference in age-specific death rates is for the middle years – for Indigenous people aged between 35 and 54 years, the rates were 5 to 6 times those for other Australians.

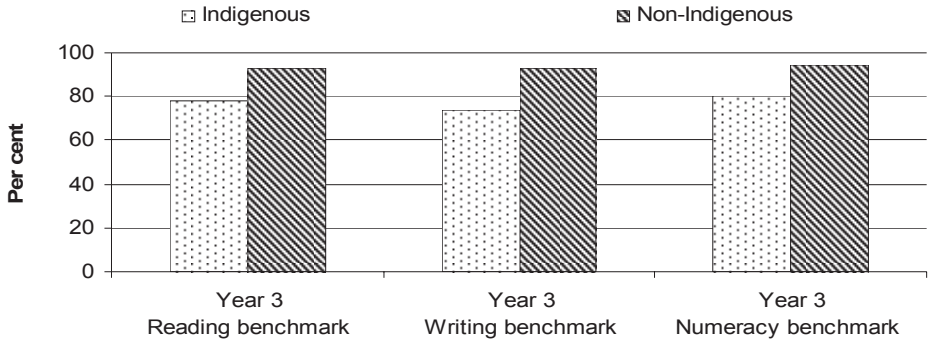
Figure 25.3 Rates of Chronic Disease are Much Higher for Indigenous People



Age standardised, 2004-05.
Data source: SCRGSP 2007b, table 3A.2.1.

Educational attainment, long seen as a key to overcoming disadvantage, is poor for Indigenous students, who are only half as likely as other students to complete secondary school, and generally have lower levels of achievement at school (Figure 25.4).

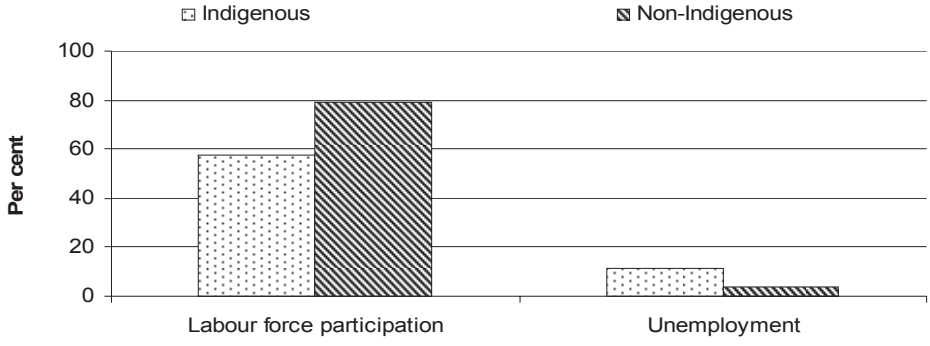
Figure 25.4 An educational Gap Appears Early



Data source: SCRGSP 2007b, table 6A.3.1; 6A.3.2 and 6A.3.3.

Poor educational outcomes are reflected in labour force statistics, with much lower participation rates (especially in private sector employment) and higher unemployment rates (Figure 25.5).

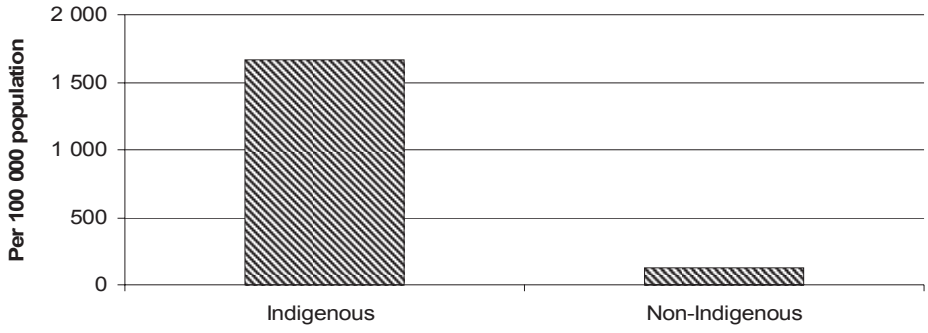
Figure 25.5 Labour Market Outcomes Differ Significantly



Age standardised; people aged 18 to 64 years, 2004-05.
Data source: SCRGSP 2007b, table 3A.5.3.

Indigenous Australians are also greatly over-represented in the criminal justice system, as both victims and offenders. In 2006, after adjusting for age differences between the two populations, Indigenous adults were 13 times more likely to be imprisoned (Figure 25.6) and juveniles 23 times more likely to be in detention.

Figure 25.6 Imprisonment rates are much higher



Age standardised imprisonment rates, 2006.
Data source: SCRGSP 2007b, table 3A.12.4.

Perhaps most distressing of all, the (reported) incidence of suicides is up to three times greater for Indigenous people than for other Australians.

Demographic Context

The intractability to date of Indigenous disadvantage could not be said to reflect a problem of numbers, as our Indigenous population currently stands at only 2.2% of the total Australian population. This is comparable to North America, but far smaller than New Zealand's Maori population (about 15 %). That said, Australia's Aboriginal and Torres Strait Islander people display some distinctive demographic features. Most notably, their lower average life expectancy, previously noted, is accompanied by a much younger population profile than for Australia as a whole. (In Australia, ageing and its impact on labour force participation and public expenditure are currently posing some major policy challenges). The younger Indigenous age profile reflects both a higher birth rate (2.1 children per Indigenous woman, compared with 1.7 for all women in Australia (Taylor 2006)), and a higher death rate, with Indigenous people on average not living long enough to be classified as 'old' even by such conventional markers as eligibility for the aged pension (Figure 25.7).

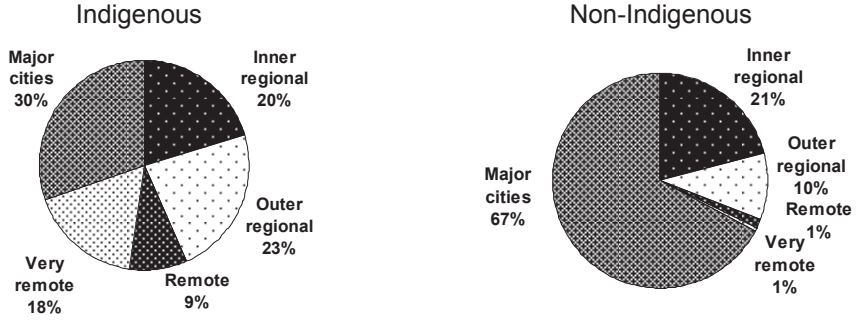
Remoteness of location is another distinguishing feature. Australia, in the words of one of our most popular poems, is a 'wide, brown land', with dispersed mining and rural activity. But it is also highly urbanised, with 90 % of Australians living in cities or inner regional centres. This applies to only 50 % of Indigenous people, however, and around 27 % of Indigenous people live in remote or very remote areas, typically in discrete Indigenous communities (Figure 25.8). In a country the size of Australia, 'remote' really does mean remote: some Aboriginal 'outstations' are several hours drive from the closest small settlement and even relatively large communities can be cut off for weeks at a time during the 'wet season'.

Indigenous people are also much more mobile than the rest of the population, with many moving between remote outstations, town camps and urban settings several times over a generation (Taylor 2006).

Figure 25.7 A Marked Difference in the Age Profile of Populations

Data Source: ABS (2001); ABS Australian Demographic Statistics (unpublished)

Figure 25.8 Indigenous People are much less Urbanised



Data for 2001.
Source: SCRGSP 2007b, figures A3.2 and A3.3.

Some Historical and Policy Background

Each people's history is unique. Nevertheless, the experiences that have marked Indigenous Australians would appear to have much in common with the Indigenous communities of many other countries – a history of conflict and dispossession, loss of traditional roles, failed assimilation and passive welfare. Last month, Australia celebrated the 40th Anniversary of the referendum that gave Aboriginal people recognition as full Australian citizens. Previously, Indigenous people were not even counted as part of the Australian population. (The Commonwealth Constitution originally stated, 'in reckoning the number of people…Aboriginal natives shall not be counted.') This statistical double

standard has an ongoing legacy even today, which will become apparent later in this paper. The referendum gave the Commonwealth Government power to legislate for Indigenous affairs, and it soon began using that power – passing antidiscrimination legislation and widening Indigenous access to workforce and welfare entitlements. It also began developing Indigenous-specific policies and institutions. To convey the shifting policy approaches over time and their mixed contributions to the outcomes that we observe today, would require more space (and knowledge) than I have at my disposal. However, it is now generally recognised in Australia that aspects of the policy approaches since the late 1960s, while well motivated and directed at desirable ends, were implemented in ways that have had some perverse, even disastrous, consequences.

In particular, equal access to statutory minimum wages and unemployment benefits effectively deprived many Indigenous people of employment, and left them dependent on welfare. As Noel Pearson, an Indigenous leader from the Cape York region, observed:

> *Everybody understood, including counsel for the Commonwealth, that this [a significant increase in unemployment in the settlements and on the Missions] would be a consequence of the equal wages decision. The Commonwealth Government's view was that the decision was right, almost regardless of the consequences. . .[but] we can never support a view, that being in a situation of passivity and dependency is the right policy. (Pearson 2003, p. 8)*

A complex array of institutions, policies and programmes have governed Australian Indigenous affairs. Since the 1967 referendum, responsibility for Indigenous affairs has been split between the Commonwealth and the States, resulting in a multilayered and fragmented mix of 'mainstream' services and Indigenous-specific services. Many Indigenous community organisations have been responsible for services that are provided by governments in non-Indigenous communities. These organisations face significant governance issues, trying to manage a large number of small scale programmes, with limited administrative capacity and uncertain funding. Several reviews of 'government governance' have found a consistent failure to acknowledge Indigenous cultural perspectives in policy design and implementation, despite acknowledgement of its importance in achieving successful outcomes.

Renewed Government Commitment and New Policy Approaches

Growing recognition of past policy failures – a recognition shared by many Indigenous leaders – together with an apparent worsening of the circumstances of many Indigenous communities, have contributed to a new commitment by Federal and State governments to overcome Indigenous disadvantage. At recent ceremonies celebrating the 1967 referendum, both the Prime Minister and the Leader of the Opposition recognised past failures and the need to do better:

> *On the 27th of May 1967, Australians said in a loud and collective voice that Indigenous Australians deserved a fair go; that the first Australians should not be second-class citizens in their own country. [But] too many of the hopes expressed so resoundingly and genuinely 40 years ago remain unrealised. . . (Howard 2007)*

> *Rather than focus on what we disagree on in this critical area – so central to our national soul – let us instead focus on what we can agree on. Let us, for example, where we can find real common ground on school retention, literacy and numeracy,*

forge a common programme. And let us work at it together – and with consistency and commitment – whoever should form the next government of Australia. (Rudd 2007)

In addition to the agreed need to forge 'whole-of-government' approaches (across portfolios and jurisdictions – see Box 1 below), key elements of the new policy approach include:

- Shifting from 'passive welfare' to 'mutual obligation', or 'shared responsibility';
- Fostering economic development and a greater role for private property;
- Improving government's ability to interact with indigenous communities in;
- Programme design and delivery;
- Improving indigenous governance;
- Recognising the need for differentiated approaches to deal with the diverse;
- Circumstances of indigenous people.

Of course, not all elements of the new approach have been universally welcomed or accepted. Some, such as changes to community land title, are highly contentious. But their introduction has been facilitated by a shared recognition by governments and Indigenous people alike that past policies and institutions have not delivered – that in important respects some have made matters worse.

Box 25.1 Some key national initiatives

The Council Of Australian Governments (COAG) is the peak intergovernmental forum in Australia, comprising the Prime Minister, State Premiers, Territory Chief Ministers and the President of the Australian Local Government Association (ALGA).

- *COAG Trials:* COAG agreed to trials of a whole-of-government cooperative approach in up to 10 communities or regions. The aim of these trials was to improve the way governments interact with each other and with communities to deliver more effective responses.
- *National Framework of Principles for Government Service Delivery to Indigenous Australians:* these principles, agreed in 2004, address sharing responsibility, harnessing the mainstream, streamlining service delivery, establishing transparency and accountability, developing a learning framework and focussing on priority areas.
- *Indigenous Generational Reform:* in April this year, COAG reaffirmed its commitment to closing the outcomes gap between Indigenous people and other Australians over a generation, and resolved that the initial priority for joint action should be on ensuring that Indigenous children get a good start in life.

Source: COAG 2007, 2004.

Just as important has been recognition by government leaders that the adverse effects of some past policies were made more damaging by governments' failure to monitor them properly and undertake timely adjustments or reform. A fundamental element of the new approach is a commitment, made at the highest political level, not only to address Indigenous disadvantage in new ways, but also to monitor and evaluate the outcomes. A key vehicle for achieving this is the *Overcoming Indigenous Disadvantage* Report.

COAG Agreement to Monitor Progress

The Report has its origins in a decision by COAG in 2002 to commission the Review of Government Service Provision to produce 'a regular report against key indicators of Indigenous disadvantage'. The Review is an inter-governmental body, comprising senior officials from central agencies of all governments. It was established in 1994 to report on the performance of a range of mainstream services across all jurisdictions in Australia. In 1997, the Prime Minister asked the Review to give particular attention to the performance of mainstream services in meeting the needs of Indigenous Australians.

The task of the new report was to 'identify indicators that are of relevance to all governments and Indigenous stakeholders and that can demonstrate the impact of programme and policy interventions'. There are, of course, many volumes of statistics detailing aspects of Indigenous disadvantage. On some counts, Australia's Indigenous people are the most researched in the world. Although valuable as sources of information, previous reports had little policy impact, and there was some initial scepticism as to whether another report could do much better in the future. What could more information contribute?

A 'Strategic' Reporting Framework

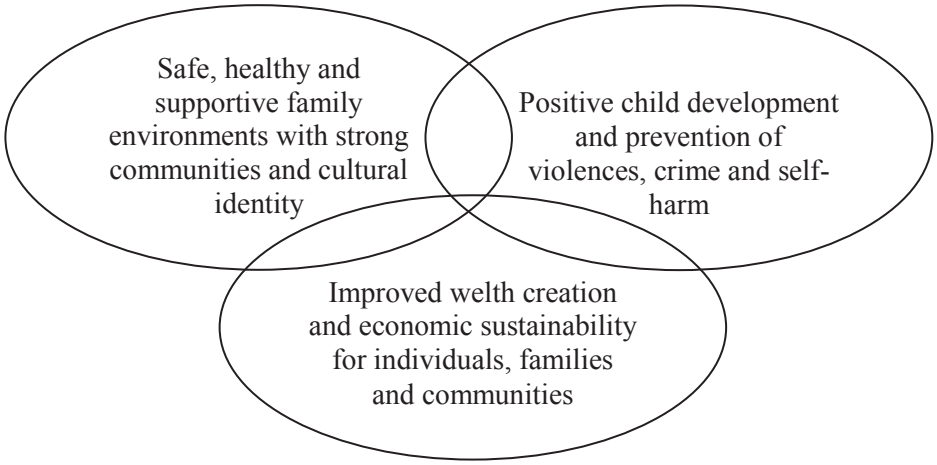
The answer lies in two features of the Overcoming Indigenous Disadvantage Report that distinguish it from other statistical compilations.

- The first is its endorsement by COAG as an ongoing vehicle for monitoring Indigenous disadvantage and the impacts of government policies and programmes. It thus has a direct link to broad policy development, which is not a feature of any other data compilation;

- The second distinguishing feature is its strategic framework (Figure 25.9). The reporting framework is based on a 'preventive model', which focuses on the causal factors that ultimately lead to disadvantage; areas where experience, evidence and logic suggest that targeted policies will have the greatest impact.

At the top of the framework are three overarching priorities, based on a report to COAG in 2000 by the Council of Aboriginal Reconciliation. The priorities are interrelated and relate to the quality of family and community life, including both cultural identity and material well-being. The vision is for Indigenous people ultimately to enjoy the same standard of living as other Australians – for them to be as healthy, as long-living and as able to participate in the social and economic life of the country.

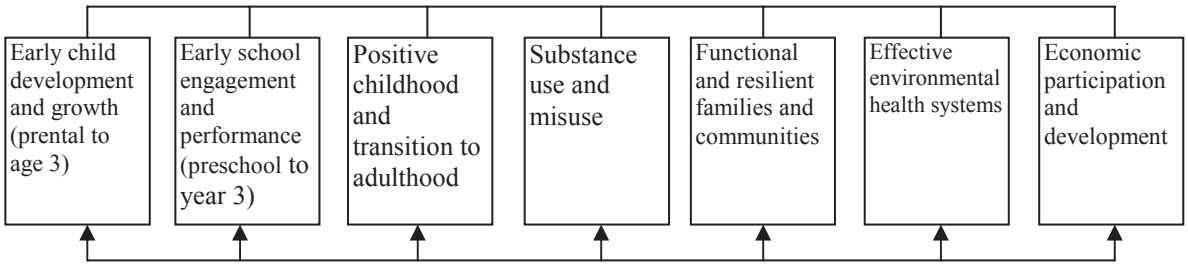
Figure 25.9 The Overcoming Indigenous Disadvantage framework



Headline Indicators

- Life expectancy - Disability and chronic disease - Years 10 and 12 retention and attainment - Post secondary education – participation and attainment	- Labour force participation and unemployment - Household and individual income - Home ownership - Suicide and self-harm	- Substantiated child abuse and neglect - Deaths from homicide and hospitalisations for assault - Family and community violence - Imprisonment and juvenile detention rates

Strategic areas for action



Data source: SCRGSP 2007b.

'Headline' Indicators Provide a Snapshot

A first tier of 'headline indicators' has been developed to provide a snapshot of how actual outcomes for Indigenous people measure up against these overarching priorities. The choice of indicators, while subjective, has generally been accepted as meaningful by Indigenous people (see below). There are a dozen indicators of social and economic status of Indigenous people relative to other Australians, a number of which I have already cited.

The framework and report could rest there, as other reports have done. However, this would not do much for policy-makers, or for those who wish to monitor programme

effectiveness. Headline indicators of this kind reflect desired longer term outcomes and most are therefore likely to change only gradually. Because most of these measures are at a high level and have long lead times (for example, life expectancy) they do not provide a sufficient focus for policy action and are only blunt indicators of policy performance. Indeed, reporting at the 'headline' level alone can make the policy challenges appear overwhelming. The problems observed at this level are generally the end result of a chain of contributing factors, some of which may be of long standing. These causal factors almost never fall neatly within the purview of a single agency of government, or even a single government.

The Innovation: 'Strategic Areas for Action'

For this reason, the framework also contains a second tier of indicators under seven 'strategic areas for action'. These areas have been chosen for their demonstrated potential to have a lasting impact on (higher level) disadvantage, and for their potential to respond to policy action within the shorter term. They assist policy makers to focus on the causes of disadvantage, with the indicators providing intermediate measures of progress. The strategic areas for action are not 'rocket science': they sensibly focus on young people, the environmental and social factors bearing on quality of life and material well-being. They – and the indicators that relate to them – have been developed with advice and feedback from governments, experts in the field and, most importantly, Indigenous people and organisations. There is broad acceptance that actions in these areas by government, in cooperation with Indigenous people, can make a difference.

A Holistic Approach

This preventive framework allows for the fact that disadvantage not only has various dimensions, it has multiple causes. That again is most obvious for life expectancy, which is the outcome of a host of influences on health and mortality across the life cycle. But the same is true for most headline indicators. For example, educational performance is shaped by a range of influences from the earliest years of life. Many Indigenous children have chronic ear infections when they first start school, which physically limit their capacity for learning. Domestic violence or substance abuse at home will clearly have a major bearing on a child's school attendance and performance. And if children are not performing adequately by year 3, they are much less likely to cope in subsequent years. This illustrates that poor educational performance, and all that flows from that, cannot be wholly laid at the door of education authorities. Responsibility for doing better needs to cross portfolios and to be at least partly borne by Indigenous people themselves.

In this sense, the Report does not promote a 'blame game'. It suggests that answers cannot be left to particular service providers to find on their own. A whole of government approach, with community support, is needed. By the same token, improvements in some individual service areas can have pervasive effects. Within the strategic area 'environmental health', for example, it is well established that overcrowding in housing contributes to adverse health outcomes, as well as domestic violence, substance abuse and, once again, school performance. It is thus an obvious target for policy action. It becomes apparent that the reporting framework has many similarities to the 'social exclusion' approach, as well as to the 'capabilities' approach championed by Amartya Sen (Sen 1983; Robeyns 2003). However, as noted, the report does not attempt to measure 'well-being' as its prime objective. Its focus is on disadvantage, a concept based on relativities – and well-being involves more than just the absence of disadvantage. That

said, the report does include some 'well-being' indicators – including self-reported feelings of happiness, and the stressors experienced by Indigenous people – and future reports may be able to include more. Reporting on Indigenous disadvantage against a framework based on causal relationships appears to be unique to this Report. While other countries have published reports comparing outcomes for Indigenous and other ethnic or racial groups within their populations, they generally have been confined to what the Review classes as headline indicators.

Broad Endorsement by Indigenous People

Although this Report was commissioned by governments, one of the key requirements was for it to be 'relevant to Indigenous stakeholders'. The Review has taken this instruction seriously, and endeavoured to involve Indigenous people at each stage of the Report's development. The draft framework was developed by a working group which included representatives from the then Aboriginal and Torres Strait Islander Commission (an elected Indigenous representative body). Before finalising the framework and collecting data, the Review conducted consultations across Australia with

Indigenous organisations, communities and leaders, resulting in some changes to the framework. Following release of the 2003 and 2005 Reports, further consultations were conducted and further changes made to the indicators. At each round of consultations, there was broad support for the Report's framework. Indigenous people endorsed the priority outcomes and generally agreed that the indicators reflected the issues that were affecting their communities and causing disadvantage (SCRGSP 2007a).

Are Things Getting Better in The 'Strategic Areas'?

The data that were available when this exercise started in 2003 were so deficient that no reliable national trends could be identified for most indicators. Two of the key functions of the report, therefore, have been to drive data improvements and to establish baseline information against which to measure future outcomes. This problem, and the relatively short period since then, has (unsurprisingly) meant that there has been relatively little movement in most of the (slow-moving) headline indicators. Where there are large changes, these tend to be off very small bases. Among the few discernable trends, some go either way. On the *positive* side:

- From 2002 to 2006, apparent school retention rates to years 10 and 12 improved — although there was still a marked drop-off between years 11 and 12 — while the proportion of Indigenous people participating in post-school education increased from 5 % in 1994 to 11 % in 2004-05;

- Between 2002 and 2004-05, the unemployment rate fell by one-third, from 21 to 13 %, and median household income increased by about 10 %.

On the *negative* side:

- From 2001 to 2004-05, there was an increase in the number of long-term health conditions for which Indigenous people reported higher rates than non-Indigenous people;

- From 1999-2000 to 2005-06, the rate of substantiated notifications for child abuse and neglect doubled for Indigenous children but rose only slightly for non-Indigenous children;

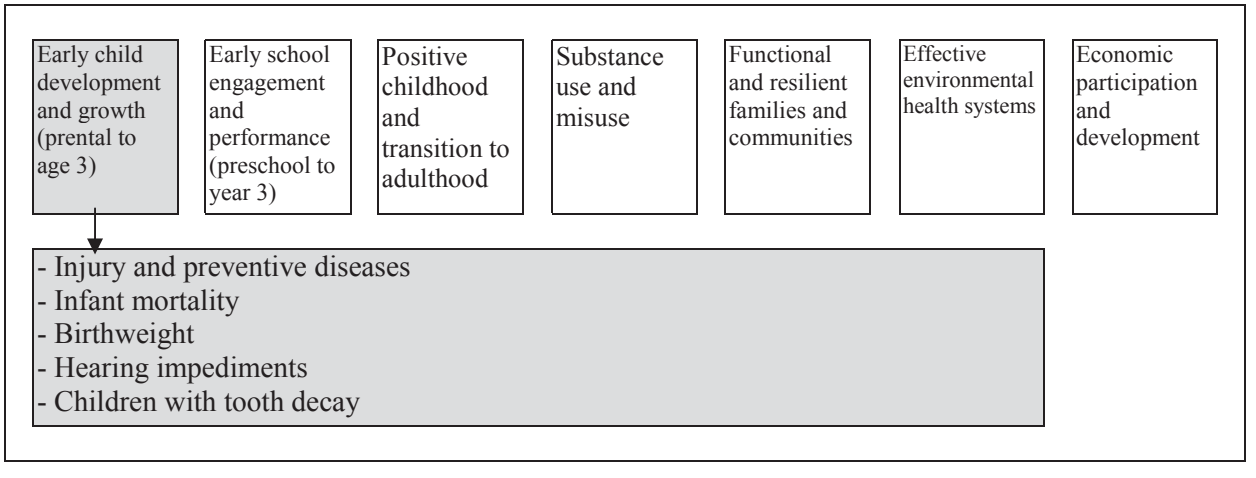
- Indigenous imprisonment rates increased by one-third between 2000 and 2006.

As noted, the 'strategic areas' and 'strategic change indicators' reported in the framework should be more amenable to early improvement. And, over time, such improvements could be expected to influence outcomes at higher levels.

Delays in data collection and publication meant that the 2007 report was largely restricted to data for 2005 and earlier, giving little scope to observe policy impacts from approaches implemented since 2003. (2006 Census data were not available for the 2007 Report, but will enable the next report to be more revealing.)

What Do the Available Data Tell Us In Each of The Seven 'Strategic Areas For Action'?

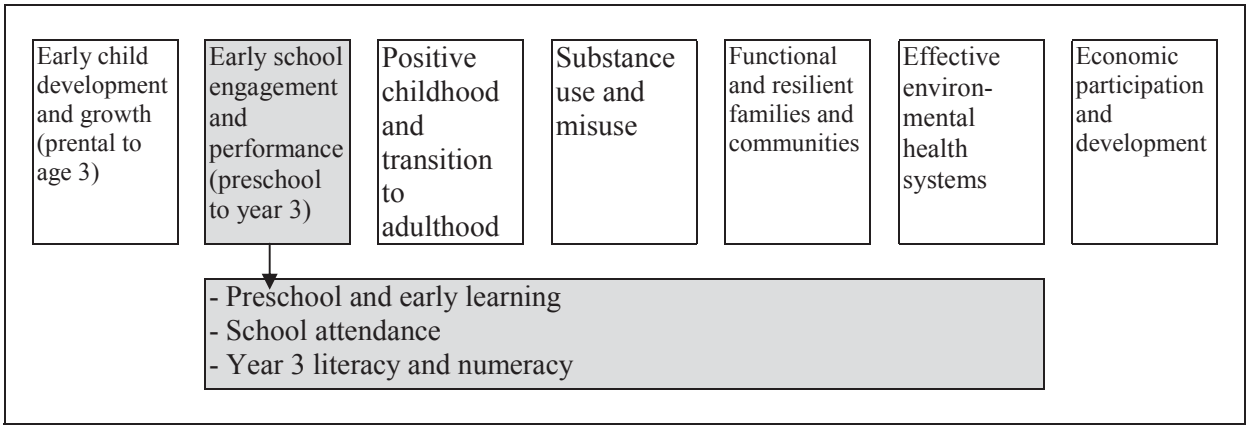
Early Child Development and Growth (Prenatal to Age 3)



The first three years of life play a crucial role in life outcomes. Stress and neglect in these early years can have significant effects on later health and educational performance, so that policy action in these early years can yield significant longer term benefits.

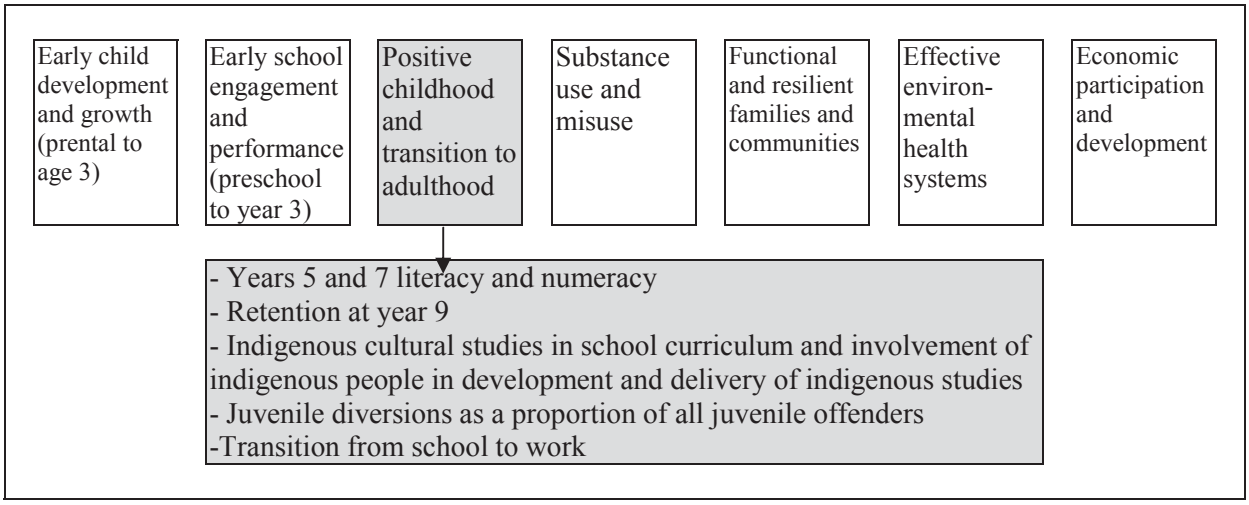
Thus far (anticipating the results in other strategic areas as well) the results have been mixed. Infant mortality rates have improved in recent years, in states and territories where data are available. Low birth-weight – correlated with subsequent health problems – did not improve. However, hospitalisations for various ear diseases declined by 25 % from 2002 to 2005.

Early School Engagement and Performance (Preschool to Year 3)



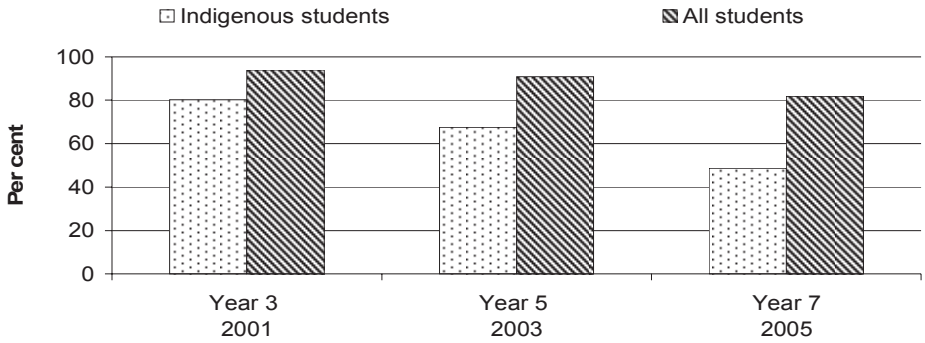
Pre-school can be particularly important to the future educational performance of children. It is therefore significant that Indigenous enrolments increased slightly between 2002 and 2005 and are at a comparable rate to those for non-Indigenous children. The story is similar for school participation rates for 5 to 8 year old children. However, it is attendance at school that really counts and, while data are lacking here – a key deficiency that has proven slow to remedy – available evidence suggests a marked disparity. This is likely to contribute to the significant gap in learning outcomes that opens up between Indigenous and other students by year 3, which had narrowed only slightly in 2005.

Positive Childhood and Transition to Adulthood



Later childhood and adolescence – and the transition to adulthood – are key points in a person's development. However, many young Indigenous people falter at this crucial stage (Figure 25.10).

Figure 25.10 Educational Disparities Progressively Widen – The Numeracy Benchmark



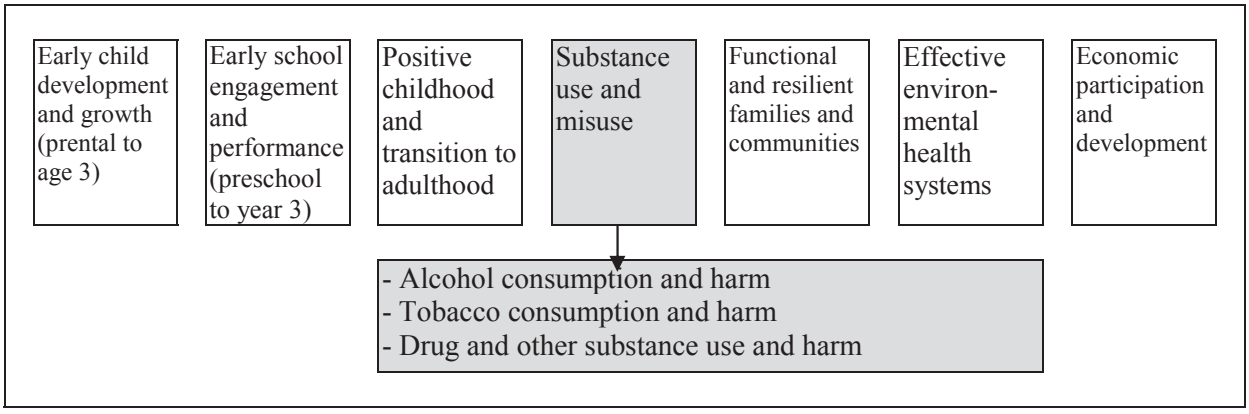
Proportions of students who achieved the numeracy benchmark. A similar cohort of students is assessed at each year level.
Source: SCRGSP 2007b, figure 7.1.15; tables 6A.3.3, 6A.3.6 and 6A.3.9.

As Indigenous students progress through school, the proportion who achieve the national minimum literacy and numeracy benchmarks decreases (while the proportions for non-Indigenous students were fairly stable, with some decline for numeracy).

In some states and territories, diversion programmes allow young offenders to be dealt with outside the traditional court processes. Despite evidence that diversionary programmes can successfully reduce re-offending, a smaller proportion of Indigenous than non Indigenous juvenile offenders had this opportunity.

In 2004-05, 40% of Indigenous people aged 18 to 24 years were neither in the labour force nor studying, compared with 11% of non-Indigenous people in the same age group, with little discernible change since 2002. Research shows that young people in these circumstances, whether Indigenous or not, are at particular risk of long term disadvantage.

Substance Use and Misuse



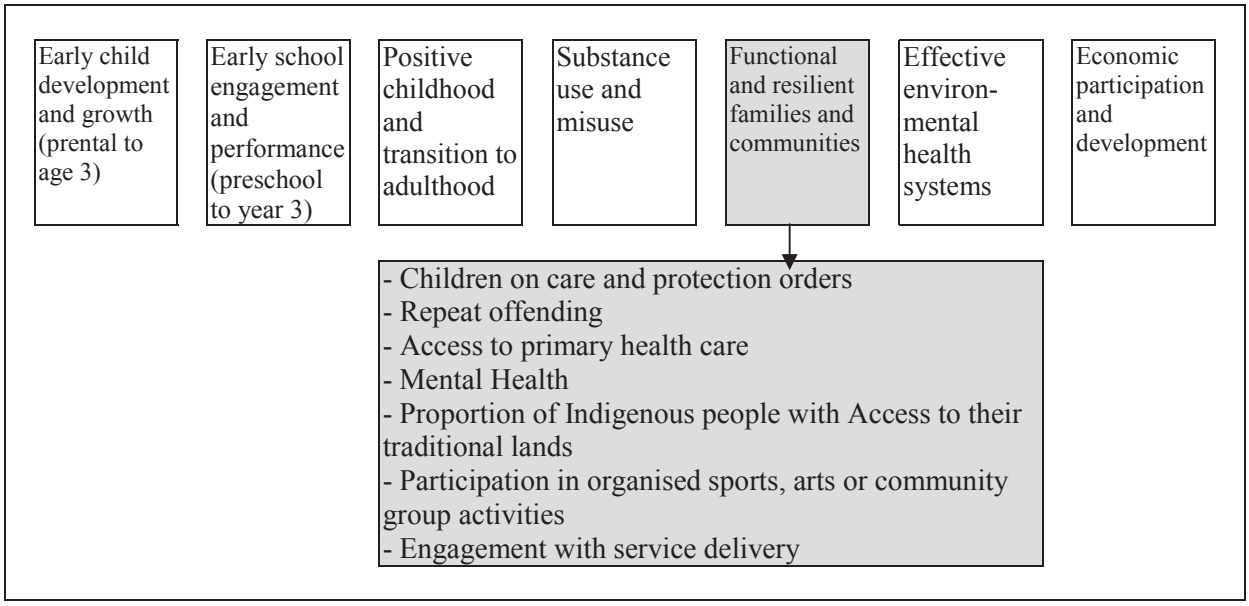
Substance use and misuse can have far-reaching effects on a person's quality of life and health, and on the well-being of those around them. Many factors have a role, including socioeconomic status, unemployment and poor education. The general situation for Indigenous people is poor. For example:

- The rate of short term 'risky' drinking for Indigenous people, at 17%, was nearly double the rate for non-Indigenous people, and had not changed since 1995.

- 28% of Indigenous adults living in non remote areas reported illicit substance use in the previous 12 months.

Fortunately, there appears to be some reduction in the scourge of petrol sniffing that has been blighting young lives in remote communities. Recent evidence suggests that the introduction of 'non-aromatic' fuels, together with promotion of alternative activities for young people, have had a major impact.

Functional and Resilient Families and Communities



Families and communities are the bedrock of any society. Indigenous leaders have argued, and research confirms, that dysfunctional families can undermine the potential for individuals to enjoy good health, educational attainment and employment. That said, the functioning of families and communities is a subjective and 'private' matter, and it is inherently difficult to develop meaningful indicators, or to collect reliable data.

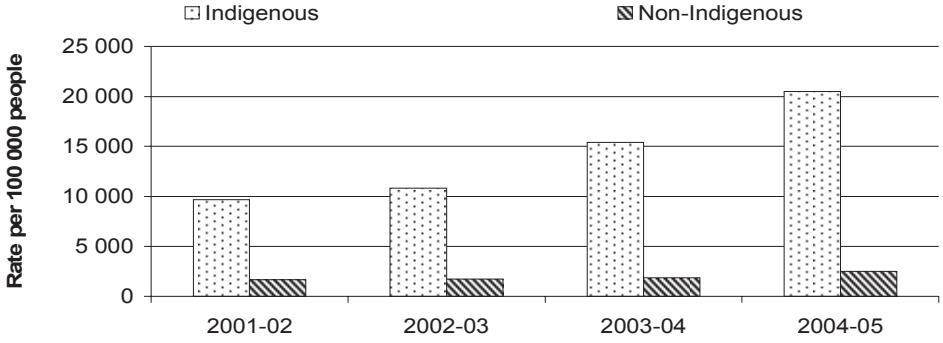
This is an evolving area of the Report, with ongoing work on how to address key elements of family and community resilience, including Indigenous cultural dimensions. However, available information does not tell a positive story. For example:

- From 1999-2000 to 2005-06, the rate of children on care and protection orders increased for both Indigenous and non-Indigenous children in all states and territories. But the rate for Indigenous children was almost six times greater.

- The impact on families and communities of high imprisonment rates, already noted, is compounded by high rates of repeat offending (74 % versus 52%). Despite the introduction of more culturally relevant sentencing options in many jurisdictions, there has been no improvement in reoffending at the national level.

- Primary health care is obviously crucial to the timely detection and treatment of illness and disease. The significantly lower expenditure on primary health care for Indigenous people (half that for non-Indigenous people), could well be contributing to the rising rate of hospitalisations for potentially preventable conditions, suggesting that more policy effort is needed in this area (Figure 25.11).

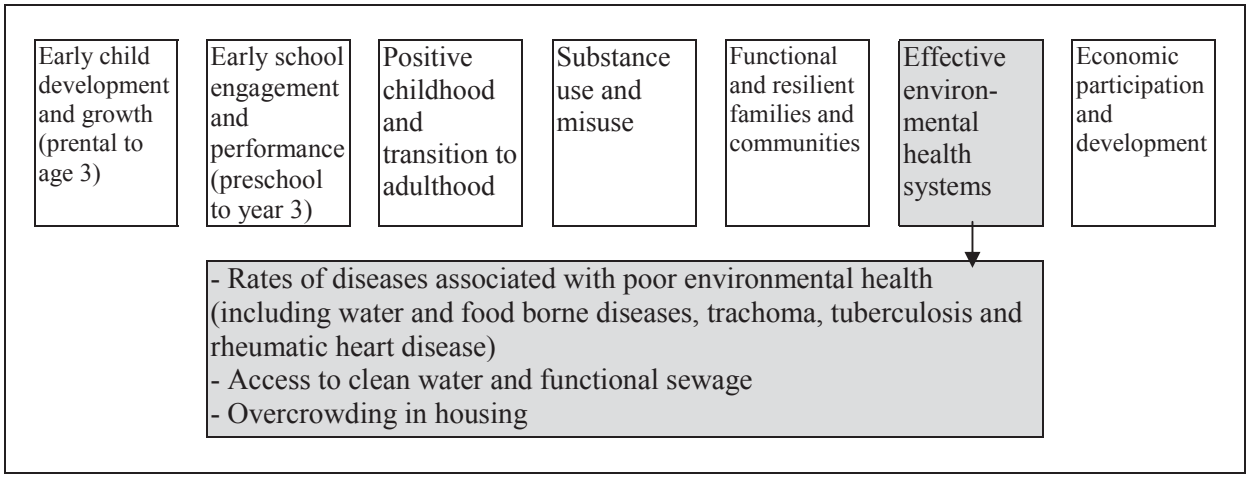
Figure 25.11 Rising disparity in Hospitalisations for Potentially Preventable Conditions



Source: SCRGSP 2007b, tables 9A.3.1 and 9A.3.2.
Data source: SCRGSP 2007b, tables 9A.3.1 and 9A.3.2.

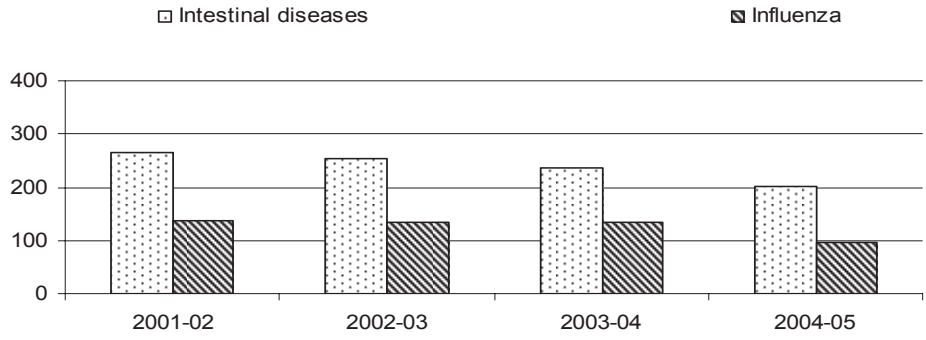
Data on the mental health of children, though limited, is concerning. 'Life stress events' are strongly associated with a high risk of clinically significant emotional or behavioural difficulties. In a large survey in Western Australia, more than one in five Aboriginal children were living in families where seven or more major life stress events, such as death, imprisonment, violence and severe hardship, had occurred in the preceding 12 months. The same survey found, however, that very few children had had contact with Mental Health Services (Zubrick *et al*. 2005). This reflects a more general issue. In 2004-05, over 26 000 Indigenous people who needed to go to hospital in the previous 12 months, did not do so, for a variety of reasons.

Effective Environmental Health Systems



The conditions in which people live and work have a major influence on their well-being and social behaviour. Sanitation, drinking water quality, disease control and housing conditions all contribute to environmental health. One indication of the relative living conditions of Indigenous people is the incidence of diseases associated with poor environmental health – up to four times as high as for other Australians. Since 2002 there has been a significant decrease in hospitalisation rates for such diseases for the 0–14 age groups (Figure 25.12), though a puzzling (slight) increase for older Indigenous people (possibly related to a greater willingness to present for treatment).

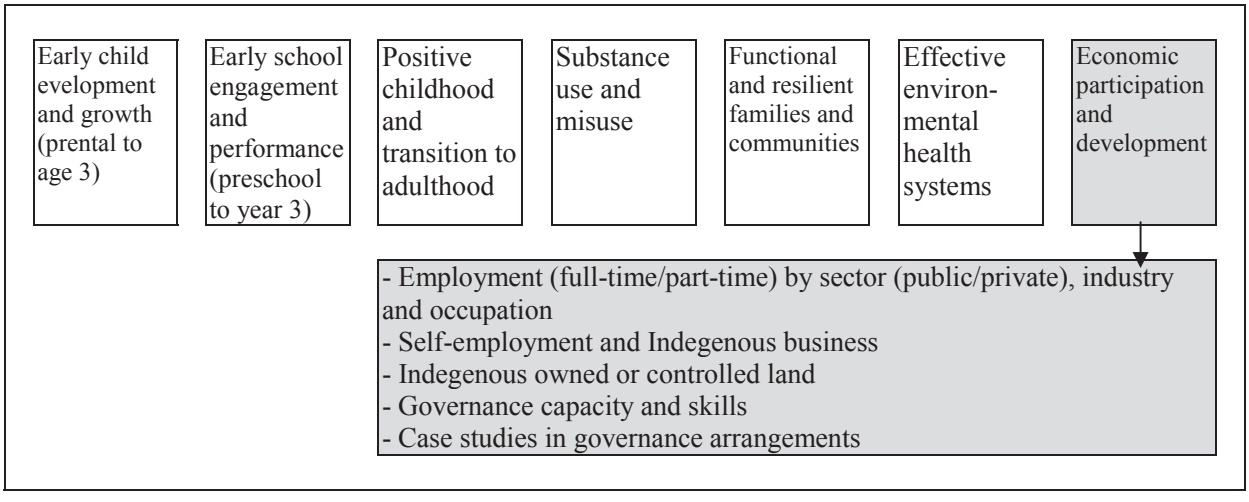
Figure 25.12 Hospitalisation of Children for Environmental Diseases has Decreased



Children 0–14 years.
Data source: SCRGSP 2007b, table 10A.1.2.

Overcrowding in housing is a particular problem, even allowing for cultural differences, and has been shown to have particularly adverse impacts on health, family violence and educational performance. Over 25% of Indigenous people were living in overcrowded housing in 2004-05 (up to 63% in very remote areas), with little change since 2002.

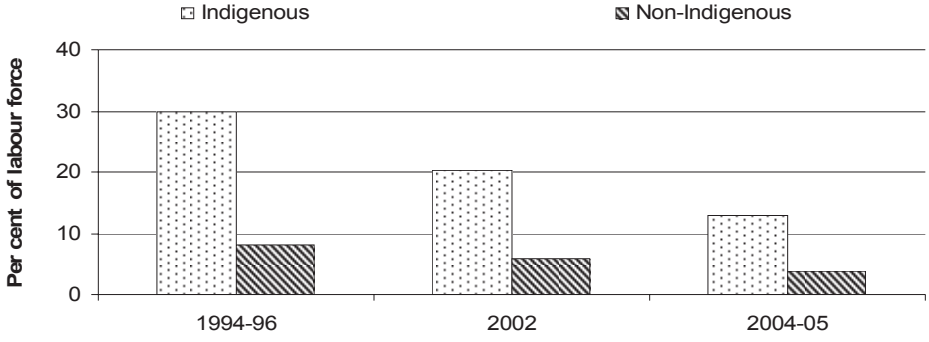
Economic Participation and Development



In the long term, all the strategic areas have a bearing on material well-being, as well as other aspects of disadvantage. However, a separate set of indicators relating to the economic participation and development of Indigenous people was seen as critical to focusing attention on what could be done in the shorter term. The extent to which people participate in economic life is closely related to their living standards and broader well-being. It also influences self-esteem, self confidence, how people interact at family and community levels, and is closely related to social exclusion.

It is heartening, therefore, that we have seen both a rise in labour force participation (more people available for work) and a decline in unemployment (fewer people not in work) over the period from 1994 to 2005 (Figure 25.13). Moreover, improvements in outcomes for Indigenous people appear to have exceeded improvements for the economy as a whole (that is, the gap in outcomes between Indigenous and non- Indigenous people has been closing).

Figure 25.13 The Indigenous unemployment rate has fallen



Non-Indigenous data for 1994-96 are total population aged 15 to 64 years, all other data are for people aged 18 to 64 years.
Source: SCRGSP 2007b, figure 3.5.9.

That said, the unemployment rate for Indigenous people is still more than three times greater on average than for other members of the workforce. And, compared to the employment market as a whole, there is a much greater reliance on publicly funded jobs – notably through the Community Development Employment Program (CDEP) a form of 'work for welfare' for Indigenous people.

This strategic area also contains an indicator relating to Indigenous owned or controlled land, in recognition of its potential economic as well as cultural value. There has been a steadily rising trend in Indigenous-controlled land. Most is in very remote parts of Australia and its potential productive value varies greatly from place to place. Scope to realise the economic potential of Indigenous land is inhibited in many cases by common property ownership and inalienable title – the effects of which have been well documented in the development economics literature.

Finally, this strategic area reflects the growing recognition of the importance of good governance to enhancing economic performance. However, capturing this in any meaningful quantitative sense is a major challenge. Initially, the Report has focused on formal training in skills relevant to capacity building in administration. In 2005, 22% of

non-Indigenous students were undertaking such training, compared to 9% of Indigenous university students. Over 42% of non Indigenous technical students (TAFE) were undertaking governance related studies, compared to 38% of Indigenous students. This appears to be an improvement from 2001, when non-Indigenous students were almost five times as likely to be studying in governance related fields.

The Report also draws on recent research into features of good Indigenous governance, based on the Harvard criteria relating to institutions, leadership, capacity building, self determination and cultural match. The Report has added the additional criterion, 'resources', and discusses the importance of 'government governance' – the relationship of government with Indigenous organisations and communities.

Our report discusses these features in some detail, drawing on numerous examples of good Indigenous governance from the national Reconciliation Australia/BHP Billiton Indigenous Governance Awards (for which I was on the judging panel last year). The winners of the 2006 Indigenous Governance Award are profiled in Box 2 below. The challenge now is to propagate these successful models.

Box 25.2 Reconciliation Australia/BHP Billiton Indigenous Governance Awards – 2006 Winners

The Indigenous Governance Awards were established in 2005 by Reconciliation Australia, in partnership with BHP Billiton. Their purpose is to identify, encourage and reward Indigenous organisations that employ sound management practices.

Gannambarra Enterprises, NSW

Develops sustainable businesses and provides opportunities for local Indigenous people to find employment in their preferred fields. Located in a regional centre (Wagga Wagga) the organisation has been instrumental in promoting reconciliation in a conservative environment. An emphasis on teamwork at all levels — Board, management and staff — and strong and candid engagement with participants has challenged stereotypes held by both Indigenous and non-Indigenous people.

WuChopperen Health Service, Qld

Formed in Cairns in 1979 as an essential health care service provider, its services include specialist clinics and chronic disease management, oral health, and social health and well-being. It also oversees the establishment of medical services and clinics in remote regions. The service manages to combine state of the art medical facilities with culturally appropriate service delivery. It is a striking model of an organisation successfully 'walking in two worlds', to the significant benefit of its community.

Source: Reconciliation Australia 2006.

In Summary

While there is still insufficient data in many areas to judge the outcomes from recent policy efforts, some of the emerging trends provide cause for hope. The pronounced gains in economic participation are particularly welcome, given the key role that employment plays in reducing poverty and social exclusion. These gains, however, are yet to be reflected in the living conditions of many Indigenous people. Importantly, there has been no real improvement in housing overcrowding, a core contributor to several aspects of disadvantage. This problem is of long standing and is an area where determined

government action should have been able to make a difference in a relatively short timeframe. It remains apparent that least progress – and even some deterioration – has occurred in those areas that are least directly amenable to government policy measures. For example, domestic violence and child abuse are difficult areas for policy intervention wherever they occur. Remoteness and the greater relative scale of these issues in Indigenous communities are additional barriers for policy intervention. This poses a major challenge for public policy simply because the answers do not depend on government alone. Corresponding efforts within Indigenous communities are also necessary. Future reports will allow us to make stronger statements about whether recent actions in these areas are having an impact on outcomes. It is already clear, however, that in all the areas identified as crucial to reducing disadvantage, outcomes fall well short of what is needed.

'Things That Work' Can Add Up

Fortunately, there is more going on in Indigenous communities than is being (or can be) captured by statistics. Our consultations across the country have brought to light many positive and successful initiatives at a local or community level. Because they are localised in their effects, they tend to be swamped in the aggregate statistics (even at the State or wider regional level). We therefore have included in the Report mini-case studies of 'things that work' in each of the target areas, to assist the dissemination of information about what is working in some communities. There is growing demand for such information – in April this year COAG agreed to establish a jointly-funded clearing house for evidence about best practice and success factors in overcoming Indigenous disadvantage.

We found clusters of 'things that work' in the areas of 'Early childhood development and growth' – reflecting government emphasis on intervening early in the life course. Many programmes focused on providing culturally relevant maternal and child health services. As noted, infant mortality rates are improving, as are vaccination rates and children's hospitalisations for preventable diseases. Several innovations have targeted school attendance. 'Things that work' include programmes linking school attendance to participation in sports activities, 'open education' programmes to support secondary school students in remote areas, and several schemes providing scholarships for Indigenous students from regional and remote areas to attend private boarding schools. Indigenous cultural studies have been introduced into some schools' curricula, with Indigenous people involved in their development and delivery. This has improved Indigenous students' self-esteem and achievement at those schools, and provides non-Indigenous students with the opportunity to learn more about Indigenous people and their perspectives.

Many 'things that work' have also emerged in the area of 'Economic participation and development', including assisting Indigenous people into jobs. Importantly, many of these have strong private sector involvement, particularly from large mining companies operating in remote areas of Australia.

Our analysis of the 'things that work', together with consultations with governments and Indigenous people, identified the following factors that many of the success stories have had in common:

- Co-operative approaches between Indigenous people and government (and the private sector);

- community involvement in programme design and decision making – a 'bottom-up' rather than 'top-down' approach;
- good governance; and
- on-going government support (human as well as financial).

Some Concluding Remarks

Indigenous disadvantage in Australia, as elsewhere, has complex causes that are inter related and cumulative. Overcoming such entrenched influences is a major challenge for public policy. Success demands a sustained effort over a considerable period of time. That has been recognised by leaders of government and Indigenous people alike, and is reflected in COAG's explicit adoption of a 'generational' perspective.

Some Indigenous leaders have gone further and argued for more explicit timeframes and targets (and recently this has been endorsed by the federal Labor Opposition) (Rudd 2007b). There is much to be said for targets as a means of galvanising action. But the value of targets depends not only on their feasibility, but also, more pragmatically, on the ability to measure progress against them. In many areas this has not been possible in the past, because data have not existed, or have lacked consistency over time or across jurisdictions. The most vexing example of this, given its importance as the lead headline indicator, is life expectancy (see Box 25.3).

Box 25.3 Measuring Life Expectancy Has Proved Difficult

While the life expectancy estimates in the Report are the best that can be compiled with currently available data, it has not been possible to present time series or trend statistics for Indigenous life expectancies, as the proportion of Indigenous people identified as such in death registrations has varied over time.

Life expectancy estimates for Indigenous Australians are sensitive to the demographic assumptions and differential quality of data across jurisdictions. The life expectancy estimates for Indigenous Australians presented in the 2005 and 2007 Reports (a 17 year gap) were not comparable to — and replaced — life expectancy estimates previous reported in the 2003 Report (a 20 year gap).

The three year difference between the newer estimates and those previously published represent improvements in methods and data quality and do not represent any changes over time in Indigenous life expectancy.

Source: SCRGSP 2007b.

As a result of COAG's commissioning of the Overcoming Indigenous Disadvantage Report, that is now changing. The Report itself does not include specific targets, but its framework would provide support for such an approach. This applies particularly to the 'strategic areas for action', which are generally recognised as being amenable to influence by governments over shorter time-frames than the headline indicators. Moreover, their narrower scope means that data definitions are simpler and measurement issues less severe, so that there would be greater confidence in the results (which is likely to be critical in the heightened politics that would surround specific targets).

Date Deficiencies Remain in Key Areas

That said, although data have improved since the first Report in 2003, information in some key areas remains poor. For example, we still do not have meaningful comparative data on school attendance, or on learning outcomes according to the degree of regional remoteness. Hospitalisation data for Indigenous people in NSW and Victoria, the two largest states, are too inaccurate to be published. A variety of other data gaps are detailed in the Report, spanning areas such as birth-weight, hearing impediments, family and community violence and environmental health. Remaining data issues mainly relate to administrative or departmental collections. Australia's national statistical agency, the Australian Bureau of Statistics (ABS), has greatly improved the extent and quality of its survey-based collections. In addition to giving greater attention to Indigenous people in the national census, a series of specific Indigenous surveys have been conducted since the mid-1990s. The 2004-05 National Aboriginal and Torres Strait Islander Health Survey was the largest national health survey of Indigenous Australians ever conducted, covering over 10 000 persons in both remote and non-remote areas. The ABS has also developed an Indigenous Community Engagement Strategy, to develop and deliver statistics in an accessible manner to Indigenous communities, and to increase their understanding of and participation in ABS collections.

At the administrative level, some laudable efforts at data improvement are also underway, assisted by expert data agencies such as the ABS and the Australian Institute of Health and Welfare. But more is needed – particularly a commitment by governments to resource the infrastructure necessary to collect and publish the data. Just as importantly, public servants and service providers need to understand the importance of the data to formulating good policies and programmes that will best assist those they are meant to serve. Many of the problems with current data collections stem from reluctance to 'ask the question'. By the same token, issues still arise with Indigenous people's willingness to self identify, because of fear of discriminatory treatment.

The need to improve administrative data has been recognised by all governments. As part of its recent Indigenous Generational Reform initiative, COAG has announced further efforts to address data gaps, including allocating significant additional funding (COAG 2007).

Some Progress in 'Embedding' the Reporting Framework

As the remaining data deficiencies are remedied, the Report will become an increasingly effective vehicle, not only to monitor change but also to drive it, by exposing where action is most needed. To realise its potential in this role, it is essential that governments integrate the reporting framework into their policy development and evaluation processes. Initially, there were some doubts about this happening, but there have been encouraging developments over the past couple of years.

Importantly, COAG has embraced the reporting framework as the centrepiece of its Indigenous Generational Reform strategy, and is building clearer links between the strategic framework, the National Framework of Principles for Delivering Services to Indigenous Australians, the COAG Reconciliation Framework and bilateral agreements between governments. The inter-governmental Ministerial Council for Aboriginal and Torres Strait Islander Affairs is also promoting use of the framework to drive action in policy and planning at the national, state and local levels. Indigenous Action Plans by

other Ministerial Councils, however, have a mixed record, contrary to the acknowledged need for whole-of-government approaches involving every portfolio.

At the federal level, the Departmental Secretaries' Group on Indigenous Affairs uses the Overcoming Indigenous Disadvantage (OID) framework to report annually on the performance of Indigenous programmes across government. Performance measures in 'Shared Responsibility Agreements' between Indigenous communities and the Government mirror the strategic change indicators in the OID framework. A number of states and territories are also incorporating the OID framework into their policy agendas and reporting. Some (Western Australia and Queensland) have produced State versions of the Report. Other jurisdictions draw on elements of the framework for strategic plans, but the links are less direct and the Report's ability to drive change in those jurisdictions is accordingly more circumscribed.

Looking Forward

Indigenous disadvantage continues to cast a shadow over Australia's otherwise undoubted economic and social achievements. Increasing affluence has made the disadvantaged circumstances of our country's first inhabitants all the more stark – and unacceptable to the wider Australian community. The now widely acknowledged failures of past policies have given impetus to new approaches. Just as importantly, past failures to correct in time policy approaches that were having adverse effects, have prompted a more integral role for monitoring and evaluation going forward.

The Overcoming Indigenous Disadvantage Report has a central place in this. Its framework of headline and strategic change indicators provides a breadth of information, helping both to measure and drive progress in overcoming Indigenous disadvantage in all its dimensions. After three reports in this series, it seems clear that, while some progress is being made, much more remains to be done.

References

ABS (Australian Bureau of Statistics) 2001, *Experimental Estimates and Projections, Aboriginal and Torres Strait Islander Australians*, Cat. no. 3238.0, Canberra.

COAG (Council of Australian Governments) 2004, *Communiqué*, 24 June 2004, http://www.coag.gov.au/meetings/250604/index.htm#indigenous. 2007, *Communiqué*, 13 April 2007, http://www.coag.gov.au/meetings/130407/ #indigenous.

Costello, Peter 2006, *Address to the Bulletin Top 100 Most Influential Australians Luncheon*, 26 June 2006, Sydney, The Honourable peter Costello, Treasurer, www.treasurer.gov.au/tsr/content/speeches /2006/012/asp.

Howard, J. 2007, *Address to 'Their Spirit Still Shines', Commemorating the 40thAnniversary of the 1967 Referendum*, Old Parliament House, Canberra, http://www.pm.gov.au/media/speech/2007/speech24341.cfm.

Pearson N. 2003, *Our Right to Take Responsibility*, Gold Leadership Paper, Institute of Public Administration Australia, Victoria. Productivity Commission 2005, *Economic Implications of an Ageing Australia*, Research Report, Canberra.

Reconciliation Australia 2006, *2006 Indigenous Governance Awards: The Winners*, http://iga.reconciliation.org.au/igawards/winners.html.

Robeyns, I. 2003, *The Capability Approach: An Interdisciplinary Introduction*, Training Course, 3rd International Conference on the Capability Approach, Italy, http://www.fas.harvard.edu/~freedoms/teaching/CAtraining20031209.pdf

Rudd, K. 2007a, *Future of Reconciliation: 40th Anniversary of the 1967 Referendum*, Speech 27th May 2007, The Honourable Kevin Rudd, Leader of the Opposition, http://www.alp.org.au/media/0507/speloo270.php.

2007b, *Commemorating the 40th Anniversary of the 1967 Referendum*, Media Statement, 27th May 2007, The Honourable Kevin Rudd, Leader of the Opposition, http://www.alp.org.au/media/0507/msloo270.php.

SCRGSP (Steering Committee for the Review of Government Service Provision) 2007a, *Framework for Reporting on Indigenous Disadvantage: Report on Consultations 2006*, Productivity Commission, Canberra.

2007b, *Overcoming Indigenous Disadvantage: Key Indicators 2007*, Productivity Commission, Canberra. Sen, A. 1983, *'Poor, Relatively Speaking'*, Oxford Economic Papers no. 35, Oxford University Press.

Taylor, J. 2006, *Population and Diversity: Policy Implications of Emerging Indigenous Demographic Trends*, CAEPR Discussion Paper No. 283/2006, Canberra.

UK Cabinet Office Social Exclusion Taskforce 2007, *Context for Social Exclusion Work*, http://www.cabinetoffice.gov.uk/social_exclusion_task_force/context/.

Zubrick, S.R., Silburn, S.R., Lawrence, D.M., Mitrou, F.G., Dalby, R.B., Blair, E.M., Griffin, J., Milroy, H., De Maio, J.A., Cox, A. and Li, J. 2005, *The Western Australian Aboriginal Child Health Survey: The Social and Emotional Wellbeing of Aboriginal Children and Young People*, Curtin University of Technology and Telethon Institute for Child Health Research, Perth.

Chapter 26
Trends in Worker Requirements and the Need for Better Information to Make More Informed Decisions in a Global Economy

Randall W. Eberts
W.E. Upjohn Institute for Employment Research, United States

Abstract

*Studies have shown that growth in per capita GDP and productivity are highly correlated with worker skills. However, countries have limited knowledge of the competencies of their workforce, especially since the skill requirements for many jobs are increasing and becoming more complex. This paper provides an overview of what we know and don't know about the change in skill requirements, what additional information is needed to make informed decisions, and how public policy should respond, including strategies for narrowing the skill gaps. Included is PISA (the OECD Programme for International Student Assessment), and various studies conducted in the United States and the United Kingdom. Also reviewed are the two systems developed by the US and the UK to track skills within occupations: O*Net and NVQ, respectively*

Introduction

Increased globalisation and rapid technological change are transforming how people view their world and see their own challenges and opportunities. Workers are concerned about job security, skill requirements, and future wage growth; businesses are scrambling to understand their competition, which could come from any corner of the world, and how to compete successfully; and governments are seeking to understand the needs of their workers and businesses in order to design effective policies to improve their comparative advantage. Unfortunately, our world is changing faster than the development of tools to understand it. National statistical accounts were not designed to account for such an integrated world in which supply chains literally reach around the globe. Even the calculation of a country's GDP and productivity growth may be biased by the way in which import prices are calculated and incorporated into national accounting frameworks (*Business Week* June 18, 2007).

Our understanding of what drives a country's economy in this new world order is also incomplete. It is difficult to determine what gives one country an advantage over another when the world economy is so closely linked and interdependent. We do know, however, that one critical area is a nation's human capital. Studies have shown that growth in GDP

per capita and productivity are highly correlated with worker skills. However, we need to know much more about the demand and supply of skills within and across countries and over time than we currently do. Workers, businesses, and governments need more information on the demand for skills, the stock of human capital, the nature of skills shortages, the causes and consequences of technological change, globalisation and demographic factors on skill supply and demand in order to make better decisions and policies. Information is emerging to help shed light on these issues, but the data are not comprehensive in scope and systematic in the way they are collected across countries. OECD's Programme for International Assessment of Adult Competencies (PIAAC), for example, offers an opportunity to provide such information, by identifying and measuring differences across countries in adult competencies, by assessing the consequences of competencies on economic performance, and by gauging the performance of education and training on generating competencies.

This paper provides a brief overview of what we know and don't know about the change in skill requirements, what additional information is needed to make informed decisions, and how public policy should respond, based on the limited information available. Space limitation does not permit a comprehensive review of these topics. Each country has its own unique perspective on these issues. The world looks different from the viewpoint of a developed country compared to that of a developing country; yet the information needs are similar. My intent is to touch on a few common themes and to use my experience in the US and my knowledge of the UK to illustrate these points.

Skills and Economic Performance

Studies at various levels of aggregation--from the individual worker, to firms, to metropolitan areas, to countries – show that skills are positively related to market outcomes. Hanushek and Kim (1995) find a statistically and economically significant positive effect of the quality of education on economic growth (real GDP per capita) in 1960-2000. They estimate that, within their 50-country sample, a one country-level standard deviation higher test performance would yield around 2% higher annual growth rate. The sample includes both developed and developing countries. They use a composite measure of the results from mathematics and science tests given to teenagers in school, obtained from OECD's Programme for International Student Assessment (PISA). This measure of educational attainment is an improvement over measures of the number of years of schooling or an earned degree or certification, used by a majority of studies prior to theirs. Hanushek and Kim show that the number of years of schooling is a poor proxy for skill acquisition. School systems differ in their curriculum and their effectiveness in providing educational services, and students differ in their ability to learn. Their estimates using test scores from PISA dwarf those using years of schooling, and illustrate the need for a more accurate measure of skills than what has traditionally been used.

Educational attainment is also highly correlated with productivity growth, which is essential for a country's future growth in wages and standard of living, and to compete successfully on the global stage. For example, the UK is particularly concerned with its lagging productivity performance (Campbell 2006). It is the world's fifth largest economy as measured by GDP but 13th with respect to global competitiveness and 18th in terms of GDP per capita. Many attribute UK's gap to relatively poor skill levels among its workers. 14% of its workforce does not have any qualifications. Only 19% have the literacy skills expected of an 11-year old and only 50% have the numeracy skills expected

of that youngster (Campbell 2006). This has become the "call to arms" in that country to improve their adult worker training system and to reform their traditional educational system.

The benefit of skills is also evident at the sub-national level. Metropolitan output and employment growth is highly correlated with educational attainment (Eberts *et al* 2006; Glaeser & Saiz 2003) for the US. In addition, education and training have paid off for both businesses and workers during the past decade or so. For businesses, increases in educational attainment were responsible for an estimated 11 to 20% of the growth in worker productivity in the US (Decker *et al.* 1997). For workers, studies estimate that an additional year of education increases a worker's annual wages by 6 to 10% (Card 1999). The earnings gains are even more pronounced when the educational curricula include an academic year of more technical and applied coursework, with returns ranging from 10 to 15%. Basic literacy and numeracy skills also have significant payoffs. Studies, for example, have shown that high-literacy workers can earn nearly three times that of low-literacy workers. In terms of annual increases, the study found that an increase in literacy proficiency raised expected annual earnings by 15% (Sum 1999). In addition, the type of skills acquired in college matters. Jacobson, Lalonde, and Sullivan (2001) estimate that an academic year of more technically oriented vocational and academic math and science courses raise earnings by about 14% for men and 29% for women, whereas less technically oriented courses yield very low and possibly zero returns. These estimates are for displaced workers receiving training at community colleges.

The returns to worker qualifications are substantial in the UK as well. For example, males with a degree earn up to two-thirds more than unqualified workers. Studies also find that the returns to academic qualifications are significantly higher than the returns to vocational qualifications. Studies in other developed countries find similar returns. For example, using the Canadian portion of the seven-country International Adult Literacy Survey (IALS), Green and Riddell (2001) found that for every ten-point increase in the literacy score earnings of Canadian workers increase 3.3%.

Non-cognitive skills are also correlated with earnings. Heckman and Rubinstein (2001) first estimated the importance of non-cognitive skill and found that high school equivalency recipients earn less than high school graduate despite being smarter. While not observed directly, they attribute this result to lower non-cognitive attributes of those who drop out. In a more recent paper, Heckman, Stixrud and Urzua (2006) find that better non-cognitive skills lead to more schooling, and to an additional earnings return. Carneiro *et al*. (2006) find non-cognitive skills measured in childhood to have similar effects in the British 1958 National Child Development Study. Non-cognitive skills should not be dismissed as superfluous in meeting the needs of business. On the contrary, the growth in the service sector, particularly those occupations that require interpersonal skills with customers, places a greater emphasis on these so-called soft skills.

Trends in Skill Requirements

Increased globalisation and rapid technological change have led people to claim that the skill requirements of jobs are much higher today than ever before. One can point to many occupations which require substantially different if not higher skill requirements than they did a few decades ago. Secretaries, for example, who once found pride in the number of words per minute they could type (on a manual typewriter) or the amount of shorthand they could take, must now master complex word processing programmes and

navigate the web. Assembly line workers have replaced strong backs with dexterous minds to operate intricate automated systems.

Some argue, on the other hand, that advances in technology have made performing jobs no more demanding than in the past and some occupations require less quantitative skills. Clerks at McDonald's fast food restaurants, for example, select images of food on their cash registers instead of remembering and punching in the price of each item. The register displays the correct amount of change to be returned to the customer instead of relying on cashiers to compute that amount in their heads. In retail stores, most items are scanned, so there is no need to record numbers or to transfer information from one form to another. Furthermore, a large share of transactions are paid through credit or debit cards and these require even fewer skills – swiping the card through a card reader and waiting for the receipt to print.

There is also a deepening of the tasks expected of workers in certain occupations. Aided by technology, workers are expected to do more and to do it better than before. Secretaries are expected to produce documents that resemble printed pages instead of cruder looking typewritten pages, which 30 years ago would have required the collaboration of graphic artists, layout artists, and professional printers. Another example is in manufacturing. All production workers in North America's newest, state-of-the-art auto engine plant share the same job classification. Each worker is expected to know how to do everything, at any time, and at any place in the plant. In other less technologically advanced engine plants, workers typically specialise in a limited number of tasks, and the workforce is distributed among a dozen or so occupational classifications.

Have the demands of jobs changed? The answer is definitely yes in terms of what tools are available to carry out the task, as we see with the use of sophisticated computers and other electronic devices rather than manual typewriters. Have the skill requirements increased within each specific occupation? The answer to that question is not as clear, particularly when skills are measured according to literacy requirements (reading and quantitative). Unfortunately, very few, if any, studies actually trace the skill requirements of specific occupations over time (Barton 2000). For those that have, they find little change. What they do find is that the fastest growing jobs (occupations) require higher skills than the fastest declining jobs. Consequently, the skill mix of the total stock of jobs has increased, and with that compositional change the demand for higher skills has also increased. There are still jobs requiring lower skill levels, but their share of total employment has declined. Some occupations totally disappear while others are created. When job requirements drastically change as a result of technological change, for example, the name of the occupation often changes, so it is difficult to follow specific occupations over long periods of time.

The same compositional change is occurring in the UK. Those sectors with the highest qualifications have grown the fastest in the past decade and are expected to continue to grow faster than sectors with lower skill requirements. For example, computer and related activities, whose workforce has the highest skill levels of the 27 broad business sectors, has grown the fastest in the UK over the past decade. On the other hand, textile and textile products, the sector with the lowest skills, has experienced the greatest decline in employment in the UK. Therefore, much of the increase in skill requirements results from a compositional change in occupations towards those with higher skill requirements.

While it is apparent that compositional changes have lead to a greater demand for skilled workers, sufficient information is not available to follow the changes in skill

requirements within occupations. In the US, the U.S. Department of Labor developed the Dictionary of Occupational Titles (DOT) in 1939, which attempted to identify the skill content of hundreds of occupational classifications, but much was done subjectively. Recently, the US Department of Labor replaced DOT with O*NET, a new classification system which surveys workers in specific occupations about the skill content of their jobs. However, neither offered a longitudinal perspective on the change in skill requirements

Some studies have followed over time the qualifications of those who hold jobs in specific occupations. Using this information to assess skill requirements assumes a good match between employer needs and worker qualifications. Some work shows that this assumption may not hold. As Barton (2000) describes "College graduates are in occupations formerly held by high school graduates, and high school graduates are in jobs formerly occupied by people without a high school diploma." (p.33). Hecker (1992) stated that "it is not possible to precisely identify and measure the number of jobs that require a college degree." Adkins found that "if we take the 1940 level of educational attainment in individual occupations and (roughly) calculate the proportion of the total number of male college graduates in 1969 that would be needed to meet 1940 educational attainment standards for occupations, we will account for only 45% of the stock of male college graduates in 1969."

But the fact remains that the US economy has successfully absorbed college graduates, even as the supply increased dramatically. Not only are college graduates finding jobs, they are also paid at higher rates than high school graduates, many of whom filled the same jobs a few years earlier. The market results favoring college graduates lead Tyler, Murnane, and Levy (1995) to conclude that, even though occupations once held by high school graduates are now held by college graduates, workers holding these positions are not over-educated. Others have argued that employers have turned to college graduates to fill their positions, not because they have the additional skills required, but because a college degree signals other employability traits. It is fair to conclude that it is not possible to track the change in skill requirements within occupations.

Therefore, sufficient data are not available to determine how much of an increase in skill requirements is attributable to increases within occupations and to compositional change. This is an important distinction when trying to design educational and training programmes to accommodate these changes.

Skill Shortage

By most accounts, the supply of skilled workers is not keeping pace with changes in demand due to technology and globalisation. This conclusion is supported by the compositional shift toward occupations that demand higher skill requirements and the wage premia of higher-skilled workers. These trends are evident in the US, the UK, and other countries. These trends also support what businesses in many countries have been saying for years. In the US for example, the 2001 American Academy of Management Association Survey on Workplace Testing reported that one in three job applicants tested by employers lacked the basic skills necessary to perform the jobs they sought in 2000. In the UK, businesses cannot fill a quarter of their job vacancies because of skill shortages (National Employer Skills Survey 2005). This skill crisis is not new, however. Businesses complained of a skill shortage during the boom of the 1990s, the recession of 2001, and today.

Manufacturing in the US and UK is hit harder than most other sectors. A recent study of skill shortages in manufacturing, sponsored by the National Association of Manufacturers in the US, revealed that 80% of companies surveyed were impacted by skill shortages (2005 Skills Gap Report – A Survey of the American Manufacturing Workforce). 80% of manufacturers who returned the survey responded that they expected skilled production workers to be in short supply over the next three years. During this period, strong basic employability skills and technical skills will be needed the most. Despite these shortages, 74% of respondents indicated that a high-performance workforce will be key to their business success. Yet, the majority of companies (64%) surveyed formally train less than 60% of their workforce.

In the UK, upwards of 45% of manufacturing firms report skills shortage vacancies, and skills gaps account for between 4 and 8% of employment. Basic metals producers and metal fabricators report over 40% of their job vacancies unfilled because they cannot find qualified candidates. Retail trade and the regulated industries of electricity and gas suppliers have the least difficulty finding qualified candidates; yet, one in five vacancies in these industries go unfilled because of skills shortages, according to the respondents.

In looking at the various sectors within the UK economy, those industries with the highest skill requirements are least affected by skill shortages, as measured by the percentage of vacancies due to skill shortages. For example, computer and related activities is ranked toward the bottom of skill shortage related vacancies (22nd) but has the highest percentage of workers with a National Vocational Qualifications (NVQ) of three or higher and is second highest with respect to the percentage of the workforce with high qualifications (equivalent to a tertiary education). Not surprisingly, the computer sector is also second highest with respect to median weekly earnings and employment in that sector has also grown the fastest between 1994 and 2004. It is expected to continue to lead the other sectors in employment growth through 2014. High wages and future growth prospects make this sector attractive to workers from other sectors and students planning a career. Evidence shows that the fastest growing sectors are attracting younger workers.

A recent survey of UK manufacturers found that only 60% of respondents provided training to their employees, and only 35% provided two days of training within the previous 12 months before the time of the interview. This same study shows a large difference in survival rates of business between those that provide training and those that don't. What is most interesting is that those reporting the highest proportion of skills-related vacancies provide the least training for their workers. There appears to be a disconnect between those sectors who need to train workers to meet their needs and their ability or willingness to do so. But those sectors also tend to experience lower productivity growth rates and offer lower wages.

Wage and employment growth by occupation is more complicated than a simple demand and supply effect. It may also result from the interaction between technological change and worker skills. Autor, Levy, and Murnane (2003) conclude that information technology complements highly educated workers engaged in abstract tasks, substitutes for moderately educated workers performing routine tasks, and has less impact on low-skilled workers performing manual tasks. This results in an increase in earnings growth that favor high skill workers and low-skill workers relative to middle-wage workers. Similar patterns are found in employment of these three groups of workers. Autor and Dorn (2007) find rapid growth of employment in the highest skilled occupations – managerial and professional specialty occupations – between 1980 and 2005, gaining

7.1% as a share of overall employment between 1980 and 2005, a 30% increase. By contrast, they show a u-shaped pattern for the 'middle skill' group of technical, sales and administrative support occupations over this period, expanding in the 1980s and then contracting to below its initial 1980 level over the next 15 years (consistent with the growing substitution of technology for routine tasks). Most strikingly, employment shares in three of the four low-skill occupations fell sharply in each decade.

Filling the Skill Gaps

There is little doubt that the US and the UK experience skill shortages for various sectors and for their overall economies. Other countries have similar experiences. However, in order to understand the relative supply and demand of skills within any one country, one has to look beyond that country's borders. Integration of the world economy has increased the supply of both low-skill workers and high-skill workers for more developed countries, with the emergence of many developing countries into the world economy and the full participation of former Soviet Union bloc countries and China and India in global trade. Freeman (2006) points out that globalisation makes forecasting skill shortages or surpluses in the US or in any country more difficult than in the past. One must understand not only the internal dynamics of a country's labor market, but also the effect of world supplies and demand on that country's domestic labor market. Weighing these factors is difficult and requires an understanding of both the skill quality of workers entering the labor market and the nature of the stock of skills of workers already in the workforce.

Measuring Worker Competencies

Workers enter a country's labor force in one of three ways: 1) as native workers after completing some level of schooling, whether it is secondary or post-secondary; 2) as native adults who have not participated for a while or perhaps never, such as working mothers who return after child rearing, or retirees who decide after a few years to go back to work, perhaps part-time, or 3) as immigrants. Assessing the skill competencies of these groups of workers is important for understanding the ability to fill the skill requirements of employers, but information available about these three groups varies.

General information exists on the schooling of workers entering the workforce after completing school. We know the number of years of schooling, the percentage enrolled in college and the degrees earned. For countries that participate in PISA and other international test surveys, literacy scores are also available. This information, combined with labor market data, such as wage premia and absorption rates, is helpful in understanding the trends that specific countries experience.

While information about schooling is useful in following general trends, it is difficult to use to compare competencies across countries and over time. The curriculum content of educational systems differs across countries, as do the expectations regarding the rate of normal student academic progress. In addition, curricula changes over time so that a college degree today may not be comparable to one 30 years ago. A more systematic approach is needed. The International Adult Literacy Survey coordinated by OECD provides such a benchmark by measuring the literacy levels across selected countries. In addition, several countries have initiated their own surveys. The UK's Department of Education performed three skill audits, in 1994, 1998 and 2003, which provided a longitudinal perspective on the skills of their workers compared with those in US, France,

Germany and Singapore. These countries were chosen because their productivity levels and economic growth rates were higher than the UK. The purpose was to allow UK policy-makers to observe how stocks of qualifications for different age groups compared with other countries. They found that UK persistently had a higher proportion of low-skill workers than the others.

Including different age groups in cross-country skills surveys of workers is important in understanding the trends in competencies of different age groups and to find ways to provide additional training for incumbent workers. For instance, projections suggest that the US will have 13% more people between the ages of 18-23 in 2050 than it did in 2005, whereas Japan and Western Europe will have 13% fewer. The US share of 18-23 year olds in advanced countries will continue to trend up. The same is true for the older cohort (15-59 years of age) and ratio of the Chinese population to the US population in this age group is expected to remain constant during this time (Freeman 2006).

Different vintage workers possess different skills, and studies show that it is initial education and training completed before individuals enter their mid-twenties that produces higher qualification levels. Once individuals reach their thirties, few if any gain academic or vocational qualifications at a level above which they have already attained (Porter & Campbell, p.42). The International Adult Literacy Survey (IALS) and other surveys of adult training show that workers with higher education and skill levels are more likely to receive training than those with lower skills. Thus, adult training has a tendency to increase the skills of the already skilled instead of enlarging the pool of highly skilled workers.

It is also important to distinguish between ethnic groups, gender, and immigrants. Because of differences in culture and asset wealth, ethnic groups and immigrants differ in their access to education and completion. The same is true for women, but for other reasons. In the US, African American and Hispanic populations have lower high school completion rates and college enrollment rates than white and Asian populations, whereas women have higher rates. These tendencies can significantly affect a nation's skill levels.

Immigration is a powerful factor in understanding skill competencies. Large numbers of immigrants (legal and illegal) cross international borders every year, and these movements can have a profound effect over time on the composition of a countries workforce. For example, projections for the US reveal that by 2050 the US population will be comprised of 26% Hispanic, 14% black and 8% Asian, compared to 5% Hispanic, 12% black and 1% Asian in 1970. Countries experiencing slow if not negative population growth rates will increasingly depend upon immigrants to fill job vacancies. Europe's population is set to decline over the next 50 years. Italy will lose 28% of its population by 2050. In order to maintain its working age population, Italy would need to start importing more than 350 000 immigrants per year or, alternatively, keep its citizens working until they are age 75 (Hall 2000).

Assessing Skill Requirements

Measuring skill competencies of the workforce is only half the equation. The other half is to understand the skill requirements of the jobs that need filled. The UK has developed a system called the National Vocational Qualifications (NVQ) that reflects the skills and knowledge needed to do a job effectively, and shows that a candidate is competent in the area of work the NVQ represents. These standards are statements of performance that describe what competent people in a particular occupation are expected

to be able to do. They cover all the main aspects of an occupation, including current best practice, the ability to adapt to future requirements and the knowledge and understanding that underpin competent performance. The qualification framework is designed to help learners make informed decisions on the qualifications they want to pursue, by comparing the levels of different qualifications and identifying clear progression routes to their chosen career. The US has developed the O*NET system of classifying job requirements but does not provide a certification process to determine whether a person meets those requirements.

Strategies to Fill the Skill Gaps

In considering strategies to narrow the skill gaps, the direction of such discussion depends upon the inherent purpose of intended policy. From a narrow economic perspective, the choice is between growing GDP, on the one hand, and improving living standards by increasing GDP per capita through greater productivity, on the other. One can always increase GDP by opening the borders to immigration; increasing GDP per capita requires investment in human and physical capital in order to increase productivity growth. To improve the latter, however, more and better schooling and job training and greater provision of occupational information are potentially critical to any nation preserving or improving its comparative advantage in high tech sectors under global competition (Freeman 2006). Expanding the discussion to include social inclusion and civic involvement underscores the need for a more highly qualified population.

The obvious strategies are to increase access and completion of schooling for those underrepresented population groups and to expand training opportunities for those in the workforce. Many policies and strategies have been proposed and implemented in a host of countries to address these issues, with varying success. These are difficult issues, and each country pursues them in the context of their own cultural and societal norms and educational systems.

However, in a closely integrated global economy the issue is how one country can differentiate itself from another. Obviously, the worldwide supply of low-wage workers coupled with globalisation makes it exceptionally difficult for relatively high-wage economies to compete on cost and price. With the mass flows of immigrants, the off-shoring of goods and services production, and the rapid increase in the supply of highly skilled workers in developing countries, the ability for a country to differentiate itself with respect to skilled workers is also becoming increasingly difficult. The growth of people with high skills is spreading across the globe, and access to post-secondary education becomes more widespread both within and across countries. China and India, the two most populous nations in the world, are expanding their educational systems to compete for high-skill work, which has huge ramifications for high-skilled workers in other countries.

The advantage is therefore on the demand side. The issue is how businesses combine knowledge in unique productive ways to engage in innovative and enterprising activities that cannot be replicated elsewhere (Brown *et al.* 2006). An increase in the supply of skills, therefore, is a necessary but not a sufficient condition for higher productivity.

The policy question, therefore, is to identify place-specific policies that can enhance a country's ability to use skilled workers in a more productive way than other countries. Philpott contends that skills must be allied with other people-oriented management practices, such as high performance work practices (Philpott 2006). He lists these

practices as including profit sharing for all employees, continuous improvement systems, job rotation, performance pay, mentoring, cross-function teams, flexible working, and annual review of employees training needs. A smarter working policy would thus seek to stimulate investment in training by tackling the underlying causes of low demand for skills. And because this would involve efforts to increase the uptake in productive people management, the policy would enable more productive use to be made of the existing stock of skills in the economy. Yet, the majority of UK organisations have failed to implement high performance workplaces.

Richard Florida emphasises amenities preferred by the creative class, which typically includes highly educated and skilled workers. Investment in new technologies and in the environment within which new innovations are created is also tantamount to future productivity growth.

Another place-specific initiative to promote worker productivity and business competitiveness is the formation of regional partnerships. During the past two decades, several countries have attempted to address deficiencies resulting from fragmented and overlapping programmes and the gap between these programmes and the needs of businesses by forming partnerships among business, workforce development agencies, economic development organisations, and educational institutions. These partnerships typically target the needs of business in specific industries within a single labor market. The purpose is to leverage existing resources by providing better coordination and collaboration among key stakeholders. They focus on promoting the proper match between a worker's competencies and employers' needs. For businesses unable to find qualified workers, regional partnerships provide clearer lines of communication between the needs of business and the educational institutions and workforce development agencies providing the needed training.

Several lessons for successful partnerships have been gleaned from case studies conducted across several countries (Eberts 2007). The more pertinent ones for promoting economic development are: 1) business, as customer, should be the common focus; 2) outcomes must be agreed upon, quantified, and tracked, 3) local organisations must become entrepreneurial and problem solvers and form strong networks among the stakeholders; and 4) strong leadership is required to help define and advocate for the common purpose and to mobilise community resources.

Another factor that can impede the competitiveness of countries is the over-qualification of workers. Surveys in the UK reveal that a third of university graduates say they are over qualified (Brown *et al*. 2006). This inefficiency robs sectors that require more highly skilled workers of the labor resources they need and wastes scarce country resources by training workers with skills they and the economy can't use. Many studies have found that over-educated workers earn less than equally educated workers who are employed in a job that matches their education, whereas undereducated workers who are employed at a job level that is higher than their level of education, earn more (*e.g.* Hartog, 2000). Over-education is often seen as a short-term problem resulting from a lack of coordination in the adjustment of schooling requirements and schooling investments between firms and individuals (Duncan & Hofman 1981). However, several studies have found that for a large group of workers over-education is a long-term phenomenon (Dolton & Vignoles 2000).

One policy strategy is to better align skill requirements with worker competencies. This requires much better information about employer needs and worker skills. Training

takes time, and students and workers need better career planning information to prepare for the current and future needs of employers.

Another policy option is for businesses to make better use of available workforce skills. This may mean that businesses need to invest in more technologically advanced equipment or economic development agencies should attract and nurture more high-value added businesses.

Of course, market forces will correct some of these misalignments. Higher wage premia for college graduates will motivate students to enroll in college and wage differentials between sectors will entice workers to move to the higher paying jobs they are qualified to fill. However, with a global economy, governments need to be cautious in constructing policy responses so that they don't set up the wrong incentive structure. And this becomes more complicated as the dynamics of supply and demand for worker skills are influenced by global factors. Consequently, all parties – governments, other policy organisations, educational institutions, workforce and economic development agencies, businesses and workers – need better information to understand the international dynamics of worker skills and to make more informed decisions.

Conclusion

In the global economy, human capital development is tantamount to a nation's competitiveness and to improving the well-being of its citizens. However, countries have limited knowledge of the competencies of their workforce. This hampers their ability to assess their comparative advantage relative to other countries, to understand the gaps in education and training that may exist, and to design effective policies to help close those gaps. As the skill requirements of jobs increase and become more complex, assessing workforce competencies becomes more difficult. Recording the years of education or terminal degrees or credentials is incomplete in capturing the richness of a person's skill set and equally important the qualifications necessary to fill a job. Without this information, workers and businesses have difficulty making a good match, and educational and training institutions are unable to design curricula to prepare workers for future skill requirements. Attempts are being made to fill the gap. The OECD has pioneered cross-country assessments of adult skill competencies and knowledge and skill acquisition by youth. Several countries have developed taxonomies of skills within occupations, such as the US's O*Net and the UK's NVQ. These are excellent starts, but more needs to be done for the world economies to informed decisions to develop their most precious resource – its workforce.

References

Adkins, Douglas , "The American educated labor force: An empirical look at theories of its formulation and composition." Higher Education and the Labor Market, Margaret Gordon, ed. Carnegie Commission on Higher Education.

Autor, David H. (2007), "Structural Demand Shifts and Potential Labor Supply Responses in the New Century," Federal Reserve Bank of Boston Conference on Labor Supply in the New Century, May

Autor, David H. and David Dorn (2007), "Inequality and Specialization: The Growth of Low-Skill Service Jobs in the United States." MIT Mimeograph, August.

Barton, Paul E. (2000), "What Jobs Require: Literarcy, Education, and Training 1940-2006," Princeton, NJ: Policy Information Center, Research Division, Education Testing Service, January.

Brown, Phillip, Hugh Lauder, David Ashton, Gerbrand Tholen (2006), "Towards a High-Skilled, Low-Waged Economy? A Review of Global Trends in Education, Employment, and the Labour Market," in Sam Porter and Mike Campbell (eds.), *Skills and Economic Performance*, Sector Skills Development Agency.

Carneiro, P., Crawford C. and Goodman, A. (2006) "Which skills matter?", Unpublished, Institute for Fiscal Studies.

Card, D. (1999), The Causal Effect of Education on Earnings, Chapter 30 in *Handbook of Labor Economics,* Volume 3a, edited by Orley Ashenfelter and David Card, North Holland Press, p. 1801-1859.

Campbell, Mike (2006), "Building Prosperity: A Skills Agenda for the UK," in Sam Porter and Mike Campbell (eds.), *Skills and Economic Performance*, Sector Skills Development Agency.

Collier, William, Francis Green and Young-Bae Kim, "Training and Establishment Survival," Skills for Business Research Report 20, March 2007.)

Decker, P., Rice, J. and Moore, M. (April 1997), Education and theEconomy, An Indicators Report, National Center for Education Statistics, U.S. Department of Labor, NCES 97-269.

Dickerson, A., Homenidou, K, and Wilson, R.A. (2006), Working Futures 2004-2014: Sectoral Report, Sector Skills Development Agency, January.

Dolton, P., Silles, M. (2003). "The Determination and Consequences of Graduate

Overeducation". In F. Büchel, A. de Grip, & A. Mertens (Eds.) *Overeducation in*

Europe. Current Issues in theory and Policy (p.189-216). Cheltenham: Edward Elgar.

Duncan, G., & Hoffman, S.D. (1981). The incidence and wage effects of overeducation.

Economics of Education Review, 1, 75-86.

Eberts, Randall, Geroge Erickcek, and Jack Kleinhenz (2006), "Dashboard Indicators for the Northeast Ohio Economy," Federal Reserve Bank of Cleveland Working Paper 0605.

Eberts, Randall (2007), "Designing Effective Employment Strategies through Local Economic Development," paper prepared for OECD LEED Workshop on Designing Effective Employment Strategies through Local Economic Development in China, April.

Freeman, Richard (2006), "Is a Great Labor Shortage Coming? Replacement Demand in the Global Economy," NBER Working Paper 12541, Cambridge, MA: National Bureau of Economic Research, September.

Glaeser, Edward and Albert Saiz (2003), "The Rise of the Skilled City," Harvard Institute of Economic Research, Discussion Paper No. 2025, December.

Green, David, and W. Craig Riddell (2001), "Literacy, Numeracy, and Labor Market Outcomes in Canada," Statistics Canada.

Hall, Ben (2000), Immigration in the European Union: problem or solution?

OECD Observer, Centre for European Reform, London, (October).

Hanushek, Eric, and Dongwook Kim (1995), "Schooling, Labor Force Quality, and Economic Growth," NBER Working Paper 5399, Cambridge, MA, National Bureau of Economic Research, December.

Hartog, J. (2000). Over-education and Earnings: Where Are We, Where Should We Go?

Economics of Education Review, 19, 131-147.

Hecker, David (1992), "Reconciling Conflicting Data on Jobs for College Graduates," *Monthly Labor Review,* July.

Heckman, James J. and Yona Rubinstein (2001), "The Importance of Noncognitive Skills: Lessons from the GED Testing Program," *The American Economic Review*, Vol. 91, No. 2, (May), p. 145-149.

Heckman, James J., Stixrud, Jora and Urzua, Sergio (2006), "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior" (February), NBER Working Paper No. W12006

Jacobson, Louis, Robert LaLonde, and Daniel Sullivan (2001), "Estimating the Returns to Community College Schooling for Displaced Workers," Federal Research Bank of Chicago Working Paper No. 2002-31.

Philpott, John (2006), "Raising Productivity: From Skills to High-Performance Working," in Sam Porter and Mike Campbell (eds.), *Skills and Economic Performance*, Sector Skills Development Agency.

Steedman, Hilary (2006), "International Comparisons of Changes in Qualification Levels," in Sam Porter and Mike Campbell (eds.), *Skills and Economic Performance*, Sector Skills Development Agency.

Sum, A. (1999), Literacy in the Labor Force, Results from the National Adult Literacy Survey, National Center for Education Statistics, U.S. Department of Education, NCES 1999-470, p. 212.

Tyler, John, Richard Murnane, and Frank Levy (1995), "Are more college graduates really taking 'high school' jobs?" *Monthly Labor Review*, December.

Chapter 27
Child Development, Human Development and the Progress of Societies[i]

Fiona Stanley
Founding Chair, Telethon Institute for Child Health Research, Australia
Australian Research Alliance for Children and Youth, Australia

Stephen R. Zubrick
Professor, Curtin University Center for Development, Australia
Head, Population Science, Telethon Institute for Child Health Research, Australia

Abstract

This paper postulates that a society's progress might best be measured via human development, primarily occurring through participation in social, economic and civic networks. The paper argues that the uneven distribution of development opportunities (and resulting benefits) underlies the substantial human development gaps in sub-populations, using as an example Aboriginal Australians vs. the rest of the Australian population. The paper highlights the need to identify the pathways of participation and ways to enrich them. To do so, the authors argue that demographic measures need to give way to better measures of developmental resources. They then present some common demographic variables along with an expanded variable framework based on a human-development perspective. Also presented is cross-sectional data that demonstrate that developmental resources within families and between societal contexts are linked.

The Australian Population: A Global Microcosm?

In Australia, relative stability in general economic and social development has operated to produce measurable advances in many of the key capabilities of Australia's population (Australian Bureau of Statistics 2004, 2006). Along with policies that govern public health and health access, policies for mandated education, more equitable access to employment, ownership of land and housing, a stable and nutritious food supply, and law and order, have helped to produce real increases in resource domains for the development of children and subsequently in their lifelong social, economic and civic participation.

i Correspondence about this article should be directed to Professor Stephen R. Zubrick, Curtin University Centre for Developmental Health Child Development, Human Development and the Progress of Societies (S.Zubrick@curtin.edu.au)

At a population level, Australia has secured a stable democracy with healthy, educated families in which to rear children allowing more (rather than less) developmental resources for children to be mobilised and directed towards their development. In addition to time and income these developmental resources also include forms of human, psychological and social capital known to affect lifecourse developmental capabilities from infancy and childhood onwards (Zubrick *et al.* 2005; Zubrick *et al.* 2000). For some, this has allowed developmental 'bootstrapping' from generation to generation as gains in human development flow from adults to children. However, beyond this broad population effect, there has been less equity for others in the distribution of the flow of developmental opportunities and benefits resulting in a substantial human development gap between sub-populations.

For Aboriginal Australians this is particularly so – giving rise within Australia to two fundamental population dynamics: a) a dominant developed-country scenario with a population that is increasingly aging, with a low birthrate, and in which their younger generation is subject to the emergent and rising prevalence of the non-communicable diseases (eg asthma, cardiovascular disease, obesity, diabetes and mental ill-health) and a slowing or reversal of developmental gains made in previous decades (Stanley *et al*. 2005); and, b) a minority developing-country scenario concentrated within the Australian Aboriginal population and characterised by an increasingly youthful population with a high birthrate, entrenched social exclusion and the concurrent emergence in this sub-population of diseases associated with the worst features of *both* developed- and developing-country status (*e.g.* cardiovascular disease, obesity, diabetes, mental ill-health *and* low birthweight, infectious diseases, and a Year One failure rate at primary school of over 60%) (Zubrick *et al*. in press; Mitrou *et al*, 2006). This represents an Australian domestic humanitarian crises that in many ways recapitulates aspects of the current global circumstance between developed and developing country outcomes in health, education, employment and social circumstances (see Cooke *et al*. 2007) (Table 27.1).

Table 27.1 Selected Characteristics of the Australian Majority and Aboriginal Populations

	Australian population	Australian Aboriginal population
Population (fertility rate)	20,561,00 (1.8)	419,600 (2.15)
Median age	36.6 years	20.5 years
Adult-to child ratio[1]	2.95	1.19
Life expectancy	78 years males 83 years females	59 years males 67 years females
Infant mortality rate	5.3 per 1000	10.6 per 1000
Fetal growth restriction	11%	20%
Mental health morbidity for <18yo	17%	24%
Low academic competence	20%	58%
Retention to Year 12 school[2]	75.7%	39.5%

[1]Persons aged 18+ for every 0-17 year old
[2]In Australia, 10 years of schooling has been compulsory

One country, two populations – one growing old, one growing young – and both increasingly dependent upon just-born or yet-to-be-born generations. This Australian demographic profile invites us to think deeply about growing old and growing young.

First, it invites us to be cautious about characterising a national population as (for example) "aging" in the presence of different demographic mechanisms for sub-population(s) within nations. In Australia, if social and economic policies are solely guided by the demographic characteristics of voters in the majority population, then policy responses will contribute to the maintenance and possible widening of the human development gap between mainstream and, in Australia's case, Aboriginal populations.

Second, whether they are growing older, or younger, populations will rely on the skills and talents of yet-to-be-born generations: In other words, tomorrow's children. Any concern we might have about supporting our populations must be concerned with how we measure and foster lifecourse human capabilities through current and future child development. In the policy response to meet the needs of increasingly aged populations, and in the accompanying competition for funds to do so, the developmental needs of infants and children are likely to be viewed complacently if not antipathetically.

It is with this mind that we put forward some general considerations in policy viewpoints that govern how the progress of a society might be measured, who gets measured, what gets measured and why.

Agency, Governments and Nations: Different Views of Society and Progress

Any approach to measuring the progress of societies confronts tensions in the views of agencies, governments, and nations with respect to what constitutes progress (Zubrick 2005).

Table 27.2 System Views of Human Development

What the health system sees	
Physical health outcomes	***Mental health outcomes***
diabetes, cancer, respiratory conditions, cardiovascular disease	Attention deficit disorder depression/suicide aggression
What other agencies see	
Developmental outcomes	***Risk behaviour outcomes***
poor attachment, poor cognitive development poor speech and language	substance use, early sexual activity, physical inactivity
Educational outcomes	***Social outcomes***
truancy early school leaving alienation	institutional care, racism imprisonment, inequity, exclusion
What the political system sees	
Participation	
economic social civic	

Whether human services are the responsibility of government agencies and/or blended with services from Non-Government Organisations the problem focus within an agency tends to be relatively circumscribed. This means that *within* agency sectors such as health, education and justice, their respective problems are viewed in terms of the prevention, detection and treatment of illnesses, or the enrolment and retention and performance of students, or in the management of crime, imprisonment, and law and order. Each of these sectors has its own reified language, specific culture of problem solving and action, associated legislative and regulatory frameworks and customary sphere of action and influence. Importantly, many of the determinants of specific problems in the missions of any single department are likely to be on causal pathways more accessible through the missions of other government departments or through political means.

Accordingly then, the problem view *across* government and non-government agencies becomes considerably wider and encompasses problems of *human development*. These include risk behaviour outcomes such as substance abuse, early sexual activity and teen parenting; academic outcomes such as truancy, early school leaving and alienation; social outcomes such as individuals in institutions, children in custody, criminal behaviour; and developmental outcomes such as poor attachment, poor cognitive development and poor speech and language. These outcomes have shared causal pathways that produce considerable burdens at the population level and diminish human capital and government wherewithal to address the many fronts that require action. These similarities in the shared causal pathways for these outcomes include: their persistence across the lifecourse, their causal complexity, and social gradient properties that underlie their distributions in populations.

Ultimately however, the global concern among nations interested in measuring the progress of societies is on forms of population *participation*: economic, civic and social. Thus, GDP and rates of employment, voting behaviour and volunteerism, educational enrolments and rates of access to social benefits feature in the interests of the governments of nations and in their measures of progress. This leads political systems to measure their progress in terms of participation, while across government departments and agencies within nations progress is measured through reductions in burdens associated with human development, and within specific departments and agencies each will confront the measurement of the performance of their specific missions by reductions in (for example) incidence and prevalence of specific health morbidities or mortality, educational failure, incarceration rates, and so on (see Figure 27.2).

In its broadest form then, the progress of societies is about human development, and human development is about participation – social, civic and economic - and its underlying capabilities:

- To be able to survive;
- To be knowledgeable;
- To have access to resources for a necessary standard of living; and to participate in the life of a community

(See Sen, 2003; ul Haq, 2003; United Nations Human Development Corporation, 2005).

These capabilities are closely tied to resources that support the development of children, adults, communities and nations. They lead us to suggest what, in any population, be it growing older or growing younger, a desirable framework and strategy for the measurement of progress in societies.

Pathways to Participation as Measure of Progress

The elements underpinning regular measurement of population progress include: a shared theoretical base for selecting some measures over others, time as a variable to allow periodic reporting and comparison, a systems approach to insure engagement, sustainability, and uptake of findings, and enduring partnerships (McQueen & Puska 2003). If it is generally agreed that the pathways to participation lie across agency sectors rather than within them, and that a human development approach provides a general unifying approach that can be used across agency sectors, then governance arrangements will be needed to provide leadership and accountability in achieving implementation and response to routine measurements of progress. Establishing a shared theoretical base could be facilitated where measurement approaches adopt a causal framework aimed at participation. In our view, this dialogue would be facilitated by more context enrichment of existing data to enable sectors (ie health, education, justice, etc) to identify those modifiable pathways to participation that would produce developmental benefit for individuals and particularly, for populations. How can greater context enrichment of pathways to participation be achieved?

A first reasonable step would be to change the basic approach in the use of demographic variables in measurement frameworks. Demographic measures need to give way to better measures of developmental *resources* that are relevant to the *contexts* of people's lives and that link to broader level (macro) social, environmental and economic indicators as well as individual health and human development outcomes of interest. In the first instance this need not lead to radical changes in data items and fields collected in, for example, routine surveillance – particularly where some level of traditional demographic description is already present in the collected data. What is required is a more explicit commitment to the human development model. This means relinquishing the belief that population developmental gain can be reduced merely to the actions of individuals themselves. From a developmental resources framework it is then possible to move towards using the data to describe contexts in which health (and other) outcomes can be measured over time. The OECD has commenced this process already in its work on the intergenerational transmission of disadvantage (D'Addio 2007).

For example, Table 27.3 shows some common demographic variables in along with an expansion of this variable framework using a human development perspective – one that acknowledges that there are "resources" that are relevant to human development. So, for example, a traditional demographic variable measuring "employment" is broadened to include a measure of time – that is, the number of hours in paid employment. Time is frequently characterised along dimensions of both quality and quantity, and it is also regularly understood by its value in economic and social terms. Individuals supply time as well as income to a variety of activities: receiving income from time spent in the marketplace and receiving "utility" from time spent, eating, sleeping, participating in other activities including caring for children. This measurement can then be used as a proxy for "time poor" in the sense that a high commitment in hours to paid (or unpaid) work may reflect little or no time to invest in one's self or others. Thus the demographic variable of employment, when modified to capture hours in paid employment, can be

used to characterise contexts that are "time poor". It is still possible to distinguish those who are employed from those who are not, however, a degree of context enrichment is gained in the process.

Table 27.3 Measures of Demography, Human Development Resources and Developmental Contexts

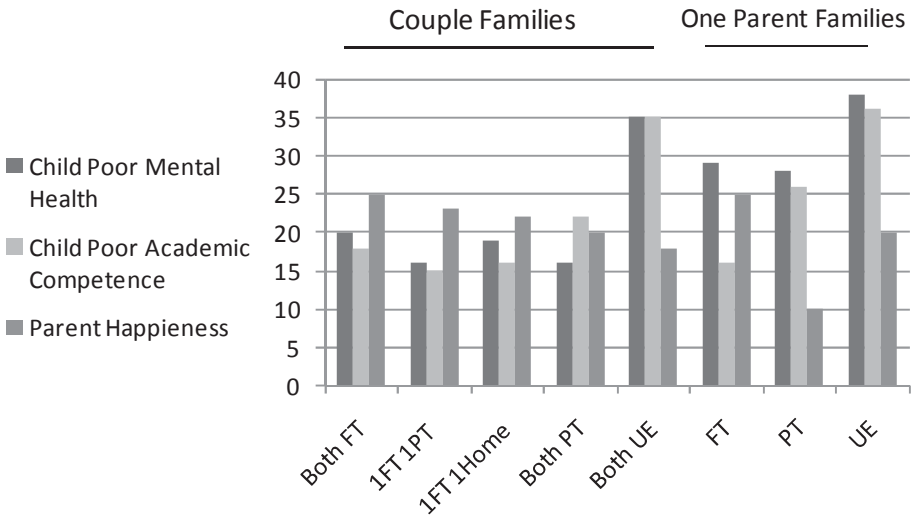
Demographic variables	Indicators of human development resources	Descriptions of developmental contexts
Employment	Time Hours in paid employment Hours for self	Job/no job Time poor
Income	Income Total income Financial strain	Cash poor
Education Marital status	Human capital Family structure Education Physical health	Partner/no partner Kids/no kids
	Psychological capital Mental health Parenting skills	Low control/high control Low stress/high stress
Race Gender Age Ethnicity	Social capital Social support Social exclusion Participation	Social support/no social support High mobility/low mobility Participation/no participation Access/no access

Many of the demographic variables shown have a reasonable analog in terms of resources for human development. Some areas, such as developmental resources corresponding to "psychological capital" are rarely represented in sets of corresponding demographic variables. Given the considerable population burden that mental health disorders pose, and their clear relationship to other non-communicable diseases, this absence in surveillance data will have to change. Similarly, the local environment of the neighbourhood, community and school do not constitute demographic variables per se, although there may be capacities for such measures to be linked to relevant small area data, or to be derived through aggregation. The point is, they provide critical contextual information about individuals and populations. This provides a description of the contexts in which people live their lives which moves beyond the traditional collection of "demographics" and requires displaying data that are more descriptive of social settings contingencies. Data described this way are also more descriptive of how human service agencies do their business.

Figure 27.1 provides a demonstration of these principals using a single time point from cross sectional data of a large random sample (N = 2737) of Western Australian families (Zubrick *et al*. 1995; Zubrick & Kovess-Masfety, 2005). The academic competency and mental health of children in these families are shown with respect to the family work contexts that occur in two-parent and sole parent families. In this case the outcomes are of interest to (at least) two sectors: health and education. Measures also include: time (mean hours of employment per week), income (the lowest income quartile), human capital (maternal education, disciplinary style and family function), and psychological capital (parental mental health and life stress events). These data show that developmental resources *within* families and *between* contexts are linked – for example poor human capital resources (ie low maternal education) occur more commonly in contexts of parental unemployment and low family income, greater levels of stressful life events, and poorer parental mental health. Children in these families are more likely to

have a continuation of these characteristics between contexts in the form of a greater likelihood of attending a disadvantaged school.

Figure 27.1 Distributions of Developmental Resources and Contexts In Australian Children Aged 4-16



Total Work Hrs/week	**96**	**67**	**48**	**23**	**0**	**42**	**16**	**0**
Lowest Family Income (%)	4	4	3	36	46	33	84	91
Low Educated Mother (%)	14	7	13	12	27	9	10	25
Poor Parent Mental Health (%)	9	11	10	17	12	19	26	20
Family Conflict (%)	17	17	18	10	34	16	27	26
Life Events>2 (%)	11	12	15	33	41	27	25	35
Coercive Dicipline (%)	5	5	6	3	11	1	6	8
Disadvantaged School (%)	12	25	31	39	47	22	37	49
Population of Children (%)	11	34	25	4	5	2	4	7

FT=Full Time PT=Part Time UE=Unemployed Home=Not in Labor Force

Several sectoral interests are represented here. Aside from the obvious interests for the health and education sectors – in the form of the mental health morbidity and academic competency of populations of children, poor maternal mental health and the distribution of disadvantage schools – the welfare and family and community sectors have their interests represented by estimates of low family income, family conflict and coercive discipline. All of this is further contextualised by the nature of the family structure and employment arrangements within these families.

The example used here is a cross-sectional snapshot and does not reveal changes over time. However, the data view is more inclusive of other jurisdictional interests and widens the potential for dialogue and exchange among sectors – an essential step in the establishment of systems approaches where determinants of outcomes in one sector (for example, health), are influenced or controlled by determinants in another (for example, education). In applying this approach, and deciding on the nature of the data to be collected and/or displayed, some basic questions need to be addressed:

- What are the principal modifiable determinants of human development or other outcome(s) of interest?
- What are the pathways that lead to and from these outcomes?
- What processes influence these paths and to what magnitude?
- Where are the shared interests?
- How are these talked about inclusively?
- Is it worth it?

As well as measuring patterns and levels of participation across and within populations measures must capture information about the modifiable determinants of lifecourse participation. Our view emphasises a commitment to firm theory and corresponding evidence to guide the selection of the measures for periodic reporting. These measures are contextual as well as behavioural and need to reflect the interests and capacities across agency sectors and urgently, across national boundaries. Greater contextualisation will require acknowledgement that there are multi-level influences on population health that transcend, or are superordinate, to individual behaviour. These macro-level influences may well cap the extent to which population participation can be improved without fundamental changes to social, cultural or political policies and practices. To fail to document these influences will not only result in blaming the individual, but in blaming agencies charged with the responsibility of effectively intervening.

Conclusion

The tension to advance better measures of the progress of society is a product of differing constituent views, theories and pragmatic demands. Our general view is that the broad tenants as outlined in human development theory offer some way forward in unifying views 1) within specific human service agencies and departments, 2) across these government sectors and 3) at the national (political) level. Measures of the pathways to participation (economic, social and civic) are worthy candidate measures of the progress of societies. These pathway measures could document modifiable means for fostering progress. Human development is about participation and the expansion of the choices that individuals have to live the type of life they value. Participation is tied to the availability, access and distribution of the "resource mix" for human development *across the lifecourse*. We believe a lifecourse population perspective to measuring the progress of societies would provide a framework capable of engaging the various levels of influence involved in advancing such progress through a focus on the pathway mechanisms that prompt, facilitate or constrain participation.

Acknowledgments

Portions of this work were made possible by a Programme Grant (353514) from the National Health and Medical Research Council of Australia.

References

Australian Bureau of Statistics (2006), Measures of Australia's Progress 2006, Australian Bureau of Statistics (Catalogue 1370.0), Canberra.

Australian Bureau of Statistics (2006), Australian Social Trends 2006, Australian Bureau of Statistics (Catalogue 4102.0), Canberra.

Zubrick, S. R., Silburn, S. R., & Prior, M. (2005). Resources and contexts for child development: Implications for children and society. In S. Richardson & M. Prior (Eds.), No time to lose: The well being of Australia's children (p. 161-200). Carlton: Melbourne University Press.

Zubrick, S. R., Williams, A. A., Silburn, S. R. & Vimpani, G. (2000), Indicators of Social and Family Functioning (No. ISBN 0 642 44985 6), Department of Family and Community Services, Commonwealth of Australia, Canberra.

Stanley, F., Prior, M., and Richardson (2005) Children of the Lucky Country. Macmillan: Sydney.

Zubrick, S. R., Silburn, S. R., Lawrence, D. M., Shepherd, C., Mitrou, F. G., DeMaio, J., *et al*. (in press). Population capability and human development models: Policy changes to improve the lives of Australian Aboriginal children and families. In G. Robinson, U. Eickelkamp, J. Goodnow & I. Katz (Eds.), Contexts of child development: Culture, policy and intervention. Darwin and Sydney: Charles Darwin University Press

Mitrou, F. G., Lawrence, D., DeMaio, J., & and the WAACHS Team. (2006). Education of Aboriginal and Torres Strait Islander children and young people. In D. Trewin (Ed.), Australian Social Trends (ABS Catalogue 4102.0). Canberra.: Australian Bureau of Statistics

Cooke, M., Mitrou, F., Lawrence, D., Guimond, E., and Beavon, D. (2007) Indigenous well-being in four countries: An application of the UNDP'S Human Development Index to Indigenous Peoples in Australia, Canada, New Zealand and the United States. BMC International Health and Human Rights, 7:9 (doi: 10.1186/1472-698X-7-9).

Zubrick, S. R. (2005). Context enrichment of health risk surveillance data for its use and application. Social and Preventative Medicine, 50 Suppl 1, S16-S24.

Sen, A. (2003) Development as capability expansion. In S. Fukuda-Parr & A. K. S. Kumar (Eds.), Readings in human development: Concepts, measures and policies for a developmental paradigm (p. 3-16). New Delhi: Oxford University Press

Haq, M ul. (2003). The human development paradigm. In S. Fukuda-Parr & A. K. S. Kumar (Eds.), Readings in human development: Concepts, measures and policies for a developmental paradigm (p. 17-34). New Delhi: Oxford University Press

United Nations Development programme. (2005) Human Development Report 2005: International cooperation at a crossroads – Aid, trade and security in an unequal world. New York: United Nations Development programme.

McQueen DV, Puska P, editors. (2003) Global behavioural risk factor surveillance. New York: Kulwar Academic/Plenum Publishers.

d'Addio, A.C. (2007) Intergenerational Transmission of Disadvantage: Mobility or Immobility across Generations? A Review of the Evidence for OECD Countries. Organisation for Economic Co-operation and Development, Paris.

Zubrick, S. R., & Kovess-Masfety, V. (2005). Indicators (Chapter 12). In H. Herrman, S. Saxena & R. Moody (Eds.), Promoting mental health: Concepts, emerging evidence and practice (p. 148-168). Geneva: World Health Organisation

Zubrick, S. R., Silburn, S. R., Garton, A., Burton, P., Dalby, R., Carlton, J., *et al*. (1995). Western Australian Child Health Survey: Developing Health and Well-being in the Nineties. Perth (WA): Australian Bureau of Statistics and the Institute for Child Health Research

Chapter 28
Human Capital and the Older Worker: The Need for Solid Indicators

Peter Ester
Professor of Sociology, Tilburg University, The Netherlands
Director, OSA - Institute for Labour Studies
Chairman, Advisory Commission on Economics & Society of Statistics Netherlands (CBS)
Member, overall CBS Advisory Board

Marcel Kerkhofs
Senior researcher, OSA - Institute for Labour Studies, Tilburg University, The Netherlands

Abstract

A big challenge for OECD economies, this paper point out, is creating indicators that effectively assess how well countries are upgrading the human capital of older workers. The authors begin with an aged-based analysis of the workforces in a large cross-section of OECD countries – including the EU member states, the United States, Japan, Mexico and Turkey. Included is an analysis of the causes of relatively early retirement and examples of what some countries, such as the Netherlands, are doing to attract, retain and train older workers. The authors analyse the findings of studies on age-related productivity and the affect of profit-sharing on training intensity and workforce exit rates. Based on the findings of these and other studies, the authors make recommendations as to what types of indicators could help policy makers effectively address the ageing of the workforce and the decline in the number of workers.

Introduction

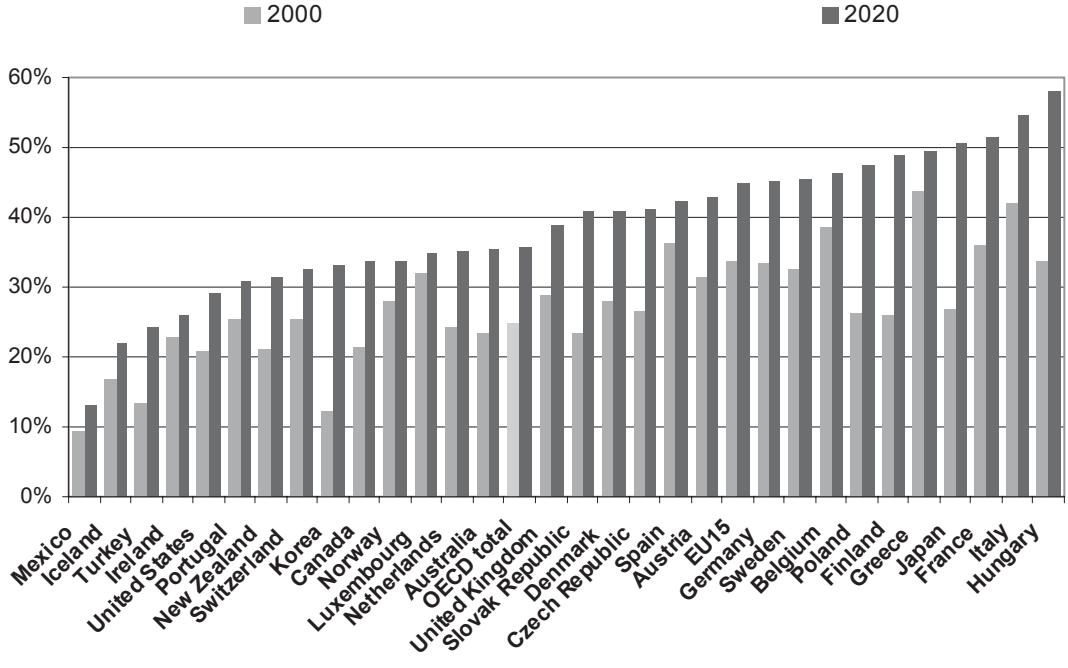
Three of the most fundamental challenges OECD countries are presently facing, are the ageing and decline of its workforce, the acceleration of globalisation processes, and the swift move towards a knowledge economy (OECD 2007). In the next two decades the demographic profile of OECD countries will rapidly and radically change: at the turn of this century about 15% of the OECD population was aged over 65; by 2030, one out of four people in OECD countries will be over 65. In the first half of the 20th century the size of the OECD working-age population expanded by about 75%, the projected growth is

only 4% for the first half of this century. Without drastic policy changes, the OECD workforce will be shrinking considerably.

Figure 28.1 shows the projected ratio of the population aged 65 and over to the labour force in the OECD member states in 2020. The ratio will be about 36% on average in all OECD countries but over 50% in countries such as Japan, France, Italy, and Hungary. The Netherlands closely resembles the OECD mean ratio. The ageing of the OECD workforce is no longer an issue for a distant future, its impact is already felt on major OECD labour markets and will manifest itself more prominently in the next years.

These profound demographic changes come at a time that OECD countries are increasingly involved in globalisation processes, not only in the economic domain but also in the social, political, and cultural realm. Global economic players are not bound by national borders, production takes place in geographic areas were it is most cost-effective, and workers also cross national borders. Multinationals effortlessly move their prime activities to regions were labour costs are low (OECD 2005a). The rapid development and worldwide diffusion of ICT is a major facilitator of these globalisation trends. The ageing of its populations and the globalisation of its member states take place in an era that OECD countries are becoming knowledge economies, meaning that its global competitiveness more and more depends on the competitiveness of its knowledge stock and innovation 1 potential, and on the basic (hard and soft) skills of its workforce.[1] Education, training, and improving and maintaining competences are more important than ever before.

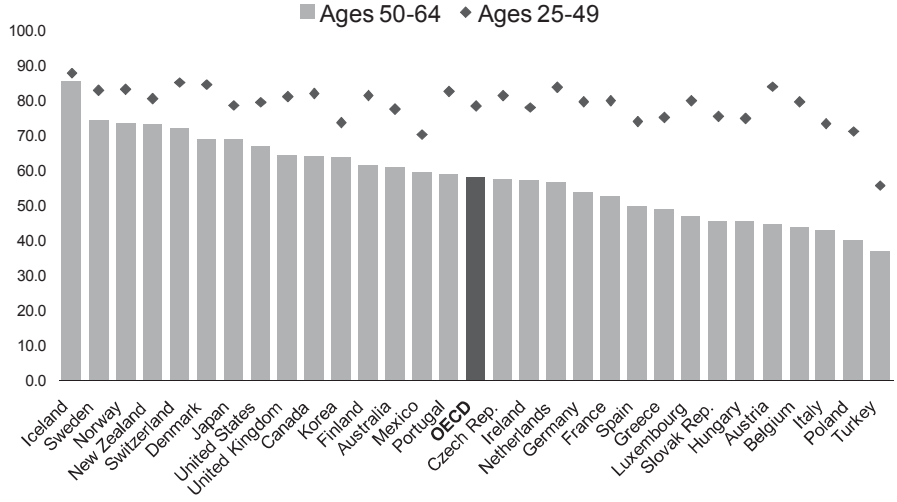
Figure 28.1 Ratio of the Population Aged 65 and over to the Labour Force (OECD)



Source: OECD

In combination these three fundamental challenges imply that most OECD countries have to play the economic game with fewer and older players, on a much larger and more competitive playground with many more teams, where knowledge of the game becomes quickly obsolete, where rules of the game change permanently, and where advanced training methods and professional upgrading of skills are key factors in being successful. The theoretically obvious scenario of bringing in younger and highly skilled players from outside (the immigration scenario), meets with little public support in various OECD countries, particularly in some of the older European member states and in Japan (OECD 2006). These countries are therefore reluctant in embracing full-swing pro-immigration policies. But the tremendous impact of globalisation forces and the emerging knowledge society mean that OECD countries need to act in one way or another, even or perhaps particularly if they reject or temper the immigration scenario. One of the most discussed alternative policy options is to promote that older employees work longer, even after the age of 65. Keeping workers longer in the workforce is an increasingly popular labour market policy in OECD countries to combat demographic trends such as the ageing of its workforce. We live longer, we have fewer children, so we have to work longer seems to be the current policy motto. But this scenario too is not an easy one. During the nineties many OECD countries massively encouraged (or forced) older workers into early retirement. A policy that turned out to be quite successful. The consequences were that work after the age of 50 became less common and that the effective retirement age is quite below the official retirement age as Figures 28.2 and 28.3 show; though substantial OECD country differences are observed (OECD 2006). The Netherlands is about at the average OECD level.

Figure 28.2 Work Becomes Less Common after the Age of 50 but with Large Country Differences: Percentage of the Polulation Aged either 25-49 or 50-64 who are Employed, 2004



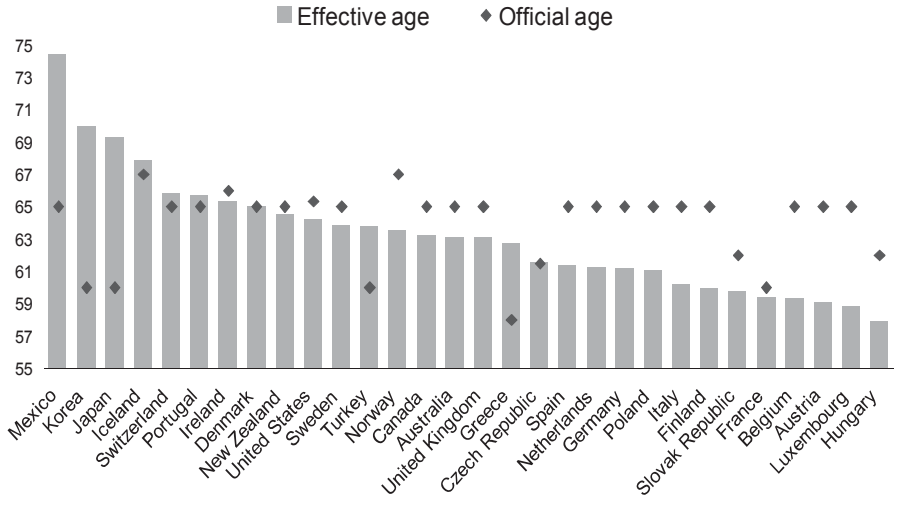
Source: OECD

The EU proclaimed two key policy goals with respect to the labour participation of older workers. The Stockholm Objective (2001) states that "at least half of the EU population in the 55-64 age group should be in employment by 2010" (European Commission 2003, p.157) and the Barcelona Objective (2002) aims to achieve a "five

year delay in the average age at which people withdraw from the labour force by 2010" (*ibid.*), *i.e.* an increase in the average labour market exit from 60 to 65 yrs. One of the paradoxical issues is that in the last two decades of the previous century, leaving the labour market well before the age of 65 became culturally and personally perceived as an obtained right of and among older cohorts of workers, which extended to next cohorts as well. Discouraging early retirement meets with psychological resistance among older employees: this formerly smooth labour market exit is quite suddenly blocked as the favourite route of ending their labour market career. Such cohort-bound reactance effects are typical for an intergenerational reshuffle of gains and losses of these final labour market transitions and may gradually subside. Several OECD economies show that promoting later retirement can indeed be successful. The Netherlands is a case in point in this respect. The labour force participation rate of Dutch older workers (55-64 yrs) rose from 24% in 1993 to 40% in 2004, an increase that is above the OECD average (Stichting van de Arbeid 2006; OECD 2005b). Along with Finland, the Netherlands has been the most successful OECD country in reversing the trend of falling participation rates among older employees.

But keeping more older workers in the OECD workforce is, of course, not enough. We need older workers that are still productive, innovative, flexible, and willing to invest in advancing their skills and competences. In view of the three main challenges facing OECD countries today (rapid population ageing, accelerating globalisation, and emerging knowledge society) older workers too have to play a significant role in addressing these issues. They have to develop and maintain their resources in facing the requirements of a rapidly globalising knowledge economy. Equally important: we need employers in OECD countries that are willing to invest in workplaces that keep older workers productive, innovative, flexible, and with up-to-date skills. What OECD economies need, in other words, is a larger number of older workers with substantial *human capital*.

Figure 28.3 Effective Age of Retirement and the Official Age



*The effective age of retirement is the average age at which workers over the age of 40 withdrew from the labour force over the period.
Source: OECD

The pressing questions to be addressed are: how are older workers doing in terms of human capital? Do they invest in updating and maintaining their skills? What about their productivity? Is it considerably below that of younger workers? Is the age-dependency of human capital and productivity time-invariant, or are human capital and productivity of older workers higher in more recent cohorts than in previous cohorts of older workers? And what about companies: do they invest in human capital maintenance of its older employees? Do they encourage training and lifelong learning for older workers? In order to address at least some of these urgent issues, countries can learn from each others experiences, which underlines the importance of the availability of a set of high-quality indicators, preferably for all OECD countries, that allow monitoring of relevant developments, international comparison and benchmarking.

Human Capital

Human capital is to the knowledge society what natural resources were to the industrial society. Angel Gurría, Secretary-General of the OECD, defines human capital as "the knowledge, skills, competencies and attributes that allow people to contribute to their personal and social well-being, as well as that of their countries" (OECD 2006, p.3). Education (both formal and informal, initial and lifelong learning) plays a crucial role in developing and securing human capital (Becker 1964). Data show that human capital investment and management differ widely across European countries with Sweden and Denmark leading the European Human Capital Index, and Germany, Portugal, Spain, and Italy at the bottom (Ederer 2006). Also at a more global level it is found that countries that invest(ed) heavily in education, in human capital, do better on a wide range of social, economic, and development indicators (Lutz & Goujon 2001).[23]

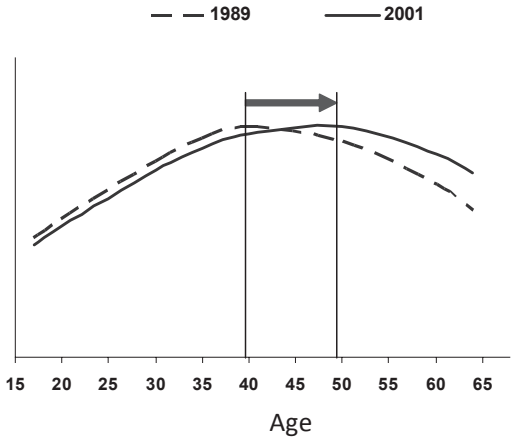
Thus, human capital is a key notion in understanding and explaining how well individuals and countries are doing. But how does human capital relate to age, to an ageing workforce? Human capital theory argues that the earning potential of a worker is dependent on the sum of knowledge and skills which the worker has acquired during his or her life course so far (Becker 1964, 1975; Ben-Porath 1967; Mincer 1958, 1962; Schulz 1964). Knowledge and competences are mostly accrued during the initial educational period. Education is the prime gateway to human capital and thus to career advancement.[2] Additional human capital can be gained during the career through experience and training. In the meantime, there is also a constant erosion of human capital. Just as the physical individual ages, the obtained human capital also ages. During periods of participation this aging – and usually until late in the career – is compensated through additional experience in the form of tenure (Román 2006). The human capital model holds that wages increase with workers' age but a decreasing rate. It is therefore rational for individuals to strongly invest in human capital accumulation at the start of their working career in order to maximise the monetary gains of this investment. The implication of the model is furthermore that older workers will be reluctant to invest in acquiring additional human capital as the payback period is relatively short (*i.e.* until their retirement). This payback rationality also holds for the firm: their interest is to invest in human capital growth of those employees that are likely to remain with the firm for a longer time (thus excluding older employees) (see Gilbert 2001). The combination of both cost-benefit arguments leads to a situation in which older workers are unwilling to invest in further training and the firms they work for are disinclined to invest in skill updating and improvement of its older workers. Such human capital erosion of older workers is of course a bad thing in the current situation that OECD countries need to

promote longer working and later retirement by older employees. But longer working by older workers may also change the training payback calculus by both older employees and firms: the longer older employees will stay in the workforce, the more it pays to update their skills and competences. As retirement age increases, so do the expected returns of investing in human capital of older workers (Riphahn & Trübswetter 2006). In the following we will look at some recent findings on the link between age, human capital, training and retirement that are relevant for OECD policy making and the construction of comparative indicators to support that.

Human Capital, Productivity and the Older Worker: The Dutch Case

The most fundamental of the questions above is how human capital – and productivity along with it – is related to age. On the basis of the theory of investments in human capital it is commonly expected that productivity rises during the early years of one's career, but that this rise gradually flattens out – due to diminishing returns to human capital investments and a falling back of the amount of these investments – which may eventually resulting a fall of human capital at higher ages. The age-productivity profile could therefore be graphically depicted as an inverse U-shaped curve. Empirical evidence for this relationship – and particularly for the falling human capital levels at later ages – is however mixed[3]. First of all there is a large number of studies pointing out that for particular skills no negative relationship with age can be found. But even if for some skills, such as the ability to perform physically demanding tasks, a negative relation with age can be[45] established this may be compensated by acquiring other skills and more work experience. It is therefore tempting to not just study the age-productivity nexus bottom-up, by starting with the skills as building bricks of productivity, but to study productivity indicators directly. This sounds easier than it is, as this line of research is typically hampered by the difficulty to measure productivity of individual workers. The use of direct productivity measures has proven to be useful at the macro level, but at the level of the individual workers most studies are restricted to case studies for work processes for which productivity can be monitored directly as it is in general hardly possible to measure the productivity of individual workers.

Figure 28.4 A Qualitative Representation of Estimated Age-Productivity Profiles



Source: the authors, based on estimates of Gelderblom and De Koning (1992, 2002).

One way around this problem is to relate productivity indicators from firm level data to workforce characteristics, or even better to information on individual workers. For the latter type of analysis one would need linked employer-employee data, which are still rare. For the former type of analysis one could however use data from firm-level surveys. In an early study along these lines, Gelderblom & De Koning (1992) have used data from a Dutch establishment survey for 1989.[5] In this study they have related the per capita value added in an establishment to the age composition of the work force and a range of other characteristics, such as the line of business and the work force size. They find clear evidence of an inverse U-shaped age-productivity profile, for which the productivity peak is located in the age range of 40-45 years. Although the analysis is not on the level of individual workers, its implication is that an increase of the share of 40-45 year old workers[6] increases the added value per worker more than for the shares of younger or older age categories. A decade later this study was repeated (Gelderblom, de Koning & Kroes, 2003) using data from the same panel for 2001. The findings of this later study are interesting in two respects. First of all, the results confirm the shape of the age-productivity profile found in the earlier study. Secondly, the age-productivity profile, although it again has an inverse U-shape, now reaches its peak for the age category 50-54 years old. In the 12 years between the observation periods in the first and second study, the peak of the age-productivity has moved about 10years to the right (see Figure 28.4 for a graphical representation).

Human Capital: Training and Experience

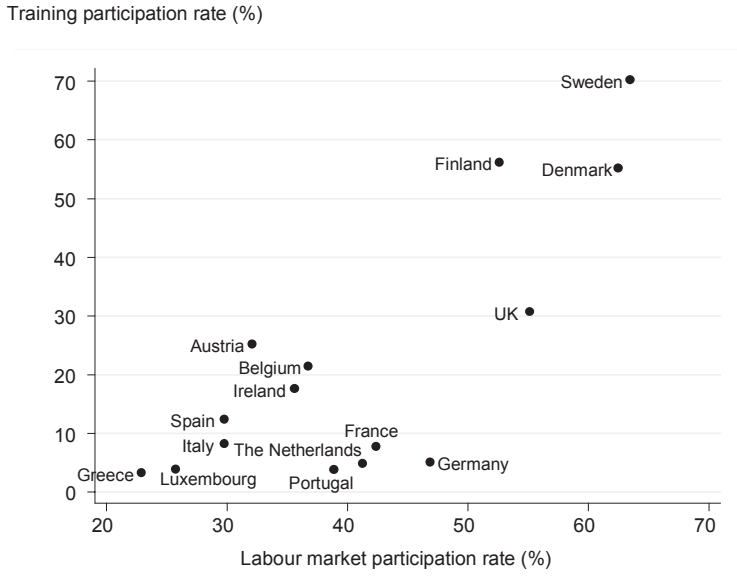
The employability of older workers can – as for all workers – be improved by raising their human capital, be it general or firm-specific. The key question here is how that can be achieved. Numerous studies demonstrate the importance of the initial level of education that workers bring to the labour market. Once at the labour market these investments are sunk and additional effort is required to raise the value of human capital or prevent its depreciation due to skills obsolescence (De Grip *et al*. 2002; De Grip 2006). The latter being particularly relevant for older workers.

Traditionally, investments in human capital by training are however much lower for the older age groups than for the younger. This phenomenon is usually explained – or one could argue summarised – by the fact that the rate of return to investments in human capital is lower for older workers. Heckman (2000) shows that investments in human capital are more effective for younger than for older workers. This is not only because the payback period for younger workers is longer, but is also due to the fact that investments at earlier ages promote investments later on. This 'skill begets skill' mechanism on the one hand emphasises the importance of investing early, but also stresses the fact that early investments raise the rate of return to investments in human capital at later ages. For the purpose of stimulating the human capital of the older workers of the future, it therefore pays to invest in younger age groups now. But additionally, earlier investments have increased the trainability of the more recent cohorts of older workers relative to their predecessors. This reinforces the plea in the previous section for indicators of human capital that not only distinguish between age groups, but also account for *intergenerational* differences and the development over time of human capital within birth cohorts.

The rate of return to investments in human capital are critically linked to the length of the payback period. Most OECD countries now foster policies to promote the labour

force participation of older workers, for example by discouraging the use of early retirement options. If effective, this in itself will prolong the expected payback period and spurt investments in the human capital of older workers. Recent research by Schils and Fouarge (2007) on data from 15 European countries[6] indeed shows that there is a positive relation between labour participation rates and training intensity of people who are in the age category of 50-64 (see Figure 28.5).

Figure 28.5 Participation in Training and Labour Market Participation of People Aged 50-64 in Europe, 1994-2001

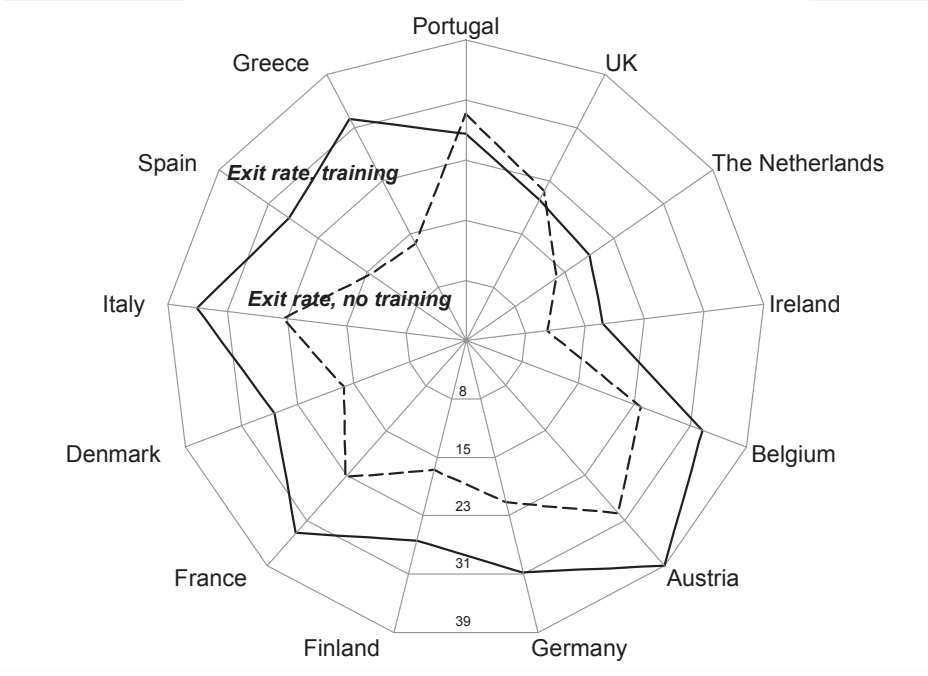


Source: Schils and Fouarge (2007).

The positive relation between working and training is evident, but the scatter plot cannot tell whether higher participation rates lead to more training or more training promotes higher participation rates. To shed some more light on this issue the authors have estimated an econometric model that relates the probability of retiring in 3 years time to whether or not an individual has participated in training activities in the preceding year.

A direct comparison between those with and those without training activities shows that the exit rates in the latter group are considerably higher (see Figure 28.6), implying a 10% higher participation rate for the 50-64 year olds who were involved in training than those in the same age group that were not. By using a multivariate panel model (using an instrumental variables approach to account for the endogeneity of participation in training) they show that the direct causal effect of training on the exit rate is to raise the labour force participation rate for people aged 50-64 by 7%. Stimulating older workers to stay in the labour force will therefore have a self-reinforcing effect, increased labour market participation leading to more training and more training in turn stimulating labour participation. As an aside, this also highlights that it is important to be aware that not only the older workers of the future will be different from older workers now or in the past, so will the institutional settings and the labour market conditions they will be in.

Figure 28.6 Labour Market Exit Rates of Older Workers (50-64) in Europe (exit in next 3 years) in Relation to Participation in Training in Previous year



Source: Schils and Fouarge (2007).

Although training can be seen as the main tool for maintaining and upgrading human capital, one should not equate human capital with formal training and schooling alone. As contended by Arrow in 1962, 'learning by doing' is another important determinant of productivity growth and is therefore a significant component of human capital. Borghans, Golsteyn & de Grip (2006) have recently proposed two survey-based indicators to measure informal learning. The first measure – 'informal learning' – is based on a subjective assessment of the time spent on tasks that the respondent considers 'instructive'. The second measure – 'skill development' – is based on a self –assessment of the respondent's own skill level (relative to what optimal performance in the job requires), and is defined as the change of this self-assessment in the last two years. Apart from the focus on informal learning, the advantage of these measures is that they refer to the skills that are relevant for the respondent's actual job and not to an abstract set of skills. The study shows the potential of these relatively simple indicators to study the importance of formal and informal learning. Due to informal learning the effect on skill development of having a full-time job for 1 year is equivalent to half a year of schooling. If only for results like this, it would be worth investigating the use of this type of indicators for international comparisons. An outcome that is relevant for human capital of older workers is that on average 31% of working time involves informal learning. For younger workers this percentage is higher (40%) and for older workers it is lower, but even at age 60 25% of working time involves informal learning. The impact of informal learning can therefore be particularly relevant for the human capital accumulation at higher ages.

Wage Flexibility

For the purpose of improving the employability of older workers, it is essential that wages mesh with productivity. If this is a problem, it can be alleviated by an increase of the human capital or by aligning wages with productivity. In The Netherlands, as in many other European countries, older workers are often perceived as relatively expensive, which is usually associated with the wage scales being too rigid and based on seniority rather than performance. During a recession this would lead to higher lay-off rates for older workers and lower probabilities of finding a new job. Recent research points at another way in which wage flexibility improves employability. If more wage flexibility improves employment stability, it would also make the expected payback period for investments in human capital longer, creating incentives for more on-the-job training.

Two recent studies have investigated this mechanism empirically. Both focus on the effect of profit-sharing on the training intensity. Azfar & Danninger (2001) argue that if profit-sharing makes wages more responsive to economic developments (cyclical), the expected payback period for investments in human capital will increase as a result of a lower probability that the workers will be laid off in the future. To test this hypothesis they use US data (NLSY panel) for the years 1988-1994. The emphasis in their analysis is on the effect of a high turnover rate of younger workers which would reduce the investments in firm-specific human capital early on in the career. The findings, however, immediately carry over to the link between the firing rate of older workers and the expected payback period of such investments, and to the payback period for investments in general human capital as well. On average the lay-off rate is 10% in their sample, but – accounting for differences in other characteristics – this rate is 4% lower for workers with profit-sharing than for the workers without. Additionally, having a profit-sharing system also reduces the number of quits. The result of this is that workers with profit-sharing receive 55% more firm-specific training – measured in weeks of company training – than workers without profit-sharing. The training incidence – participation in training activities in a given year – is raised by 25%.

Gielen (2007) extends the theoretical argument by noticing that profit-sharing may also create an incentive to put more effort into making the additional human capital pay, thereby raising the returns to training directly. Moreover, this study more specifically focuses on the impact for older workers and uses a different method to account for self-selection into training or jobs in which more training is offered. The data that are used are from the British Household Survey for the period 1998-2003. The effect of profit sharing on training, is again statistically significant, but is smaller than in the earlier study. On average the participation in training activities is 25% higher for workers with profit-sharing than for workers without, but after eliminating the effect of self-selection, the causal effect of profit-sharing is to raise training incidence by 5% on average. Interestingly, this effect is smallest for workers aged 30-49 (only 1.6%), somewhat higher for the younger workers (5.9%) and largest for workers aged 50-65: their participation in training activities increases by 15% if they are paid a profit-related wage.

In conclusion, both studies provide evidence for the hypothesis that more wage flexibility increases the rate of return of investments in human capital, particularly by extending the payback period for these investments. This mechanism adds to the self-reinforcing effect mentioned in the previous section: lowering the exit rate by reducing the lay-off probability stimulates training, which in turn lowers the exit rate due to (early) retirement, which further stimulates training, and so on.

Conclusions

Increasing the labour force participation of older workers will be one of the main targets of socio-economic policy for the decades to come. For that purpose the role of investments in human capital is essential. In the words of the OECD: "This is a challenging task. It has proved particularly difficult to reduce inequalities in training participation by age and skill" (OECD 2006, p.12). For this task we need a solid scientific understanding of the mechanisms involved and indicators for monitoring the relevant developments and for international comparisons and benchmarking.

The overall picture that emerges from the research surveyed in this paper underlines the importance of training and particularly stresses the self-reinforcing effects of training and work: more training promotes labour force participation, which in turn raises the rate of return to investments in human capital, creating an incentive for further training. It also stresses the important role of informal learning – however hard to measure – arguing that policies to promote labour force participation by reducing unemployment and non-participation can themselves be seen as policies to support human capital accumulation. The role of wage flexibility opens another vista, specifically for the employability of older workers, which is generally thought to be weakened by a relatively unfavourable relation between wages and productivity. If wage flexibility leads to greater employment stability, it would also stimulate training.

Human capital is an intrinsically dynamic process. The returns to investments in human capital are not immediate and are reaped over a longer period of time. The returns to these investments moreover depend on earlier investments and on the circumstances and career perspectives at the time these investments are made. The need for and potential of these investments change over time. In that respect, it is important to realise that the older workers of the future are different from the current and past cohorts of older workers, as are the circumstances they will be in. To better understand and monitor these dynamics, human capital indicators should therefore reflect the composition of the human capital stock by age and birth cohort. Additionally, the research on informal learning demonstrates the importance of having indicators of this component of human capital that can be compared across countries and allows tracking of its development over time.

In view of demographic developments, OECD economies will increasingly depend on older workers. The challenge is to make sure that these older workers are and stay competitive in terms of their human capital. Advancing the human capital of older workers is one of the major challenges facing OECD countries. In order to evaluate the success or failure of such a fundamental human capital policy, we need solid indicators. This paper made a plea to develop comparative human capital indicators that show the composition of human capital by age groups and by generation and that measure informal learning and skills acquisition. In developing such indicators, a next OECD World Forum could assess how well OECD economies have done in polishing the human capital of older workers.

In an era of population ageing, OECD countries can no longer afford to waste the human capital accumulated by older workers. In order for OECD economies to stay competitive, they simply need to invest in older workers. Enabling older workers to update and secure their human capital, is the OECD's next frontier.

Notes

1 In 2004 less than 60% of the population aged between 50-64 yrs in OECD countries had a job (OECD, 2006)

2 See Education at a Glance 2006, OECD.

3 See for instance the overviews in McEvoy & Cascio (1989) and Gelderblom (2006).

4 See for example Prendergast (1999) for an overview of studies of the relationship between productivity and performance-related pay.

5 This was the first wave of a biennial panel dataset from the population of all Dutch establishments with 5 or more employees and was conducted by the OSA Institute for Labour Studies. See for more information: www.uvt.nl/osa.

6 The data used in this study are individual data from the European Community Household Panel (ECHP) for the period 1994-2001.

References

Azfar, O. & S. Danninger (2001). Profit-sharing, employment stability, and wage growth. *Industrial and Labor Relations Review*, Vol. 54, No. 3. (Apr., 2001), p. 619-630.

Becker, G.S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy* 70, 5: 9-49).

Becker, G.S. (1964). Human capital: a theoretical and empirical analysis with special reference to education. New York: National Bureau of Economic Research.

Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. Journal of Political Economy, 75(4), part 1, 352-365.

Borghans, L., B. Golsteyn & A. de Grip (2006). Meer werken is meer leren: Determinanten van kennisontwikkeling, [Dutch, Working more is learning more, determinants of skill development]. Den Bosch: CINOP.

De Grip A. (2006). Evaluating human capital obsolescence. ROA Working Paper 2006/2E, Maastricht: ROA.

De Grip A., J. van Loo & K. Mayhew (2002). The economics of skills obsolescence: Theoretical innovations and empirical applications. Amsterdam: Elsevier Science Ltd.

Dekker, R. & A. Kleinknecht (2003). Flexibiliteit, technologische vernieuwing en de groei van de arbeidsproductiviteit, [Dutch, Flexibility, tecnological innovation and the growth of labour productivity]. Tilburg: OSA, Organisatie voor Strategisch Arbeidsmarktonderzoek.

Ederer, P. (2006). Innovation at work: The European Human Capital Index. Brussels: The Lisbon Council.

European Commission (2003). Employment in Europe 2003. Luxembourg: Office for Official Publication of the European Communities.

Evers, G. & T. Wilthagen (Red.) (2006). Arbeidsproductiviteit en arbeidsmarktdynamiek, [Dutch, Labour productivity and labour market dynamics]. Tilburg: OSA, Organisatie voor Strategisch Arbeidsmarktonderzoek.

Gelderblom, A. (2006). The relationship of age with productivity and wages, in: European Commission (2006): Ageing and employment: Identification of good practice to increase job opportunities and maintain older workers in employment, Brussels: European Commission.

Gelderblom, A. & J. de Koning (1992) Meer-jarig, minder-waardig? Een onderzoek naar de invloed van leeftijd op productiviteit en beloning, [Dutch, Higher in age, lower in value?]. OSA report V39, The Hague: OSA Institute for Labour Studies.

Gelderblom, A., J. de Koning & H. Kroes (2003) Leeftijdssamenstelling en bedrijfsprestaties, [Dutch, Age composition and firm performance]. OSA report A201, Tilburg: OSA Institute for Labour Studies.

Gielen A.C. (2007) Performance pay, training and mobility. Working paper Tilburg University.

Gilbert, L. (2001). Human capital and labour market transitions of older workers. Applied Research Branch, Strategic Policy Human Resources Development Canada. Quebec.

Heckman, J. (2000). Policies to foster human capital, Research in Economics, 54(1), p. 3-56.

Kanfer, R. & P.L. Ackermann (2004). Aging, adult development, and work motivation. The Academy of Management Review 29, 3: 440-458.

Lepak, D.P. & S.A. Snell (1999). The human resource architecture: Toward a theory of human capital allocation and development. The Academy of Management Review, 24, 1: 31-48.

Luty, W. & A. Goujon (2001). The world's changing human capital stock: Multi-state population projections by educational attainment. Population and Development Review 27, 2: 323-339.

Mayhew, K. & B. Rijkers (2004). How to improve the human capital of older workers or the sad tale of the magic bullet. Paper presented at the joint EC-OECD Seminar on Human Capital and Labour Market Performance Brussels, Dec. 8.

McEvoy, G.M. & W.F. Cascio (1989). Cumulative evidence of the relationship between employee age and job performance. Journal of Applied Psychology, 74: 11-17.

Mincer, J. (1958). Investment in human capital and personal income distribution. Journal of the Political Economy, 66(4), 281-302.

Mincer, J. (1962). On-the-job training: costs, returns and some implications. Journal of the Political Economy, 70(5) Part 2, 50-79.

OECD (2003). Beyond rhetoric: Adult learning policies and practices (see also the 2005 update: Promoting adult learning). Paris: OECD.

OECD (2005a). Measuring globalisation: OECD Economic Globalisation Indicators. Paris: OECD.

OECD (2005b). Ageing and employment policies, the Netherlands. Paris: OECD.

OECD (2006). International migration outlook 2006. Paris: 2006.

OECD (2007). Human capital. How what you know shapes your life (by Brian Keeley). Paris: OECD.

Prendergast, C. (1999). The provision of incentives in firms. Journal of Economic Literature, 37, 7-63.

Riphahn, R.T. & P. Trübswetter (2006). Population ageing and continued education. Bonn: IZA, The Institute for the Study of Labor.

Román, A.A. (2006). Deviating from the standard: effects on labour continuity and career patterns. Amsterdam: Dutch University Press.

Schils, T. & D. Fouarge (2007). Participation in training and its effect on the decision to retire early. Working paper OSA, Tilburg: OSA.

Schultz, T.W. (1961). Investment in human capital. American Economic Review, 51(1), 1-17.

Stichting van de Arbeid (2006). Arbeid & leeftijd. Aanbevelingen ter bevordering van de arbeidsdeelname van ouder worden werknemers, [Dutch]. The Hague: Stichting van de Arbeid.

Chapter 29
Societal Well-being after Experiencing Trauma at the Hand of "Others": The Intertwining of Political, Economic and Other Visible Factors with Hidden Psychological Processes Affecting Victimised Populations

Vamık Volkan, M. D.
Professor Emeritus of Psychiatry, University of Virginia, United States
Senior Erik Erikson Scholar, the Austen Riggs Center, United States [i]

Abstract

The author presents ways to analyse the health and well-being of societies that have endured trauma, regardless of the causes – from bad economics, to natural calamities, to authoritarian regimes and war-like conditions. The paper focuses on "soft" or psychological issues that disturb a society's well-being, arguing that although these soft issues are less-well researched, they are just as important as "hard" macro-level disturbances, such as high interest rates, budget cuts and tight monetary policies. Two cases are analysed: (1) the impact that the dictatorial rule of EnverHoxha (from 1944 to 1985) had on Albanians' response to the 1997 collapse of a massive pyramid scheme, wiping out half of Albania's GDP; and (2) how the long-term effects of apartheid continue to be reflected in the relatively high crime rate in South Africa during 2006, a decade after the Truth and Reconciliation Commission proceedings.

> *"We may insist as much as we like that the human intellect is weak in comparison with human instincts, and be right in doing so. But nevertheless there is something peculiar about this weakness. The voice of the intellect is a soft one, but it does not rest until it has gained a hearing. Ultimately, after endlessly repeated rebuffs, it succeeds. This is one of the few points in which one may be optimistic about the future of mankind."*
>
> *Sigmund Freud*

Since I am a physician and a psychoanalyst, I consider the well-being of a society from a medical point of view. Long ago, the constitution of the World Health Organisation (WHO) defined health as "a state of complete physical, mental and social well-being and not necessarily the absence of disease or infirmity" (World Health

i www.austenriggs.org/Senior_Erikson_Scholar/

Organization 1946). It is interesting that this statement includes a reference to social well-being which directly refers to the main theme of this Second OECD World Forum. This statement describes an idealistic and perhaps unattainable view of health and raises many questions. With the theme of this conference in mind, I ask the following: What is social well-being? How do we measure or foster it? Finding answers to these questions is very difficult.

The first reason it is difficult to assess and measure the "health" of a society is because of the very long list of factors that can cause social disturbances in the first place. A list of factors that disturb the well-being of a society would include: bad economics, unfair trade, mismanagement of human rights, criminal activities, lack of education, destructive religious or political ideologies, natural calamities, endemic diseases, pollutions, man-made accidental disasters, assassinations of political leaders, dictatorships, the collapse of political or ideological "empires," terrorism, wars and war–like conditions.

Another reason for our difficulty in defining and assessing the well-being of a society is that very often no single professional discipline can fully understand and evaluate factors that cause social problems. An interdisciplinary approach then becomes required to investigate the causes of the problems and subsequent social progress. Building up interdisciplinary teams takes time and often cannot be achieved, mainly because of competition between disciplines. Accordingly, some key aspects of what causes disturbances, and potential progress that grows from knowledge and understanding, may remain illusive.

During the last few decades, "globalisation" has become the buzzword in political as well as academic circles that personifies a wish for prosperity and well-being of societies by standardising economic and political elements and by bringing democratic freedom everywhere in the world. The tragedy of September11, 2001 and the Western World's – especially the United States' – response to it, the wars in Iraq and Afghanistan, and war-like conditions in Africa and elsewhere, in my mind, at the present time make an idealised version of globalisation an illusion. In the twenty-first century once more we are witnessing the amazing ability of the human mind to create incredible technological achievements, while the aggressive aspect of human nature remains the same. Furthermore, globalisation that includes prejudice, racism and an indifference to national differences (Liu & Mills 2006; Morton 2005; Kinnvall C. 2004; Ratliff 2004) never brings about the well-being of the affected societies.

At the beginning of this presentation I referred to Freud's statement where he encouraged us to continue to use the human intellect to better life, in spite of the strength of human instincts and the role they play in creating massive disasters. This OECD World Form is a platform to utilise the human intellect and look at factors that disturb the well-being of societies at the first place. I am sure that some of the items I listed before will be discussed at this meeting in depth. Each item on this list, even if it is closely related to some of the others, requires specific consideration in order to turn its negative impact on societies into positive societal processes. Some of them will benefit more from input formulated by economists, others by educators, still others by those from the legal profession, and so on.

Global Transformations: "Soft" and "Hard" Factors

It is generally thought that global transformations take place along three major dimensions: economic, political, and social. Such transformations not only affect the adult populations, but also children and youth, the future generations (Kaufman *et al.* 2002). When there are crises in these transformations, we usually look to a number of "hard" and macro-level factors to explain the causes of social problems. For example, in an economic crisis what comes to our minds first are visible factors such as austere budget cuts, high interest rates, and strict monetarist policies. Bolitho, Carr and their colleagues (2003) remind us that there are softer and micro-level processes hidden behind the "hard" and macro-level considerations, and they include psychology among them. They state that "soft" micro-level factors have been increasingly recognised by economists since the days of Adam Smith (1723-1790). The Bolitho-Carr school of thought focuses on the importance of communicating human perceptions and motivations, both within and between community and organisational groups affected by economic crisis. Such communications will improve efforts to combat downward spirals into poverty. In spite of the optimism implied in Bolitho, Carr and their colleagues' paper, too often macro-level factors of a society's problems overshadow the micro-level factors, and the later are overlooked. This happens despite the reality that often micro-level factors must be taken into consideration if there is to be a reversal in conditions within the society. By analogy, this is like a big machine that requires tiny screws to be in their proper places for it to function well, or even function at all.

In this very brief presentation, my illustration of the intertwining of macro-level and "soft" psychological issues as they support or disturb a society's well-being, is necessarily limited. First, I will expand on the idea of using psychology in diagnosing and suggesting solutions for an economic collapse by including the effects of a past and shared massive societal trauma on a society which made communication itself between subgroups very difficult. Second, I will focus on societal trauma at the hand of "Others/Enemies" in a more general way, and describe some hidden shared psychological processes that take place in the affected societies that disturb the well-being of people for decades and even centuries to come. Recognising these processes and then dealing with them can make the work on visible macro-level factors, such as economic and political changes, more effective. The type of psychology I refer to in this presentation is the psychoanalytically informed large-group psychology.

Post-Enver Hoxha Albania: A Meeting Place of Economic and Psychological Issues

From 1944 until his death in 1985, Albania was ruled by dictator Enver Hoxha. A few years after his death, communist rule in Eastern Europe collapsed and the effects of this were felt deeply in Albania. Albania's first free elections since the 1920s took place in 1989, and by 1996 a large number of Albanians thought that they were rolling in new wealth. In that year 85 000 Mercedes cars were registered in Albania. But this wealth was imaginary. Twenty pyramid schemes, which grew into murky empires, were the foundation of this imaginary wealth, and in March 1997, pyramid operators one by one began tumbling down, swallowing an estimated $1.2 billion in savings. This amount was about half of Albania's gross domestic product. This economic disaster touched "everything from Liberian shipping companies and German salami plants, to New York bank accounts" (Frank 1988). The pyramid schemes were so embedded in Albanian

society that their collapse initiated a national uprising in which 1 500 people were killed, and outside military intervention was necessary to bring order to the situation.

Auditors from PricewaterhouseCoopers and Deloitte & Touche, as well as other experts in economics from the World Bank, Italy, Greece and Turkey, arrived in Albania to scrutinise "hard" macro-level factors in an effort to unravel the pyramid finances. Their success was very limited. It was said that some powerful and influential Albanians, who were themselves connected with the missing money, stonewalled and interfered with the investigation. The failure of the pyramid system ushered in new social disorganisation.

In the second half of 1997, the Atlanta-Georgia based Carter Center in the United States, under the leadership of former President Jimmy Carter, wanted to help the social disorganisation in Albania mainly through focusing on macro-level economic factors. It proposed to assist Albania in the development of a National Development Strategy (NDS) through a broad "participation" process. The Carter Center was following the World Bank's definition of the term participation: "Participation is a process through which stakeholders influence and share control over development initiatives and the decisions and resources which affect them" (*The World Bank Participation Sourcebook,* 1996, p. xi). The World Bank was outlining methods for stakeholders' consultation and various outreach techniques to garner public input. A few years earlier, the Carter Center's National Development Strategy in Guyana had been a success story. Thus, it expected similar success in Albania when a former diplomat Tom Forbord, who had worked on the Guyana project, went to Albania as the Carter Center's representative.

Armed with well-prepared plans, Forbord was to identify influential Albanians, choose participants for meetings to design economic strategies, and, once these strategies developed, get feedback from the Albanian government and opposition, help to revise and improve plans, pass them along to authorities, and influence the economic recovery in Albania. But Forbord immediately noticed that something was seriously wrong. Even with the prestige of a former U. S. President behind him and access to Albanian government and opposition officials at the highest levels, it was difficult for the diplomat to gather a group of influential Albanians willing to participate in serious discussions on economy or the future of the country's social disorganisation.

In 1998 The Carter Center sent a small group to Albania to learn what went wrong: Joyce Neu, then the assistant director of the Carter Center's Conflict Resolution Program and a linguist; Norman Itzkowitz, a historian from Princeton University who is one of the leading authorities on Ottoman history and who is also trained as a psychoanalyst; and myself, then a member of the Carter Center's International Negotiation Center (INN). Previously the three of us had worked extensively in other troubled spots of the world with other colleagues from various disciplines to understand and deal with psycho-political and psycho-social issues (Volkan 1988, 1997, 1999; Volkan & Itzkowitz 1994). We learned that the enduring psychological effect of societal trauma, which happened decades ago, on the Albanian society was a major factor in the Albanians' difficulty in utilising the Carter Center's offer to develop a strategy for economic recovery and the well-being of the society. The so-called "soft" psychological issues were like little screws that needed to be placed in their proper places before the Albanian big machine worked. Our findings were primarily based on many psychoanalytically informed interviews we conducted with Albanians from different backgrounds.

Enver Hoxha had transformed Albania from a semi-feudal remnant of the Ottoman Empire into an industrialised economy. But improvement in economy does not always make for the well-being of a society, just as being a millionaire does not guarantee the

physical and mental heath of an individual. As a psychoanalyst I have treated poor individuals as well as very wealthy individuals with depression or other psychological disturbances. In truth, Enver Hoxha's society was a "sick" society.

Albania has a relatively homogenous population in spite of religious divisions and two major subgroups: the Gegs and the Tosks. But during Enver Hoxha's long reign, the country was divided into those who were followers of the leader and those who had a "black spot" on their family. If someone was perceived to be against Enver Hoxha and his brand of communism, even if they complained about not finding freshly baked bread in the bakery or were caught playing backgammon, this person could face a "black spot," or a likely prison term. There appeared a rigorous division of goodness and badness, loyalty and disloyalty. "Bad" people would often face torture or exile to a certain areas in the country. A person could try to remain "good," but a relative who became "bad" could ruin the entire family. For example, young people who had someone with a "black spot" in their family would not be able to obtain a university education. Officials were sent to elementary schools regularly to interview children about whether their parents spoke against Enver Hoxha in the privacy of their homes. When democracy came to Albania the psychological effects of the decades-long societal split did not disappear. If someone had a son in a high school, this person would wonder if the headmaster was the son of a person who had tortured his or her grandfather.

Psychologically speaking, a healthy society is one where the citizens trust one another while interacting among themselves under culturally accepted guidelines and democratic legal rules. Using psychoanalyst Erikson's (1985) term, "basic trust" among citizens disappears when a society is not healthy. Basic trust is a concept that describes how children learn to feel comfortable putting their own safety in a caretaker's hands; by developing basic trust, children discover, in turn, how to trust themselves. In a healthy society, adults also depend on trusting themselves and others to remain functioning citizens.

In 1998 Albanian society continued to suffer from a severe societal split, even though it was hidden. This was the reason why it was difficult to put a group of influential Albanians in a room, expect them to trust one another, discuss societal problems, and come up with mutual understanding for developing economic strategies. The severe societal split between good and bad families was also reflected in severe splits in other large-group processes. For example, the Albanian main opposition party and the ruling party were so far apart, that the opposition's elected members would not even attend the regular meetings of the Albanian parliament.

We were lucky to find an Albanian organisation, a kind of "think tank," that was co-chaired by two intellectuals – one came from a family who had been loyal to Enver Hoxha, and the other one came from a family with a "black spot." Both of them were aware of the severe psychological division in their country, and they had formed their organisation to study and do something about this situation. We began making plans to bring a group of influential Albanians, including those knowledgeable in economics, to this "think tank" and *first* help them remove the effects of the social splitting and develop "basic trust." Then we would support them in their efforts to make an economic recovery plan for Albania. We also met with the leader of the opposition Sali Berisha who was a former prime minister. Interestingly we found him receptive to our ideas; he himself was aware that the severe social division that remained psychologically from the days of Enver Hoxha was paralysing Albania. In fact, after he and his people talked with us,

Berisha ordered his party deputies to resume attending parliament. This was short-lived without more input from us.

We could not put our plans to start a psychologically informed economic recovery process in action in Albania due to events that took place in the former Yugoslavia. Ethnic Albanians in Kosovo began to suffer Serbian attacks; the NATO bombing of Serbia followed. In the summer of 1999, some 450 000 Kosovar Albanians sought shelter in Albania. When these refugees flooded Albania's disused factories, sports stadium, and city parks, it looked like the chaos would reach its peak. This tragedy however, in a peculiar and unexpected way, had a significant impact on both the economy and psychology of Albania.

When I revisited Albania at the end of 2000 I learned that taking in the refugees during the upheaval in Kosovo had improved the economy (officially, by 8%) as various officials from foreign countries, journalists and others poured in from abroad, stimulating the Albanian private and governmental security industries. Interestingly enough, I observed that having an external enemy in the Serbs began to heal splits within Albania and appeared to remove the malignant large-group psychology that interfered with work on macro-level social, economic and political issues.

I briefly described my study of post Enver Hoxha Albania to illustrate that there are many causes for a society's well-being or its entry into a paralysing chaos. Some of these causes are visible, while others require psychological assessments in order to be noticeable and definable. My emphasis is on the effect shared societal trauma has on a society's well-being after the trauma ends. In Albania a group of people inflicted trauma on another group of people of their own kind due to a political ideology and the personality organisation of a political leader. Elsewhere, shared societal traumas occur due to other causes. In South Africa racism was the factor responsible for Apartheid.

What Does South Africa Teach Us Ten Years After The Truth and Reconciliation Commission Proceedings?

In November 2006 I had the honor of giving the keynote speech at the University of Cape Town as part of a meeting reflecting upon ten years of South Africa's Truth and Reconciliation Commission and celebrating Archbishop Desmond Tutu's life of peaceful justice (Volkan 2006a). It is clear in my mind that the Truth and Reconciliation activities played a significant role in preventing bloodshed South Africa. There might have been unimaginable tragedies in this country without the character of new leadership and the work of the Truth and Reconciliation Commission. Nevertheless, ten years later and "basic trust" has not been achieved. Driving through Cape Town or many other cities in South Africa, one sees fences, barbed wire, or signs on walls in residential neghbourhoods that declare, "Protected by Armed Response." Eagle and Watts (2002) make reference to 1996 data in South Africa describing most disturbing statistics: An average of 52 murders a day, a rape committed on average every 30 minutes, a car stolen every 9 minutes and an armed robbery committed every 11 minutes. Ten years after the Truth and Reconciliation work, in 2006, a report informed us that during the previous year in South Africa 1 200 children were murdered, 1 500 children were the victims of attempted murder, 24 000 children were assaulted and 2 200 children were raped (*Cape Times* editorial, "Suffer the Children," November 22, 2006, p.10).

Psychologically speaking, many blacks in South Africa have turned the frustration and rage they experienced under Apartheid onto themselves (Volkan 2006a). In

November 2006 I had an opportunity to make some observations in the Cape Town Township of Langa, which is separated from a very plush golf course by a super highway. What made the greatest impression on me was not the unbelievably bad physical conditions and unemployment rate in this place, but the many black students I saw singing and dancing in a classroom at the township's school, which is also surrounded by a fence. While watching them I remembered that one of their male teachers had been murdered only a week before, and the news had terrified the kids. I also knew that every other girl and one boy out of five, who were so beautifully and gracefully singing and dancing in front of me as if they had no worries, were rape victims. For an outsider like me it was impossible to reconcile their traumatised selves with the personalities they were outwardly exhibiting. I had to think that they were showing an adaptation to *constant* trauma, the understanding of which was beyond my emotional comprehension.

What South Africa's situation shows us – a full ten years after the Truth and Reconciliation Commission Proceedings – is how the effects of a societal trauma persist after it is over, above and beyond real-world issues pertaining to economy and politics. My belief is that any macro-level strategy to deal with South Africa's present economic and political problems may face resistance, unless it first takes into account the shared psychological processes that exist in the society, and unless further meaningful communications between the white and black communities take place. The sense of humiliation much of the population still feels must be removed.

The Psychological Impact of Past Events

When a large group of persons, such as blacks in South Africa during Apartheid, are oppressed by "Others," the victimised group experiences a shared sense of shame, humiliation, and even dehumanisation. They cannot be assertive, since the expression of direct rage toward the oppressing group threatens their livelihoods and even their lives. Their "helpless anger" interferes with their mourning over losses that touch every aspect of their lives, ranging from their dignity, to their property, relatives or friends (Volkan 2006b). Shared unfinished psychological tasks are passed from generation to generation. The existence and the effects of such transgenerational transmissions are clearly demonstrated in psychoanalytical studies of the descendants of Holocaust survivors (For a review of such studies see: Volkan *et al.* 2001).

Societal traumas at the hand of "Others," obviously can also be inflicted during armed struggles. Their effects are different from those of massive traumas due to natural causes, mainly because they, like oppressive or racist regimes, induce shared shame, humiliation, helplessness, and dehumanisation and make societal mourning complicated, or even impossible. We can recall the well-known American novelist William Faulkner's lines: "The past is never dead, it's not even past" (*Requiem for a Nun*, Act I, Scene III). As I write this paper in the spring of 2007, Argentine courts are once again confronting the horrors of a brutal Cold War regime under Augusto Pinochet that executed thousands of dissenters. The three-decades-old wounds will be reopened. When Slobodan Milosevic was alive and in power he, with the help of some Serbian academicians and the Serbian Orthodox Church, re-enflamed the "memories" of the Battle of Kosovo that took place in 1389. Elsewhere I describe the details of their work and how it rekindled shared emotions in Serbia that were connected with a 600-year-old wound, and the part this activity played in creating an atmosphere for most brutal actions against Bosniaks and Kosovar Albanians (Volkan 1997).

At the present time it is difficult to imagine peace in Iraq. But, let us imagine it. First macro-level issues involving politics, economy, legal issues and environment need to be taken care of. This is natural. Once more I will use a medical analogy. If someone comes to the emergency room with a bleeding stomach, the first thing to do is to stop the bleeding. But the heath of this individual will not be looked after properly until the hidden cause of the stomach bleeding is found and repaired. Similarly, for the well-being of Iraq, after the hoped-for peace arrives, it will be necessary to heal the psychological scars in order to create a firm foundation for the macro-level changes. At present, generally speaking, we do not yet have established guidelines for studying these types of psychological factors in previously traumatised societies that would instruct us in implementing such tasks as necessary. However, programmes for fostering the progress in societies should be aware of shared psychological processes that potentially exist in them.

Concluding Remarks

In this presentation I tried to illustrate how macro-level economic, political, legal, and environmental factors are intertwined with rather hidden shared psychological processes in traumatised societies, even long after the trauma is over. These are the societies in need of improving their well-being. I emphasised the role of interdisciplinary teamwork in order to assess what such societies require to foster progress and opportunities for material prosperity, but also to create an atmosphere where the citizens can maintain their "basic trust" in relating to one another and their governments.

Shapiro and Carr (2006) state that the attempt to understand society is a daunting prospect, and that it may be "a defense against the experience of despair about the world, a grandiose effort to manage the unmanageable" (p.256). I join them, however, in their suggestion that to make some efforts nonetheless is essential for societies' psychological well-being and even survival. My colleagues and I have developed a method called the "Tree Model" (Volkan 2006b) that is designed to support co-existence between enemy groups. In this presentation I will not be able to give details about this method except to say that with some modification, its principles can also be used, to one degree or another, in scenarios such as those I have discussed today: between those who were victimisers and those who were victimised and their descendants.

References

Bolitho, F. H., Carr, S. C., Chuawanlee, W., Chaijukul, Y., *et al*. (2003). Poverty and economic crises. In *Poverty and Psychology: From Global Perspective to Local Practice*. Ed. S. C. Carr and T. S. Sloan, p.205-225. New York: Kluwer Academic/Plenum Publishers.

Eagle, G., and Watts, J. (2002). Psychodynamic contributions to formulations of the impact and treatment of traumatic stress incidents. Part 1. *Psycho-Analytic Psychotherapy in South Africa*, 10: 1-24.

Erikson, E. H. (1985). *Childhood and Society.* New York: W. W. Norton.

Frank, R. (1988). Jungle accounting: Auditors in Albania pick through rubble of pyramid schemes. *The Wall Street Journal,* August, 6, p.1,7. Kaufman, N. H., Rizzini, I., Wilson, K., and Bush, M. (2002). The Impact of global economic, political, and social transformation on the lives of children: A framework for analysis. In *Globalisation and Children: Exploring Potentials for Enhancing Opportunities in the Lives of Children and Youth*. Ed. N. H. Kaufman and I. Rizzini, p. 3-18. New York: Kluwer Academic/Plenum Publishers.

Kinnvall, C. (2004). Globalisation and religious nationalism: Self, identity, and the search for ontological security. *Political Psychology*, 25: 741-767.

Liu, J. H., and Mills, D. (2006). Modern racism and neo-liberal globalisation: The discourses of plausible deniability and their multiple functions. *Journal of Community and Applied Social Psychology*, 16: 83-99.

Morton, T. L. (2005). Prejudice in an era of economic globalisation and international interdependence. In *The Psychology of Prejudice and Discrimination: Disability, Religion, Physique, and Other Traits,* Vol. 4. Ed. J. L. Chin, p. 135-160. Westport, CT: Praeger.

Ratliff, J. M. (2004). The persistence of national differences in a globalising world: The Japanese struggle for competitiveness in advanced information technologies. *Journal of Socio-Economics,* 33: 71-88.

Shapiro, E., and Carr, W. (2006). "Those people were some kind of solution": Can society in any sense be understood? *Organisational & Social Dynamics*, 6: 241-257.

The World Bank Participation Sourcebook (1996). Washington, DC: IBRD/The World Bank.

Volkan, V. D. (1988). *The Need to Have Enemies and Allies: From Clinical Practice to International Relationships.* Northvale, NJ: Jason Aronson.

Volkan, V. D. (1997). *Bloodlines: From Ethnic Pride to Ethnic Terrorism*. New York: Farrar, Straus and Giroux. (*Bluts-Grenzen: Die historischen Wurzeln und die psychologischen Mechanismen ethnischer Konflikte und ihre Bedeutung bei*

Friedensverhandlungen. Trans. by Klaus Kochmann. Zurich: Scherz Verlag, 1999), (*Kan Bağı.* Istanbul: Bağlam Yayınevi, 1999).

Volkan, V. D. (1999). *Das Versagen der Diplomatie: Zur Psychoanalyse nationaler, ethnischer und religiöser Konflikte. (The Failure of Diplomacy: The Psychoanalysis of National, Ethnic and Religious Conflicts).* Giessen: Psychosozial-Verlag.

Volkan, V. D. (2004). *Blind Trust: Large Groups and Their Leaders in Times of Crises and Terror.* Charlottesville, VA: Pitchstone Publishing. (*Blindes Vertrauen: Grossgruppen und ihre Führer in Kriesezeiten*. Giessen: Psychosozial-Verlag, 2005), (*Körükörüne Inanç.* Istanbul: Okuyanus, 2005). Volkan, V. D. (2006a). The next chapter: Consequences of societal trauma. *www.austenriggs.org/Senior_Erikson_Scholar/*

Volkan, V. D., (2006b*). Killing in the Name of Identity: Stories of Bloody Conflicts*. Charlottesville, VA: Pitchstone Publishing. (*Kimlik Adına Katillik*. Trans. by Banu Büyükkal. Istanbul: Apha, 2007).

Volkan, V. D., Ast, G., and Greer, W.F. (2001). *Third Reich in the Unconscious: Transgenerational Transmission and its Consequences.* NY: Brunner-Routledge.

Volkan, V. D., and Itzkowitz, N. (1994). *Turks and Greeks: Neighbors in Conflict.* Cambridgeshire, England: Eothen Press. (*Türkler ve Yunanlılar.* Trans. by Banu Büyükkal. Istanbul: Bağlam Yayınevi, 1999).

World Health Organisation (1946). Constitution of the World Health Organisation. *Chronicle of the World Health Organisation*, 1: 29-43.

Chapter 30
The Living Planet Index and Ecological Footprint: Tracking the State of Global Biodiversity and Human Pressures on the Biosphere

Jonathan Loh
WWF International
Honorary Research Associate, Institute of Zoology, Zoological Society of London

Abstract

WWF's Living Planet Report has been published biennially since 1998. It was developed as an attempt to quantify and communicate the gravity of the crisis faced by the earth's biodiversity and the challenge for humanity to live within the capacity of the biosphere. Two indicators have been used to measure the state of and pressure on the earth's ecosystems. The Living Planet Index (LPI) tracks the state of global biodiversity by aggregating trends in thousands of populations of vertebrate species living in the world's terrestrial, freshwater and marine ecosystems, starting from a baseline set to 1.0 in 1970. The LPI database has now been expanded to the extent that it is possible to produce sub-global LPIs on particular biomes, taxonomic groups, regions or countries. The Ecological Footprint measures humanity's demand on the biosphere through the consumption of natural renewable resources and the burning of fossil fuels. It can be used to compare countries, regions, or humanity as whole against the biological capacity of the earth. These two indicators have been adopted by the Convention on Biological Diversity (CBD) to measure progress towards achieving its target of reducing the rate of biodiversity loss by 2010.

Living Planet Report

The indicators discussed in this paper were developed for the WWF Living Planet Report (see http://www.panda.org/news_facts/publications/living_planet_report/index.cfm) over the last decade. The LPR began ten years ago as part of a campaign leading up to the new millennium. Its aim was to attempt to quantify, monitor and communicate the declining state of the world's biodiversity and the growing human pressures on the biosphere from the consumption of natural resources. The first LPR was published in 1998 and has since been updated biennially. The measure of the state of biodiversity is called the Living Planet Index and was developed by the author and colleagues based at the World Conservation Monitoring Centre specifically for the Living Planet Report. The

indicator of human pressure on the biosphere is called the Ecological Footprint, and was developed originally by Mathis Wackernagel and colleagues and has been used and subsequently modified in the Living Planet Report since 2000. The last report was published in 2006 and was a collaborative effort with the Zoological Society of London, WWF's partner now working on the Living Planet Index, and the Global Footprint Network, based in Oakland, California, which is responsible for producing the Ecological Footprint.

The 2010 Biodiversity Target

In 2001, at a meeting of the Convention on Biological Diversity (CBD), 190 of the world's governments adopted a target to "achieve by 2010 a significant reduction of the current rate of biodiversity loss at the global, regional and national level as a contribution to poverty alleviation and to the benefit of all life on earth". This target was endorsed the following year at the World Summit on Sustainable Development in Johannesburg. Such public commitments mean that, for the first time, the public can hold its leaders accountable for their success or failure in meeting measurable and quantifiable objectives on this critical issue. WWF and other NGOs are monitoring their progress and, wherever possible, contributing to the achievement of this target. Equally, they will not fail to point out where nations are falling short of their stated aims and will continue to call for much-needed action.

Focal Areas

Within the context of the CBD, governments agreed to monitor and report on progress under seven focal areas relating to the overall 2010 target. These are:

- Status and trends of the components of biological diversity;
- Sustainable use;
- Threats to biodiversity;
- Ecosystem integrity and ecosystem goods and services;
- Status of traditional knowledge, innovations and practices;
- Status of access and benefit-sharing;
- Status of resource transfers.

This paper focuses only on the first two of these focal areas.

2010 Target Indicators

For each of the seven focal areas, indicators have been selected. In reality, only a few of the selected indicators are sufficiently well developed or have sufficient data available to enable them to measure progress towards the focal area of the 2010 target. Some indicators, therefore, have been chosen "for immediate testing" and others for "further development". Because the target explicitly calls for a *reduction in the rate* of

biodiversity loss, the data used to measure progress must go back many years prior to 2010, or preferably decades. Only a few of the selected indicators currently have datasets available that allow changes over decadal timescales to be measured.

Under the focal area 1, *status and trends of the components of biological diversity*, there are good datasets for four of the selected indicators:

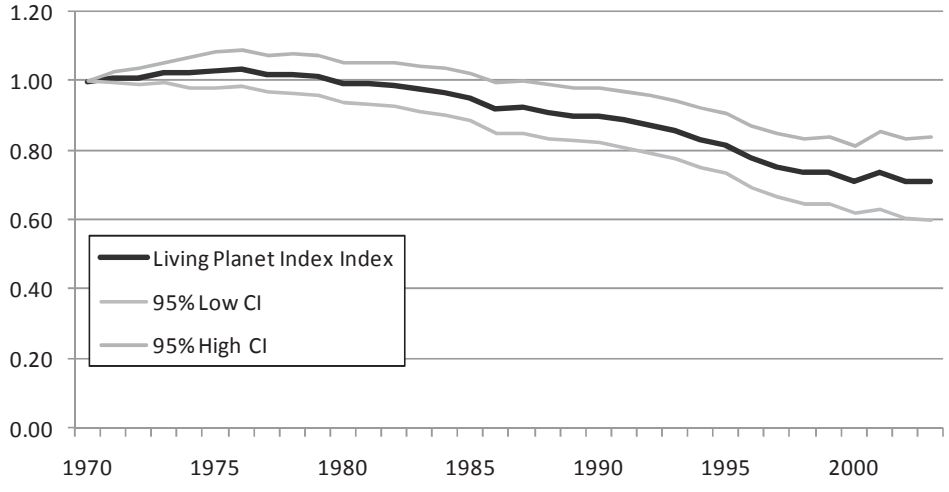
- Trends in the extent of selected biomes, ecosystems and habitats – although currently time-series data at the global scale exist only for forests, and possibly coral reefs;
- Trends in abundance of selected species (living planet index);
- Coverage of protected areas, although this does not in fact measure the status of global biodiversity but one of the responses to its decline; and
- Changes in status of threatened species (red list index).

Under focal area 2, *sustainable use*, global-level time-series data exist only for the Ecological Footprint.

The Living Planet Index

The Living Planet Index (LPI) is an indicator of the state of the world's biodiversity; it aggregates trends in thousands of populations of vertebrate species from all around world into a single index, starting at 1.0 in 1970. Separate indices are produced for terrestrial, marine and freshwater species, so the LPI can also be considered to be a measure of trends in ecosystem health for different biomes or habitats. Although vertebrates represent only a fraction of all known species, it is assumed that their population trends are typical of global biodiversity as a whole.

Figure 30.1 Living Planet Index



The LPI is based on trends in about 5 000 populations of nearly 1 500 species. It is calculated as the average of three separate indices that measure trends in populations of 752 terrestrial species, 389 freshwater species and 310 marine species (see Table 30.1). Between 1970 and 2003 (the most recent year for which sufficient data are available) the LPI fell by about 30%. This is also the case for the terrestrial, marine and freshwater LPIs, which confirms that the earth is losing its biodiversity at a rate unprecedented in human history.

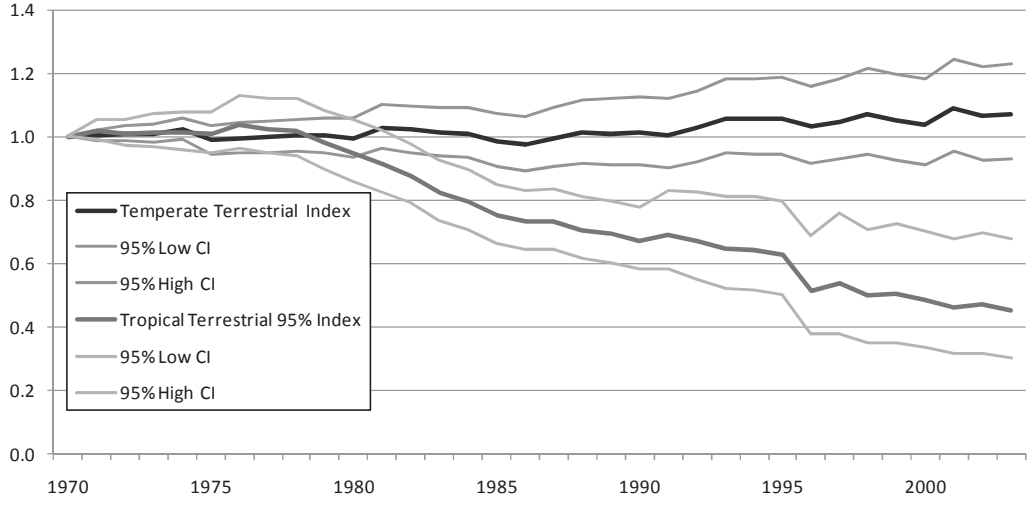
Table 30.1 Number of Species in the LPI

Species	Terrestrial	Freshwater	Marine	Total
Fishes	0	111	129	240
Amphibians	2	75	0	77
Reptiles	14	25	7	46
Birds	531	164	121	816
Mammals	205	14	53	272
Total	752	389	310	1451

Temperate and Tropical Terrestrial LPIs

The decline in terrestrial species populations by about 30% between 1970 and 2003 masks a significant difference between species in temperate regions and species in the tropics. While temperate species populations remained reasonably stable on average, tropical species declined by more than 50%. This rapid fall in tropical species populations reflects the rapid conversion of natural habitat to cropland or pasture over the last fifty years, driven ultimately by the growth in human population and increasing world demand for food, fibre and timber. The conversion of natural habitat to farmland in temperate regions, on the other hand, largely occurred long before 1970, and the consequent decline in species populations is not therefore reflected in the temperate index. The LPI does not say that the current state of biodiversity is worse in the tropics than temperate regions, but that the trends over the last three decades have been worse.

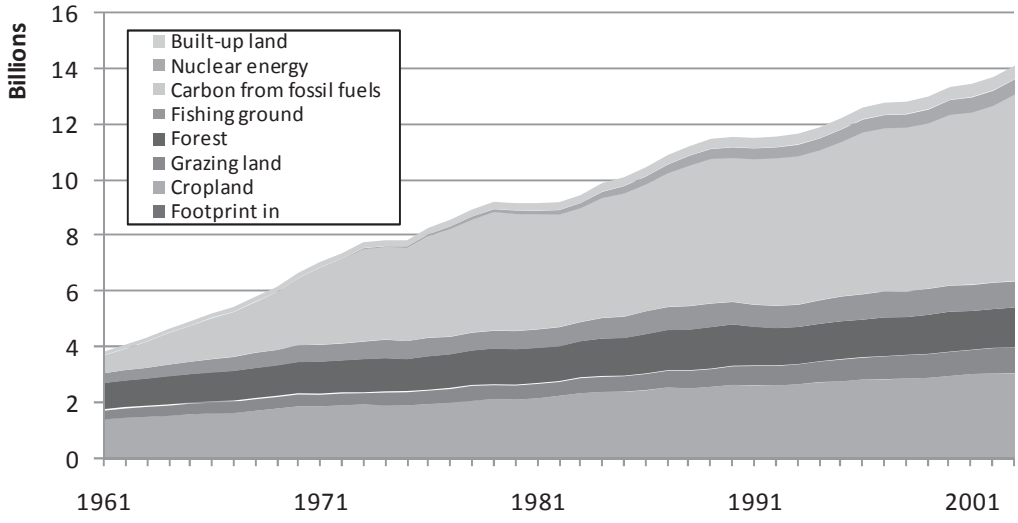
Figure 30.2 Temperate ad Tropical Terrestrial Indices



Ecological Footprint

The ecological footprint measures the area of biologically productive land and sea required to sustain the resource consumption of a given population. The world's ecological footprint tripled between1961 and 2003 to over 14 billion hectares, while world population approximately doubled. It is composed of the cropland, grazing land, forest and fishing grounds needed to produce food, fibre and materials; the urban land occupied by buildings and infrastructure; and the biologically productive land that would be required to absorb the carbon dioxide emitted from burning fossil fuels. The carbon footprint was the fastest growing component from 1961 to 2003 and now comprises about half of humanity's total footprint.

Figure 30.3 Ecological Footprint by Component, 1961-2003

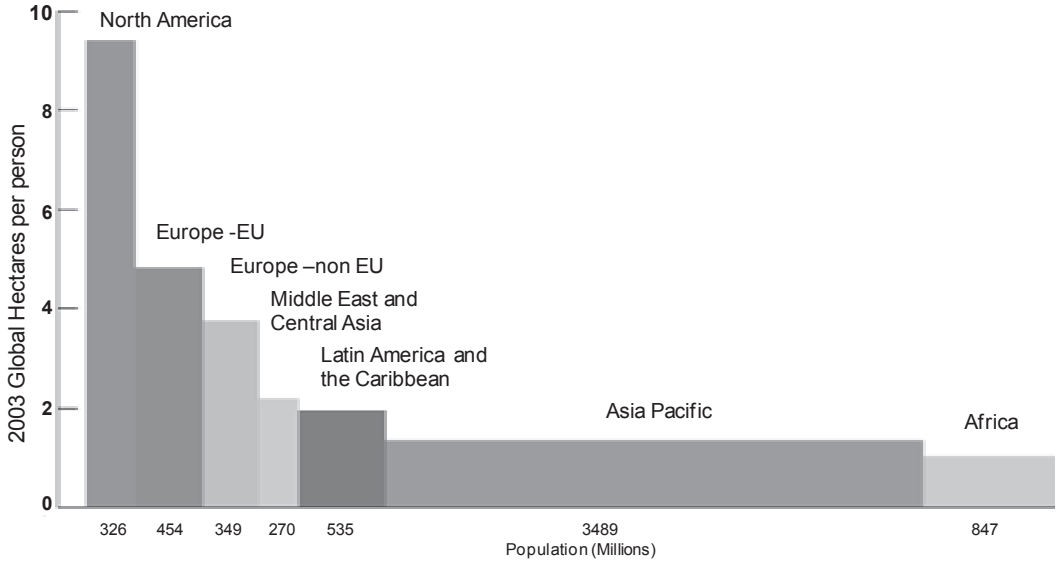


Our demands on the biosphere can be compared with supply, or the earth's biologically productivity. About a quarter of the earth's surface, or just over eleven billion hectares, is biologically productive. Humanity's ecological footprint is therefore about 25% larger than the earth's biological capacity, or biocapacity, a situation described as overshoot. Overshoot is, by definition, unsustainable, and will result in the deterioration of resource stocks and increasing carbon dioxide concentrations in the atmosphere and ocean.

Another way to look at overshoot is on a per capita basis. Eleven billion hectares of biologically productive space divided between more than six billion people equates to less than two hectares per person. A regional breakdown of the global ecological footprint shows that the North American footprint is about 9.4 hectares per person, the EU's is about 4.8 hectares per person, while the Asia-Pacific and African per capita footprints are 1.3 and 1.1 hectares respectively. In the figure below the height of each rectangle represents a region's per capita footprint, and the width represents its population. The area of the rectangle is proportional to a region's total ecological footprint. In order to achieve a sustainable future, in which humanity's resource consumption remains within the biological capacity of the earth, the world average per capita footprint must fall below 1.8 hectares per person. However, this ignores future population growth, and ignores the

needs of other species which also depend on the earth's productivity. Ultimately we may have to live on less than one hectare of biocapacity per person. The challenge for the coming century will be to reduce humanity's global footprint and, at the same time, enhance the biological productivity of the earth in order to sustain 10 million of our own species, as well as ensure the survival of the myriad others with which we share this planet.

Figure 30.3 Ecological Footprint by Region, 2003



Chapter 31
The State of Ecosystems and Progress of Society

Odeh Al-Jayyousi[i]
Regional Director, IUCN West Asia, Central Asia and North Africa Regional Office, Amman- Jordan

Abstract

This paper intends to shed some light on the global assessment of ecosystems and its implications to the progress of societies. Although assessment of ecosystem services is best addressed at micro-scale, the UN Millennium Ecosystem Assessment (MEA) reveals alarming trends in terms of depletion of renewable resources and threats to our natural capital. Progress of societies will be studied in terms of transformations in the perceptions and valuations of the social and natural capital.

Key words*: ecosystems, society, sustainable development, biodiversity.*

Introduction

Minimising uncertainty and understanding the external environment characterise the purpose of human effort in advancing scientific research and development. Science and research strives to explain, control and predict phenomena around us from material sciences to natural and behavioural sciences. The history of human civilisation is one of converting natural resources (or natural capital) to human-oriented uses.

Between 1700 and 1980, 1.2 billion hectare of agricultural land was gained at the expense of a roughly equal amount of forest (Richards 1990). Such conversion involves loss of species and biological diversity which in turn limits the provision of eco-system services. The problem is further magnified due to pollution as a result of land use and ecosystem changes.

The conversion of natural areas to agriculture has brought benefits in terms of food production, but it has come at a cost of lost species and changes in ecosystem functions. Part of the problem is that the benefits of agriculture are commercial, whereas the benefits of species diversity do not show up in the form of financial forms. The challenge is that economic approaches are faced with the question of what are the benefits of conservation and what are the economic value of ecosystem services and species. There is a critical

i E-mail: odeh.al.jayyousi@iucn.org

need to explain why biodiversity needs to be conserved and what does this mean to the progress of society.

There is a sign that there is a global political concern about biodiversity loss, since most nations ratified the 1992 Convention of Biological Diversity – with the exception on the United States – However, the funding available annually in conserving the world's biodiversity through conventional means as establishing and managing protected areas is limited and does not address the ecological threats (MEA 2005). There is a gap between the scale of the problem and the will of politicians and citizens to tackle it.

Economists tend to be divided between those who regard loss rates as potentially catastrophic and those who find it hard to believe that much of economic value resides in diversity per se. The goal defined by economists for resource allocation is based on utility maximisation and the pursuit of happiness for the greatest number of people. Issues related to environmental conservation are addressed by "welfare economics" which deals with human well-being and what underpins that well-being.

One key strategic option is to adopt economic rationality and reasoning since it is economics that drives land use decisions, scale and location of development.

On the other hand, it is welfare economics that has to inform the policies needed to slow the rate of biodiversity loss and land use change (Jaffe & Al-Jayyousi 2002).

One dimension to the problem of un-sustainability is the population growth and the un-sustainable resource use in North America and Europe. Besides, the emerging economies in China and India impose more pressures on natural resources worldwide in terms of timber and fuel. Hence, there is a need to re-frame and re-define the notion of sustainable development to adequately address progress of societies. The concept of sustainable development and the future of sustainability were debated by The World Conservation Union –IUCN so as to evolve a new paradigm (IUCN 2006).

What Did We Learn from the Millennium Ecosystem Assessment (MEA)?

In 2005, the MEA study was concluded with a contribution of about 1 300 experts from 95 countries over five years. This marked the end of the unique global survey that was called for by the UN. The key message from this global assessment is that humans over the last 50 years are putting such a strain on the planet's ecosystems that their capacity to sustain future generations can no longer be taken for granted. However, the MEA also shows that it is possible to do something about the problems. This requires substantial changes in policy and practice and the conceptualisation of a new paradigm for sustainable development.

In a series of seven reports the MEA has assessed the state of global ecosystems and their role for human well-being. Ecosystems provide food, wood, textiles and medicine. Specifically, ecosystems are a source of aesthetic, spiritual, cultural, and recreational values; they support crop pollination, maintenance of water quality and soil fertility, and regulate regional climate, natural hazards and pests. The MEA demonstrated the interactions between social and ecological systems and should incite policy makers to incorporate ecological issues in all sectors.

The key findings of the MEA indicate that about 60% of the ecosystem services that support life on Earth are being degraded or used unsustainably. It is believed that if nothing happens in the present, the deterioration consequences of this degradation could

become significantly worse in the next 50 years (MEA 2005). Given the scale of human action and impact, it is no longer possible to deal with these issues at a local or national scale. International agreements are essential, and their impacts need to be improved by purposefully co-ordinating between environmental agreements and other international economic and social agreements.

One of the main messages of the MEA is that the international community needs to make environmental conservation a top priority if it wants to meet the UN's Millennium Development Goals (MDGs). The MEA asserted that any progress achieved in addressing the goals of poverty and hunger eradication, improved health, and environmental protection is unlikely to be sustained if most of the ecosystem services on which humanity relies continue to be degraded. Despite the progress achieved in increasing the production and use of some ecosystem services, levels of poverty remain high, inequities are growing, and many people still do not have a sufficient supply of or access to essential ecosystem services. The poor are generally those most vulnerable to the deterioration of natural systems. The regions facing the worst problems of ecosystem degradation – sub-Saharan Africa, Central Asia and parts of South and Southeast Asia- are also those having difficulties in achieving the MDGs.

What is critical is that humans have sound understanding for the environment and how it works, so that they can make the necessary decisions to protect our natural capital. To co-create a sustainable future, we need to devise adequate means to value our natural capital and human resources. Scenarios explored by the MEA predict progress in eliminating hunger, but at too slow rate to halve the number of people suffering from hunger by 2015. The improvements are slowest in South Asia and sub-Saharan Africa, where the problems are biggest. Environmental conditions, in particular, climate change, soil degradation, and water availability, influences this goal through its effect on crop yields and effects on the availability of wild food sources. However, the MEA warns that changes in ecosystems such as deforestation are going to influence the abundance of human pathogens like malaria and cholera and other possible new diseases. Malaria accounts for more than 10% of the disease burden in Africa. Had it been eliminated 35 years ago, the continent's gross domestic product would have been $100 billion larger today. A major result of the MEA study is that ecosystems and the services they provide are economically far more significant than many realise. It is cheaper to conserve ecosystems rather than pollute and clean afterwards. For example, the MEA shows how an intact wetland can be worth $6 000 per hectare whereas one cleared for intensive agriculture is worth only around $2 000 per hectare. Similarly, a mangrove forests in Thailand is worth at least $1 000 per hectare versus about $200 per hectare when cleared for shrimp aquaculture. This implies the need to integrate ecology and poverty alleviation. Sachs and Reid (2006) argue for the need to invest in ecological infrastructure in poor countries and establish a periodic assessment of the benefits that people obtain from ecosystems. In sum, The main conclusions of MEA are: 1) environmental degradation is a major barrier to the achievement of the MDGs. Out of 24 ecosystem services, only the productivity of 4 had been enhanced over the last 50 years. More than 70% of the 1.1 billion poor people surviving on less than $1 per day live in rural areas, where they are directly dependant on ecosystem services; 2) Investing in environmental assets and equitable strategies are vital to achieve national goals for relief from poverty, hunger and disease; 3) Reaching environmental goals requires progress in eradicating poverty.

To transform the existing situation, no scenario in the MEA represents business as usual. Major changes are needed in consumption patterns and education paradigms (Al-

Jayyousi 2001, 2003). Also, there is a need to re-think notions of green technology, Green Fund for mainstreaming environment in national development strategies, and periodic assessments of MEA and higher prices for exploiting ecosystems could reverse the degradation of ecosystems services over the next 50 years.

The following section will address the value and significance of incorporating our natural capital to measure the progress of societies. It also seeks to shed some light on the linkages and synergy between ecology and culture in the future thinking of sustainability and progress of societies.

Our Natural Capital - Learning From Ecology and Culture

Our perception of reality is governed by the metaphors. Conceptualising society as an ecosystem (not a machine) informs our worldviews and mind-sets. As argued by Kiuchi and Shireman (2002), nature is a source of knowledge, a research and development lab with 3.8 billion years of product development.

In a knowledge economy, businesses and societies run not just on fossil fuels and raw materials but also on ideas, information and inspiration, they begin more and more to resemble the creative systems of nature – systems like the prairie, the coral reef and the rainforest. The patterns and cycles of he rainforest, for example, can inform us how phases we see in nature can evolve towards more creative, value creating, life affirming patterns.

In an attempt to link environment and culture, Islam (as will be outlined later) views humanity as part of nature. Also, human life cycles are analogous to ecological processes. Societies can get inspirations and information from nature and ecology. A society founded on information – one whose lifestyle and ideas reflect the idea that knowledge and design are the root source of value – is where the emerging economies like China and India should be heading. Seeing unity within diversity among disciplines and themes of ecology and culture can open new possibilities for harnessing the best of our financial, human and natural capital.

The heavy emphasis of the industrial economy on replication tends to cultivate false beliefs about the source of value. According to the machine model of business and society, it is natural to think we create value by taking in resources so as to develop products along with pollution and waste. But a business does not create profit by consuming raw materials. All profit, all value, is created by design. Trees are made mostly of air, cars and computers are made of rock. What makes them tress, cars and computers is their design, their structure, the way their parts are arranged. The Arabic language and the Islamic term to describe this: *thaher (visible)* and *batin (invisible)*. The former is the seen part of an object. The latter is its back or invisible side. This duality of *thaher and batin* explains the substance and form of the underlying reality.

As we leave the machine economy behind and cultivate a living economy, we discover new drivers for societal values. For example, diversity promotes sustainability simply because *diversity is choice*. The more diverse the organisms in an ecosystem, the more types of resources are available to deal with a challenge, the greater the likelihood of success.

Broadly speaking, in the cultures of the north and west, in Europe, industrialised America and Asia, material affluence is vast. Yet the loss of connection to community, and the alienation it brings with it, can erode our own foundations and undermine our

culture's capacity to sustain itself. In the cultures of the south and east, parts of Middle East, Latin America, Africa and Asia, we are reminded of the value of human connection. Yet many lack the material security they need.

Living systems, such as the water cycle, are regulated by such limiting factors as seasons, weather, sun and soil, each of which is governed by feedback loops. Feedback in nature is continual. Such elements as carbon, sulphur, and nitrogen are constantly being recycled. Societies and business may not only be informed by nature but also reformed by nature and ecological processes.

The following reflects on various logics to draw analogies and contrasts between ecological and economic systems:

Socio-logic

There are no correlations between economic indices and social and ecological ones. For example, In the United States in 1996, a year when the stock market hit new highs, the Fordham University's "index of social health" did not. The index, which tracks problems like child abuse, teenage suicide, drug abuse and child poverty, had fallen 44% below its 1973 high.

In a world where a billion workers cannot find a decent job or any employment at all, humans cannot by any means create a sense of value and dignity in people's lives when we are simultaneously creating a society that clearly has no need for them. Social wounds cannot be saved nor the environment "saved" as long as the economic model adopts the concept of using more natural capital and employ fewer human beings.

With view to meeting the needs of a sustainable future therefore, according to Hawken *et al*. (1999), contemporary business economics is the equivalent of pre-Copernican in its outlook. A society that squanders its resources wastes its people and *vice versa*. Just as overproduction can exhaust topsoil, so can over-productivity exhaust a workforce.

Of the $9 trillion spent every year on consumption and investment in the U.S., at least $2 trillion, Hawken *et al*. (1999) asserts, is wasted. An example of waste familiar to everyone is sitting in a traffic jam on a congested freeway.

Eco-logic

As confirmed by the findings of the MEA, humans are destroying the most productive systems ever seen on earth. Economics and technology cannot function as a reliable guide until natural capital as well as human capital are placed on the balance sheets of companies, countries and the world (Hawken *et al*. 1999).

A sustainable society, community or ecology needs an accurate balance sheet that reflects the status of the ecosystem services and enhances the flow of financial, manufactured, human and natural capital. The currents economic metrics do not account for the value and benefit for the ecosystem services.

Archi-logic

Green architecture advocates that 'organic' building that would 'integrate art, natural and local materials, sunlight, green plants, energy conservation, clean water', not to

mention happy employees. Fundamentally then, green buildings are superior to ordinary structures as a result of the same sort of design integration that makes hypercars better than ordinary cars. The key to sustainability according to Hawken *et al*. (1999) is to seek that all four capital stockholdings (financial and manufactured, natural and human) are as prudently stewarded.

The next industrial revolution then, after the digital revolution for is an ecological. This ecological state of society can best take place in a newly integrated form of democratic, market-based system of production and distribution, integration and restoration.

The key implications and requirements for this ecological revolution include increased resource productivity, enhancement of flow and service and reduction or elimination of waste.

Reducing and eliminating the very idea of waste, can be accomplished by redesigning industrial, or utility based, systems on biological lines. This serves to change the nature of industrial processes and materials, enabling the newly sustainable company to maintain and develop a constant use of materials in continuous closed circles. This entails a new perception of value, a shift from the acquisition of goods to the purchase of services, whereby quality, utility and performance is continually sought to promote natural and social well-being.

Through a restorative economy, a reversal of worldwide planetary destruction is required, through reinvestments in sustaining, restoring and expanding stocks of natural if not also social capital. As a result, the biosphere will be able to produce more abundant ecosystem services and natural resources. This is the realm, in fact, with which the World Conservation Union -IUCN is most particularly concerned.

Ecosystems and the Progress of Societies

Progress of societies is a process rather than a product and it is evolving through a balanced realisation of all capitals including natural capital, social capital, financial and manufactured (technology). Measuring and defining (or re-defining) progress of societies relies mainly on a societal embodiment of a dream that integrates matter and spirit. In this section I am attempting to induce a paradigm shift with respect to the definition of the western goal as the pursuit of happiness. My aim is at evolving and reviving a culturally-based notion for the progress of societies that have resonance to norms and values of the Middle East; i.e, Islam. Paradoxically, Islam is currently viewed as a threat rather than a remedy or cure from the fragmentation of knowledge, matter and spirit and human and the community. In this paper, I am arguing that Islam can offer a fresh and renewed model for a re-definition of progress of societies. This is attained through the articulation of what constitutes a good life (Hayat Tayebah) and also part of a larger process of civic renewal (Al-Jayyousi 2000).

The end of the millennium witnessed three *independent* processes, i.e, the information technology revolution; the economic crisis of capitalism; and the emergence of cultural social movements, such as human rights, feminism and environmentalism. Castells and Blackwall (2000) argued that the interaction between these processes, and the reactions they triggered, brought into being a new dominant social structure, the network society; a new economy, the informational/global economy; and a new culture. The logic embedded

in this economy and society (in this case Islam) underlies all institutions in an interdependent world.

It is valuable that the discourse and knowledge of biological diversity with its ecological networks and feed back processes inform and reform our views to cultural diversity and the notions of progress of societies. If we view society as an ecological system, hence societies can learn many lessons from ecology and biological systems. This means a genuine transformation of the metrics of measuring sustainability and progress of societies.

Transformation of Gross Domestic Product (GDP)

In liberal economic theory, happiness was already an economic measurement used interchangeably with utility as well as general welfare. Economists attempt to quantify happiness through measurements in consumption and profits. The underlying assumption is that when societies consume more of a product it means that it is good and generate happiness. It is this equating of high consumption levels with happiness that has been challenged by proponents of gross national happiness (GNH). GNH depends on a series of subjective judgements about well-being to address the limitations of gross domestic product (GDP).

Taking care of the earth (our natural capital) is currently viewed a key component in realising sustainable development. After the results of the MEA, there is stronger evidence that ecological sustainability is fundamental to safeguarding human as well as planetary health.

However, gross domestic product (GDP), remains the dominant indicator of national performance for most developed countries. Measuring progress covers a wide range of attributes which include health, social and spiritual dimensions. A number of experts challenge the GNP as an adequate measure for assessing the progress of societies (Korten 1995). As a result, genuine progress indicator (GPI) and happy planet index (HPI) were developed as a refined version of GDP to address other dimensions like the state of environment and other social and health aspects.

The GPI is an index that intends to reflect the sustainable economic welfare. The formulation of GDP was based on the basics of GDP but modified to address the value of household and community work, and deducts the costs of commuting, pollution, land degradation and industrial accidents. However, the GPI does not account for either the accumulation or decline of human capital (health, skills, knowledge and experience) or social capital as sound and stable institutions and supportive communities.

The GPI advocates argue that the GDP commands such overwhelming attention that needs to be replaced by a similar, single measure, which is underpinned by considerations of sustainability.

On the other hand, other indices include the Happy Planet Index (HPI) which is a measure that shows the ecological efficiency with which human well-being is delivered. The index is built from three different indicators, two of which are objective: life expectancy and ecological footprints and the third is people's well-being or "life satisfaction" which depends on the world views and the mental and physical health. Human Development Index (HDI) is another index developed in 1990 and consists of three elements; standard of living, life expectancy, and knowledge.

In sum, all what these measures attempt to do is to ensure that markets tell us the ecological and social truth.

Islam and Re-definition of Progress

Having the venue of this conference in Turkey sheds some lights on the value of harnessing this forum for conveying new notions of progress of societies and sustainable development between East and West. Turkey can play this role by provide a synergy between both biological diversity and cultural diversity. It is enlightening to realise the historical role of Islamic city in promoting co-existence, socialisation of knowledge, and dialogue between cultures (Al-Jayyousi, 2004). This role is dearly needed in a world that is deeply divided by the imagined clash of civilisations.

Islamic thought looks upon the challenges of the twentieth century as a crisis of values. It realises the need to re-examine the foundations on which the entire structure of society is built. The crisis in economic, ecological and political relations is the natural outcome of values and institutions that characterise modern civilisation. It is believed that through a thorough understanding of the social values of Islam, set in the context of sustainable development, value-oriented communities can develop (and transform) a creative and innovative approach to the challenges confronting humanity today. Paradoxically, the dominant worldview (in the media) projects Islam as a potential global threat, the intent is to transform this mis-conception and present Islam as a source of remedy and a way to redefine progress of societies.

Islam covers all aspects of human life. It regulates the relationships between God, Human, and Nature. It is based on the recognition of the unity of the Creator and of human's submission to His will. Everything originates from the One God, and everyone is responsible to Him. Human is viewed as a trustee (*khalifah*) and a witness (*shahed*). His role and responsibility is to ensure that all resources, physical and human, are utilised in a reasonable, equitable, and sustainable manner.

Nature is created by God (*Allah*) for the benefit of humans. The relationship between humans and Nature is based on harmony, since all creatures obey the laws (*sunan*) of God. Humans are urged to explore and utilise the natural resources in a sustainable manner. It is through submission to the Will of God that peace is brought about. Harmonisation of human's will with the Will of God leads to a responsible and balanced life. Every human activity is given a transcendent dimension; it becomes sacred, meaningful, and goal-centred.

Islam does not approve the useless cutting of trees and bushes. Man can use their fruits but he has not the right to destroy them. Nor does Islam allow waste among even lifeless things, to the extent that it disapproves of wasteful use of water, even if there is no scarcity of water. Its vowed purpose is to avoid waste in every conceivable form and to make the best use of all resources.

Humans have been endowed with countless powers and faculties. They possess intellect and wisdom to achieve balance in this Universe. Their very life and success depends on the proper use of these powers. God has also provided humans with all those means and resources to make their natural faculties function and to achieve the fulfilment of their needs. The environment and surroundings contain resources that are harnessed for humans. Human beings should cooperate and exchange knowledge (Hikmah) to establish a better and prosperous life. The proper use of human's powers implies the attainment of benefits to public or what referred to as "construction of the world (Emmartu al Kawn).

Every other use of resources which results in waste or destruction is wrong and unreasonable.

A summary of the main Islamic guiding principles and axioms with reference to resource management and the utilisation of knowledge are outlined below:

1. Assimilation of knowledge (Hikma): Islam adopts the notion that knowledge and wisdom is of universal nature. Members of the Muslim community are encouraged to utilise and benefit from the accumulated experiences of other nations. Socilaisation among cultures is identifed as a key message for all nations to understand and know each other.
2. Quality of performance: Islam urges believers to perfect their deeds and to be kind and merciful to other creatures (including animals and natural resources). The key purpose of human is to do well deeds and support good causes.
3. Conservation: It was stated in both Quran and Hadith that waste in all forms is not acceptable.
4. Social responsibility: Caring and sharing the community's concerns are emphasised in Islam. Members of the society, are urged to be concerned about public affairs and issues. They should undertake consultations in all matters. These cover a wide array of civic life and resource management aspects.
5. These notions from Islam can constitute new parameters for progress of societies. Transforming Islam to a knowledge-based and diverse worldview will help both our ecosystems and the global community.

Conclusions

There is a sign that there is a global political concern about biodiversity loss, however, the funding available annually in conserving the world's biodiversity through conventional means as establishing and managing protected areas is limited. There is a gap between the scale of the problem and the will of politicians and citizens to tackle it. One dimension to the problem of un-sustainability is the population growth and the un-sustainable resource use in North America and Europe. Besides, the emerging economies in China and India impose more pressures on natural resources worldwide in terms of timber and fuel. Hence, there is a need to re-frame and re-define the notion of sustainable development to address progress of societies.

The global ecosystem assessment asserted that any progress achieved in addressing the goals of poverty and hunger eradication, improved health, and environmental protection is unlikely to be sustained if most of the ecosystem services on which humanity relies continue to be degraded.

To transform the existing situation, major changes are needed in consumption patterns and education paradigms. Also, there is a need to re-think notions of green technology, Green Fund for mainstreaming environment in national development strategies, and periodic assessments of global ecosystems and higher prices for exploiting ecosystems could reverse the degradation of ecosystems services over the next 50 years. A sustainable society, community or ecology needs an accurate balance sheet that reflects the status of the ecosystem services and enhances the flow of financial, manufactured, human and natural capital. The currents economic metrics do not account for the value and benefit for the ecosystem services.

It is valuable to evolve and revive a culturally-based notion for the progress of societies that have resonance to norms and values of the Middle East; i.e, Islam. Paradoxically, Islam is currently viewed as a threat rather than a remedy or cure from the fragmentation of knowledge, matter and spirit and human and the community. It is argued that Islam can offer a fresh and renewed model for a re-definition of progress of societies. This is attained through the articulation of what constitutes a good life (Hayat Tayebah) and also part of a larger process of civic renewal.

The essence of progress of societies, from an Islamic perspective, is that human purposefully cooperate and exchange knowledge (Hikmah) to establish a better and prosperous life. The proper use of human's powers implies the attainment of benefits to public or what referred to as "construction of the world (Emmartu al Kawn) in a sustainable and equitable manner. The ecological revolution can be informed and reformed by notions from Islam to create unity within diversity and integrate all forms of capitals.

References

Al-Jayyousi, O., (2001). Capacity building for desalination in Jordan: Necessary conditions for sustainable water management, Desalination, Vol. 141, p.169-179.

Al-Jayyousi, O. (2003). Scenarios for public-private partnerships in water management: A case study from Jordan. Water Resources Development. Vol. 19(2), p. 185-201.

Al-Jayyousi, O. (2004). Knowldeg creation in the water sector: towards a learning organisation.Water Resources Development. Vol. 20 (2), p.165-175.

Castells, M. (1996). Rise of a network society. Oxford, Blackwell.

Hawken, P. and Lovins A. and Lovins L. (1999). Natural Capitalism, Rocky Mountains Institute. CO, United States.

Jaffe M. Al-Jayyousi, O. (2002). Planning models for sustainable water resource development. Journal of Environmental Planning and Management, Vol. 45(3), p.309-322.

Kiuchi, T. and Shireman (2002). What we learned in the Rainforest: Business lessons from nature. Berrett-Koehler Publications Inc. San Francisco.

Korten D. C., (1995). When Corporations Rule the World. Berrett-Koehler Publishers, Kumaian Press.

Millennium Ecosystem Assessment (MEA), (2005). Ecosystems and Human Wellbeing. Synthesis. Island Press. Washington, DC.

Richards, J., 1990. Land transformation. In the *Earth as Transformed by Human Action. (Eds. B. Turner, H. Clark., R. Kates., J. Richards., J. Mathews & W. Meyer), p. 163-78. Cambridge University Press, Cambridge.*

Sachs, J. and Reids D. (2006). Investments Towards Sustainable Development. Science, 312 (2776): 1002. The World Conservation Union – IUCN (2006). The Future of Sustainability. Gland, Switzerland.

Chapter 32
How Reliable are Climate Change Projections: A Statistician's Perspective

Dennis Trewin
Statistical Consultant, Australia

Abstract

This paper argues that the work of the International Panel for Climate Change (IPCC) contains some potentially serious statistical flaws that will tend to exaggerate the extent of future climate change (beyond 20 years). The author presents the IPCC's seven different climate-change scenarios, which depend on assumptions about economic growth, population and energy sources used, among other things. The author then describes four major concerns about these scenarios: (1) the Castles/Henderson criticism centered on the use of exchange rates to compare the size of economies; (2) the economic convergence assumption; (3) population growth rate assumptions; and (4) lack of estimates as to the robustness of the climate change projections. The author then presents his perspective on global warming, including which of the IPCC scenarios is the most statistically viable, and suggests how the statistical flaws could be addressed.

Introduction

Climate change is one of the most important, if not the most important, global policy issues facing us today. Quite rightly it is the subject of considerable scientific work and much policy debate. I am going to be critical of some aspects of the work of the International Panel for Climate Change (IPCC) in my comments. Please don't interpret that as me being one of those climate change sceptics. I am not. I do believe it is an issue that our societies must address. But the policy interventions should be based on the best possible evidence and I believe the IPCC's work on climate change projections falls short of that standard. Furthermore, some of the commentary based on their work exaggerates their findings.

Having made that critical statement I would like to recognise the strength of much of IPCC's work. It is wonderful that it has been possible to being together so many experts to try to develop a collective view on a very important topic. It is one of the most outstanding examples of international collaboration. Furthermore, their work has had a substantial influence on national and international policy. And the outputs of the work of the IPCC will provide an excellent base for much more scientific work and will confirm

to do so. I should also commend them on the transparency of their work even though it makes it easier to criticise, this work.

Unfortunately the experts gathered together for the IPCC's work did not include statistical experts. This has resulted in some potentially serious flaws in the statistical work of the IPCC which I will describe shortly. These flaws will tend to exaggerate the extent of future climate change. These upward biases are not sufficient to suggest there is not really a climate change problem but they may be sufficiently important to influence the adaptation policies of individual countries. If unnecessary policies are introduced, public funding could be wasted possibly at the expense of additional funding for items such as health, education or other environmental issues, other than climate change.

You could think of policy interventions on climate change as a dichotomy – those that are aimed at reducing the growth in greenhouse gas emissions and those aimed at adapting to the impact of future climate change.

Starting with the first issue, in most countries, there is general agreement that the level of greenhouse gas emissions has to be reduced with much debate on appropriate policy interventions. The work of the IPCC has been a major influence. Their work has shown without doubt there has been global warming, increases in sea levels, etc. They have also shown fairly convincingly that human activity is a major contributor to the increase in greenhouse gases. I have no quibble with this work. My concern is with the projections of possible future climate change. Whilst, in my view, there may be upward biases in their estimates of future climate change, they are not sufficient to suggest that climate change should not be addressed with some urgency. But these biases may be important for determining appropriate responses or adaptation to future climate change. For example, the policy response would be quite different if the increase over the rest of this century is likely to be 2 degrees rather than 6 degrees. Governments and others deserve the best possible evidence base – there is a lot at stake.

And it is not just governments that might be impacted. One private sector example is insurance. Insurance companies are likely to use IPCC projections to help them assess risks. If the risks are judged to be higher than what they might be, those taking out insurance are likely to pay too high a premium.

The public will also be disadvantaged if public spending is sub-optimal and strong needs, such as those in health, education, water and waste management for example, are not met by the public purse because funds are diverted to climate change adaptation policies that may go beyond what is unnecessary.

Of course I recognise there is considerable uncertainty in the predictions. And that current and future policy responses, or lack of responses, may lead to decreases or increases in greenhouse gas emissions compared with current scenarios. There will be other external factors at play which also impact on the future. The rapid decline in population growth over the last decade or so was not predicted by most when I was doing university studies some 40 years ago!

You might ask what are my credentials for making the criticisms I am about to make. First, I was the Head of the Australian Bureau of Statistics for nearly 7 years. Second, I am a past President of the International Statistical Institute, the peak global body for statisticians. Third, I am Chairman of the Global Executive Board of the International Comparison Program. This Program, which is hosted by the World Bank, is responsible for providing up-to-date estimates of purchasing power parities (reference period 2005-2006) and it is the non-use of power purchasing power parities which will be one of my

criticisms. Fourth, I have been a member of the Australian State of the Environment Committee since 1998. This Committee is charged with providing an independent report on the State of the Environment to the Australian Government.

I have recently retired as Australian Statistician so my views should not be interpreted as representing those of the Australian Government.

The Conclusions of the IPCC

These are not always well understood. Many commentators have a tendency to exaggerate the likely impact, using the upper limit of the IPCC projections, and causing some public alarm.

According to the IPCC's Summary for Policy Makers, temperatures are expected to increase by about 0.2 degrees per decade over the next two decades for a range of scenarios. That is, temperatures are expected to increase by nearly 0.5 degrees by 2030.

A lot of the projected increase is due to past actions and an increase of about 0.3 degrees would be expected even if greenhouse gas concentrations could be maintained at 2000 levels.

Beyond the next two decades, different outcomes result from the different scenario groups. The 'worst case' A1FI scenarios projects temperature increases of 2.4 to 6.4 degrees (with a best estimate of 4.0 degrees) by the end of the century. Sea levels are projected to increase by 0.26 to 0.59 metres over the same period. Their 'best case' B1 scenarios project temperature increases of 1.1 to 2.9 degrees (with a best estimate of 1.8 degrees) by the end of the century. Sea levels are projected to increase by 0.18 to 0.38 metres. Recent public announcements suggest many countries are trying to reduce the level of their greenhouse gas emissions. The speed at which these changes are made is also important because the accumulative affects of greenhouse gas emissions are very important.

Climate Change Scenarios

The scenarios are based around six basic storylines.

A1 – This describes a future world of very rapid economic growth, global population that peaks in mid century and declines thereafter, and the rapid introduction of new and efficient technologies. Economic convergence among the regions is assumed with a substantial reduction in regional differences in per capita income. There are three subgroups within the A1 family distinguished by the direction of their technological emphasis.

A1FI – fossil intensive energy sources

A1T – non-fossil energy sources

A1B – balanced across all sources

A2 – This describes a very heterogeneous world with emphasis on self reliance and preservation of local identities. Continually increasing population is expected. Economic development is primarily regionally oriented and per capita economic growth and technological change more fragmented than in other storylines.

B1 – Similar to A1 but with a rapid change in economic structures toward a service and information economy with reductions in material intensity and the introduction of clean and resource efficient technologies.

B2 – This storyline has an emphasis on local solutions to economic, social and environmental sustainability. Population continues to increase but at a slower rate than A2. Intermediate levels of economic development are assumed with less rapid and more diverse technological change than in the A1 and B1 storylines.

Part of my argument will be that the statistical biases are such that some scenarios with poorer outcomes are much less likely than those with better outcomes. The scenarios should not be treated as equally likely. This impacts the longer term projections only. Over the next 20 years projected increases are much the same for all scenarios. It is over the longer term that major differences from the scenarios emerge.

If policy interventions lead to even more optimistic scenarios than those described in the box above, projected outcomes would be even better over the longer term. I will not comment on this aspect as it is beyond my brief. But my comments will suggest that the B1 scenario family may be most likely of the scenario families but with a lower level of economic convergence.

What are my Concerns?

They are summarised below and expanded in the following paragraphs.

1. The first criticism is the so called Castles/Henderson criticism. Because exchange rates rather than purchasing power parities are used to compare the size of economies in the climate change models, the economic growth of developing countries (where exchange rates tend to be undervalued compared with purchasing power parities) has an upward bias where economic convergence is assumed as is the case in many of the scenarios. (see next point)
2. The second criticism is the economic convergence assumption in many of the scenarios. For example, in the A1 scenario average incomes are predicted to converge by nearly 2% per annum. For the A2 scenario the convergence is a much more realistic 0.5% per annum. (For the B1 and B2 scenarios, assumed convergence is somewhat higher than that assumed for the A2 scenario but a lot lower than for A1.) While there is likely to be some convergence move in this direction, the most likely outcome is that most of the developing countries of today will still be behind the more developed countries and hence the global growth rates will be overstated.
3. The median population growth rate assumptions used in the climate change models reflect the growth rates of 20 years ago not those of today. During this time, global population growth has decreased from 2.5% per annum to 1.2% per annum.
4. There are no estimates of the robustness of the climate change projections to the various assumptions in the models.

In the third IPCC report too much emphasis was on presenting the results as an estimated range of 1.6 to 5.8 degree increase by 2070 implying this is a rectangular distribution with many commentators interpreting this as meaning each outcome is equally likely. This is not the case. The median projection, based on the full range of

scenarios, is towards the lower end of the range. The upper level of the range is a scenario based on more extreme assumptions. Yet, in my country, we get many media reports suggesting Australia's temperature will increase by 6 degrees by 2070. The IPCC did not say this but the way they presented the results in documents most accessed by the public did not discourage this sort of response.

I am pleased to say that in the fourth IPCC report, this problem has largely been addressed. There is much more information to enable the distribution of the scenarios to be assessed. Importantly, they have provided a best estimate for each of the scenario groupings.

The first criticism is the one that has had most public debate. I will try to explain what happens as a result of using market exchange rates rather than purchasing power parities. The main impact is at the starting point for the models – 1990. Because market exchange rates are used to compare the size of developing economies, there is a tendency to under-estimate the size of developing economies compared with the developed economies which in turn are over-estimated. This is because market exchange rates tend to under-estimate the 'true' comparative costs of developing countries (ie goods and services seem relatively cheap in developing countries). So, with a lower base than there should be and the economic convergence assumptions, growth rates tend to be exaggerated for developing countries. The impact is clearly shown for scenarios A1 and B1 in Table 32.4 below where more rapid economic convergence is assumed and the growth rates seem very high.

Table 32.4 Comparative Annual Growth Rates in GDP (on a per capita basis)

Scenario/Source	Annual Growth Rate (%)
A1	3.1
B1	2.4
US Department of Energy	1.9
B2	1.8
World Bank	1.8
International Energy Authority	1.5
A2	1.0

The economic growth rates for the A1 and B1 scenarios are much greater than for the US Department of Energy's, World Bank's or IEA's economic projections. I believe the reason is what I have outlined above ie the non-use of purchasing power parities, combined with convergence assumptions for many scenarios. In fact the data set above indicates that the economic convergence assumption rather than the non-use of PPPs may be the main reason for the high level of economic growth implicit in these scenarios. Why not use the authoritative World Bank comparative growth assumptions in at least one scenario set? I find the arguments that this estimate is not in the peer reviewed literature very hollow – it probably is more authoritative than most of those in the peer reviewed literature.

The differences are greatest for the so-called Africa/Asia/Latin America/Middle East region (which comprises most of the developing countries) where the growth rates in the A1 and B1 scenarios are clearly above the upper end of the projections from the DoE, World Bank and IEA whilst A2 and B2 scenarios fall near the centre. This reinforces my argument that the non-use of PPPs, together with the convergence assumption, may be leading to an upward bias in the economic growth assumptions in the A1 and B1

scenarios. This, in turn, will lead to estimates of excessive growth in energy demand and emission levels.

With respect to the second criticism, the World Bank also provides regional growth projections which could be used instead of convergence assumptions. Indeed these estimates imply a level of convergence between the regions.

The impact of these flaws may not be as great as perceived. I understand the bias affects both the 'numerator' and 'denominator' in the models. For example, the amount of technology change leading to energy efficiency is also based on economic growth rates. Also a high proportion of the greenhouse gases are already in the system. Independent sources suggest the overall impact may result in an upward bias of up to 0.5 degrees in the best estimate from the relevant scenarios. Although this does not change the overall conclusions of the climate change modelling, it may still be important and retaining this statistical flaw in the models provides ammunition for the critics and undermines the overall credibility of the work of the IPCC. And suggestions have been made for practical ways on incorporating PPPs into the climate change models by Nordhaus, for example, in his paper for the IPCC Seminar on Emission Scenarios.

Turning now to the third criticism about population growth assumptions, this may be the most important of my criticisms. The official population projections produced by the international community are not used because they are not in peer reviewed literature. Yet they are probably more authoritative and up-to-date than most in the literature. Population growth has slowed remarkably in recent years and many of the peer reviewed projections would not have caught up with more recent trend. This has been driven by changes in Africa and Asia in particular. In both continents, fertility rates have declined substantially and in Africa mortality rates have increased because of the AIDS epidemic. Mortality has also increased in the former Soviet countries chiefly because of lifestyle reasons. The IPCC have noted the difference, in their report and they do plan to incorporate lower population projections in the scenarios being developed for the fifth report.

The annual growth rate from the medium United Nations population projections is 13 to 19% lower than those in B2 and will be even lower than those in A2. However they are broadly consistent with those in A1 and B1. In fact they show peak slightly earlier than the medium UN projection. Scenarios A2 and B2 assume continually increasing population whereas A1 and B1 assume a population that peaks mid century.

It is interesting to note that difference between the best estimates for A1B and A2 as well as B1 and B2 is 0.6 degrees. This gives an order of magnitude estimate of the impact of the different population assumptions.

Why not develop a set of scenarios based around the authoritative economic and population growth assumptions?

With respect to the fourth criticism, it would be very useful to have estimates showing a robustness of the scenarios to changes in the key assumptions. Taking the A1 family as an example,

- What would happen if the World Bank's projections of economic growth were used instead of those currently used?
- What would happen if higher population growth assumptions were used?

In working out how to adapt to climate change, governments and others are engaging in a form of risk management. It would help them considerably if they were better able to answer a range of 'what if' questions such as those posed above.

Validity of the Scenarios

In light of the above analysis, I would make the following comments on the likelihood of various scenario groups.

A1 – Unlikely to occur because economic growth assumptions are unrealistically high.

A2 – Unlikely to occur because population growth assumptions are unrealistically high

B1 – This is the most likely to occur although it probably overstates the degree of economic convergence (but this appears to be less important than the population growth assumptions)

B2 – This is unlikely to occur because of unrealistically high population growth assumptions.

It may be timely to remind ourselves of the most important climate change projections for B1. The starting point I am using is 2007 so have made appropriate adjustments to the IPCC estimates which use 1990 as its starting point.

Table 32.5 Best Estimates and Likely Range

	Best Estimate	Likely Range
Temperature change to 2030	0.5	0.3 – 0.75 dgs
Temperature change to end of century	1.5	0.8 – 2.6 dgs
Sea level rise to end of century	n.a	0.15 – 0.35m

What Would I Advise Policy Makers?

My competence is as a statistician so I should not provide advice beyond this competency. Given this, I would advise the following:

1. Climate is extremely variable. It is unwise to read too much into short term variations. For example much of south and east Australia is going through a very dry period but there have been two similar dry periods since 1900. This dry period may or may not be due to climate change.
2. There has been definite global warming since 1900 and associated rises in sea levels. These increases are statistically significant. Furthermore statistical models have demonstrated that human activity (especially greenhouse gas emissions) has been a major contributing factor to global warming and associated impacts.
3. Although not yet statistically significant, there are clear signs of distributional impacts from global warming. In Australia, the northern areas appear to be becoming wetter and the southern and eastern areas becoming dryer. Cyclonic activity seems more frequent in the northwest and less frequent in the north east. There also appear to be more frequent extreme events. Although these trends are

not statistically significant, and may be due to natural climate variability, the distributional impacts are consistent with climate change models.

4. Without steps to address greenhouse gas emissions further climate change and rises in sea levels could be expected. The impacts are cumulative which increases the importance of addressing this issue before the accumulations become too great.

5. The best estimate of temperature increase to 2030, under a range of scenarios, is an increase in temperature of 0.5 degrees (range 0.3 – 0.75) and an increase in sea levels of 8 centimetres. Higher estimates in this range are unlikely because they are based on scenarios with unrealistically high growth rates. The distributional impacts and possible increase in extreme events may be important from a policy perspective and some allowance should be made for this.

6. Looking further out there is a greater degree of uncertainty in the projections but the B1 group of scenarios is the most consistent with authoritative assumptions on economic growth and population growth. According to this range of scenarios temperature change from now to the end of the century is estimated to be 1.5 degrees (range 0.8 – 2.6) with the range of sea level increases from 15 – 35 centimetres.

7. Given that there is still a lot of uncertainty on what might happen, an adaptive approach might be considered. For example, policies could be based on best estimates at present with some hedging for risks but regular reviews of the situation as more information on climate becomes available.

Conclusions

My conclusions are:

1. There are statistical flaws in the climate change models which tend to exaggerate the extent of future climate change.

2. The IPCC has been reluctant to correct these flaws but seems to be considering revised scenarios on population projections in their Fifth Assessment. In the meantime, a trusted body such as the OECD should be commissioned to assess the impact of these flaws so governments have the best evidence base on which to develop climate change policy. Specifically, it should develop climate change projections using scenarios based on authoritative estimates of economic and population growth with the use of PPPs to when aggregating economies along the lines suggested by Nordhaus.

References

IPCC (2007), Climate Change:The Physical Science Basis, Summary for Policymakers

IPCC (200), Working Group III report for the Fourth Assessment Report

Nordhaus, W. (2006), "Alternative Measures of Output in Global Economic – Environmental Models: Purchasing Power Perity or Market Exchange Rates.", prepared for IPCC Expert Meeting on Emission Scenarios

Part Eight
Gender: A Key Dimension of Societal Progress

Chapter 33
Social Institutions and Gender Equality: Introducing the OECD Gender, Institutions and Development Data Base (GID-DB)

Denis Drechsler; Johannes Jütting; Louka T. Katseli
Development Centre,OECD

Abstract

The OECD Development Centre's Gender, Institutions and Development Data Base (GID-DB) is a new cross-country research tool with comprehensive measures of gender equality. It improves upon existing sources because it is the only data base on gender that systematically incorporates indicators of social norms, traditions and family law. The GID-DB thereby permits analysis of hypotheses that link cultural practices to gender equality, human development and economic growth. A cross-country comparison of the data indicates that inequalities in social institutions are particularly pronounced in countries with low female literacy rates. Similarly, our analysis suggests that inequalities in social institutions are not directly linked to a country's level of income.

Introduction

Gender equality has been defined in terms of "equality under the law, equality of opportunity … and equality of voice" (World Bank 2001); as such it is a policy objective of nearly universal acceptance. Gender equality moreover has instrumental value because it enhances the long-term growth prospects of countries. For these and other reasons, its pursuit is one of the eight Millennium Development Goals promulgated by the United Nations. Measuring the status and tracking the progress of gender equality is consequently an important undertaking, but a difficult one given the various dimensions along which discrimination against women occurs (Klasen 2007).

This paper contributes to the identification and development of valid measures of gender equality by presenting the OECD Gender, Institutions and Development Data Base (GID-DB).[1] The GID-DB provides data covering all key dimensions of gender equality, including information on social norms, traditions and family law (collectively referred to herein as social institutions). It thereby allows researchers to test various hypotheses on the determinants of gender equality, to analyse the effects of gender inequality and to build composite indicators of gender equality.

The GID-DB is an important complement to existing data compilations. Commonly used sources like the *Human Development Reports* of the United Nations Development Programme (UNDP), the United Nations' *World's Women Surveys* and the World Bank's data base *GenderStats* mostly provide sex-disaggregated data in the areas of health, education, employment, political participation and income. They nevertheless fail to include information on social institutions, which have been identified as an important determinant of the well-being of women and the long-term prospects for economic growth (World Bank 2001; Antecol 2003).

The paper is organised as follows: the section, "Construction of the GID-DB: How to Measure Social Institutions?", presents the main building blocks of the data base and discusses in detail how the new social institutions variables are measured and coded. Following section "Illustrating the Value of the Data Base"illustrates the usefulness of the new data base, highlighting patterns of gender inequalities related to regions in the world and across income groups. It also provides an example of how the GID-DB can be used to address important policy questions by looking at the determinants of female labor force participation. One important part of the data base is devoted to report a country's performance in gender equality measured through composite indicators like UNDP's Gender-related Development Index (GDI) or the Gender Gap Index which was recently introduced by the World Economic Forum. The specific merits and limitations of these different indices compared to our newly established Social Institutions and Gender Index (SIGI) are discussed in the section "Composite Indicators: A comparative Assessment". The last section concludes.

Construction of the GID-DB: How to Measure Social Institutions?

The Gender, Institutions and Development Data Base is a comprehensive data collection that presents 60 indicators related to gender equality. All world regions and country income categories are well represented in the GID-DB. There are 161 economies in the data base, though the number for which measures of social institutions are available is generally lower. With a few exceptions, economies with fewer than 1 million inhabitants are not considered.

The GID-DB is built around five major blocks that either determine and/or measure gender equality: social institutions, women's access to resources (*e.g.* health and education), political empowerment, the economic status of women and composite indicators measuring gender equality. Given the data base's focus on gender-related differences rather than the absolute values of a particular indicator, many variables are presented as ratios. Thus the GID-DB not only includes the percentage of female students enrolled in primary, secondary and tertiary education, but also provides the corresponding female/male enrolment ratios. Similar rationales apply to health data (*e.g.* life expectancy) and information on the economic role of women (*e.g.* economic activity rate): in both cases, variables are presented in absolute terms and relative to the corresponding value for men.

So far, the literature on the determinants of gender equality has primarily focused on either measuring women's socio-political status (*e.g.* education, health and political rights) or on constructing composite indicators (*e.g.* Klasen 2007). The importance of social institutions for gender inequality, however, has been largely overlooked. This is an important omission as deeply enshrined norms, values and attitudes may arguably be the most important determinants for gender equality: they have generally been in existence

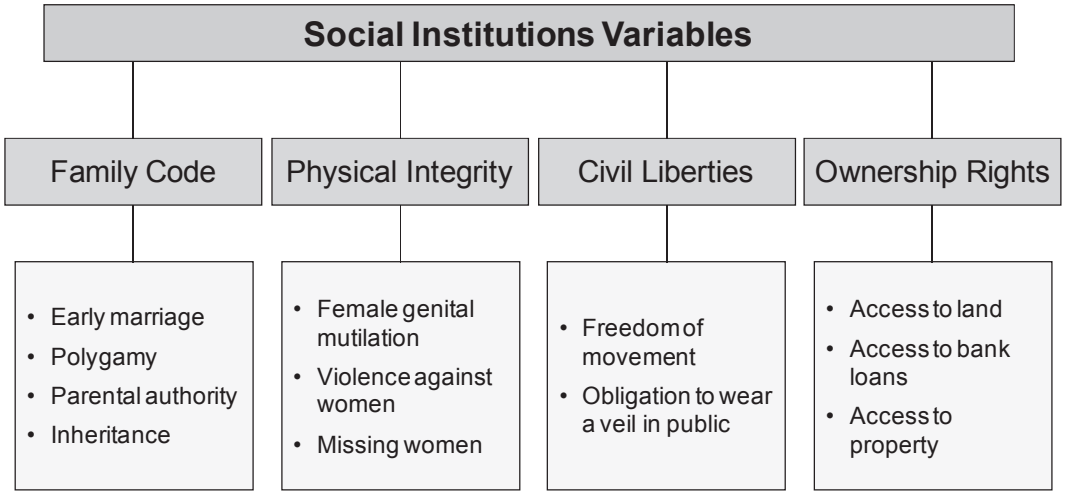
for centuries, are extremely difficult to change and frequently override formal laws and regulations (Sen 2007).

In order to give a broad overview of gender (in) equality that is rooted in social institutions, we distinguish twelve single indicators that we group into following sub-categories: (i) the prevailing family code; (ii) women's physical integrity; (iii) women's civil liberties; and (iv) women's ownership rights (Figure 33.1). Each of these variables is coded between 0 (equality) and 1 (high inequality). Ratings of social institutions variables generally consider the extent of inequality[2] as well as the size of the female population that suffers from its application. For example, a very discriminatory institution might be prevalent in a social group that constitutes only 40% of a country's population; then for that observation the value of the indicator will be

$$1 \times 0.40 = 0.4$$

where the first term indicates the highest level of discrimination and 0.40 takes account of the fact that only 40% of the population is affected.

Figure 33.1 The GID-DB's Social Institutions Variables



Source: Own Illustration.

Family Code

Four variables[3] are grouped in the "family code" sub-index[4]: early marriage, polygamy, parental authority, and inheritance practices. Inheritance practices measure whether bequests are equally shared between male and female offspring. Depending on the degree to which regulation is in favor of male heirs, the variable is coded between 0 (equal treatment of sons and daughters) and 1 (inheritance is only given to male offspring). Parental authority is coded 1 for a society where fathers, as a rule, have complete control over their children and 0 where they evenly share authority with their children's mothers. Whether parental authority is granted equally to men and women and whether women are discriminated against in inheritance practices is documented in a study commissioned by the French Parliament (Lang 1998).

A social institution of special relevance is that of early marriage: where very young women are married, parents (fathers) and not young women themselves have the power to make important decisions about marriage and household formation. Moreover, within households, the generally older husbands have disproportionate authority and decision-making power. We use the percentage of women married before the age of 20 reported by the United Nations Development Programme (2004) to construct our early marriage indicator, which varies between 0 (early marriage does not exist) to 1 (all women have been married before the age of 20 at least once).

In the absence of any comprehensive overview of the worldwide prevalence of polygamy, the GID-DB focuses on the extent of legal or customary recognition of this social institution. Our polygamy variable is therefore not an estimate of the percentage of polygamous households, but an indicator of the acceptance of polygamy within a society, which is easily comparable across countries. The value 1 (0) indicates the general approval (rejection) of polygamous practices in a society. Our ratings are based on country case studies and particularly Lang (1998).

Physical Integrity

Three variables are grouped in the "physical integrity" sub-index: the prevalence of female genital mutilation, the existence of legislation punishing acts of violence against women, and the percentage of women that are "missing" due to sex-specific abortions or unfavorable living conditions.

Our principal sources for female genital mutilation (*e.g.* Amnesty International, WHO, UNDP) are prone to estimation errors; definitions across national authorities, however, do not differ. In the case of large variations between the data, we consider the most reliable source (*e.g.* according to date of publication, number of references, and potential bias of information source[5]). We directly translate the share of women who have undergone genital mutilation into our 0 to 1 coding system; *i.e.* 18% corresponds to a value of 0.18.

For violence against women, we quantify information provided by the United Nations Development Fund for Women (2003) on the existence of laws against (i) domestic violence, (ii) sexual assault or rape, and (iii) sexual harassment as follows: 0 if specific legislation is in place, 0.25 if legislation is in place but of general nature, 0.5 if specific legislation is being planned, drafted or reviewed, and 0.75 if this planned legislation is of general nature; 1 captures the absence of any legislation concerning violence against women.

The existence of laws and regulations is not necessarily an indication of violence against women. Hypothetically speaking, societies without violence against women do not need laws and regulations to punish it. Our indicator would thus penalise those societies in which violence against women is specifically not a problem. Similarly, the mere existence of laws and regulations concerning violence against women does not mean that these laws are successfully applied. On the contrary, their existence might even indicate that violence against women is a particular acute problem.

Despite these shortcomings, focusing on the legal aspects of violence against women appears to be a necessary substitute in view of the absence of any reliable and comprehensive overview of the actual extent of this important dimension of gender inequality.[6] What is more, information on laws and regulations is comparable across

countries and is – with the exception of Hong Kong and the Occupied Palestinian Territory – available for all countries of the data base. As a reference, the GID-DB also features information from the Demographic and Health Surveys (DHS) which have been conducted by the MEASURE DHS project. Specifically, we include data on the percentage of women between 15 and 49 who agree that a husband is justified in hitting or beating his wife for at least one reason. We nevertheless do not consider this variable in the physical integrity index as it is only available for 36 countries.

The "missing women" variable is largely inspired by the work of Sen (1990) and is coded depending on the relative prevalence of this phenomenon in a country. Specifically, we use estimates on the number of missing women from Klasen and Wink (2003), who report the difference between the number of women that should be alive (assuming gender equality) and the actual number of women in a country. We assign the value 1 to the country with the highest percentage of missing women relative to the total number of women (*i.e.* Afghanistan with a share of 9.3%). All other countries are assigned values between 0 (no women are missing) and 1 accordingly. For cases in which only aggregate estimates are reported (*e.g.* Sub-Saharan Africa), we replicate the calculations of Klasen and Wink for individual countries. It is important to note that all of these figures are rough estimates. Due to data constraints, estimations on the number of "missing women" cannot adequately take into account important factors such as migration, which have an important impact on the sex ratio of certain countries (*e.g.* countries in the MENA region receive disproportionately high inflows of male migrants).

Civil Liberties

Two variables comprise the "civil liberties" sub-index: women's freedom to leave the house independently and the requirement that women wear a veil in public. Both indicators are primarily coded based on Lang (1998). For freedom of movement, our indicators capture various degrees of restriction ranging from 0 (indicating no restriction) to 1 (signifying total dependence on male authority). Regarding the veil, women either have an obligation to wear it or they do not, and thus this variable is coded as 0 or 1. As with other social institutions, some of these restrictions may nevertheless only apply to certain groups in the population, in which case the value of the indicator is adjusted depending on the relative size of the group subject to this social institution.

Ownership Rights

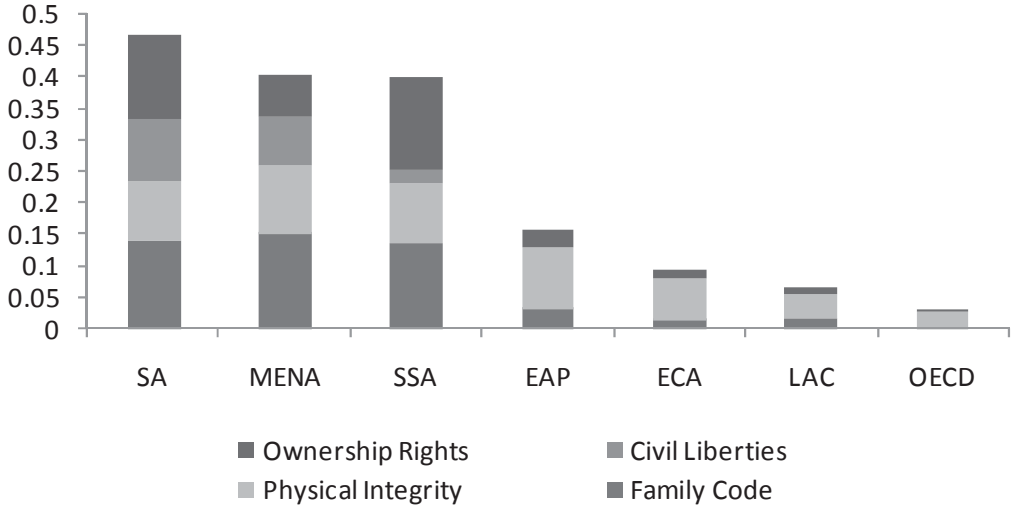
Three variables constitute our "ownership rights" sub-index: women's access to bank loans, their right to acquire and own land, and their right to own property other than land. Variations between 0 and 1 indicate the extent of restrictions and the size of the female population for which the restrictions are relevant. As before, 1 signifies high inequality (*i.e.* it is impossible for women to hold property, own land or access bank loans).

We use data and information from a variety of sources, including Amnesty International, BRIDGE (a research and information service of the Institute for Development Studies specialised in gender and development), WIDNET (the Women in Development Network), AFROL (a news agency that concentrates on Africa) and Lang (1998). Whenever possible, we compare and contrast observations with one another to cross-check the validity and reliability of information (refer to the Annex 33.1 for details on the construction and the sources of the social institutions variables).

Illustrating the Value of the Data Base

The GID-DB provides new information on the well-being of women and men across societies. The four sub-indices of social institutions described in the previous section – physical integrity, family code, ownership rights and civil liberties – indicate the depth of discrimination of women that is grounded in social norms, traditions, and family law. Plotting the aggregated country indices for the different regions in the world demonstrates striking differences between South Asia (SA), Sub-Saharan Africa (SSA), and the Middle East/North Africa region (MENA) on the one hand and East Asia and the Pacific (EAP), Europe and Central Asia (ECA), Latin America and the Caribbean (LAC), and member countries of the Organisation for Economic Co-operation and Development (OECD) on the other (Figure 33.2). Recall that our indices range from 0 (equality) to 1 (high inequality), so higher bars in Figure 33.2 indicate higher inequality.

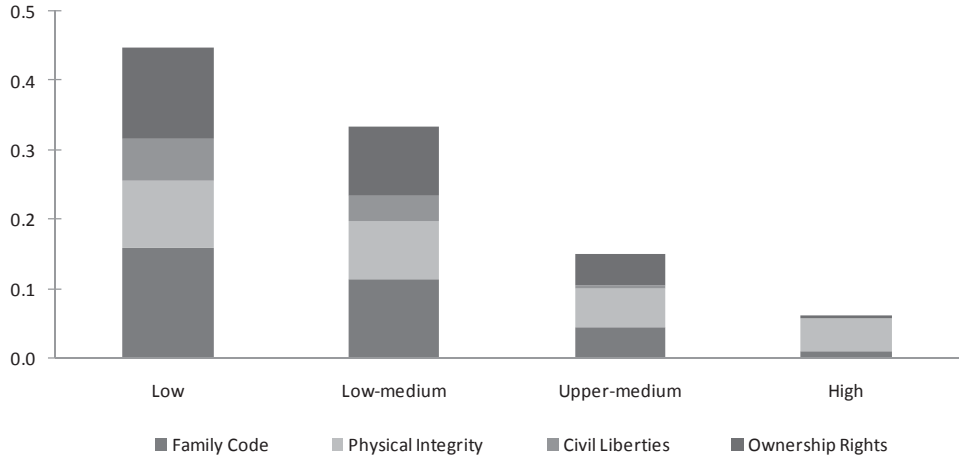
Figure 33.2 Regional Indices of Gender Inequality through Social Institutions



Source: Constructed from GID-DB. Scale: 0=equality; 1=high inequality.
Note: Also including countries for which not all data points were available.

The importance of education for gender equality is illustrated in Figure 33.3, which plots the aggregate country indices of social institutions for various levels of female education. Specifically, countries are categorised depending on the percentage of women above the age of 15 who can read and write. "Low" indicates a percentage of below 50% of the female adult population (25 countries); "low-medium" captures literacy rates between 50 and 80% (29 countries); "upper-medium" contains all countries with female adult literacy rates between 80 and 95% (36 countries), and "high" indicates literacy rates between 95 and 100% (26 countries). Clearly, gender inequality that is based on social institutions is particularly pronounced in countries with low female education. Discrimination of women through social institutions decreases with the level of education.

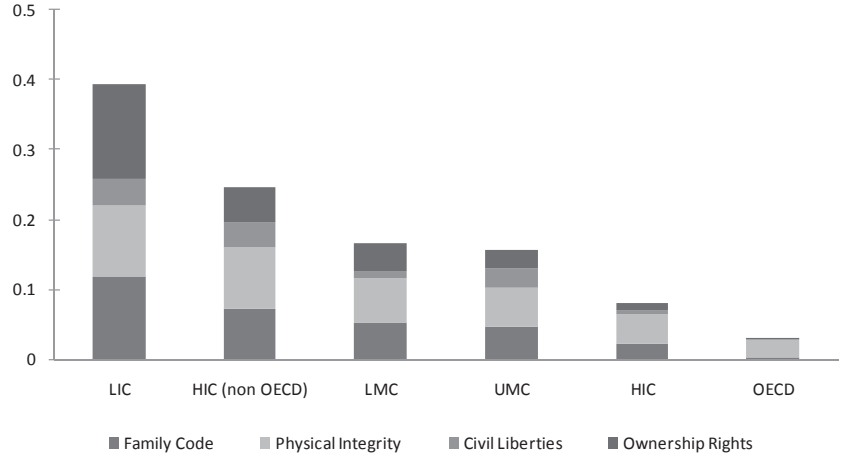
Figure 33.3 Gender Inequality through Social Institutions and Level of Education



Source: Own Illustration. Scale: 0=equality; 1=high inequality.
Note: Also including countries for which not all data points were available. "Low": 0-49% female adult literacy; "low-medium": 50-74% female adult literacy; "upper-medium: 75-94% female adult literacy; "high": 95-100% female adult literacy.

The GID-DB also allows an analysis of the relationship between economic growth and gender equality, a highly controversial topic (Forsythe *et al*. 2000). The results of a simple tabulation of countries grouped into different income categories and the degree of discrimination related to social institutions seems to support the *modernization-neoclassical thesis* of gender inequality, which suggests a reduction of inequalities in the process of economic development (Figure 33.4). Gender inequality in social institutions is most significant in low-income countries (LIC), whereas the high income (OECD) countries all have low values.

Figure 33.4 Gender Inequality through Social Institutions in Various Income Categories



Source: Own Illustration. Scale: 0=equality; 1=high inequality. Also including countries for which not all data points were available.

However, Figure 33.4 also illustrates that gender equality is not entirely linked to the level of economic development. Specifically, high-income countries (HIC) that are not a member of the OECD have notably higher values than OECD countries. Similarly, the difference between lower middle-income countries (LMC) and upper middle-income countries (UMC) is only marginal. The analysis of the relationship between economic growth and gender equality requires a more comprehensive understanding of the role that social institutions play in the development process. The GID-DB provides the necessary information to address such research questions.[7]

Composite Indicators: A Comparative Assessment

The Gender, Institutions and Development Data Base can also be used to construct composite indicators of gender equality, and exercise which has recently gained considerable attention. In addition to longer-established measures of gender equality, of which the UNDP's Gender-related Development Index (GDI) and the Gender Empowerment Measure (GEM) are the most prominent and widely used, newer indicators have emerged such as the Gender Gap Index proposed by the World Economic Forum (2006) and the Gender Equity Index (GEI) by Social Watch (2005).[8]

Table 33.2 illustrates the conceptual differences among the four main composite indicators of gender equality and highlights their particular strengths and weaknesses. The table also presents the Social Institutions and Gender Index (SIGI), a simple arithmetic average of our social institutions sub-categories, which complements and in certain aspects improves existing measures. We refrain from a thorough methodological analysis of the pros and cons of sex-disaggregated measures, gender gap indices and gender-sensitive aggregate measures and refer the interested reader to Klasen (2007).

The conventional and most widely-used indicators GDI and GEM have the advantage of being embedded into the larger debate of measuring "Human Development". In fact, the GDI is merely a gender-sensitive extension of the well-established Human Development Index (Klasen 2007). While their focus on very few issues of gender equality allows a relatively easy measurement and calculation on an annual basis, both indices arguably ignore important dimensions that are instrumental for gender equality such as the impact of economic opportunities and social institutions. They furthermore suffer from a number of methodological problems related to the measurement and interpretation of some of their variables.

In the case of the GDI, which penalises countries by lowering HDI scores for existing gender inequalities in life expectancy, education, and income, most of the difference between the GDI and the HDI seem to be driven by the "earned income" component. This is problematic as the measurement of this variable is based on some debatable assumptions. For example, it solely considers wages of the non-agricultural formal sector, which in many developing countries is rather small and not representative of the entire economy. Furthermore, it assumes that differences in earned income reflect gender gaps in consumption of human development-related goods. Countries are penalised less by the GDI for differences in life expectancy, on the other hand, which arguably deserve more attention from a human development perspective.

Although uniting important aspects of gender empowerment – *i.e.* political and economic participation and decision making, and power over economic resources – the GEM is a very specialised measure. It should thus be seen as a complement to other indices that have a broader focus on gender equality. Again, the earned income

component is problematic. In contrast to the GDI, the GEM uses income levels without a logarithmic transformation to determine inequality. This means that a poor country can never achieve a high score in the GEM ranking although it might have achieved equality in earned incomes. Conversely, a rich country might score much better then its relative income levels of men and women would suggest (Klasen 2007).

The recently established Gender Equity Index (GEI) and the Gender Gap Index are conceptually much broader than the UNDP measures and consider important additional information. However, they too fail to include variables capturing the underlying causes of gender equality. Specifically, the GEI of Social Watch focuses on socioeconomic opportunities, grouping ten well-known single variables into three dimensions: education, economic participation and empowerment. In addition to neglecting underlying causes of gender equality, the GEI also ignores other important dimensions of gender equality such as health. The Gender Gap Index introduced by the World Economic Forum is the most comprehensive measure of gender equality, combining information on 14 single indicators into four dimensions: economic participation and opportunity, political empowerment, health and survival, and educational attainment. This multidimensional approach captures many dimensions of gender equality, but it also raises the question of how this information can and should be compiled in a single indicator. As a result, the calculation of the index is relatively complex and uses weights that depend on each variable's standard deviation. Finally, although capturing more than 90% of the world's population (World Economic Forum 2006), the index has a strong focus on developed countries,

We propose the Social Institutions and Gender Index (SIGI), which presents a wide range of new dimensions and variables that are not considered by other indices. It is the only index that focuses on the underlying sources of gender equality, grouping its 12 single indicators into the four dimensions of gender inequality discussed above. In this respect it offers welcome additional information, which complements – not substitutes for – existing measures. In fact, the Social Institutions and Gender Index solely concentrates on social norms, traditions and family law that affect gender equality. It is therefore particularly relevant for developing countries where social institutions are of great importance.

This brief discussion indicates that gender equality may be too complex to be adequately captured by one "magic indicator". On the contrary, a careful selection of available indicators should be driven by the precise interest and research question. Data providers should spell out more clearly the limitations of their indicators and encourage a multi-indicator approach. Furthermore, they might want to use different approaches in order to allow for a sensitivity analysis.

Conclusion

The GID-DB compiles – for the first time in an easily usable fashion – measures of social norms, traditions and family laws that affect gender equality. This information can be used for several purposes. It allows a descriptive analysis of variations in gender equality across regions and between country income categories. Furthermore, the data base can be used to develop composite indictors of gender equality and provides evidence to undertake in depth causal analyses of the determinants of gender equality and its link to growth and development.

Addressing the underlying causes of gender discrimination is of crucial importance to improve gender equality. If social institutions prevent women from working outside of the house, increasing the enrolment rate of girls will not have a sizeable effect on female participation in the labor market. If men do not accept that women exercise authority, higher education of girls will not increase the number of women managers. More research is needed to identify policies that can successfully alter gender discrimination through social institutions. The Gender, Institutions and Development Data Base and the Social Institutions and Gender Indicator are important new tools to find answers to these challenging questions.

Notes

1 The Gender Institutions and Development Data Base can be accessed from the OECD Development Centre's web-page at: www.oecd.org/dev/gender/gid.

2 In the absence of a general threshold for the level of inequality, ratings are generally based on the relative score of a country compared to other countries.

3 Interesting information on the prevailing family code is equally given by Humana (1992) who rates the "equality of sexes during marriage and for divorce proceedings". These data are included in the GID-DB for reference. However, we do not consider the data for the calculation of our family code index as they were already published 15 years ago.

4 All sub-indices described in this section were calculated by taking the arithmetic average of the single components, also including countries for which not all data points were available.

5 Non-governmental organisations that specifically fight for the abolition of female genital mutilation, for example, may over-report its prevalence as an advocacy strategy.

6 Coomaraswamy and Kois (1999) provide an overview about the various forms of violence against women. Even for many OECD countries, data on violence against women are not available.

7 One of these research questions concerns the relationship between religious affiliation and social institutions which are discussed in detail by Jütting and Morrisson (2005).

8 The African Center for Gender and Development (ACGD 2005) introduced a regional indicator of gender equality: the African Gender and Development Index.

Annex 33.1 Comparison of the Main Composite Indicators of Gender Equality

	Gender-related Development Index (GDI) – UNDP	Gender Empowerment Measure (GEM) – UNDP	Gender Equity Index (GEI)– Social Watch	Gender Gap Index – World Economic Forum	Social Institutions and Gender Index (SIGI)
Focus	Human Development 4 single indicators of human development grouped into three dimensions: *A long and healthy life* Life expectancy ratio *Knowledge* Literacy ratio School gross enrolment ratio *A decent standard of living* Estimated earned income ratio	Empowerment 4 single indicators of economic and political empowerment grouped into three dimensions: *Political Participation and Decision Making* Percentage of seats in parliament held by women *Economic Participation and Decision Making* Female legislators, senior officials and managers Female professional and technical workers *Power over Economic Resources* Estimated earned income ratio	Socioeconomic Opportunities 10 single indicators grouped into three dimensions *Education* Literacy ratio Primary, secondary and tertiary net enrolment ratio *Economic Participation* Women in non-agricultural paid employment Estimated earned income ratio *Empowerment* Women in parliament Women at ministerial level Female legislators, senior officials and managers Female professional and technical workers	Multidimensional 14 single indicators grouped into four dimensions *Economic Participation and Opportunity* Labor force participation Wage equality Estimated earned income Female legislators, senior officials and managers Female professional and technical workers *Political Empowerment* Women in parliament Women at ministerial level Female heads of state *Educational Attainment* Literacy ratio Net primary enrolment Net secondary enrolment Net tertiary enrolment *Health and Survival* Life expectancy ratio Sex ratio at birth	**Social Institutions** 12 single indicators grouped into four dimensions *Family Code* o Early marriage o Polygamy o Parental authority o Inheritance *Physical Integrity* o Female genital mutilation o Violence against women o Missing women *Ownership Rights* o access to land o access to bank loans o access to property other than land *Civil Liberties* o Obligation to wear a veil in public o Freedom of movement
Coverage	136 countries	75 countries	154 countries	115 countries	**117 countries**
Evaluation	Pros: large country coverage; indicator availability; linked to well-known Human Development Index. Cons: simple extension of the HDI with a gender touch, earned income problematic variable with significant weight	Pros: Comprehensive measure of empowerment; linked to well-known HDI Cons: very specific measure; primarily a complement to other indices; indicator availability; small country coverage	Pros: large country coverage; comprehensive list of indicators of gender equity Cons: Omission of important dimensions (*e.g.* health)	Pros: comprehensive list of indicators and dimensions Cons: strong focus on developed countries, complicated calculation of indicator weights	**Pros**: innovative indicators, measurement of underlying reasons of gender equality, inclusion of social institutions **Cons**: very specific measure, primarily a complement to other indices; measurement problems.

References

ACGD (2005): The African Gender and Development Index, United Nations Economic Commission for Africa and African Center for Gender and Development, Addis Ababa: ACGD.

Antecol, H., (2000): An Examination of Cross-Country Differences in the Gender Gap in Labor Force Participation Rates, *Labour Economics*, 7(4), 409-426.

Antecol, H., (2003): Why is there Cross-Country Variation in Female Labor Force Participation Rates? The Role of Male Attitudes toward Family and Sex Roles, *Claremont Colleges Working Papers*.

Coomaraswamy, R. and L. M. Kois (1999): Violence against Women, in: K. Askin and D. Koenig (eds): *Women and International Human Rights Law*, Vol. 1, p.139-177.

Forsythe N., R.P. Korzeniewicz, and V. Durrant (2000): Gender Inequalities and Economic Growth: A Longitudinal Evaluation. *Economic Development and Cultural Change*, Vol. 48(3), p. 573-617.

Humana, C. (1992): *World Human Rights Guide*. 3rd ed., New York: Oxford University Press.

Jaumotte, F., (2003): Female Labour Force Participation: Past Trends and Main Determinants in OECD Countries, *OECD Economics Department Working Papers*, No. 376. Paris: OECD.

Jütting, J. and Morrisson, M. (2005): Changing Social Institutions to Improve the Status of Women in Developing Countries. Paris: *OECD Development Centre Policy Brief*, No. 27.

Killinsworth, M.R. and J.J. Heckman, (1986): Female Labor Supply: A Survey. In O. Ashenfelter and R. Layard (eds.), *Handbook of Labor Economics, Volume 1*, Amsterdam: North-Holland.

Klasen, S. (2002): Low Schooling for Girls, Slower Growth for All? Cross-Country Evidence on the Effect of Gender Inequality in Education on Economic Development. *The World Bank Economic Review*, Vol. 16(3), p. 315-373

Klasen, S. and C. Wink (2003): Missing Women: Revisiting the Debate, *Feminist Economist*, 9(2-3), p. 263-299.

Klasen, S. (2007): Gender-related Indicators of Well-being, in: McGillivray, M. (Ed.): *Measuring Human Well-Being: Concept and Measurement.* London: Palgrave, p. 167-192.

Lang, J. (1998): *Enquête sur la Situation des Femmes dans le Monde*. Paris: Assemblée Nationale.

Sen, A. (1990): More than 100 Million Women are Missing, *The New York Review of Books*, December 20.

Sen, G. (2007): Informal Institutions and Gender Equality, in: Jütting, J., S. Bartsch, D. Drechsler, I. de Soysa (Eds.): *Informal Institutions: What Role in Development?*, OECD Development Centre Studies, Paris: OECD (forthcoming).

Social Watch (2005): *Roars and Whispers. Gender and Poverty: Promises versus Action*, Montevideo: Social Watch.

United Nations Development Programme (2004): *Human Development Report 2004*, Technical Note 1, New York: UNDP.

Wood, A. (1991): North-South Trade and Female Labour in Manufacturing: An Asymmetry, *Journal of Development Studies*, Vol. 27(2).

World Bank (2001): *Engendering Development through Gender Equality in Rights, Resources and Voice*. Washington D.C.: World Bank.

World Economic Forum (2006): *The Global Gender Gap Report*, Geneva: WEF.

Statistical Sources

Afrol News: http://www.afrol.com/index.php

Amnesty International (1997): Female Genital Mutilation in Africa: A Human Rights Information Pack. Amnesty International Report ACT 77/07/97. London: AI.

Asian Development Bank (ADB): Gender and Development, http://www.adb.org/Gender/default.asp

Bridge: Development and Gender, http://www.bridge.ids.ac.uk/

Canadian International Development Agency (CIDA): http://www.acdi-cida.gc.ca/index-e.htm

International Labour Organisation (2002): Key Indicators of the Labour Market 2001-2002. Geneva: ILO.

Lang, J. (1998): Enquête sur la Situation des Femmes dans le Monde. Paris: Assemblée Nationale.

United Nations (2005): Millennium Development Goals Indicators Database. New York: UN.

United Nations Statistics Division (2006): The World's Women 2000: Trends and Statistics. New York: UN.

United Nations Development Programme (2000): Gender in Development. Statistics and Indicators. New York: UNDP.

United Nations Development Programme (2002): Human Development Report: Deepening Democracy in a Fragmented World. New York: UNDP.

United Nations Development Programme (2004): Human Development Report: Cultural Liberty in Today's Diverse World. New York: UNDP.

United Nations Development Programme (2005): Human Development Report: International Cooperation at a Crossroads: Aid, Trade and Security in an Unequal World. New York: UNDP.

United Nations Development Fund for Women (2003): Not a Minute More: Ending Violence against Women, New York: UNIFEM.

United Nations Development Fund for Women (2006): http://www.unifem.org/

WHO: http://www.who.int/topics/female_genital_mutilation/en/

Women in Development Network: http://www.focusintl.com/widnet.htm

World Bank (2005): World Development Indicators 2005, Washington, D.C.: World Bank.

World Bank, GenderStats, http://devdata.worldbank.org/genderstats/

Part Nine
Measuring Societal Progress for Local Communities

Chapter 34
The Newfoundland and Labrador System of Community Accounts

Mr. Alton Hollett; Dr. Cory Giles
Newfoundland and Labrador Statistics Agency, Government of Newfoundland and Labrador, Canada

Dr. Douglas May
Professor, Department of Economics, Faculty of Business Administration, Division of Community Health and Ethics, Faculty of Medicine, Memorial University, Canada

Abstract

The Newfoundland and Labrador System of Community Accounts (CA) is a free access Internet-based (www.communityaccounts.ca) knowledge management (KM) information communication technology (ICT) which provides a knowledge commons approach to locating, using, sharing and understanding data and knowledge which is inextricably linked to the well-being of the Province of Newfoundland and Labrador and its citizens. Based upon a determinants of determinants well-being model, the CA provides value beyond simply storing and disseminating data. Instead, the CA has established an environment that ensures that people work, communicate and pursue broader goals together by utilising and sharing key sources of policy related data and information. Supported through this framework, a process has begun that turns data into information and knowledge, allowing government and citizens to become more engaged in the choices and decisions that affect the progress of the Province. As a result, users of the CA, from various inter-related fields, are better able to develop benchmark measures and support subsequent monitoring of the social and economic progress of the province, its regions, and communities through the use of strong statistical evidence and measurable results. This in turn has accelerated policy research, thus enabling faster decision making, better quality decisions and a greater level of public service accountability and transparency.

Introduction

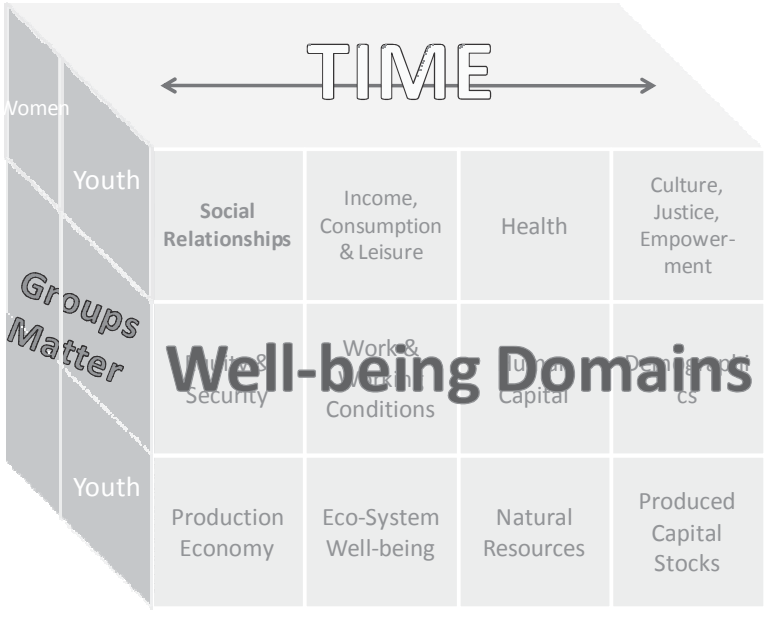
Good public management demands that public servants, communities, and citizens, work together in pursuit of important, often cross-cutting goals, such as lowering crime, improving health, and developing sustainable economies. Many government officials and stakeholders express concern, however, over the effect of "silos thinking" in preventing

the achievement of these goals. The problem with this type of thinking is that it results in people focusing only on their own specialisms or department. This, in effect, accentuates a disconnect between inter and intra-departmental activities, collaborative relationships and overall organisational goals. A large part of this sectarian behaviour is reflected in the lack of availability and sharing of important policy related data and information. Indeed, this factor alone forms a significant barrier to establishing an organisational environment that embraces an integrated, evidence-based approach to policy and programme development within and across government departments, and between social and economic sectors. To overcome this, the Newfoundland and Labrador Statistics Agency (NLSA), in partnership with Memorial University,[1] has been committed to aiding in the development of a new community of practice[2] that ensures that people in various interrelated areas are working, communicating and pursuing broader goals together through the sharing and use of data and information. A major achievement of this commitment was the development of the Newfoundland and Labrador System of Community Accounts (CA), a free access Internet-based (www.communityaccounts.ca) knowledge management (KM) information communication technology (ICT) which provides a knowledge commons approach to locating, using, sharing and understanding data and knowledge. Specifically, the CA platform allows stakeholders to benchmark and measure social progress with respect to their values and goals and to better understand and define the quality of life and well-being in communities as it relates to various social and economic determinants such as health, income, social networks, the environment and natural resources.

System of Community Accounts: An Overview*Measuring Well-being: A Determinants of Determinants Framework*

The development of the CA is formally nested in the following vision: "Individuals living in vibrant communities and neighbourhoods in which each person has an equal opportunity to pursue and enhance his/her individual and collective well-being to the fullest extent possible." For individuals, well-being emphasises healthy, educated, prosperous, and socially complete individuals, both in absolute and relative terms, that feel empowered and supported to making full use of their innate potential. Individual well-being, however, is intimately associated with collective well-being which recognises the social, cultural and psychological needs of people, their family, institutions and communities as a whole (Wilkinson 1991). Due to the subjective nature of well-being, in both cases, the concept is difficult to measure. As a result, researchers have focused more on objective indicators and how they determine the outcomes of single domains such as income, health status, crime, employment or education. An example of such a determinants model is Statistics Canada Population Health Model. From the perspective of well-being, however, there are many layers or domains that determine it, each with its own correlation of determining factors and dimensions (see Figure 34.1). Well-being research, therefore, should be thought of as a determinants of determinants model since increases in well being are the ultimate outcome and the motivation for all single domain determinant model interventions.

Figure 34.1 The Dimension of Well-being



The domains of well-being are represented by many dimensions including people, groups, space and time
Source: The System of Community Accounts, 2007

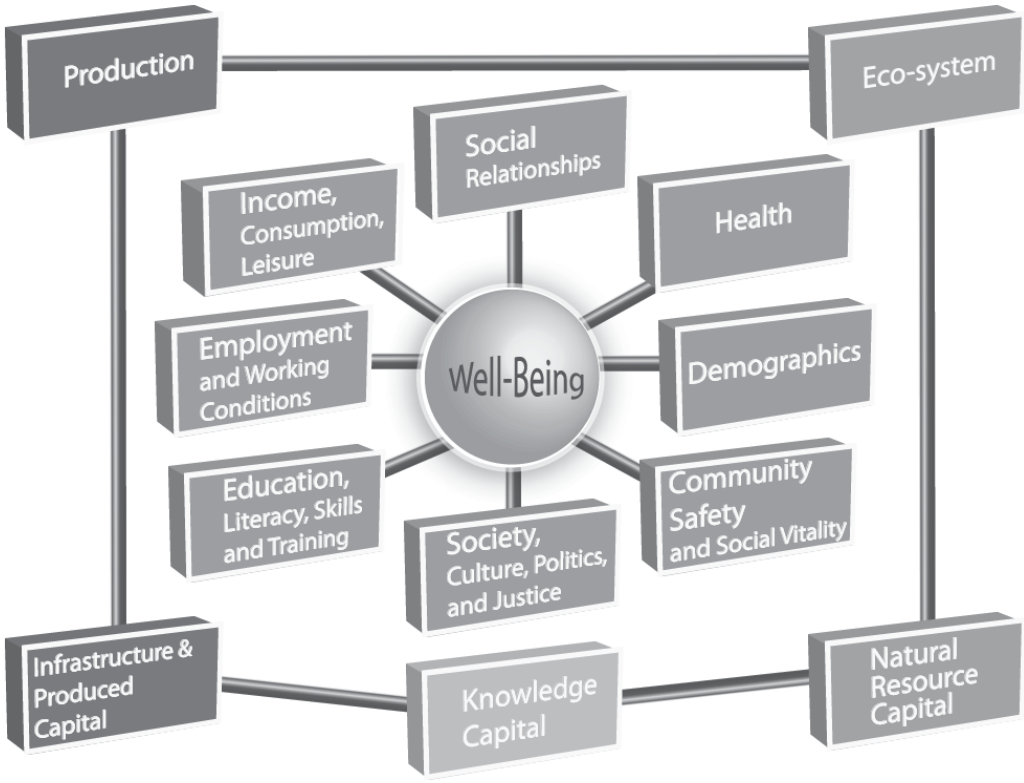
If governments, communities and citizens are going to achieve social objectives in an attempt to improve well-being they must have some idea of the status of well-being. As well, they must have some idea of where they what to go, why they want to go there and, finally, how to get there. To do this, they must engage in a culture where statistics, statistical series, and all other forms of evidence enable them to assess where they stand and where they are going with respect to their values and goals, and to evaluate specific programmes and determine their impact (Bauer 1966). Doing so will better enable them to:

- Improve the quality of life of individuals and families;
- Realise what is possible/ the potential to improve;
- Improve relationships with others communities, regions, provinces, and nations that have baring on the possibilities;
- Identify strengths, weaknesses and root causes of problems;
- Plan and implement strategic investments in people and communities; and
- Monitor social progress.

The CA is designed with these purposes in mind. It is a social accounting system based upon a determinant of determinants framework that consists of linked domains with a focus on measuring individual and collective well-being. By design, as productive and natural assets impact the social and economic landscape, indicators contained with each

of the domains gauge change and allow for evolving patterns to be measured and visualised. As a result, the relative strengths of the various linkages between the domains and how together they determine well-being can be better understood (see Figure 34.2). This infers that the well-being of citizens, communities and the province can be determined from the interactions of multiple social and economic factors and that the characteristics of each of these factors, in turn, are a function of each other. To begin to achieve this, the system provides for the storage of data within eleven interrelated accounts (*i.e.*, Household Spending, Income, Labour Markets, Production, Demographics, Social, Crime and Safety, Health, Education, Resource/Wealth, Environment and Well-being) for over 1 300 geographies including 400 communities, 80 Census Consolidated Subdivisions, 20 Economic Development Zones, 9 Rural Secretariat Regions, 4 Health Authorities, 6 Community Health Regions, 8 Institutional Health Boards, 5 School Districts, 15 Human Resource and Social Development Canada (HRSDC) Regions, 4 HRSDC Districts, 200 urban neighbourhoods, 613 schools, provinces, territories and Canada.

Figure 34.2 The Organisation of the System of Community Accounts



Source: a Community Accounts Data is organised into a system of accounts that highlights relationships among the data which have an even more profound impact on well-being than each account in isolation.
b The red bars represent the statistics that describe the nature and strength of the determinant relationships.Source: The System of Community Accounts, 2007

The range of indicators found in each account provides the basis for a major analytical framework for needs identification, planning, monitoring, research, and

evaluation. As a result, stakeholders are now better able to develop benchmark measures and support subsequent monitoring of the social and economic progress of the province, its regions, and communities through the use of strong statistical evidence and measurable results. To further build upon this the CA has begun the process of measuring the relative strengths of the various determinants through the use of more sophisticated regression techniques (*i.e.* logit). Current examples include working models describing the determinants of quality of life and health. This, in turn, has accelerated the level of understanding of the factors that influence successful social and economic development within the province thus enabling faster and better quality decision making and a greater level of public service accountability and transparency.

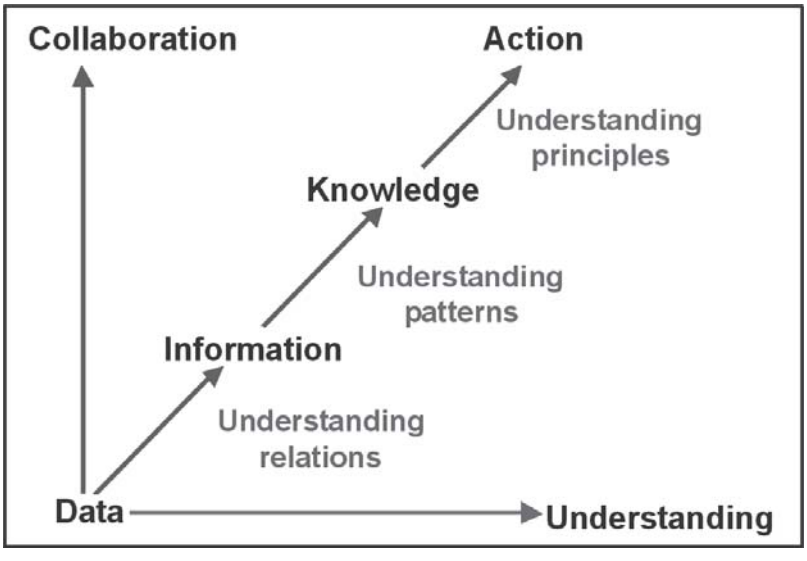
Knowledge Commons/Knowledge Management

In the traditional sense, a "commons" is a term that refers to a resource shared by a group of people. A knowledge commons is an extension of this concept, whereby the shared resource is a repository of intellectual history, collective action, and free exchange of ideas, data and information in the interest of the common good (Hess & Ostrom 2006). Today, the availability of this type of resource appears ubiquitous in nature, especially with the worldwide infiltration of digital technologies. However, due to a complex array of geographical, functional and organisational barriers, existing data, information and knowledge are, at best, only weakly integrated in a form that promotes more effective governance and evaluation of social, economic and environmental issues. To overcome this, Memorial University and the Government of Newfoundland and Labrador are establishing a knowledge commons by cultivating a climate of cooperation that facilitates new conceptual uses of authoritative community level data, information and research metrics. This, in turn, has enabled the shared collection, construction, and recombination of data and information and the development of innovative research frameworks. By establishing knowledge as a commons (*i.e.*, as a shared resource), stakeholders, such as governments, researchers, community agencies and citizens, are better able to understand and monitor the varied social and economic landscapes of the province. In keeping with this, the fundamental innovation of the CA is the provision of unrestricted access to a single reliable source of key social health and economic indicators to government and public users alike. By this virtue alone, the CA has facilitated the removal of barriers to collecting, sharing and analysing policy related data and information for decision making. Viewed through the lens of a "knowledge commons", the CA structures and democratises information and facilitates the formation of a new culture of informed, actively engaged citizens who value individual and collective responsibility for decision making and development.

Also of particular significance, the CA framework inherently supports a KM culture that aids in the systematic development, management and use of data and information that describes the Province of Newfoundland and Labrador and its people. In particular, the CA promotes the transformation of data (*e.g.*, unorganised discrete facts) into information (*e.g.*, properly selected and organised data that show relationships) and knowledge (*e.g.* collections of information that shows patterns and allow for the development of principles and to undertake useful actions) by integrating three integral KM components: *People, Processes* and *Technology* (Plunkett 2001). These components are essential to policy development and public accountability and together provide a synergy that facilitates the development of well defined communities of practice that increase organisational collaboration and capacity for understanding. *People* and organisational culture stimulate

and nurture the location of data, development of information and construction of knowledge such that open stakeholder collaboration is the norm, not the exception; *processes* ensure the effective location, acquisition, organisation and analysis of data and information; and *technology* provides the media for storage, analysis and dissemination and allows people to work together without being together. As a result, the CA promotes collaboration and increases capacity for understanding. In turn, governments and citizens are better able to engage in making informed choices and decisions that affect the social progress of the Province (see Figure 34.3).

Figure 34.3 The Transformation of Data to Information and Knowledge



The transformation of data to information and knowledge is a function of the level of organisational collaboration and capacity for understanding which is facilitated through a well defined communities of practice that represent a culture of knowledge management.
Source: adapted from Ackoff (1989) and Bellinger, Castro and Mills, 2004

Key to promoting this community of practice is ensuring that the CA is always relevant to stakeholder needs and capabilities. This is achieved through the incorporation of participatory approaches in the CA concept. The main tenet of participatory approaches is that stakeholder's influence and share control over the development of a project and the decisions and resources which affect them. With regard to the CA, stakeholders are, and continue to be, collaborators at every stage of development. Thus, participative methods have generated a sense of ownership of decisions and actions that drive the planning, the design, the implementation, the monitoring and the evaluation of the CA. A product of this approach was the development of a diverse array of analytical tools to help users summarise and interpret data as well as to enhance the participatory exchange of information and knowledge. For example, users can monitor the social and economic landscape by defining and custom building community and regional summary statistic profiles and make comparisons with other communities, regions, or the province on any combination of social, economic, and demographic information. Users can also utilise relative measures of well-being allowing them to identify communities and regions at risk without delving into hours of analysis. Another important feature evolving from this process was the development and implementation of educational support to users in

the form of online and telephone support, teleconferencing, regular classroom training sessions and online tutorials (For a time line of major development benchmarks see Appendix 34.1).

Strategies for Success

A key to the success of the CA is the placement of its construction and maintenance within the NLSA. First, the NLSA is strategically placed in the Department of Finance which offers strong leadership support in coordination with the Executive Council. Secondly, the NLSA is governed by the "Statistics Agency Act," which provides the Agency with the statutory authority and responsibility to collect, compile, analyse, abstract and publish statistical information and methods relating to the social, economic and general activities and conditions of the province and its citizens. Under this authority, the NLSA is able to successfully initiate the development of the Data Domain Advisory Committee (DDAC) to aid in the identification and negotiation of long term sharing and exchange of key provincial and national sources of existing data on health, labour market, social, crime and safety, education, demographics, income/consumption and productivity (see Appendix 34.A3 for an overview of the DDAC). The DDAC is made up of over 30 individuals representing governments, academia and communities. To date, the DDAC has resulted in the development of data-sharing partnerships with Provincial Government Departments and Agencies, Memorial University, the Newfoundland and Labrador Centre for Health Information, Human Resources Development Canada and Statistics Canada, Government of Canada, Royal Newfoundland Constabulary, and the Royal Canadian Mounted Police. The DDAC also functions to identify gaps in existing data and makes recommendations for methodologies to rectify these inconsistencies. For example, gaps in health, labour and demographic information resulted in the development and administration of the Newfoundland and Labrador Adult and Community Health Survey, Labour Activity Survey, Employer Survey, In-Migrant Survey, Out-Migrant Survey and Market Basket Measure of Low Income.

Resulting Changes

The CA is an advance in the development of social capacity. The above key innovations have shaped and refined the CA into a system that provides value beyond simply storing data. Stakeholders are now using this knowledge to understand those things that make up and affect their lives, to measure how their communities and themselves are doing and how they compare to others. The CA has initiated a process that aims to establish, within the public consciousness, an atmosphere of interest and excitement about the concept and determinants of well-being and the use of quantitative methods. This is evidenced by the increasing level of usage the CA site has received since its release in 2001, including approximately 30 000 unique visitors, over 600 000 hits and approximately 2000 data enquiries (*i.e.*, phone/e-mail). Interest in the CA has also resulted in the provision of training to over 3 000 public servants and citizens and over 100 information sessions hitting an estimated total audience of approximately 3 000 stakeholders worldwide. As well, the process of developing and sharing a common data set has contributed to the practice of governments and citizens becoming more engaged in the choices and decisions that affect the progress of their provinces, regions and communities. To this end, the CA achieves the following:

- Incorporates data and information into a user-friendly, intelligent, analytical data storage and retrieval system that can be used as a collaborative tool to enable and facilitate stakeholders to share and exchange information;

- Provides all stakeholder groups, including governments, community agencies and citizens, with free uninhibited access to a single reliable, independent source of provincial, regional, community and neighbourhood level data and information;

- Overcomes geographical, functional and organisational barriers, which inhibit stakeholders' access to data and knowledge;

- Aids in the identification of gaps in existing data and the development of methodologies, partnerships and collaborative arrangements to fill these gaps;

- Provides a baseline measure of well-being and supports the subsequent monitoring of the social and economic progress of the province, its regions, and communities;

- Promotes collaboration and enhances the use of authoritative data and information thereby accelerating policy research and enabling faster, better quality decisions; and

- Promotes and supports greater accountability and transparency.

- Major examples of uses of the ca include, but are not limited to, the following (see appendix 34.a2 for a more extensive list):

- Data focal point for the strategic social plan and production of the social audit landscape document "from the ground up" (strategic social plan, executive council, government of newfoundland and labrador);

- Communities targeted for "youth-at-risk" programmes and the development of personal home care rates based on cost of living data (department of health and community services, government of newfoundland and labrador);

- Funding allocation based on community need (community kids eat smart program);

- Location for new family resource centres (labrador strategic social plan committee, government of newfoundland and labrador);

- School amalgamation patterns (regional school boards);

- Determination of the required number of non-pandemic influenza shots, based on regional demographic profiles (western health care region);

- Providing business clients with community-level information such as education and local labour market conditions (network newfoundland and labrador, a public-private partnership developed to provide companies with information support);

- Providing individuals and businesses with better information in order to produce more accurate and effective business plans (st. John's business resource centre);

- Primary health care needs/capacity assessment for urban st. John's (st. John's primary health care working group);

- Determination of optimal location of offices to best meet the need of social assistance clients (department of human resources labour and employment, government of newfoundland and labrador);

- Data focal point for developing a province-wide poverty reduction strategy and the canada/newfoundland and labrador labour market development agreement (cnl-lmda). The purpose of cnl-lmda is to work with individuals, communities and other stakeholders to meet measurable and achievable educational and employment goals. (department of human resources labour and employment, government of newfoundland and labrador);

- Data focal point to promote the well-being of rural newfoundland and labrador through comprehensive and coordinated participatory partnerships aimed at integrating economic, social and cultural aspects of rural and regional development (rural secretariat, cabinet secretariat, government of newfoundland and labrador).

Sustainability

An overarching goal of the CA is to engage community-based and publicly driven initiatives where governing bodies only assume the role of facilitator or enabler. A key step towards promoting this shared accountability of citizens and governments was the incorporation of stakeholder-identified societal level values into the concept and definition of the community well-being and its associated determinants. As a result, the CA is viewed as more relevant in the public mind and is therefore sustainable. Understanding the context in which technology was introduced, in terms of attitudes toward information and technology, was vitally important. Initially, through a process of participation by consultation between government and citizens, stakeholders were better able to support the introduction of the technology, and believe in its benefits. As stakeholders became accustomed to the CA, the use of functional and interactive participation became more commonplace. Under this process, stakeholders have begun to participate in joint analysis by forming groups to meet predetermined objectives related to a project. Key to this success is the inclusion of rural communities in the communication process and the promotion of citizen support of locally produced solutions. Ultimately, the CA aims to engage "self-mobilisation" where stakeholders participate by initiating positive development and change independent of external institutions.

As an ICT, user accessibility of the CA has always been a major threat to sustaining its use. To ensure that stakeholders, including communities and citizens, have Internet access the Government of Newfoundland and Labrador, in partnership with the Government of Canada, has developed 180 rural and urban community Internet access centres throughout the province under the Canada/Newfoundland and Labrador Community Access Program (CNL-CAP). As well, under the Federal/ Provincial "Broadband for Rural and Northern Development Pilot Program" over 80% of Newfoundlanders and Labradorians have broadband high speed Internet access, with a commitment by both governments to have 100% of citizens connected.

Transfer and Recognition

In recognition of the CA, the NLSA received the 2003 Newfoundland and Labrador Public Service Award of Excellence, the 2003 Institute of Public Administration (IPAC) Award for Innovative Management and in 2006, the CA was chosen as a semi-finalist for the prestigious United Nations Public Service Award. Due to its ease of implementation and use, the CA is now being adopted in other jurisdictions. The Government of Nova Scotia, with the assistance of the CA Team and the transfer of the CA template, has developed a CA for their province. More recently, the CA Team developed the ACCA prototype for the Canadian Atlantic Region that was launched in August 2005. Also, talks are currently underway to develop other Atlantic Canada versions representing the Provinces of Prince Edward Island and New Brunswick. In addition to gaining provincial and national recognition, the success and transferability of the project is also being publicised internationally attracting interest from the Government of Australia and the Organisation for Economic Co-operation and Development (OECD).

Lessons Learned

In addition to the issue of Internet accessibility, there were a number of obstacles that were overcome in order to achieve the development of the CA concept. These include reviewing the capacity of users and establishing effective communication and educational frameworks to facilitate input from stakeholders and build community awareness and participation; breaking down inter-departmental, agency and sectoral behaviours that blocked data sharing; development of procedures to ensure security of data and data access; assuring a sustainable, integrated and compatible technology platform; and financial and human resource constraints.

The Provincial Government's appointment of the NLSA as the lead agency for implementing the initiative was the first step in effectively overcoming these obstacles. As the single statutory authority on statistical matters for the province, the NLSA reports to the highest levels of government. This enabled coordination of the long term strategic leadership required of the Premier, his ministers and other senior officials in devising and promoting the vision and strategies needed for the successful development and implementation of the project, particularly on issues related to the development of data sharing relationships between government departments and agencies. As well, the expertise of NLSA ensures standardised processing, storage and dissemination of methods and data that meet national and provincial standards of data security and protection of citizen privacy. For data sharing outside the Provincial Government, the formation of the DDAC, which represents multiple stakeholder sectors, aided in bridging any pre-existing divides in data sharing behaviour. These relationships have also led to the development of cost sharing arrangements between stakeholders and data partners to overcome the high cost of data licensing and vulnerability assessments associated with the secure warehousing and storing of data and information.

Central to the success of this ICT was the concept that stakeholders are experts in many domains, particularly citizens. As a result, continuous efforts are made to ensure that the communication and education frameworks function to accept and facilitate input from multiple disciplines by providing an amiable culture for interaction. This includes adjusting the level and mode of communication and educational formats based on the needs and educational background of the stakeholder. To overcome geographic and human resource barriers, which inhibit effective communication between stakeholders,

the CA Team used a variety of established networks. These include the extensive regional networks developed by various government departments and agencies, *e.g.*, Rural Secretariat Regional Planners, Regional Treasury Board Training Division Networks, as well as regular newspaper advertisements and brochure mail outs. In addition, through its collaborative partnerships with Memorial University, non-governmental organisations, communities and citizens, the CA Team participates in numerous local, regional, provincial and national conferences, speaker series and events throughout the year with a focus on community development. To ensure that the CA is sustainable and compatible with existing technologies, the CA Team works in collaboration with the Office of the Chief Information Officer, Executive Council, Government of Newfoundland and Labrador. This has resulted in overcoming issues related to browser compatibility, software and hardware upgrades, security features and off site access privileges during regional and community stakeholder sessions.

Notes

1 Dr. Douglas May, Professor of Economics, Faculty of Business Administration, Division of Community Health and Ethics, Faculty of Medicine, Memorial University, is the senior advisor for the System of Community Accounts.

2 Community of practice refers to a group of individuals who are not a part of a formally constituted work team but share a common work practice over a period of time. Communities of practice cut across traditional organisational boundaries and enable individuals to acquire new data, information and knowledge faster (Wenger and William, 2000)

References

Ackoff, R. L. "From Data to Wisdom". Journal of Applies Systems Analysis. Volume 16, 1989

Bauer, R. (ed). Social Indicators. Cambridge, MA: M.I.T. Press, 1966.

Bellinger, G., D. Castro and A. Mills. "Data, Information, Knowledge, & Wisdom". http://www.systems-thinking.org/dikw/dikw.htm, 2004.

Hess, C. and E. Ostrom (eds). Understanding Knowledge as a Commons: From Theory to Practice. Cambridge Massachusetts: MIT Press, 2006.

Plunkett, P. T. Managing Knowledge at Work: An Overview of Knowledge Management. Knowledge Management Working Group of the Federal Chief Information Officers Council, 2001.

Wenger, E. C, and M. S. William. "Communities of Practice: The Organisational Frontier". Harvard Business Review, January-February, 2000.

Wilkinson, K.P. The Community in Rural America. Westport, CT: Greenwood Press, 1991.

Appendix 34.1 Major Benchmarks for the System of Community Accounts

1998

- The NLSA was mandated to develop a government and public-wide system that embraces an integrated, evidence-based approach to policy and programme development through collaboration within and across government departments, citizens and socio-economic sectors.

- A Data Domain Advisory Committee (DDAC) was formed and tasked with identifying and negotiating for the sharing and exchange of key sources of existing data and to develop methodologies to fill data inconsistencies.

- A prototype for the CA framework was developed.

1999

- The CA prototype was replaced by a web page format.

2001

- The CA web page was changed from Excel platform to SQL database platform. This provided faster and more dynamic querying of data and uploading of information to the website and larger volumes of usage.

- Similarly, the CA web page was changed from an HTML format to ASP (active server pages) format. This extension is more advantageous because it alerts the server that there are server-side chores to be done before sending the page out to the end user. This ultimately increases speed of uploading and database querying, an advantage for analog internet users.

2002

- The CA was released to the public and government.

2003

- The development of the neighbourhood geographies was initiated.

- Consultations began with the Government of Nova Scotia to develop a version of the CA for their province.

- The CA received the 2003 Newfoundland and Labrador Public Service Award of Excellence as well as the National Institute of Public Administration (IPAC) 2003 Award for Innovative Management for the theme "*In the Know: Managing Knowledge*".

2004

- The web page was redesigned, *e.g.*, more user-friendly 3-step process for data querying, based on user feedback, public consultation and the use of external consultancy services.

- The Determinants of Quality of Life logistic modeling tool was developed and incorporated into the CA.

- In cooperation with the Faculty of Education, Memorial University of Newfoundland, a number of online interactive training modules were developed to inform users of site navigation and data analysis.

- An advanced version of the Productivity Accounts was initiated.

2005

- The Chart Director Software was incorporated into the CA. Chart Director is a built-in graphics engine specially designed for high performance server side graphics for fast and efficient use of online images and charts.

- The CA prototype for Atlantic Canada, *i.e.*, Atlantic Canada Community Accounts (ACCA), was developed and released.

- Consultations began with the remaining Atlantic Provinces to develop similar CA prototypes.

- CA team initiated development of a new CA prototype that will incorporate the Nesstar software, an advanced data management programme.

- Public consultations and development of the framework for the Crime and Safety Account was initiated.

2006

- The CA was selected as a semi-finalist for the United Nations Public Service Awards for the theme "Application of information and communication technology (ICT) in government: e-government."

- Discussions began with the Government of Australia on the possible transfer of the CA technology to their country.

Appendix 34.2 Examples of Major Uses of the Community Accounts

Government

Rural Secretariat

- The Rural Secretariat used the Community Accounts to develop profiles of the nine Regional Council regions in the province. Regional Councils were provided with information on social and economic status in their regions to inform their deliberations on a future vision for each region.

- The Burin Peninsula Regional Council of the Rural Secretariat have used the Community Accounts many times to access all types of data that are used in meetings to provide a background and understanding of the Burin Peninsula region.

- A Regional Planner used the Community Accounts in doing a workshop on proposal writing for the Family Resource Center's Provincial Conference, Training Conference 2006 to educate volunteers and employees on how to access statistical information for an evidence-based rationale to support an organisation's need for funding.

- A Rural Secretariat Planner trained a group of aboriginal front line workers on using the Community Accounts to help them increase their research knowledge and ability to use data.

Economic Development

- By using the Community Accounts, the Business Resource Centre in St. John's was able to provide individuals and businesses with better information to produce more accurate and effective business plans.

- A regional office of the Department of Innovation, Trade and Rural Development used the Community Accounts to help verify population trends in the region as it pertains to Business Development.

Health

- The Department of Health and Community Services used the Community Accounts to determined communities where youth were determined to be "at risk" to enable them to target programmes more effectively. Managers in the Department indicated that the quality and timeliness of staff analysis improved markedly by accessing the Community Accounts.

- The Division of Child, Youth and Family Services, Department of Health and Community Services has used the Community Accounts to identify Communities "at risk" to determine where Family Resource Centres would be located. In Labrador their partners, the Strategic Social Plan Regional Steering Committee, used the Community Account's dynamic output features to determine the best location in that region and, as a result, two Family Resource Centres were established.

Regional Boards

Regional Economic Development Boards

- The Regional Economic Development Board's Economic Development Officer used Community Accounts data in Central Labrador to determine aboriginal populations, employment and education.

- The Long Range Red Board and Marine and Mountain Zone Corporation used the Community Accounts to assist with rationale for various proposals as well as for their most recent business plans required by their funding partners.

- The Schooner Regional Development Corporation used the Community Accounts in developing its annual strategic plan for the corporation; to assist commercial clients with business plans and to help communities and community based organisations to develop economic development projects and funding proposals.

- The Humber Economic Development Board used the Community Accounts to conduct research to assist communities with proposal development.

- The Economic Development Officer in Marine and Mountain Economic Development Board to research labour market occupational distribution by gender in the region.

- The Emerald Zone Corporation (REDB) used the Community Accounts to develop Community Strategic Action Plans for 42 communities by preparing community profiles as marketing tools for the town and for background information on proposals for federal and provincial economic development funding. The development of these plans increased community capacity through teaching community groups (municipal economic policy committees) how to develop, write, and implement strategic plans.

- The Exploits Valley Economic Development Corporation uses the Community Accounts to gather statistics to include in their work plan, to justify activities in certain sectors such as education, mining, and others; for background information in proposals; and to prepare presentations for meetings with municipalities to demonstrate population shifts, education levels, and income levels.

- Exploits Valley Economic Development Corporation used the Community Accounts to develop maps and charts on population, education, income levels, and labor market data (*e.g.* employment rates, participation rates, and social assistance rates over time and geography) and compare regional statistics to provincial statistics for clarifying information in presentations and reports.

- The Long Range Regional Economic Development Board in Stephenville used the Community Accounts to determine the sample size for interviews with seniors in the region.

Health

- Eastern Health used the Community Accounts during the Burin Peninsula Community Health Needs Assessment.

- The Community Accounts was used by the Integrated Health Authority, Health and Community Services Western's Child Care and Youth Services to inform their planning, programme development and service delivery for the region.

- The Central Health Board uses the Community Accounts for strategic planning, service delivery planning, compiling regional/community profiles, projecting service needs through demographic monitoring, determining resource needs for particular services (*e.g.* to help determine need for long term care home beds and personal care home beds), monitoring population health status and other factors related to health.

- Because of ease of access, the Western Health Care Board was able to increase cost-effectiveness by determining the number of non-pandemic influenza shots to order. Previously, the Western Health Care Board was unable to make such informed decisions because of the time and cost involved in retrieving information.

- Health and Community Services Eastern used the Community Account to provide information on the region to enable staff in planning and implementing programmes and services.

- The St. John's Primary Health Care Working Group (PHCWG) used the Community Accounts, and in particular the neighbourhood data, to become informed about the health care needs to complete a capacity assessment for urban St. John's. They plan to use the neighbourhood data to help assess the determinants of health in this area including educational attainment, employment rates, income levels, and incidence of social assistance. The Working Group also plans to use the Community Accounts' electronic map of urban St. John's that breaks down 95 neighbourhoods to help to identify specific areas of need and gaps in services.

Education

- A social worker in a rural school used the Community Accounts neighbourhood data to inform neghbourhood groups in their design of prevention initiatives for children entering and in school who have not been part of the formal protection services section.

- Regional school boards used the demographic information on the Community Accounts that provided population projections, to make decisions regarding school amalgamation as populations of 5 to 18 year olds declined.

SSP Regional Steering Committees

- The Central SSP Steering Committee used the Community Accounts to assess the number of children under five years of age in communities in the region when developing their early childhood development initiatives.

- The Northeast Avalon SSP Steering Committee used the Community Accounts to complete a needs assessment for their early childhood development strategy. The strategy compiled regional, community and neighbourhood information in areas such as numbers of 0 to 6 year olds, health status, family income and poverty, and family composition. The information was used to determine areas where children were "at risk" and to provide a baseline for future monitoring of progress.

Communities

- In Arnold's Cove the Community Accounts data was used to develop a profile of the community for community development planning. It allowed the municipality to understand the status of the community compared to other communities and to plan where the community was hoping to make progress based on its status.

- The Town of Conception Bay South used the Community Accounts to complete a neighbourhood enumeration. They have enhanced the capability of the town to strategically plan for the future. The incorporation of this new neighbourhood level information into the Community Accounts provided a broader view of socio-economic status of the town and allowed the town and its citizens to track progress.

- The Recreation Liaison for the City of Corner Brook has requested access to neghbourhood data to enhance community recreation planning and implementation of recreation programmes and services.

- The Town Manager in Stephenville uses the Community Accounts for municipal planning of infrastructure and programming.

- The mayor of Belleoram used the Community Accounts for writing proposals for organisations in the community such as the Church, Loyal Orange Lodge, and the Harbour Authority. The town also referred the company who will be operating the Rock Quarry to the Community Accounts for information on the community.

Voluntary, Community-Based Sector

- Rural Development Associations have used the Community Accounts to develop proposals and to identify needs for a particular community or area.

- The Community Education Network in Stephenville uses the Community Accounts to do research on communities and to assist with their proposal writing.

- Baie Verte and Area Development Association uses the Community Accounts on a regular basis to determine the aging population, education levels, medical requirements in each community on the Peninsula. The information is also used to develop proposals for funding.

- The Community Services Council used the Community Accounts to develop a profile of two localities where their case studies on research collaborative governance was being completed.

- The Community Youth Networks used the Community Accounts for statistical analysis and to develop funding proposals related to youth employment in regions.

- Bay St. George Status of Women Council and Gateway Status of Women Council use the Community Accounts when writing funding proposals, to document changes and compare statistics on women, and to identify future challenges based on trends.

- The Exploits Valley Community Coalition (Family Resource Center Programs) used the Community Accounts to look at numbers of children in certain areas for programming purposes, to gather information for a business plan, to do up Power Point presentations to the Municipal Council and to gather information, to gather information on the communities currently served for evaluation purposes and report writing, and statistics for writing proposals for funding.

- The St. John's Boys and Girls Club uses the neighbourhood data in the Community accounts to assess where to offer their programmes in the city.

- The Kids Eat Smart Program uses the Community Accounts to access community data to assess community needs. This allows them to determine their funding allocations to each community.

- The Working Group on Poverty, a community-based committee, prepared a social and economic landscape on poverty in the Northeast Avalon using the Community Accounts. They developed this profile for the region by selecting key indicators and compiling data by neighbourhoods in St. John's and by communities in the region as baseline information for identifying areas of concern and as a basis for their planning.

- The YMCA used information in funding proposals to plan for future recreation services and for their delivery. They also used them to prepare a presentation to the national YMCA Board to explain the financial and local situation related to YMCA development.

- The Baie Verte and Area Chamber of Commerce used the Community Accounts to gather information on the population of different areas of the region and to determine the potential usage of high speed internet during the implementation of Broadband in their region.

Private Sector/Public Corporations

- The head office and regional managers of the Newfoundland and Labrador Liquor Corporation use the Community Accounts at the head office level and in regions to determine the locations of retail liquor outlets across the province.

- NETWORK Newfoundland and Labrador, a public-private partnership developed to provide companies with information to determine the placement of contact centres in the province, used the Community Accounts to provide companies with the community-level information they needed, such as education levels and local labour market conditions.

- An entrepreneur in Labrador used the Community Accounts to see how many children there were in a local area between the ages of 6 and 12 years to determine if there was a market for an after school daycare programme.

- An entrepreneur used the Community Accounts to do research for a rental business.

- A local media writer used the Community Accounts to obtain background demographic and labour market data for stories on rural Newfoundland and Labrador.

- The Community Business Development Corporation - Emerald used the Community Accounts for information on unemployment rates, workforce participation, social assistance dependency, education, income, and population when preparing annual business plans and for assisting clients to prepare their business plans.

- A Labrador entrepreneur used the Community Accounts research local housing statistics on rentals vs. ownership.

Post-Secondary Education

- Researchers at Community Medicine, Memorial University studying the role of social capital in economic development, used the Community Accounts to select communities with high and low levels of social capital and high and low levels of economic resiliency after the fisheries closure in the early 1990s so that comparisons could be made in differing communities.

- Memorial University professors in the Faculty of Business Administration use the Community Accounts as a resource when teaching students in marketing and labour markets.

- Professors in the Faculty of Medicine at Memorial University use the Community Accounts when teaching students in Community Medicine.

- A graduate student in Applied Social Psychology, Memorial University used the Community Accounts in his co-op placement to gather information on neghbourhoods served by a school.

- A Political Science student from University of Ottawa on a summer placement with the Western School Board Partnering Committee used the Community Accounts to study poverty and policy by preparing a profile of the area.

Appendix 34.3 Community Accounts Data Domain Advisory Committee

Table 34.A3.1 Community Accounts Data Domain Advisory Committee[a]

Domain	Members
Health	Dr. Roy West, Division of Community Health, Memorial University Dr. Doreen Neville, Division of Community Health, Memorial University Dr. Cory Giles, Newfoundland and Labrador Statistics Agency
Labour Market	Andrea Dicks, Manager, Strategic Planning, Department of Industry, Trade and Rural Development Patti Powers, Independent Consultant Gerald Crane, Director, Rural Secretariat, Executive Council Fred Phelan, Human Resources Development Canada Jean-Pierre Voyer, Social Development Research Corporation, Human Resources Development Canada
Social	Beverley Clarke, Assistant Deputy Minister, Department of Health and Community Services Dr. Donna Hardy-Cox, Professor of Social Work, Memorial University Patti Powers, Independent Consultant Marvin McNutt, Director of Adult Corrections, Department of Justice Marilyn McCormack, Deputy Advocate, House of Assembly
Education	Carmel Wyse, Department of Youth Services and Post-Secondary Education Dr. Sherry May, Director, Mathematics Learning Centre, Memorial University Terry Quinlan, Community Accounts, Newfoundland and Labrador Statistics Agency
Demographics	Dr. Roy West, Division of Community Health, Memorial University Dr. Byron Spencer, Professor, Department of Economics, McMaster University Dr. Frank Denton, Professor, Department of Economics, McMaster University Dr. Stan Winer, Professor, School of Public Policy and Administration, Carleton University Dr. Dane Rowlands, School of International Affairs, Carleton University
Income/Consumption	Robert Reid, Community Accounts, Newfoundland and Labrador Statistics Agency Gwenda Murphy, Senior Statistician, Newfoundland and Labrador Statistics Agency Dr. Urvashi Dhawan, Human Resources Development Canada Dr. Cory Giles, Newfoundland and Labrador Statistics Agency
Production	Robert Reid, Community Accounts, Newfoundland and Labrador Statistics Agency Gerald Crane, Director, Rural Secretariat, Executive Council Heather Tizzard, Manager, Industry, Trade and Rural Development,
Other	Andrew Sharpe, Executive Director, Centre for the Study of Living Standards Dr. Alice Nakamura, Professor, Department of Finance and Management Science, University of Alberta School of Business

a Dr. Douglas May (Community Accounts Advisor; Professor of Economics, Faculty of Business Administration, Division of Community Health and Ethics, Faculty of Medicine, Memorial University) is Chair of the Data Domain Committee.

Chapter 35
Developing Indicators for Local Communities: The New Zealand Experience

Abby Thornley[i]
Geograpy, Regional and Environmental Team, Statistics New Zealand, New Zealand

Abstract

New Zealand, although relatively small, is a diverse country. Local communities may vary greatly from one another in terms of their economy, demography, land area and the resource management issues that they face.

This paper provides an overview of New Zealand, its communities, its legislative framework and initiatives that are being developed for local community monitoring and reporting, including the development of indicator sets.

Local communities in this paper refer to not only the legal or formal communities that are governed at the local level, but also communities of interest such as urban and rural communities. The paper also discusses the challenges faced in developing indicators that meet the diverse needs of local communities while enabling comparability and creating a national picture to inform our responses to global challenges such as climate change and our need for sustainable development. For instance, census data for very small areas (meshblocks with very small populations) may be confidentialised.

Introduction

New Zealand Communities

New Zealand is an island nation in the south-west Pacific with two main islands, the North Island and the South Island, as well as numerous smaller ones such as Stewart Island and the Chatham Islands group. The capital, located in the North Island, is Wellington. Also in the North Island is Auckland, the leading port and largest city with a population of 1 303 068. New Zealand is an independent sovereign state, its parliament

i Any enquiries about the paper should be directed to: Abby.Thornley@justice.govt.nz

sets all national laws in the absence of any federal structure and the country was a foundation member of the United Nations Organisation.

The majority of New Zealanders are of European descent (67.6%) (Statistics New Zealand 2005a). New Zealand's indigenous Māori population makes up around 15% of the population. New Zealand's land mass is similar in size to Japan or Great Britain, with a population of approximately 4 million. Compared with other Organisation for Economic Co-operation and Development (OCED) countries, New Zealand has a small population (fourth smallest), it also has a higher than average fertility rate (fifth highest in the OECD), and a higher than average life expectancy at birth (twelfth highest) (OECD 2005). New Zealand has a relatively low ratio of people aged over 65 years compared with the labour force (eighth lowest).

Local communities in New Zealand can be defined by legal boundaries, such as regions or districts, or they can be defined by less formal criteria, such as rural and urban communities, or non-geographically linked communities with strong ethnic ties. In this paper, both formal and non-formal communities are addressed, using legal boundaries for local communities based around the New Zealand Geographic Frame. The Geographic Frame identifies of small units of variable geographical area called meshblocks, which are made up of approximately 100 households**.** The location of the meshblocks determines the local authority boundaries. The robust structure of the New Zealand Geographic Frame has meant the formal boundaries of local communities in New Zealand have been able to be consistently monitored and updated since their implementation in the 1960s.

Table 35.1 Regional Attributes in New Zealand, 2006

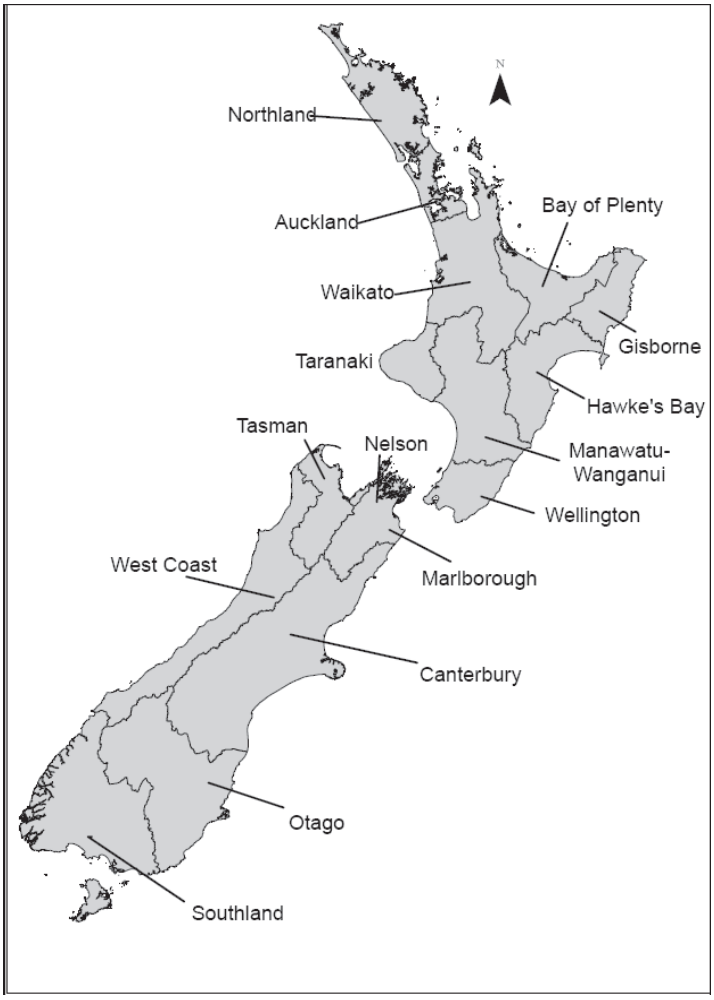
Region	Population	Percentage of NZ Population	Percentage Mäori	Population density People per Square Km
Northland	148,470	3.69	31.70	10.8
Auckland	1,303,068	32.35	11.08	215.3
Waikato	382,716	9.50	20.96	15.9
Bay of Plenty	257,379	6.39	27.53	21.0
Gisborne	44,499	1.10	47.27	5.3
Hawke's Bay	147,783	3.67	23.51	10.5
Taranaki	104,124	2.59	15.76	14.3
Manawatu-Wanganui	222,423	5.52	19.62	10.0
Wellington	448,959	11.15	12.77	55.2
Tasman	44,625	1.11	7.09	4.6
Nelson	42,888	1.06	8.67	96.7
Marlborough	42,558	1.06	10.47	3.9
West Coast	31,326	0.78	9.67	1.3
Canterbury	521,832	12.96	7.22	11.7
Otago	193,800	4.81	6.56	6.2
Southland	90,876	2.26	11.78	2.8

Note: Data in table sourced from Statistics New Zealand, 2006, *2006 Census Regional Summary Tables*

New Zealand has a diverse range of local communities, they are highly mobile, and in many cases the local communities have very different social, economic, environmental and cultural attributes which can make them unique. For example, Auckland, one of the smallest regions in terms of land, is home to approximately one third of the population, which highlights the uneven distribution of New Zealand's population (see Table 35.1 and Figure 35.1). Table 35.1 shows Auckland has the highest population density

(215.3 people per square kilometre), compared with the West Coast which has the smallest population density (1.3 people per square kilometre). Another illustration of difference is that whereas approximately 31% of Northland's population is Mäori, around 12% of Wellington's population is Mäori. These differences can be further demonstrated when looking at New Zealand's two main islands separately. Three out of four New Zealanders live in the North Island. The increase in population in the last ten years has also been greater in North Island towns and cities with a 12.6% increase recorded in North Island regions, compared with 7.6% in the South Island (Statistics New Zealand 2006).

Figure 35.1 Regions in New Zealand



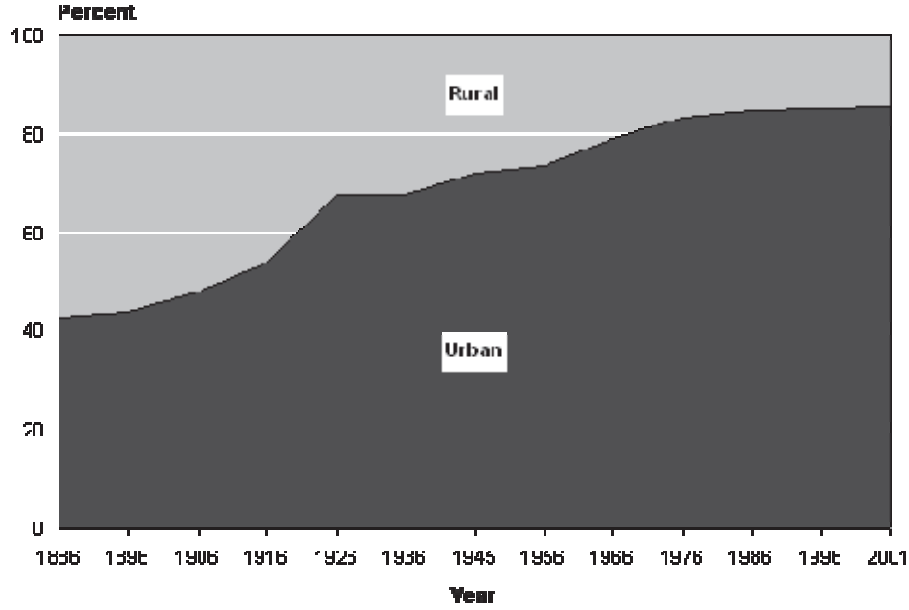
Source: Statistics New Zealand, 2007, map created for this paper.

The Changing Face of New Zealand's Local Communities

In 1881, New Zealand was predominantly rural, with nearly 60% of the population living in rural areas but the momentum behind urban development was rapidly gathering pace.

Between 1906 and 1916, the proportion of the population living in urban areas overtook those living in rural areas. Between 1886 and 2001, the population of urban New Zealand increased by over 1 500%, compared with an increase in rural areas of 83% (see Figure 35.2).

Figure 35.2 Proportion of People Living in Urban and Rural Areas in New Zealand



Source: Statistics New Zealand, 2005, New Zealand: An Urban Rural Profile

The composition of urban areas has also changed considerably, particularly since 1950. Ethnic diversity has increased in urban areas, first with the urbanisation of Mäori beginning in the 1950s, then with the rise of immigration from the South Pacific islands in the 1960s and 1970s. The 1990s experienced significant inward migration from Asia.

Over the course of the twentieth century, New Zealand was transformed from being composed of largely rural communities into a society comprised of highly urbanised ones. The nature of the economy also changed, although approximately half of New Zealand's merchandise exports continued to be of primary produce by the end of the century.

Working on the land is no longer the occupation of the majority. According to the 2001 Census of Population and Dwellings, less than 10% of the New Zealand workforce reports the agriculture, forestry or fishing industries as their occupation (Statistics New Zealand 2005a). However, the changes that have occurred in local communities have not only been driven by economic factors, but demographic factors as well. For example, the growth in New Zealand's population by approximately 23% since the 1980s (Statistics New Zealand 2005b) has meant that population density has increased, particularly in the main urban areas. In addition, the increase in New Zealand's population over the past ten years has been driven by natural increase[1] (70%) as opposed to net migration (30%). This means that often changes in local communities are driven primarily by internal migration movements rather than by significant changes in migration flows.

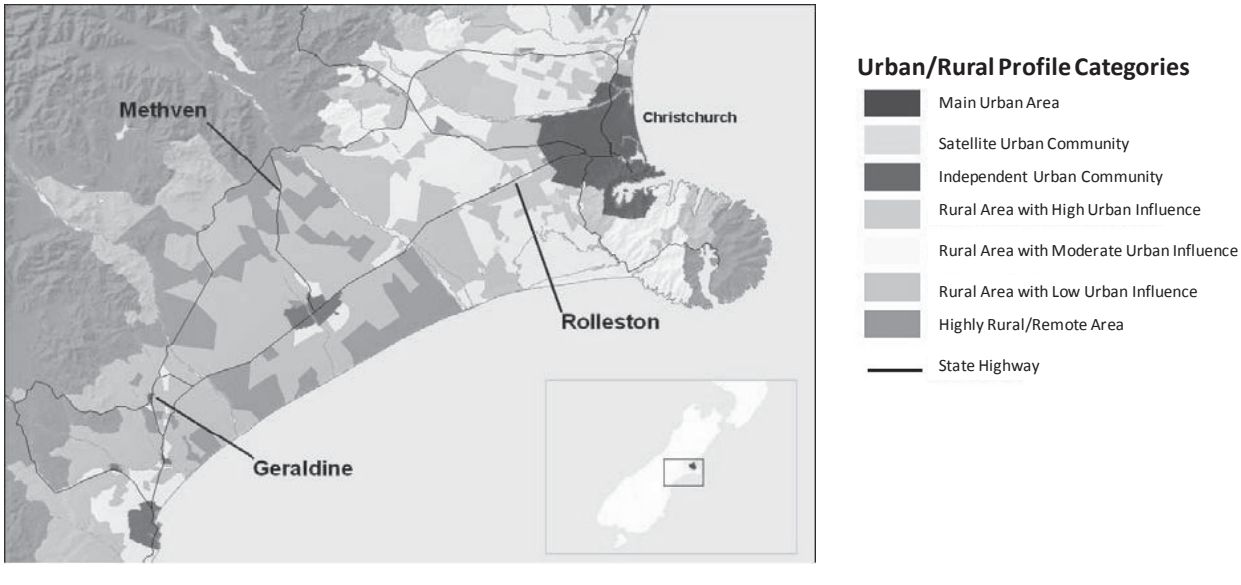
The differences between local communities in New Zealand have become increasingly apparent over the last thirty years driven by a number of economic changes adopted since the 1980s. These changes include: the removal of agricultural subsidies, the significant reduction in tariffs on imports, the liberalisation of financial market controls on capital flows, the privatisation of former state-owned assets such as telecommunications, and the introduction of inflation targeting as the primary objective of monetary policy. As a result, New Zealand became more connected with the global economy and New Zealand businesses were both able to better participate in global markets while also, of course, facing increased competition from offshore.

The changes initiated a structural change in the economy, shifting the focus from industry to services, which along with the arrival of retail chains serving both Australia and New Zealand and an increasingly cosmopolitan hospitality industry led to a new 'café culture' (Smith 2005). Globalisation has changed the face of New Zealand business – it is now more competitive internationally and there is an increasing focus on productivity, growth and innovation. Currently, New Zealand has one of the lowest unemployment rates in the OECD.

Although the 1980s saw an increase in urban communities and a decrease in rural community populations there has been a recent renewed interest in rural areas, albeit with a different driver. More recently, people have been moving to the more rural areas around cities[2] owning 'lifestyle blocks' or small scale farms and commuting to work in the city centres.

Figure 35.3 illustrates the urban rural profile of the Canterbury region and Appendix 35.1 and 35.2 illustrate the full distribution of New Zealand's local communities in an urban/rural context in 2001.

Figure 35.3 Example from the Christchurch/Canterbury area



Source: Statistics New Zealand, 2005c, *New Zealand: An Urban Rural Profile*

Sustainability and Well-being

The concepts of sustainability and well-being are important to New Zealand both at the national level and at the sub-national or local community level. Well-being is defined as the satisfaction of human preferences. The better human preferences are satisfied, the greater is well-being (Neumayer 2004). Well-being, in New Zealand, is defined in terms of four themes – economic, social, enviornental and cultural. These themes also commonly underpin the concept of sustainable development in New Zealand. Sustainable development is defined as "development which meets the needs of the present without compromising the ability of future generations to meet their own needs" from the Brundtland report (World Commission on Environment and Development 1987). Sustainability is important in determining the well-being of individuals, communities and for New Zealand as a whole. Given the changes that have occurred in New Zealand's demography, economy and environment it is becoming more important than ever to ensure the well-being of New Zealanders. It is also becoming more important to monitor these well-beings (through the development of indicators) especially as New Zealand's demography, economy and environment continues to change.

Long-term sustainable strategies are, therefore, required for our ongoing success as a nation. New Zealand is committed to developing a sustainable nation and to developing suitable internationally comparable measures of the non-economic aspects of societal, cultural and environmental progress to complement the measures of economic performance that are already commonly accepted internationally. These measures will need to be reportable at national, regional and local levels. Communities will need to monitor changes in their local areas, and to benchmark against each other. This will also enable New Zealand to be benchmarked and compared with other countries.

Sustainability is also important to the current New Zealand Government. For instance; when outlining her priorities for 2007, the Prime Minster said "New Zealand's future is dependent on long term sustainable strategies for our economy, society, environment, culture and way of life" (Clark 2007).

Governance of Local Communities in New Zealand

New Zealand's Governance Structure for Local Communities

New Zealand has a unicameral parliament. In addition to central government, New Zealand also has a structure of local government. Local government is responsible for decision making and reporting on the sustainability and well-being of their local communities. The top tier of local government is called a *region* and there are 16 regions in New Zealand. Of these 16 regions, 12 are governed by regional councils, while four are governed by the second tier of local government, *territorial or local authorities* (territorial authorities which perform both roles are known as *unitary authorities).* In total there are 86 local authorities in New Zealand.

Legislation

New Zealand has a range of legislation that assists in monitoring local community sustainability and well-being. However, two key legislative frameworks empower local

community decision making and monitoring; the Local Government Act (LGA) 2002, and the Resource Management Act (RMA) 1991.

The purpose of the LGA 2002 is to provide for a responsive and effective local government that recognises the diversity of New Zealand's communities. To balance this empowerment, the law promotes local accountability, with local authorities accountable to their communities for decisions made. The LGA 2002 therefore facilitates local authorities to take a broad role in promoting community well-being, while taking a sustainable development approach to the use and management of natural resources.

Under the LGA 2002 local governments are required to develop long-term community council plans (LTCCP); these require local authorities to develop a set of well-being indicators for monitoring well-being and to ensure sustainable development. These are required to be reported on at least once every three years.

The RMA 1991 is the main piece of legislation that outlines the processes to manage our natural environment. Its purpose is to enable the effects of economic, social and cultural activities on the natural environment are managed in such a way as to ensure the environment does not suffer irreparable damage.

The basic concept of the RMA is to promote sustainable management and efficient use of our natural resources, using The Brundtland Report (World Commission on Environment and Development 1987) definition, namely "development which meets the needs of the present without compromising the ability of future generations to meet their own needs" as a background concept. In New Zealand, district and regional plans are important aspects of the RMA 1991. These plans set out guidelines for the public to follow on how they can and cannot use the environment.

Measuring Local Community Well-being

There have been several initiatives at different levels which measure and monitor community well-being. A number of the recent key initiatives are outlined below (in chronological order).

Quality of Life Project Report

The Quality of Life Project was established in 1999 by representatives from Auckland City Council, Christchurch City Council, Manukau City Council, North Shore City Council, Wellington City Council and Waitakere City Council to provide social, economic and environmental indicators of quality of life (or well-being) in New Zealand's six largest cities.

It was initiated in response to growing pressures on urban communities, concern about the impacts of urbanisation and the effects of this on the well-being of residents. The purpose of the report is to provide information to decision-makers to improve the quality of life in major New Zealand urban areas.

The project is run by a team made up of representatives from the participating local authorities. The Quality of Life Project has since expanded to include twelve territorial authorities and there have been two reports to date, in 2001 and in 2003 (Big Cities). The project has also developed a questionnaire which can be used by other local authorities not included in the report to measure the quality of life in their area whilst ensuring comparability with those areas in the report.

The indicators monitored in this report cover topics such as: people; knowledge and skills; standard of living; economic development; housing; health; the natural environment; the built environment; safety; social connectedness; and civil and political rights. Economic growth is an indicator that is reported on under the economic development topic; Gross Domestic Product (GDP) is used as the measure for this. The data used to produce the indicators of quality of life are Official Statistics and taken mostly from Statistics New Zealand's information releases.

Long-term Community Council Plans (LTCCPs) and Community Outcomes

As outlined previously, the LGA 2002 requires local authorities to state their LTCCPs to assess progress towards the achievement of community sustainability and well-being outcomes.

Many local authorities have carried out considerable community consultation through surveys and meetings to determine their outcomes. There have been several innovative approaches to this process and a willingness to share experiences and knowledge across local government. The first monitoring report for local authorities is due in 2009 and reporting will continue on a three-yearly basis.

One example is Future Path Canterbury (FPC), an initiative to develop a long range strategic planning programme, which was developed through a partnership between the regional and city councils. The goal of FPC has been to 'develop and agree upon a broad-based community view of what we want the future to look like, and pursue the achievement of it'. For this to be successful the Future Path project considered that 'Visions only work when communities are involved in developing them' (Local Government New Zealand 2004).

Linked Indicators Project

In 2003 the government initiated a programme of work for sustainable development, called the Sustainable Development Programme of Action (SDPOA), following the World Summit in Johannesburg. One initiative that came out of the SDPOA was the Linked Indicator Project (LIP) which aimed to facilitate the monitoring of New Zealand's sustainability and well-being.

The LIP is the main top-down community indicators project. It was a collaborative project between New Zealand's statistical agency, Statistics New Zealand, and other central and local government agencies. It identified a set of regional indicators that would fill gaps in information to better inform local decision making. The term 'linked' in the terms of the LIP refers to data being linked from the national to the regional levels. An indicator was defined as a parameter that can be measured (eg, distance from a goal, target, benchmark) to show reliable trends or sudden changes in a particular condition. This linking is useful because, in order for decentralised decision making to be effective, policy makers need to have access to reliable and transparent tools/data.

Varying degrees of data currently exist, with the greatest amount of data available for reporting on indicators at the national level. This diminishes rather quickly when disaggregated to the sub-national level (see Table 35.2). Social well-being is the only complete set of indicators for which there is data available at both the national and sub-national level, whereas the environmental well-being only has data available to produce 45% of the indicators at the national level, this decreases to only 27% for the indicators at

a sub-national level. The current focus is on filling the gaps in the indicators particularly at a sub-national level.

Social well-being covers themes of civil and political participation, health, housing, leisure and recreation, safety, skills/knowledge and social connectedness. Economic well-being covers themes of economic growth, international connection, standard of living, work, infrastructure and innovation. An example of the indicator measure used to report on economic growth is GDP.

Table 35.2 Coverage of Linked Indicators

Well-being	Total number of Indicators	Percentage available nationally	Percentage available sub-nationally
Culture	5	80%	60%
Economy	14	100%	50%
Environment	11	45%	27%
Society	13	100%	100%

Source: Statistics New Zealand, 2005, *Linked Indicators*

Other Local Community Indicator Programmes

There are a range of other programmes for monitoring sustainability and well-being, however, many of these programmes focus on reporting at a national level, therefore the focus looking to the future is on disaggregating these indicators where feasible so that they can be reported at the local community level.

Statistics New Zealand has a number of initiatives underway, many in collaboration with other central government agencies, which have contributed to improving sub-national statistics, including:

- Programme of Social Statistics (POSS);
- Regional GDP;
- Local population estimates and projections;
- Regional Stocktake;
- Quarterly Regional Review;
- New Zealand's geographic frame (meshblocks by which small area data is collected for the Census. Meshblocks are also used as the basis for local community boundaries and in local body elections).

Other initiatives from central government include reports from the Ministry of Social Development, *The Social Report,*[3] the Ministry of Economic Development, *Growth through Development: Economic Indicators* (2005) and the Ministry for the Environment, *State of New Zealand's Environment* (1997).

A common understanding and monitoring of outcomes over time will be essential for comparing New Zealand's progress between communities, against past performance and against that of other countries.

Challenges

There is increasing demand for information at the sub-national level with the introduction of the LGA 2002 and the LTCCPs. However, in meeting this demand there are some key challenges. New Zealand has a small population, which means large samples need to be used to produce quality sub-national estimates. Small population sample surveys cannot always deliver reliable estimates of small communities or small groups. For instance, census data for very small areas (meshblocks) are confidential. There is also the issue of respondent burden; consideration must be taken to ensure that high quality official statistics are produced without excessive load on data providers. Respondent burden is particularly challenging given New Zealand's small population. To help overcome these challenges with data availability, confidentiality and respondent burden, a focus has been placed on the effective coordination of datasets, both survey-derived and administrative, through the Official Statistics System (OSS), a programme led by Statistics New Zealand to provide formal structures for the efficient coordination of official statistics across the entire public sector.

Another key challenge is the level of capability and capacity within local communities to undertake statistical activities. As shown previously, the scale differs greatly between local authorities and this impacts the resources that different areas have available to develop and analyse indicators for community outcomes. Statistics New Zealand, as leader of the OSS, is working in conjunction with local government and other central government agencies to overcome these challenges.

While there are programmes to monitor sustainability and well-being at the national level, there are considerably fewer sub-national ones. One of reasons for this is the comparability of sub-national datasets in New Zealand. Although local authorities collect and monitor their own sustainability and well-being, the information is often collected using different methodologies, then it is analysed and reported on in different ways. This means that it is difficult to aggregate to a national picture and often complicates comparability across regions.

Conclusions

New Zealand, although relatively small, is a diverse country. Local communities may vary greatly from one another in terms of their economy, demography, land area and resource management issues. This diversity requires local government to manage and monitor well-being and sustainability at sub-national level.

Local government has been empowered to make its own decisions about how to achieve its community's outcomes through the LGA 2002. This act requires local authorities to develop LTCCPs which are used to report on the sustainability and well-being of those local communities. In addition to the bottom-up approach to local community well-being reporting such as the LTCCPs, New Zealand has initiated top-down programmes led by central government, such as the Linked Indicators Project.

To date the main challenge faced is one of data availability. The majority of the data is available at the national level. There is some sub-national data, however, it is often collected and analysed using different methodologies, so there is a lack of consistency across local authorities.

The focus now, for New Zealand, is strengthening the degree of consistency in measuring and monitoring well-being at the local community level. This will be facilitated by the OSS through capacity and capability building at all levels of government.

Notes

1 Natural increase is defined as births minus deaths.

2 Classified as rural areas with high or moderate urban influence.

3 The Social Report, reports on social indicators at the national, regional and local authority level. Some of the indicators reported on are the same as in the quality of life report. However, the Social Report has a narrower and more detailed focus on New Zealand's society, whereas the Quality of Life report has a broader but less detailed focus on New Zealand's society, economy and environment.

References

Big Cities, Quality of Life Report, New Zealand, Available: http://www.bigcities.govt.nz/

Clark, H., 2007, Prime Minister's Statement to Parliament, Available: http://www.beehive.govt.nz/ViewDocument.aspx?DocumentID=28357

Local Government New Zealand, 2004, Engaging Communities over Outcomes, Available: http://www.lgnz.co.nz/library/files/store_005/CommunityOutcomesPaper.pdf

Local Government Bill, 2001, Available: http://www.dia.govt.nz/Pubforms.nsf/URL/191bar1.pdf/$file/191bar1.pdf

Ministry of Economic Development, 2005, Growth Through Innovation: Economic development indicators, New Zealand, Available: http://www.gif.med.govt.nz/aboutgif/indicators-2005/

Ministry for the Environment, 1997, The State of New Zealand's Environment, New Zealand, Available: http://www.mfe.govt.nz/publications/ser/ser1997/

Ministry of Social Development, 2006, The Social Report, New Zealand, Available: www.socialreport.msd.govt.nz/

Neumayer, E, 2004, Sustainability and Well-Being Indicators, United Nations University, Tokyo.

Organisation for Economic Co-operation and Development (OECD), 2005, New Zealand: Country Statistical Profile, Available: http://stats.oecd.org/wbos/viewhtml.aspx?queryname=326&querytype=view&lang=en

Smith. P, M, 2005, A Concise History of New Zealand, Cambridge University Press, Melbourne:

Statistics New Zealand, 2006, 2006 Census Regional Summary Tables, New Zealand, http://www.stats.govt.nz/census/2006-census-data/regional-summary-tables.htm

Statistics New Zealand, 2005a, New Zealand in the OECD, New Zealand, Available: http://www.stats.govt.nz/products-and-services/nz-in-the-oecd

Statistics New Zealand, 2005b, Demographic Trends Report, New Zealand, Available: http://www.stats.govt.nz/analytical-reports/dem-trends-05/default.htm

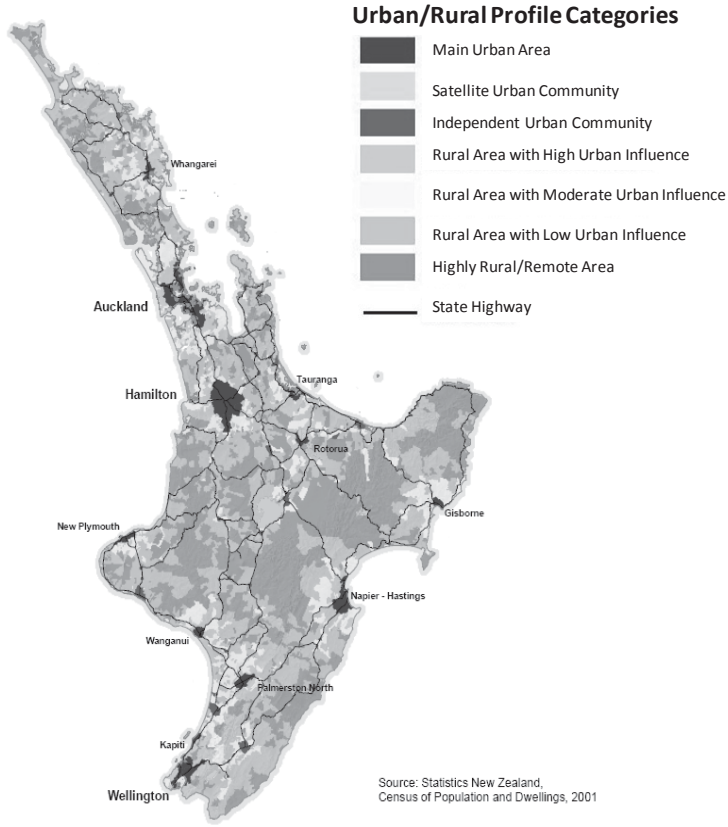
Statistics New Zealand, 2005c, New Zealand: An Urban Rural Profile, New Zealand, Available: http://www.stats.govt.nz/urban-rural-profiles/default.htm

Statistics New Zealand, 2005d, Linked Indicators, New Zealand, Available: http://www.stats.govt.nz/analytical-reports/linked-indicators/default.htm

United Nations, 2001, Indicators of Sustainable Development: guidelines and methodologies, Available http://www.un.org/esa/sustdev/publications/indisd-mg2001.pdf

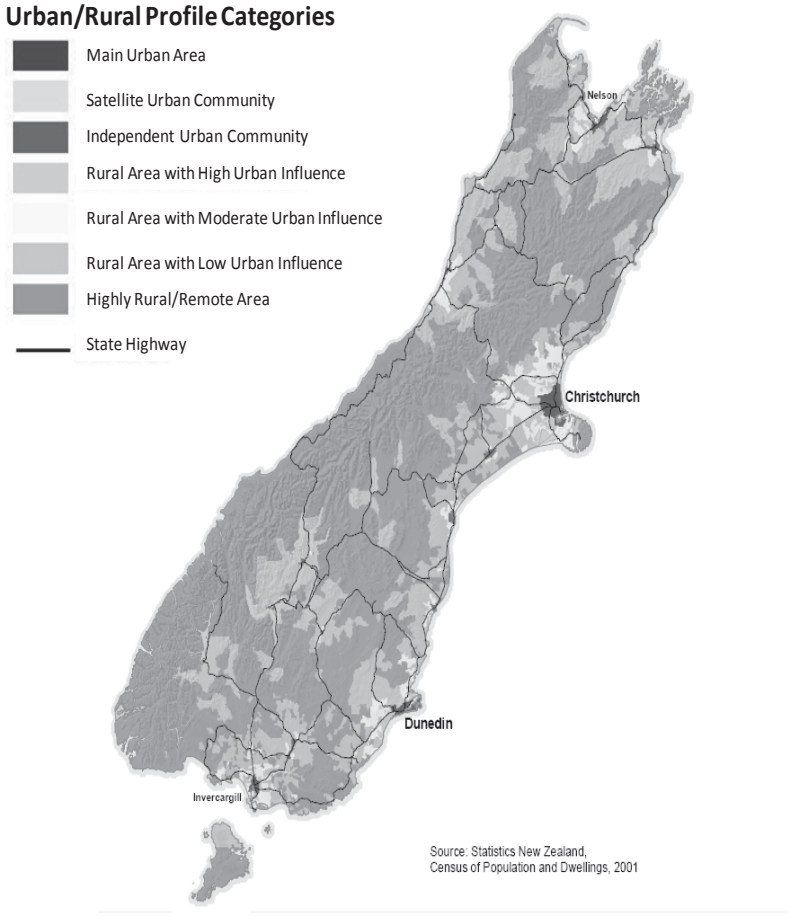
World Commission on Environment and Development, 1987, The Brundtland Report, Oxford University Press, Available http://www.anped.org/media/brundtland-pdf.pdf

Appendix 35.1 Urban Rural Profile of New Zealand's North Island local communities, 2001



Note: Data based on 2001 Census of Population and Dwellings.
Source: Statistics New Zealand, 2005c, *New Zealand in the OECD*

Appendix 35.2 Urban Rural Profile of New Zealand's South Island Local Communities, 2001



Note: Data based on 2001 Census of Population and Dwellings.
Source: Statistics New Zealand, 2005c, *New Zealand in the OECD*

Chapter 36
Achieving Outcomes From Indicators: Community Assessment Projects in the United States, A Focus on Santa Cruz County

Susan Brutschy
President, Applied Survey Research, USA

Deanna Zachary
Project Manager, Applied Survey Research, USA

Abstract

Applied Survey Research (ASR), a non-profit social research firm located in California, has completed dozens of community assessment projects (CAP) in California, Alaska, Arizona, Georgia, Nevada and Washington. Many of the CAP projects evaluate the quality of life in a particular geographic region in six subject areas: the economy, education, health, public safety, the social environment, and the natural environment. Other projects focus on particular populations such as the homeless, the elderly, children and farmworkers. ASR works with each community and each specific population to use the data as a catalyst to improve the lives of children and families.

One project in Santa Cruz County, California, was chosen in 2007 as an example of one of the best community indicator projects in the United States; the project won first place in the 2007 Community Indicators Consortium Innovation Awards sponsored by the Brookings Institution in Washington D.C. The Santa Cruz County CAP is also profiled in a new book about best practices in community indicator projects throughout the world[1]*. ASR's community assessment projects that focus on the homeless have been listed by the United States Government Department of Housing and Urban Development as a best practice.*

ASR designed the Santa Cruz County CAP based on some of the most successful and oldest community indicator projects in the United States, including the Jacksonville, Florida model and projects in Washington and New Mexico.

ASR has researched and written the report for the Santa Cruz County Community Assessment Project (CAP) every year since it began in 1994 and this project has continued through 2007 and is expected to continue indefinitely. The Santa Cruz CAP accomplishes four major goals: 1) to assess the quality of life in the county; 2) to educate community members about the economy, education, health, public safety, natural environment and social environment of the region; 3) to act as a catalyst for social and political change; 4) And to help evaluate the impact of social and political change in the county. The CAP process works both from the "bottom up" and from the "top down." In

2006, over 1 000 individuals came together to develop new goals for the CAP for 2010, as well as new indicators and survey questions for the resident telephone survey. Non-profit organisations, government agencies and businesses have adopted the CAP goals and created action strategies to attain the goals, thereby allowing the CAP process to become infused throughout the county.

This paper focuses on the Santa Cruz County CAP, how the CAP reveals the quality of life in the county and educates community members, the successes of the CAP in acting as a catalyst for change in the community, and recommendations for how community indicator projects might increase their ability to make political and social change. The paper also includes some discussion of how ASR has taken the CAP model and used it in community assessment projects in Alaska, Arizona and many California counties as well as among more specialised populations including homeless individuals, the elderly, children, and farmworkers in California, Arizona, Georgia, Nevada and Washington.

Introduction

California is divided into 58 counties, or geographical divisions. Santa Cruz County is located 65 miles south of San Francisco, a region which is historically famous for its fishing and the lumber it supplied to rebuild San Francisco after the great earthquake and fire of 1906. It is now widely known for its agricultural industries, tourism and surfing. There are 262 312 residents, the majority of whom are Caucasian (64%), followed by an increasing Latino/Hispanic population (30%) (State of California, Department of Finance 2004). The county is one of the most expensive places to live in the United States with a median home price of $625 000 in 2005, due in part to its proximity to Silicon Valley (National Association of Home Builders [NAHB], 2005), one of the nation's major technology centers.

The Santa Cruz County Community Assessment Project (CAP) was started in 1994 by the United Way of Santa Cruz County, a non-profit organisation, when it convened a collaborative of over 25 agencies ranging from public and private health organisations, to education, human services, and civic organisations. Dominican Hospital, one of two hospitals in the county, was an especially critical partner in the project. The hospital became involved due to California state legislation (SB697) that required non-profit hospitals to demonstrate the benefits they offered to their communities. The collaborative that formed in 1994 still exists in 2007, with the addition of several new partners totaling approximately 35 partner agencies directly involved in the CAP.

Applied Survey Research (ASR), a non-profit social research and evaluation firm, was contracted by the United Way to be the researcher for the CAP and to incorporate best practices from other assessment efforts across the nation. The Santa Cruz County CAP partners looked at dozens of community assessment projects and incorporated best practices especially from two highly regarded projects: the Jacksonville, Florida model, and an assessment project in Albuquerque, New Mexico. Jacksonville has one of the oldest, most successful community indicators projects, having started in 1985. The CAP initially modeled their face-to-face survey methodology of community residents from the Snohomish County Counts project in Washington. The CAP looked to Jacksonville and Albuquerque, New Mexico's community assessments to model their Technical Advisory Committees (TACs) where experts from different fields come together to discuss

indicator selection and to refine indicators. In 2005, ASR looked to a newer community indicator project known as Truckee Meadows Tomorrow in Nevada; Truckee Meadows encouraged non-profit organisations and businesses to "adopt and indicator" such that their county's political and social change efforts would permeate through their political and business institutions.

The goal of CAP is to continually improve the quality of life for Santa Cruz County residents by: raising public awareness; providing accurate information to guide decision making about the creation, management and redesign of community programs; setting community goals using measurable quality of life indicators; establishing collaborative action plans to achieve these community goals; and evaluating the impact of social and political change.

The Applied Survey Research (ASR) CAP System

When ASR conducts a community assessment project in a particular geographic region, ASR works with the local community members to choose goals and indicators that are best suited for their purposes and their region. One of ASR's goals is to include as many citizens as possible in the CAP process including community leaders, elected officials, government agency staff, non-profit organisations and volunteers. For example, in Alaska, community members were especially interested in tracking domestic violence, substance abuse, housing prices, and employment rates. In Arizona, ASR worked closely with large numbers of Native American Indians who were most interested in the health status of their residents, homelessness, the use of the Hospital Emergency Room, as well as their poverty rates and educational attainment. Each CAP that ASR conducts includes both primary and secondary data for indicators. Often, the primary data include a telephone survey to local residents and a face-to-face survey. Increasingly, ASR is using face-to-face surveys due to the fact that many people use cellular phones instead of land lines while other community members may not own phones or may be homeless.

The Santa Cruz CAP is based on primary data and credible secondary data that are gathered for 125 indicators in six areas: Economy, Education, Health, Public Safety, the Natural Environment, and the Social Environment. Economic indicators include such things as agricultural production, unemployment rates, poverty levels, and housing affordability. Education indicators include school enrollment, expenditures, testing scores, dropout rates, and university attendance. Health indicators include access to health care, insurance coverage, birth rates, prenatal care, substance use, mental health, disability status, obesity, and leading causes of death. Public Safety indicators include crime rates, child/elder abuse, and family violence. The indicators for the social environment include voting, racism, discrimination, hate crimes, homelessness, and charitable giving. Indicators for the natural environment include open space, park use, farmland, pesticide use, water quality, beach closures, water use, air quality, roadway congestion, transportation, and endangered species.

Some indicators are chosen as "key" indicators. The key indicators are not necessarily the most important quality of life indicators, but are most indicative of the overall condition of that particular subject area. If key indicators are improving or worsening, it is likely that trends in the whole are moving in the same direction. Some key indicators in the Santa Cruz County CAP in 2006 included: job growth, household income, housing affordability, educational test scores, child care enrollment, access to

health care, crime rate, family violence, voting, racism and discrimination, hate crimes, youth activities, open space acreage, coastal wetlands, and water use.

Goals and Purposes

The ASR CAP model encourages communities to choose 5-year goals for their regions and action plans to attain those goals. ASR works with communities to determine appropriate goals and encourages community and business leaders to "adopt a goal" so that there are multiple champions for each goal, and so that the community as a whole supports the work of the CAP. In Alaska, for example, community members are working to create "Safe Families" by reducing domestic violence and child abuse and neglect. The Santa Cruz CAP includes a list of community goals for improvement in each of the six CAP areas. The goals include such things as universal health care coverage for children, higher rates of education, higher levels of employment and housing, and lower obesity rates for children. Since 1994, the United Way of Santa Cruz County, a non-profit organisation, brought together over one hundred organisations on several political and social change efforts, including reducing teen substance abuse, increasing health care coverage for all children, and reducing childhood obesity. In 2006, over 1 000 community members throughout the county were involved in CAP goal setting for 2010. As of 2007, the former collaborative efforts of over 100 organisations convened by the United Way are still at work in the areas of teen substance abuse, health care and obesity, but in addition, there are more decentralised efforts to make community change by non-profit organisations, government agencies and businesses. Santa Cruz County community leaders and businesses are now adopting individual CAP goals to create deeper social and political change. For example, the County Office of Education is leading other organisations for the CAP goal of providing preschool for all children. An environmental group called Ecology Action is leading dozens of organisations to promote environmentally friendly "Green" building.

ASR encourages communities to have yearly celebrations where they honor community heroes who have made a difference in promoting the CAP goals in their region. Some communities, like Truckee Meadows Tomorrow, have a black tie dinner to celebrate organisations that have made a difference in the community. The Santa Cruz CAP has an annual press conference and luncheon where approximately 25-30 individuals are honored for the work they have done to promote the community goals.

Data Sources for the CAP

ASR encourages communities to use multiple data sources, including telephone surveys of residents, face-to-face surveys of residents and secondary data from a variety of local, county, state and federal sources. As previously described, ASR is using more face-to-face surveys in their community assessments, due to the proliferation of cellular phones, and the number of people who may be too poor to own a phone or are homeless.

The central source of primary data in the Santa Cruz CAP is a telephone survey of a sample of Santa Cruz County residents, in both English and Spanish, with over 700 randomly selected county residents. The intent of the survey is to measure the opinions, attitudes, desires, and needs of a demographically representative sample of the county's residents. Secondary data for the Santa Cruz CAP are collected from a variety of sources, including but not limited to: the U.S. Census; federal, state, and local government agencies; academic institutions; economic development groups; health care

institutions; libraries; schools; local police, sheriff and fire departments; and computerised sources through online databases and the Internet. Two key California sources are the California Health Interview Survey and California Healthy Kids Survey.

Philosophy behind the CAP

ASR encourages communities to use Results Based Accountability (RBA), a methodology created by Mark Friedman to make program or community change by developing shared outcomes and working backwards from outcomes to particular action strategies. The United Way of Santa Cruz County and Applied Survey Research (ASR) had both been working extensively with RBA and incorporated RBA strategies into the work of the ASR CAP model and the Santa Cruz CAP. RBA includes several steps: Identifying shared program or community outcomes; identifying indicators to measure current and future conditions; collecting baseline data; determining the story behind the baselines; determining the partners that have a role to play in improving outcomes; determining the best ideas to impact outcomes; and developing an action agenda to attain the outcomes. ASR also relies upon the United States federal standards known as Healthy People 2010 goals, which sets goals for the whole nation to achieve by 2010.

Publication of the CAP and Media Coverage

Communities must make it a priority to share their data widely with their residents. There are multiple ways to do this including press conferences, media coverage in television, radio and print press, and publications of the data in various forms. Each year, the Santa Cruz County CAP is published in several forms including a large compendium of the 125 indicators with data for the last decade for each indicator, where available. This larger document consists of approximately 330 pages. A shorter executive summary is prepared by the local hospital and is sent to every household in the county. The CAP acts as a rich data resource for the media, elected officials, community leaders, county agency staff, non-profit organisations, businesses and the public. The Santa Cruz County CAP is frequently covered in the daily county newspaper and one of the weekly papers, as well as on television and radio. Many public officials and/or their staff members attend the yearly press conference and are provided with free copies of the CAP. A CAP Stakeholder Survey showed that several legislators said that they relied on the CAP data. Anecdotally, legislators tell CAP stakeholders that they use the CAP data for speeches, talking points, and as a resource for creating legislation.

Outcomes of the Santa Cruz County CAP

The Santa Cruz County CAP has achieved its greatest success in the areas of teen substance use and universal health care for children. Early CAP data about the high levels of teen substance abuse and low levels of health insurance for children acted as a catalyst for change in the county. The United Way used the data to bring dozens of agencies together, including non-profit organisations, county agencies, elected officials and private citizens to develop social and political strategies to impact the problems.

Community Change for Youth and Substance Abuse

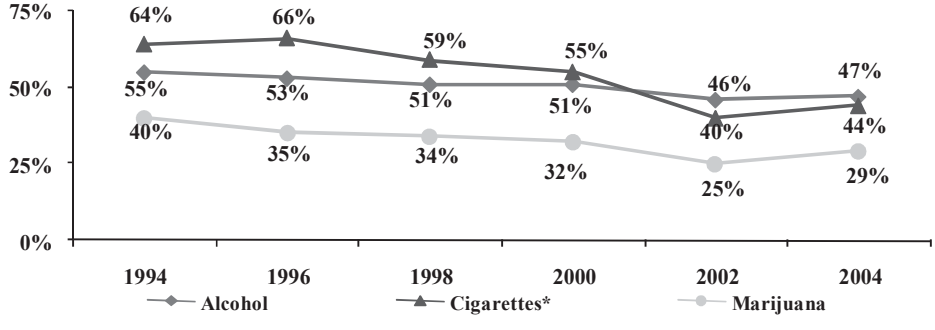
The first CAP published in 1995 showed alarming rates of youth who self reported using marijuana and alcohol in the last 30 days. The Santa Cruz County rates were higher

than in the state of California. For example, in 1994, 51% of Santa Cruz County 9th graders used alcohol in the last 30 days as compared to 44% of California 9th graders. Similarly, 40% of Santa Cruz County 11th graders reported using marijuana in the last month as compared to 26% of California 11th graders.

In 1995, the United Way used the CAP data to mobilise a coalition of 110 agencies and individuals to seek solutions to youth substance abuse. The group developed seven strategies including new laws, public education, a grand jury report, youth leadership training, referral and home visiting programs, new teen centres and treatment services, and a public policy panel on youth access to alcohol. Since 1997, over $1 million was raised to support activities devoted to teens. By 1998, the previously increasing trend lines in teen substance use began to decline. Since then, subsequent CAPs have shown consistent decreases in overall alcohol and drug use among 9th and 11th grade students. Substance use has also decreased in the state of California as a whole, but the decreases in Santa Cruz County are larger than the decreases in the state.

Figure 36.1 below shows a dramatic decline in the use of alcohol, marijuana and cigarettes for 11th graders since the first CAP was released in 1995.

Figure 36.1 Percent of Santa Cruz County 11th Graders Reporting Substance Use in the Last 30 Days



Source: The American Drug and Alcohol Survey, 1994-1998. California Healthy Kids Survey, 2000-2005.
* Cigarette use is the percentage of students reporting "ever using" cigarettes.

Overall use declined dramatically, but data also showed a trend towards more binge drinking among those youth who did use alcohol. Those data led to a new effort to stop binge drinking, which will be described in the next section of this report under New Initiatives.

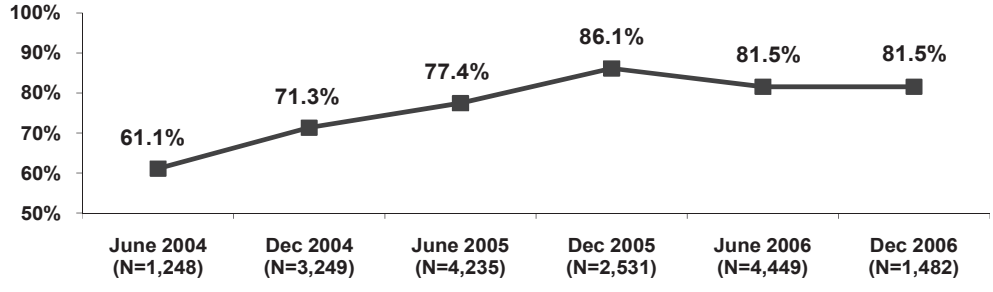
Community Change for Children and Health Care

The CAP data findings in the late 1990's also led to a change in the way health care is provided to children in the county. CAP data from the early reports showed that children's access to health care was a major issue for the county, especially for low-income and immigrant children. As a result of those early findings, there was a county-wide effort to improve health care for children. Now, all Santa Cruz County organisations that serve children are working toward guaranteeing health access for all children either directly or indirectly. Many programs offer direct assistance to accessing insurance and

health care such as providing outreach and enrollment in children's health insurance, and other programs offer referrals. In 1999, there were 1 023 children enrolled in a low-income health care plan for children; that number increased by 376% to 4 868 children in 2006.

There has been a special effort to provide health insurance to children under five years old. Figure 36.2 below shows the increase in the percentage of children ages five and younger who were covered by health insurance in 2006 as compared with 2004.

Figure 36.2 Percent of First 5 Child Participants with Health Insurance, at Six-Month Reporting Intervals



In 2004, health insurance coverage data from CAP also supported the creation of a coalition of over twenty organisations to launch a new initiative to create an insurance program for undocumented immigrant children. As in California as a whole, there are many children in Santa Cruz County who do not have legal immigrant status. There is an especially large number of Latino families in the southern part of the county who are agricultural workers, including some who are migrant farmworkers. Local community members wanted to develop an insurance program to cover those children. The new insurance program began in July 2004 with 28 children ages 0-5 and by January 2005 had 1 050 children enrolled in a program that offers medical, dental, vision and mental health coverage. Santa Cruz County is one of the first counties in California to offer its own health insurance program to alleviate health disparities, especially for low-income Latino immigrant children.

Catalyst for New Initiatives: Binge Drinking

In addition to the community change in the areas of youth substance abuse and children's health insurance, there are many new initiatives which have emerged from CAP data. The youth substance abuse figures showed a steep decline from 1994 to 2004, but there was a simultaneous troubling finding in the area of youth binge drinking. Youth overall were using alcohol less, but the youth that did use alcohol were bingeing more, where they drank five or more drinks within a two hour period. Santa Cruz County is the 6th worst county in California for binge drinking rates for residents ages 18-34. The alarming binge drinking rates motivated the youth substance abuse coalition to turn to the topic of binge drinking. The coalition applied for and received a three-year grant from the California Department of Alcohol and Drug Programs to provide a needs assessment and identify and implement environmental prevention strategies for binge drinking.

CAP Efforts Lead to Initiative on Childhood Obesity

CAP data in 2002 showed especially alarming percentages of obesity among children and adults. The county ranked at the bottom (63rd) of California's 66 counties and health jurisdictions for overweight children aged 5 to under 20 years old. These data findings galvanised the United Way of Santa Cruz County and several other organisations to convene a group of over 150 members in a new countywide initiative to take action including providing culturally appropriate health education, and cooking classes, promoting physical activity, integrating nutrition information at school, assisting schools in providing healthy food, and advocating for better local and statewide legislation such as more pedestrian friendly cities. Staff of the Initiative are working closely with the California State Assemblymember for Santa Cruz County to craft new state legislation to encourage healthier eating under government food programs, as well as making it easier to apply for food stamps and maintain eligibility. One recent California legislative success was to allow food stamp recipients to use their Electronic Benefit Cards at farmers' markets so that recipients could buy good quality, organic and local produce. The collaborative is providing outreach to low income Latino communities in the county that have especially high rates of obesity and diabetes. One strategy includes working with owners of mobile food vans that supply fast food to farmworkers, in order to offer more healthy food alternatives. Additionally, the group created a cookbook in Spanish for how to make traditional recipes with lower fat methods. There is also a Latino Five a Day program that encourages Latino community members to eat from 5 to 8 servings of fruit and vegetables each day.

Homeless Census and Survey

CAP data revealed higher levels of homelessness than were expected in Santa Cruz County. In response to those higher rates of homelessness, county agency staff and representatives of non-profit organisations wanted more detailed data about the homeless population, including how long they had been homeless, what factors led to their homelessness, the demographics of the population, as well as their family status and health status. The desire for more data led to two new efforts, one was to conduct a countywide homeless census and survey, and the second was to add questions about homelessness to the CAP resident survey, which had been conducted every year.

The county, therefore, embarked on its first systematic effort to provide a detailed census and survey of the homeless population in 2000. ASR was hired as the research partner in the effort and created a unique methodology to hire homeless individuals to help in the effort. The belief is that homeless people have the greatest knowledge about the overall homeless population including the places where homeless people sleep at night, the encampments they create, and more detailed information about their lives. ASR hired and trained homeless individuals to participate in both the census of the homeless population as well as a detailed survey of the homeless. The census was conducted at one point in time, by using homeless workers and volunteers to canvass the entire county to count all the homeless people they observed. That point in time visual count was supplemented by telephone calls to hospitals, shelters, treatment centres, and the jails to get accurate counts of their homeless populations. Subsequently, homeless workers and social service providers conducted individual one-on-one surveys with homeless individuals to determine the demographics of the homeless population, whether the individual was working, their health status, their educational attainment, their family status, whether they had children staying with them, their use of substances, their mental

health status, and whether they had suffered from domestic violence. The study provided this detailed information to allow city and county planners and social service providers to have a more accurate picture of the homeless population in order to provide better services. The homeless data helped to provide the impetus for new shelter services, including a new residential homeless shelter for individuals and families, new funding for transitional housing, more permanent supportive housing, and the Homeless Person's Health Project of Santa Cruz County. There have been approximately $1.5 to $2 million more dollars for homeless funding since the 2000 homeless census and survey. The homeless census and survey was repeated in Santa Cruz County in 2005 and 2007.

ASR Homeless Census and Survey: Best Practice Recognised by the US Department of Housing and Urban Development.

The ASR methodology has been listed by the United States Department of Housing and Urban Development as a best practice and ASR has conducted similar studies in other California counties, including San Bernardino, San Mateo, Santa Clara, Mendocino, Monterey, and the largest municipal housing study in the country, in Los Angeles County in 2005 and 2007. ASR has also worked with the municipalities of Atlanta, Georgia, and Las Vegas, Nevada to conduct homeless census and surveys in those regions.

ASR Children's Reports.

ASR has applied the CAP model to children's issues in several California counties, including Contra Costa, Santa Cruz, Santa Clara, San Mateo, and Nevada County. The Children's Reports focus on children ages 0-18 and highlight several general areas of health and well-being including physical and mental health, child safety, proficiency in school, and economic stability. Some of the health indicators include infant mortality, low birth weight, pre-natal health of mothers, breastfeeding, access to health care, physical fitness and teen sexual health. Some of the child safety indicators include child abuse, foster care, injury hospitalisation, drug/alcohol/cigarette use, and death rates. School measures include kindergarten readiness, child care availability, reading proficiency, high school dropout rates, and college readiness. Economic measures include child poverty rates, participation in free lunch programs at school, and housing affordability. ASR is working with community residents to create a Children's Agenda for Santa Clara County, California, one of the major counties for the high tech industry of Silicon Valley. ASR completed a Children's Report for the county, and then worked with non-profit organisations and government agencies to create a Children's Summit where nearly 500 individuals gathered to choose the most important outcomes and indicators for children. The Children's Agenda collaborative is currently in the process of refining the outcomes, indicators and action plans for the county. ASR also completed multiple children's reports for nearby San Mateo County, which together with Santa Clara County has launched an interactive website with the children's data gathered by ASR. The site is funded by the Lucille Packard Foundation for Children's Health and is found on the web at http://www.kidsdata.org.

ASR Farmworker Surveys: The Largest Farmworker Survey in United States History Using the Promotores Model

ASR is using the CAP model in its work with Washington State farmworkers. ASR trained farmworkers to survey approximately 3 000 other farmworkers about their quality

of life, health status, housing, employment history, and level of civic engagement. ASR trained both farmworkers and promotores to interview other farmworkers as they picked crops in 15 different Washington counties, representing over 90% of farmworker production in the state. The method is based upon the Promotores model where community members promote health in their own communities. Promotores provide leadership, education and resources and they have a trusted status that allows them access to farmworkers who traditional health interviewers would not have had access to. The study includes migrant workers, settled workers, and indigenous Indian workers from Mexico. Health survey questions include: health status, range of health issues/diseases, recent doctor/dentist visits, mental health, domestic violence, health care access, insurance coverage for adults/children, pesticide exposure, results of pesticide exposure, and housing status (including water quality, pests, plumbing, heating). There were extensive questions about worker housing and desired housing that are being used to inform government policy makers about where new farmworker housing should be built. Surveys took place in public places, including grocery stores, flea markets, farmworker encampments and at special events. As a result of the survey, the state of Washington has already doubled its funding for farmworker housing.

Data Gaps Lead to Other New Studies

In addition to the homeless study, Santa Cruz County CAP data led to other new survey efforts in Santa Cruz County in the areas of farmworkers, children who witness domestic violence, rape victims and the disabled. Santa Cruz County has a large population of farmworkers and migrant farmworkers; however, county staff and the staff of non-profit agencies did not have an adequate picture of the lives of farmworkers. This gap in the data led to a call for further research into the farmworker population, a study of which was completed in 2001.

The CAP revealed other data gaps such as with victims of domestic violence and their children and more recently for people with disabilities. ASR, the research partner on the CAP, was also conducting evaluations in the surrounding counties about the co-occurrence of domestic violence and child maltreatment. Historically, those two populations had been treated separately through separate county and non-profit agencies. But recent research shows that in a national survey of 6 000 American families, 50% of men who frequently abused their female partners, also abused their children (Peled, E., Jaffe, P.G. & Edleson, J.L., 1995). Further, children who witness abuse between their parents frequently experience behavioral and emotional problems, such as aggression towards others, withdrawal, low self-esteem, and lower school achievement. Santa Cruz County did not have detailed data about the co-occurrence of child maltreatment and domestic violence, so CAP stakeholders decided to add indicators to the CAP to capture those data, including new questions on the resident surveys.

In 2005, crime data from the CAP showed an overall declining trend line in all violent crime in the county, except for rape. In 2005, the CAP findings for rape were especially alarming and drew the attention of the media as well as more calls for research by advocates. This led to a new study of rape that was released to the Santa Cruz City Council in the Spring 2006.

Summary and Conclusions

ASR developed a CAP model after researching the Jacksonville, Florida community indicator project and projects in New Mexico and the state of Washington. Over the last thirteen years, ASR has made changes and refinements to the model and has used it in many California counties, as well as in Alaska and Arizona. The ASR Santa Cruz County CAP is one of the oldest CAP projects in the country, having started in 1994. ASR also uses the CAP model with specialised populations, such as the homeless, the elderly, children, and farmworkers. ASR has completed over a dozen homeless projects, many indicator projects that focus on children ages 0-18, and the largest farmworker survey in the history of the United States.

When ASR began this work, there was no national organisation to facilitate research and best practices in the field. ASR could only look to other successful indicator projects, such as Jacksonville, Florida. Fortunately, now there is a national organisation known as the Community Indicator Consortium that holds a yearly conference and has a newsletter to bring together experts in the field of community indicator work.

CAP projects are essential contributors to political and social change in the lives of children and families. For example, in Santa Cruz County, the CAP acted as a catalyst in creating new effective programs in the areas of teen substance use and offering universal health care coverage to all low-income children in the county. To summarise, the Santa Cruz County CAP has led to:

- Reduced substance abuse among youth;
- The development of universal health care coverage for all children in the county, regardless of immigration status;
- An initiative to reduce youth binge drinking;
- An initiative to promote nutrition and physical activity to reduce childhood obesity;
- Research on specific vulnerable populations such as the homeless, agricultural workers, rape victims, and people with disabilities;
- An executive summary that goes to every home in Santa Cruz County;
- Awards for Community Heroes (individuals and organisations serving the county and working towards the Community Goals) at an annual celebration and data sharing event.

The Santa Cruz County CAP was chosen in 2007 as an example of one of the best community indicator projects in the United States; the project won first place in the 2007 Community Indicators Consortium Innovation Awards sponsored by the Brookings Institution in Washington D.C. The Santa Cruz County CAP is also profiled in a new book about best practices in community indicator projects throughout the world.[2] The United States General Accounting Office (GAO) has mentioned the Santa Cruz County CAP project in one of their reports about best practices for indicator reports and staff of the GAO interviewed CAP stakeholders and ASR staff about the history of the CAP. In 2006, ASR received a prestigious national award for its work with the homeless and for

other social problems. The award was given to ASR by the Association for Applied Clinical Sociology. To view the range of community assessment work that ASR has conducted please visit our website at http://www.appliedsurveyresearch.org.

Notes

1 Sirgy, J., Phillips, R., Rahtz, D. (2007). Community Quality of Life Indicators, Best Practices III. The International Society for Quality of Life Studies, (ISQOLS).

2 Sirgy, J., Phillips, R., Rahtz, D. (2007). Community Quality of Life Indicators, Best Practices III. The International Society for Quality of Life Studies, (ISQOLS).

Part Ten
Using Indicators to Make Governments Accountable

Chapter 37
Measuring and Improving the Performance of Governments

Prof. Dr.Ömer Dinçer
Former Undersecretary of the Prime Ministry, Turkey

Abstract

Measuring and improving the Government's performance is a function of managerial and societal culture. By this aspect it could be seen as a change management project in public management. Performance oriented management entails institutionalisation that in turn envisages data and information based management. Measuring performance in public administration should be structured as an approach in which all the administrative levels, from international to national and local, are taken into account in an integrated manner.

In recent years, internal dynamics as well as international commitments of Turkey have triggered series of significant improvements ranging from publication of certain basic sets of statistical indicators to fundamental institutional and legal changes, thus increased importance of transparency and accountability.

Introduction

Measuring and improving the Government's performance could be interpreted as a systems and a socio-cultural phenomenon.[1] This area is a dependent variable of the managerial culture in the public sector. Government policies have direct and indirect impacts both on the private sector and the public at large. In a broader sense, therefore, it can be suggested that the performance of the Government is the function of the societal culture. In fact, the notion of "performance culture" per se is mentioned.[2]

It is useful to consider measuring and improving the performance of the Government as separate concepts. While the former is relatively a technical and methodological issue, the latter defines a more comprehensive and multi-dimensional intervention. Put differently, measurement provides information about outcomes of the Government's policies, decisions and actions. On the other hand, the way how certain outcomes in certain areas could be achieved and how relevant factors could be altered to get desired outcomes shape the method and the content of the improvement.

Outcomes borne by Government's actions are complex issues that have multi-dimensional determinants. Many external factors affect Government's performance, thus outcomes. Measuring Government's performance independently isolation from uncontrolled factors poses difficulties. It requires a comprehensive systems design and a

change in managerial culture to measure Government's performance, which translates it essentially to a change management project.

In the last few years measuring and improving the performance of the Government has become a priority area in Turkey. Comprehensive public sector reforms have been realised to this end.

Measuring and Improving the Performance of Government as a System

Globalisation causes convergence among countries in terms of public administration and its reform. On one hand, demand by governments and international actors to compare the development courses of nations have been on the rise. Nations have had an ever increasing need to learn from and compare themselves with others.[3] Developments in information and communication technologies have substantially altered and facilitated research, learning and benchmarking and rendered them more accessible. On the other hand, the concepts, notions and tools in this area evolve to uniformity. The performance management has been adopted by many governments, independently from the level of development as well as socio-cultural, economic and political conditions.

Performance management needs a systems design in which the relevant mechanisms are established and an effective organisation is set up. In this area there are various policy and implementation options ranging from the improvement of the existing indicators to radical structural transformation. As measuring and performance are integrated and multi-dimensional approaches, their management too requires a multi-sectoral and multi-disciplinary approach. It is clear that the ultimate performance of the Government is a dependent variable of the civil servants. Therefore the personnel regime may also be regarded as a component of the system.

Data and information based management requires the development of a perspective encompassing administrative hierarchies, from international to national and local levels. Performance management may be more effective by producing and using data and information at these levels and by ensuring their integrity and consistency. Systems and levels of measurement are not independent of one another. In the case of transparency, for instance, there are different effectiveness indicators to be monitored by different administrative levels. A government which claims accountability should ensure integrity and coherence that range from international transparency indices to the information provided by non-governmental organisations at the local level.

The design of such a system aims at achieving an assessment of present statistics within a framework, acquisition of data in unmeasured areas and the coordination of indicators as well as measurement and evaluation systems. The systems approach forms the parts of a whole as regards data and information which are used for measuring progress. They should be the expressions of a political and technical design and reflect a meaningful scenario as a whole. This may only be achieved through data and knowledge based management by information, by putting forth refined objectives to be reached through policies, at the Government and agency levels. Ensuring the coherency of indicators as well as the system is another critical factor. In addition to the traditional indicators, there might be a need to introduce new measures into the system. The composition that may best represent the progress must be created.

International organisations and non-governmental organisations manage data on the performance of governments in a multitude of areas. In addition to the conventional data

published by organisations such as the United Nations, the World Bank and the EU, there is data available such as the Corruption Perceptions Index[4] and the Democracy Index[5] published by the Transparency International and Economist Intelligence Unit respectively. In certain circumstances governments find themselves in a position to be "accountable" to the international actors, further supporting the argument that structuring the system of measurement should commence from the international level.

There are efforts by national NGOs in this area as well. For example, some non-governmental organisations in Turkey carry out studies on measuring citizen satisfaction on a systematic basis and in line with international standards. These studies can be utilised as cross control mechanisms vis-a-vis official measures.

As mentioned above, measuring does not have a meaning and existence as a sole technocratic and bureaucratic action. From this point of view, the question of "For whom must the performance of the Government be measured?" becomes legitimate. Organised and coordinated data and information collection is a category, its presentation is that of another. The data and information do not only concern the international and national organisations but also citizens. Various derivatives at various complexities and details of the same data may be provided for citizens, experts, decision makers, parliament and international society.

Critical Factors in Measuring and Improving the Government's Performance

First, a clear definition of performance, which is one of the basic concepts of the new public management, is an important issue. In this regard, short and long term performances must be discussed. Short term relief that may cause problems in the longer run should not be considered as a performance. Within the democratic system, a concept of the performance based on sustainable progress has to be substituted for populism. This approach will not only support the democracy but also provide useful contribution to the development process.

Determining the indicators is a critical factor that should be taken into consideration in measuring performance. Wrong, timeless and insufficient indicators may disrupt the whole system.

Second, the first level considered when measuring governments' performance is governmental level itself that represent the political will. Elections are ultimate and indispensable mechanisms that measure the performance of the governments in democratic order. However, they cannot be sufficient in a democratic culture to take elections solely as a performance criterion as they are general and held in long time intervals. In order to continuously monitor and evaluate the performance of governments the representative democracy has to be enhanced along participatory lines. This approach is salient as it increases the opportunities of citizens to hold the Government accountable for its actions as it also allows governments to make self-assessments and accordingly, changes in their policies.

To measure the performance, above all, governments must put forth measurable objectives when they take office. In other words, the performance must be defined in advance and declared to the public clearly. Rather than abstract and immeasurable programs, tangible and watchable ones should be demanded from governments.

Third, it is required that a strong link be established between declared policies and the allocation of resources for their implementation. In this context, it is important to

institutionalise a comprehensive and transparent budget process. Strategic perspective in the policy making and allocation of resources based on sound information and analyses must be strengthened.

Fourth, measuring the performance of the political decision-makers alone does not suffice. Accountability in all administrative hierarchies that starts with top managers must be clarified. Power and flexibility should be granted to the managers to enable them to take initiatives, solve problems and create innovations, in return for accountability. In other words, the complementary link between authority and responsibility should be established. One who does not have authority cannot be held accountable and authority should not be granted to the one who cannot assume accountability.

Fifth, not only the performance of the central but also that of the provincial organisations and local administrations should be measured. In order to increase the administrative quality and acting on the principle of competence in government, transparency and accountability mechanisms should be established effectively. The first step is articulating the objectives and priorities at agency and business unit levels and monitoring by indicators of performance. Ultimate measurement of the performance should be carried out at the individual employee level and should be seen as the integral part of the personnel regime.

Sixth, the mechanisms to measure performance should be designed independently from decision makers and those who carry out activities. One who measures should not be confronted with the one who is measured. Naturally agencies may set up internal measurement systems for self-assessment, which, however, do not suffice. Self-assessment should be regarded simultaneously with the independent evaluation and rather than an egocentric understanding, measurements sensitive to environmental perceptions should be conducted.

Seventh, the duty of the individuals and institutions measuring performance is not to judge but guide and support by feedback. Performance measurements, being different from the legal compliance evaluations, should be assessed not in terms of legal but of administrative results. It should be remembered that it is possible to work extremely inefficiently while complying with procedures perfectly. The history of the public management is full of examples of waste and failure that were produced in conformity with the rules. The public management should be able to work both in compliance with the law and in a result-oriented manner. The law which hampers the efficiency or efficiency which breaches the law is a dilemma. Both of these basic parameters should be taken not as opposite but complementary factors for a sound and human-focused management.

Eighth, a single measurement or measurements by the public sector alone should not be deemed sufficient. NGOs and the private sector should monitor and measure the government performance in their own areas of interest. The existence of NGOs that the society trusts and that are capable of conducting objective analyses without daily political concerns would be very useful. While these measurements contribute to the democratic culture provide the opportunity for the citizens to obtain information outside the public sector, they would also enable governments to hear about their performances from different perspectives.

Ninth, monitoring the perceived developments in addition to measuring the performance with objective data is important. Data collection is not sufficient. Simultaneous field studies which display the perceptions should be carried out. To

display the quality and as for different public services and institutions, should be focus on citizen satisfaction researches. These works to be conducted regularly would contribute to the "objective oriented culture" to take root. In addition to measuring service beneficiaries' satisfaction, gauging the satisfaction of the employees would contribute to the institutional performance as well.

Tenth, besides the macro-economic indicators (growth, inflation, budget, borrowing, foreign trade etc.) specific performance criteria in certain areas, which gain periodical importance, should be put forth and monitored. Not only the economic indicators but also the democratic, social and environmental indicators should be watched. Albeit a difficult area, the performance in foreign policy and relations should be monitored as well.

Finally, while analysing the performance, global trends around the world should be benchmarked and the developments in comparable countries should be taken into account. Various reliable and periodic international indices should be followed and the public should be informed accordingly. Internationally comparable and reliable measurements in various topics such as competitiveness, social development, freedom, environmental pollution and educational performance have critical roles in this respect. The importance of the organisations such as the OECD arises at this point. Countries should contribute to the development of the analyses at global level and some measures should be taken in order to facilitate the way the public is informed about the results.

Measuring and Improving the Performance of the Government in Turkey

Internal dynamics and accession process to the EU have led serious reforms to be realised recently in Turkey. Significant developments have been achieved in democratisation, human rights and transparency. Considerable changes have taken place that focused on increasing performance in financial and administrative areas. Transparency and accountability are the main principles of the deep-rooted public reforms realised in the last few years. Change in institutions, more specifically in managerial and behavioral approach has been targeted, all serving directly or indirectly to measure and improve the governmental performance. On the other hand, efforts towards change have been confronted with a strong resistance. Suffering from the lack of technical and analytical perspectives, this resistance materialises on an ideological basis that undermines causality.

One of the important reforms is the overhaul of the public financial management and control system. Planning and budgeting system has been improved by introducing medium term expenditure framework. The foundation of the result-oriented management and budgeting has been laid. The reform has been shaped in line with the principles and need of achieving better management, developing governance and accountability.

New bold steps have been taken about producing basic statistics as well. State Statistics Institute has been restructured. Life Satisfaction Survey has been conducted annually since 2003. This survey measures and monitors in time the satisfaction of the individuals on topics such as subjective happiness perception, health, social security, formal education, working life, income, personal security and justice services and personal development. The survey has revealed interesting results so far. The data about the satisfaction of our citizens have given us the opportunity to review our general ideas and assumptions. According to this research, in 2003, 2004 and 2005 the rate of the people who consider themselves happy and very happy reached to 60%. Those

developments would cause deep-rooted changes in public agencies and supply of services.

The Law on Right to Access to Information and approval of the Code of Public Ethics[6] are other important developments. Measuring and evaluation of performance has been realised not only in a supply-oriented manner but by the perspective of strengthening the governance through introducing the citizenship awareness and participation into the system. Law on Right to Access to Information has become effective following its adoption by the citizens in a short time with the support of the non-governmental organisations and academics. This is a strong indicator that the society would react positively to the participatory public management in Turkey.

Within the framework of our commitments to the EU, the development indicators at statistical regional units (NUTS) level have now been reported. Using information and communication technologies effectively, the Ministry of Interior has developed the local performance indicators at the national level and implemented the Performance Measurement Project in Municipalities (BEPER)[7] which allows for a comparable and consistent monitoring of the performance of local governments. On the other hand, monitoring and evaluation capacities of the central agencies should also be strengthened as determination of indicators individually by a municipality has a local meaning and function. To be able to detect their role in a system and to measure their contribution to broader national goals and objectives require the development of central capacity. Strengthening the local governments increases the need to develop monitoring and evaluation capacity at the center in turn.

Legal infrastructure on regulatory impact analysis which ultimately aims at impacts on society of regulations has been established.

Conclusion

Measuring and improving the performance of the Government is not solely a technical practice. Establishing management based on performance and improving it requires cultural change at administrative and individual levels. The cultural change may be facilitated by designing the performance phenomenon as a comprehensive system. The performance of the governments should be handled on a hierarchical structure from international level to the local level in terms of both measures and determinants of performance. The substance should be considered as a change management project.

In Turkey some serious structural changes have been realised recently. The experience has showed that basic building blocks of a comprehensive performance management have been formed, there exists a social and cultural potential which will make performance management possible and functional, and that the management of reform is critically important.

Notes

1 Though the Government here refers to the central government in a narrower sense, it should be noted that local administrations are also included in this framework.

2 Diamond, J. (2003a,b)

3 Even the methods developed for measuring budget programs and public agencies' performance can be imported by the governments. An interesting example for this is the integration of the Program Assessment Rating Tool, developed for evaluating USA federal programs by Office of the Management and Budget, into the financial management systems of Scotland and Thailand. (http://www.innovations.harvard.edu/awards.html?id=7496)

4 http://www.transparency.org/policy_research/surveys_indices/cpi/2006

5 http://www.economist.com/media/pdf/DEMOCRACY_INDEX_2007_v3.pdf

6 http://www.basbakanlik.gov.tr

7 http://www.beper.gov.tr

References

Diamond, J. (2003a). "Performance Budgeting: Managing the Reform Process." IMF Working Papers No. 03/33

Diamond, J. 2003b. "From Program to Performance Budgeting: The Challenge for Emerging Market Economies." IMF Working Papers No. 03/169

TÜİK (2005) "Yaşam Memnuniyeti Araştırması". Ankara: TÜİK

TÜİK (2004) "Yaşam Memnuniyeti Araştırması". Ankara: TÜİK

Chapter 38
AfriMAP: On Not Using Indicators to Score Progress in Governance

Bronwen Manby
Senior Programme Adviser, AfriMAP, Open Society Institute

Abstract

This article describes the methodology and experience of AfriMAP, a project of the Open Society Institute's network of African foundations, in working with civil society organisations in different African countries to carry out audits of government performance. It explains why AfriMAP has decided not to use its questionnaires as the basis for scoring countries on their performance, but rather focused on qualitative analysis. Finally it indicates which quantitative indicators are useful to AfriMAP, and what other indicators would be useful.

The Origins & Working Methods of AfriMAP

AfriMAP, the Africa Governance Monitoring and Advocacy Project, was established in 2004 as an initiative of the Open Society Institute network's four African foundations to monitor the compliance of member states of the African Union (AU) with the standards that the AU has adopted in relation to good governance, democracy, human rights, and the rule of law. In particular, AfriMAP is a response to the AU's adoption of the New Partnership for Africa's Development (NEPAD) and its associated African Peer Review Mechanism (APRM). AfriMAP aims both to promote engagement in a coherent way with these standard-setting exercises by the continental institutions and contribute to their implementation at national level.

NEPAD and the African Peer Review Mechanism

The New Partnership for Africa's Development is a strategic framework setting out a 'vision for Africa's renewal', initially adopted by African heads of state in 2001. Though it had independent origins and has its own secretariat based in South Africa, NEPAD is a programme of the AU.

In July 2002, the AU summit supplemented NEPAD with a Declaration on Democracy, Political, Economic and Corporate Governance. According to the Declaration, states participating in NEPAD 'believe in just, honest, transparent, accountable and participatory government and probity in public life'. The Declaration

committed participating states to establish an African Peer Review Mechanism, by which heads of state and government would themselves assess adherence to and fulfilment of its commitments by their fellow leaders.

In March 2003, the NEPAD heads of state and government implementation committee adopted a memorandum of understanding (MOU) on the APRM. This MOU effectively operates as a treaty, and states that do not sign it are not subject to peer review. As of February 2007, 27 countries had done so, just over half the member states of the AU. The meeting agreed to the establishment of a secretariat for the APRM, also based in South Africa, and the appointment of a seven-person 'panel of eminent persons' to oversee the conduct of the APRM process and ensure its integrity.

The APR Secretariat, functioning by late 2003, developed a questionnaire based on a wide range of African and international human rights treaties and standards to guide the assessment of participating states' compliance with the principles of NEPAD. Its questions are grouped under four broad thematic headings: democracy and political governance, economic governance and management, corporate governance, and socio-economic development. The questionnaire was formally adopted in February 2004, in Kigali, Rwanda, by the first meeting of the APR Forum, made up of the heads of state or government of all states participating in the APRM.

The APRM process in each country results in three important documents. The first is a 'self assessment' by the country concerned against the APRM questionnaire. The development of this self-assessment is supposed to be highly participatory and not controlled by the government; a 'national dialogue' about the challenges the country faces. The eminent persons and the APR Secretariat review this self-assessment, and develop a separate 'country review report' that takes into account the self-assessment but also independently collected information. Based on the self-assessment and the draft country review report, a programme of action (PoA) is developed by the government concerned in consultation with the eminent persons. The PoA includes a detailed log frame presentation of costed activities and targets to achieve. The country review report and PoA are ultimately presented and approved by the APR Forum, at which point they are made public. Three reports have been published so far: Ghana, Kenya and Rwanda. The country review reports of South Africa and Algeria are due for consideration in the margins of the AU summit in mid-2007.

Aid-for-Governance? The APRM and Development Assistance

The APRM was, at least in part, adopted as a result of a deal between African leaders and the G8: the G8 promised in its 2002 Africa Action Plan that more development assistance would be made available, in return for African action to improve domestic governance to ensure that the new aid would not be wasted. This promise was elaborated and repeated at successive summits, most importantly in 2005 at Gleneagles, which followed the publication of the report of the Commission for Africa appointed by British Prime Minister Tony Blair.

The APRM has, since it really began its work in 2004, taken on its own life in Africa and operates not necessarily with reference to development assistance. But insofar as the quid pro quo of action on governance for increased aid is still operational, there appear to be distinctly different interpretations from the G8 and African sides of what the APRM should achieve.

From the G8 (and OECD more generally), the view seems to be prevalent that the purpose of the APRM and similar assessments should be to score African countries on their governance performance: if a country achieves a certain standard, then it should be rewarded with additional aid. This approach is made explicit in the European Commission's governance profiles, which are intended to generate scores against a set of criteria developed by the Commission. (However, both the UK and France appear to be taking somewhat different, less score-based directions.) From the side of African governments, meanwhile, the understanding is rather that the purpose of the APRM is for each country to decide for itself what its main challenges are, and that whatever the challenges identified, development assistance should be awarded to support the process of addressing them. There is a very strong resistance to scores that rank countries on the basis of rich country experts' opinions, whatever the strength or otherwise of the methodology used.

In the context of this (perhaps overstated) dichotomy, AfriMAP has taken the view that the African analysis of the most useful role for the APRM is likely to be more productive. Its methodology was developed accordingly.

AfriMAP Methodology

AfriMAP works with the Open Society Institute's Africa foundations – the Open Society Foundation for South Africa and the multi-country Open Society Initiatives for Southern, West and East Africa (OSISA, OSIWA and OSIEA) – and national civil society organisations to conduct systematic questionnaire-based audits of government performance that can feed into or complement the APRM process.

AfriMAP's methodology was developed while the APRM was itself being set up, and evolved during a period of discussion over two years among both OSI staff and external experts that among other things considered and rejected the idea of developing 'shadow reports' based on the APRM's questionnaire. Instead, it was decided to focus on areas of particular interest to OSI. The three AfriMAP reporting themes ultimately crystallised as: justice sector and the rule of law; political participation and democracy; and effective delivery of public services. These three themes very broadly cover the classically described three arms or functions of government: the judiciary, courts and access to justice; electoral processes, citizen participation in government and the functioning of the legislature; and the effectiveness of the executive in delivering what should be the fruits of democracy to citizens.

In each case, we are focusing on institutions and processes rather than outcomes: in this, we are broadly following the UNDP definition of 'governance' as 'the exercise of political, economic and administrative authority in the management of a country's affairs at all levels. Governance is a neutral concept comprising the complex mechanisms, processes, relationships and institutions through which citizens and groups articulate their interests, exercise their rights and obligations and mediate their differences' (UNDP 1997).

Detailed research questionnaires for each theme were developed through extensive background research on appropriate standards and existing efforts to monitor their implementation, and in a series of dedicated two-day expert meetings. In terms of general methodology, one of the more important sources drawn on was the questionnaire used for the International IDEA 'State of Democracy' reports.

The justice sector questionnaire was the first completed, followed by political participation and democracy, and then effective delivery of public services. The justice sector questionnaire was the easiest to develop, since the relevant standards are the most clearly established at international and African levels (including for example, the exhaustive Guidelines and Principles on the Right to a Fair Trial and Legal Assistance in Africa, adopted by the African Commission on Human and Peoples Rights), and many other surveys have already examined the relevant issues, even if not with exactly the same focus. The second theme is also well-established in relation to election standards in particular, and functioning of the legislature to a lesser degree (where there tend to be more best practices than established standards), and we tried to extend it to the wider question of participation in decision making processes endorsed at the African level by the Charter on Popular Participation in Development (which is quite fuzzy, but establishes the principles clearly). This includes respect for freedom of expression and association, as well as wider opportunities for public engagement in policy debate. The African Charter on Democracy, Elections and Governance adopted in January 2007 adds greater force to these standards.

The third theme, on delivery of public services, was the most difficult to develop and is perhaps the most innovative. The questionnaire goes beyond the relatively well-trodden path of reporting on the implementation of the measures provided in the UN and AU anti-corruption treaties, to draw, among other documents, on the Charter for the Public Service in Africa, which sets out comprehensive guidelines on merit-based recruitment and promotion, declaration of assets, professionalism and courtesy etc; on the work of such organisations as the International Budget Project, which has led the way in developing standards of transparency and accountability in budget-making; on developing norms in relation to decentralisation; and – of most interest to this meeting – on the initiatives of PARIS21, UNSTATS, UNECA and others in relation to strengthening African statistical capacity. We have written the questionnaire with the intention that researchers may use it to assess the performance of the set of institutions involved, from the national ministry down, in delivering one particular public service – in most cases we expect this to be education or health – rather than the whole executive arm of government. In this way, we hope to prompt more focused research that could be used for effective advocacy on one key service which government should be ensuring that citizens can access. The questionnaire thus does not look at outcomes in education or health (literacy, disease prevalence etc) but rather at the capability of government machinery to take steps to improve poor outcomes. (And one of the basic functions of government that must work in order for politicians or citizens to plan and monitor public service delivery effectively is the production of reliable national statistics.)

All three questionnaires include a focus on development assistance and relations with foreign countries or international institutions, and all three have already been revised on the basis of feedback from the users.

We started work in five 'pilot' countries: South Africa, Malawi, Mozambique, Ghana and Senegal. The choice was determined by the OSI foundations. Work on an AfriMAP survey in Kenya was launched in April 2007. We also anticipate beginning work in Lesotho, DRC, Zambia, Benin, and Sierra Leone during 2007 and 2008. The AfriMAP countries have all signed up for the APRM (though we do not exclude the possibility of working on non-APRM countries in future), and are countries where there is a certain level of democratic openness so that the methodology is more likely to be able to have an advocacy impact in affecting the actual practice of government.

All the research for the AfriMAP reports is carried out by national partners, usually existing grantees of the OSI Africa foundations, and AfriMAP and the foundations have put a great deal of effort into ensuring that the research and reporting process has included a wide range of voices at national level. We have engaged in many preliminary meetings with important national stakeholders during the preparatory phase, and then held round table workshops at which the reports' findings have been discussed with relevant stakeholders. Several of the reports have also had technical advisory teams who have commented on findings during the research process. This commitment to process means that the time frame for production of our reports is at least 18 months to two years per country; longer than we expected at the outset.

Through these surveys we are trying to mobilise (and provide financial and technical support resources for) national civil society organisations and experts to produce high quality research products that analyse the workings of government in a qualitative way and – critically – provide recommendations on how to address the problems, around which civil society can mobilise. One of the key audiences for this research is the APRM, African governments' own model of audit, which we hope to strengthen both by example of quality research and by enabling more effective citizen participation in its processes. In this respect, our interaction with the OSI foundations funding in each country is crucial for successful follow-up advocacy based on the AfriMAP reports. In both Malawi and Mozambique, for example, AfriMAP research on the justice sector has been twinned with a grant to the university faculty of law by OSISA, and distribution of the AfriMAP reports among both academics and justice system staff will form a key information resource.

In some areas, AfriMAP's work has clear commonalities with surveys of sector performance by other institutions: the World Bank's assessments of the justice sector in different countries, for example. However, AfriMAP is trying both to get into more overtly 'political' areas, including by focusing on democracy, and in relation to public service delivery to cross the usual divide between health or education specialists and the governance and democracy organisations.

The intention is to repeat the surveys at intervals, though as updates on new developments rather than fully detailed assessments; the periodicity will depend also on decisions relating to numbers of additional countries that are added to the AfriMAP portfolio. As the number of reports builds up, AfriMAP will also start producing comparative analysis across countries – and we will be establishing a useful database of information on these issues that others can use for the same purpose.

The Merits and Demerits of Scores or Analytical Reports

Although the idea is that the AfriMAP audits will all follow a standard template that will enable countries to compare their experiences and learn from each other, we are not giving countries a score or ranking them against each other – and we have resisted suggestions that we should do so.

Why Does Afrimap Not Generate Numerical Governance Indicators From Its Surveys?

AfriMAP made a very early decision that it would not seek to develop a scoring method to use its questionnaires to create a governance ranking based on the opinions of

international or national experts, or even focus groups of 'ordinary people'. We came to this decision for the following reasons:

Our experience as activists working on the African continent is that the reaction to the Freedom House rankings (in particular) is generally to treat them as comic, and our knowledge of how the rankings were generated did not encourage us to challenge this view.

Some civil society rankings do have a certain immediate advocacy impact – for example, Transparency International's Corruption Perception Index – but rarely take you beyond the obvious in terms of indicating which countries are doing well or badly. It is impossible to draw any real distinctions between countries that rank close together on the list.

Although governance rankings with a more systematic, carefully constructed and transparent methodology – such as the Global Integrity Index of the Center for Public Integrity – could over time possibly provide a competitive incentive for improvement among neighbouring countries, this would require the creation of an all-Africa dataset, repeated annually or biannually, with a very major input of resources for an output of One of the main objectives of the AfriMAP audits was to provide an analysis of government systems that could help Africa address the crisis of implementation that it currently faces. A generalised governance ranking system based on indicators does not provide recommendations on reform measures.

- In particular, the AfriMAP surveys aim to inform the grant-making carried out by the OSI foundations in Africa. The foundations require qualitative analytical research to inform useful allocation of resources, rather than one-line rankings.

- It made most sense to construct a system of governance audit that meshed with the surveys being conducted by the APRM, since our aim was to provide constructive input into the APRM process and provide the sort of quality information that could strengthen the APRM's country review reports. (We did not adopt the APRM questionnaire itself as the basis for research, on the grounds both that it has significant deficiencies – for example, it does not request information on the state of media freedom – and that it ranges so widely that it would be difficult to provide quality inputs across the issues covered, while some areas went beyond the core interests of OSI.)

- We wanted to retain flexibility to vary the questionnaires and their application over time and in each country, in order to take account of our own learning processes and changing contexts: a scoring system valid over time requires complete consistency in the questions asked and a rigid set of rules to guide the calibration of responses.

- Governance rankings generated by intergovernmental institutions such as the World Bank's Country Policy and Institutional Assessments (CPIAs) that consider the inner workings of government in quite some detail – though they have their own faults (see for example OECD, 2006) – also depend on privileged access to government information, which as a civil society initiative we would not have.

What Quantitative Governance Indicators are Useful?

This is not to say that some score-based information sources and rankings are not useful – and we use them in our reports. AfriMAP uses both qualitative and quantitative information to prepare its audits of governance performance. The issue is therefore not so much the concept of governance indicators but the ways they are generated and how they are used.

In addition to the outcome indicators for poverty reduction, education and health for the Millennium Development Goals and similar international targets, the indicators that AfriMAP has used include, in the governance field:

- The index of press freedom prepared by Reporters sans Frontières;
- The comparative questionnaire-based surveys by the International Budget Project on budget transparency;
- The reports on government responsiveness to requests for information produced by the Open Society Justice Initiative; and in particular;
- The opinion poll results tracking public opinion in Africa on a range of issues related to democracy and governance published by Afrobarometer and DIAL.

The usefulness of these surveys is that they are based on indicators that provide 'real' information in a very clearly defined field. In the case of the press freedom, IBP and OSJI reports, the focused format is transparent in what it is measuring, the questions being asked have clear yes/no or quantifiable responses, and it is obvious both how to obtain an improved ranking and why (if you agree with the basic premise that citizen access to information is a good thing) it would be desirable to do so.

In the case of Afrobarometer and DIAL, the questions are more open-ended, without quantifiable responses on an individual basis, but the reports they produce do not claim to be providing a 'true' picture of respect for democracy, the level of corruption etc, but are rather a snapshot of citizen views based on a demonstrable and standard methodology. Importantly, the information they provide is kept disaggregated, so that what you learn are separate stand-alone reports on what percentage of citizens think (for example) that multi-party elections are a better way of choosing a government than military coups or inheritance; or that corruption is widespread; or that the police are helpful. These percentages tell you interesting detailed information about the interaction of a broad range of citizens (and not just an elite) with government, rather than being wrapped into one aggregate score that provides a superficially attractive simplicity but no information that could actually be used by policy makers or activists to guide work on specific reforms. The World Bank's country Governance Diagnostic Surveys fit in this same more useful category.

We also think that indicators that are context-specific rather than universal are more likely to be useful. While there is no doubt that the more established western democracies have contributed a great deal to theories and best practices regarding governance, universal indicators have tended to emphasise the value systems and norms of organisation that are dominant in those more established democracies. For example, in measuring a specific aspect of governance such as electoral systems there is need to take into account a country's electoral and political history. A numerical index purporting to

reflect the independence of a country's electoral body may fail to capture the normative framework within which governance institutions are constructed. Similarly, indicators relating to the media too often assume that the economic context is unproblematic, yet the economics of information differ greatly among countries and provide very different challenges: from the cost of newsprint or broadcasting equipment to domination by private sector conglomerates that drown out independent voices.

For this reason, AfriMAP has attempted to work in a way that moves away from the perception that measurement of governance is often externally driven and a cousin of the conditions that are often imposed as part of aid packages. To the extent possible, measurement of progress is thus linked to the value systems (including AU treaties, NEPAD, APRM, etc) that have been explicitly endorsed by countries that are being assessed.

What Other Indicators Does Afrimap Use?

AfriMAP's reports use other relevant numerical indicators when we can find them – and indeed national researchers are often generating useful primary information on indicators of various types, or collating it from disparate and hard-to-access sources (including interviews with government officials), especially in less-well documented countries (that is, most places in Africa apart from South Africa).

In addition to the commonly used statistics on the criminal justice system or conduct of elections, these include, for example, data on the gender breakdown of judges; on the average time to bring a case to court; on what percentage of courts have access to a complete set of legislation; on how many laws are passed by a national legislature each year; on the number of civil servants disciplined or prosecuted for corruption offenses; on (under)staffing of technical positions in key ministries. Even if these data are available in-country, in annual reports of ministries for example, the reports are often not online, and the data are not collated and analysed in such a way that allows cross country comparison. AfriMAP is not, however, able to collect exactly the same information in each country, since there does usually have to be at least some basic collection of statistics by a national authority for us to report.

What Quantitative Indicators are Missing that would be Useful for AfriMap?

AfriMAP does not have the human, financial or technical resources to carry out on our own account major national information-collection activities nor detailed opinion poll research. In some areas of interest to us, this has meant that our national partners have had to respond to questions in the questionnaire with the response that information is not available to enable an effective response to the question asked. Areas on which we would like to report and have found that quantitative information is lacking in at least some countries, and which would provide important levers for civil society to use to put pressure on government to improve performance, include:

Statistics about the functioning of the justice system that could be (and in some cases are) easily collected by the courts, such as the percentage of cases in court where litigants or accused don't have legal representation; the number of cases that are adjourned for absence of court officials (judges, registrars, interpreters etc); or the attrition rate from number of reported crimes, through arrests, prosecutions and convictions.

Systematic monitoring of issues related to the right to information and equitable access to broadcast media, beyond the existence of legislation supporting such principles (percentage responses to requests for information; percentage airtime given to different parties, etc).

Improved indicators about compliance with best practices in public financial management (PFM), including clear reports on budget execution figures, the percentage of procurement contracts that are advertised, etc. The World Bank does use its CPIAs to monitor PFM against a range of indicators, and produces scores on this basis that it shares with governments, but these scores have only very recently been made public, still lack transparency in their construction, and are very poorly publicised or understood at national level. It is hard for the non-technical-expert to tell from the scores why and where in particular the country has fallen down; and some of the scores do not seem to be based on fact, but rather opinion. The IMF's Reports on Standards and Codes (ROSCs) also provide useful information but are highly technical and not available for most African countries – and their publication is at the discretion of the government concerned. Moreover, it is not possible for civil society to replicate the assessments since much of the information on which they are based are only accessible to government or institutions such as the World Bank that have privileged access. There is a need for much greater publicity and education about these measures, as well as serious consultation at national level on what should be included and how it should be applied: there is natural suspicion at the ideological content of assessments from the international financial institutions. In this regard, the recently established Public Expenditure and Financial Accountability (PEFA) joint program of a wide range of donors is encouraging, since it provides a framework of greater transparency.

- Indicators along the same lines as those applied to PFM relating to the funding of political parties: including the existence of legislation regulating party finance and its enforcement, requirements of transparency, auditing of party accounts, and the provision of equal access to state-funded benefits if any (including access to the broadcast media, use of state vehicles etc, as well as state funding of parties in itself).

- Similarly, reports on a check list of statistical capacity indicators, ideally carried out by government itself in conjunction with international agencies, would educate national civil society and legislators about the importance of this key government function. Statistical capacity indicators have been developed at international level, but are little known outside the specialist technical agencies, are available for few countries in Africa, and are not presented in a way that would allow engaged citizens to understand what exactly it is that needs to be done to improve the situation.

- In the area of general civil service reform, the information that would be useful is much less technical, but still hard to find. Monitoring government performance in improving public service delivery would be made easier by requirements to collect and publish information such as: current staffing complement for the public service, disaggregated to the individual unit and with information on whether approved positions are vacant; the percentage of civil service positions filled through public advertisement; the pay rates in cash terms for all public employees; the percentage of public employees who are subject to and actually receive performance evaluations; the number of personnel days lost each year through strikes or absenteeism by category of employee, etc.

- In the context of surveys by Afrobarometer or others, a set of questions (in addition to those already included) around the attitude of African citizens to traditional leadership and justice structures and how they should best be integrated into the post-colonial state; and others on the usefulness of different mechanisms for seeking dispute resolution (including courts of various kinds, constitutional bodies such as national human rights institutions, street committees and similar, vigilante groups, police…).

- To take account of the realities of African countries' access to resources, an indication for each government department of the amount requested by the department for its budget, and the funds actually available. (In many cases this is technically available, but not in a commonly reported way.)

There is a need for much more effective monitoring (and policing) from the OECD side of donor compliance with the Paris Declaration on Aid Effectiveness. Notoriously, OECD member states have resisted disaggregation by country of information collected by the Development Cooperation Directorate (DAC) on implementation of the Paris Declaration. The joint document published in 2005 by the UN Economic Commission on Africa and the OECD DAC on Development Effectiveness in Africa is notably fuzzy in the indicators set, and needs further elaboration and monitoring to turn it into a more useful instrument – in particular, clearer standards on untying aid. However, even if the OECD produced regular and detailed reports on the degree to which the existing commitments have been met by each OECD country, it would be a demonstration of seriousness from the donor side.

Conclusion

In general, it seems to us that assessments of institutional performance (for example, the degree to which the judiciary is independent, the strength of legislative oversight of the executive, or the effectiveness of the national audit institution) are extremely important, but not easily reducible to quantitative indicators. Nor is it apparent what attaching a number to an assessment of this type – for example, that parliamentary committees are weak – would achieve.

Governance indicators are used for different purposes and different users will have different priorities. An investor who wants to assess the level of risk in several countries may find that scoring and ranking provides a good basis for decision making. For most development agencies, the professed goal is to make significant contribution to improving governance so it is in this regard that a more in-depth assessment of progress or lack thereof is warranted. Governments – the focus of all this attention and ultimately responsible for implementing the insights these measures are supposed to provide – will have yet more specific requirements to inform their decisions.

Rather than wrestling with the methodological problems of providing numbers for things that cannot in reality be counted, it would make more sense to invest time and resources in ensuring that the things that can usefully be counted are done so according to the best methodologies. And that this information is made public. If quantitative indicators are to be established, they are much more useful when they are based on information that naturally lends itself to numerical reporting, and shows clearly where the deficit is and how it can be improved. If aggregated scores are generated from these indicators, then the disaggregated numbers must also be freely available.

Where quantitative indicators are not appropriate, qualitative analysis will be necessary and can make a much more useful contribution to policy reform. Even where quantitative indicators are completely appropriate, they will need qualitative interpretation and analysis to lead to useable recommendations.

Finally, in regard to all such governance measurement initiatives, there is a great need for much wider outreach and consultation by the relevant institutions on what the appropriate measures are and how to apply them. This is especially true when it comes to areas that are traditionally the realm of 'experts' – especially public financial management, statistics etc. In addition to national governments, this includes, most obviously, parliaments: the role of national legislatures is too often overlooked, both as an institution that should be included in measures of governance quality, but also as a forum for explaining and seeking the input of the citizens' representatives on how their government's performance should be rated. The input of civil society organisations, especially those working on human rights and democracy, will also be needed if the evaluation of government performance is to bite at national level. No governance measurement system is likely to be effective or useful for internal as well as external actors unless it is devised and informed by a process that involves those same actors in its development.

References

AfriMAP documentation, including questionnaires, available at http://www.afrimap.org.

Christian Arndt and Charles Oman, *Uses and Abuses of Governance Indicators*, OECD, 2006.

Development Effectiveness in Africa: Promise and performance: Applying mutual accountability in practice, OECD-DAC and UNECA, 2005.

Governance for sustainable human development: A UNDP policy document, United Nations Development Programme, January 1997.

Chapter 39
National Progress and the Effectiveness of International Aid

Rob Swinkels
Senior Statistician, Ministry of Foreign Affairs, Netherlands

Maarten Brouwer
Director for Effectiveness and Quality, Ministry of Foreign Affairs, Netherlands

Abstract

The author discusses the impact that the Millennium Development Goals (MDGs) have had on development statistics and on the work of development agencies, specifically the Netherlands Development Cooperation Agency. Presented are both the positive effects of the MDGs, as well as the problems, especially the lack of reliable statistics with which to quantify outcomes. The paper outlines what is required to set up better in-country data collection systems that monitor the progress of national, sub-national and sector plans. Included is a discussion of Poverty Reduction Support Plans (PRSPs) and Performance Assessment Frameworks (PAFs) and the Results, Resources and Partnership approach created by the World Bank and OECD/DAC.

Introduction

Chair, ladies and gentlemen thank you for this opportunity at this magnificent forum. In the next ten minutes or so I will discuss my experience working for the Netherlands Development Cooperation Agency and being a user of development statistics. "Does aid work? Can you show me the results? Where are the numbers? Do you have baseline data?" Just a few questions I have been struggling with for the last two and half years without being able to provide real good answers. Still, the questions are justified. I am sure my desparation is felt by many more, also in this audience. I would just like to share a few experiences with you in our attempts to come to better answers.

The introduction of MDG's was really a dream come through. Finally, clear targets, more focus and a set of measures against which performance could be measured. But the dream quickly turned into a nightmare. For us policymakers quantifying outcomes without having the statistics at hand to provide answers, is just short of a nightmare. It is not that I don't like the MDG's, not at all. But parliamentarians, partners, people start asking nasty questions about progress, achievenments, results, valua for money. So the MDG's proved to be a powerful tool of change in the world of development. Numbers

became important, as were results. Avoiding bureaucracy, unnecessary coordination and wasteful overlap entered the agenda with considerable force. These days, donors spend much of their time trying to make aid more effective through harmonisation of donor and country procedures, alignment to country systems and applying (potentially) more effective aid instruments such as budget support. This is very important. But it is also important to remember that these are just a means to an end.

The ultimate goal is to achieve concrete development results. And when I say results, I do not refer to a new plan or a strategy or a policy or even a new road. By results I mean improvements in the living conditions of the population or of specific groups that are yet to reap the fruits of progress in our partner countries. Yet tracking progress along the objectives countries have set for themselves to see whether results are being achieved appears much more difficult than foreseen. As someone from a donor agency, I need to know what progress is being made in the countries we operate in. But I face major difficulties in doing that. Why is that so? What was the powerful change MDG's embodied? The simple answer is that it forced donors and development partners to focus on results. Moving away from inputs, from planning, from arranging throughput and care for outputs, outcomes and even impacts is a hugh effort. But I feel it is rewarding to do so, with all difficulties it entails. We should thank the MDG's just for this change of doing bussiness. But how can we practise what we preach?

Netherlands Experience with Results Measurement in Partner Countries

We recently published our second biennial 'Results in Development' report which I think provided a valuable experience in the measurement of development progress in our partner countries. The report provides an overview of (joint) results achieved by the Netherlands development cooperation efforts. We write this report to inform parliament and to account for the four and a half billion euro we annually spent on official development assistance. We want to assess and demonstrate what difference we have made as a medium sized bilateral donor.

This second edition of the results report uses the same methodological approach as the first edition. This means that we have deliberately refrained from a causality explanation of the results achieved. The causality between Dutch support and outcome/impact results is undermined by problems of multidonor cooperation, time delays and multiactor approaches. In stead we have chosen to identify results in the priority fields of our bilateral assistance programme in each of the coutries we work with closely. Of those results we clarified our ontribution and those of others. Thus we presented impacts/outcomes, outputs and inputs without identiying the precise link between them. This we did two years ago and again two months ago.

But in the last report we added so called *result-chains* for each of the key sectors and thematic areas where we are active in our partner countries. The result chains provide an ex ante analysis of how a set of country and donor inputs could lead to concrete outputs, which, in turn, could lead to number of sector outcomes. In this way the intervention logic per sector or theme was identified and each of the results collected were classified in the links of the results chain. What matters most of course are the outcomes. They show how joint country and donor efforts eventually lead to improved lives of the country's population. Outcomes are usually outside the control of the government, because many different actors are involved in transferring input and outputs into outcomes. Governments can play initiating roles, but intermediate institutions and civil

society structures determine in the end the specific shape those outcomes take. They show for instrance how government outputs, such as social service delivery, are really working for the people.

The way we approached the data collection was to gather information on outputs and outcomes from each of our embassies in our partner countries for each of the sectors and thematic areas in which we operate. Information on outputs was usually based on reporting systems in our partner countries, whereas information on outcomes was taken from international databases, such as those on the MDGs. We also used information gathered through evaluation studies conducted by our own inspectorate agency. The report provides a comprehensive overview of the activities we have helped to undertake, the outputs we have helped to produce and the results we have helped to achieve. It was well received by Dutch parliamentarians and the public. It thus helped to maintain their support for development cooperation by the Netherlands government.

But I have to admit here that developing this report was a real challenge. At the level of our partner countries result chains often do not exist, especially at the sector (or thematic) level. And data collection systems are not well-linked to development plans and are weak. This is especially so in sectors such as health, agriculture, forestry, and thematic areas such as governance and environmental protection. As experts in our embassies rely on country monitoring frameworks and data collection systems, many of them had difficulties to report concrete progress in the sectors or thematic areas they operate in. This maybe worrying to us as a donor agency, but it should be much more worrying for our partner governments. If the facts are unknown, planning becomes rather obscure. Donors and partners share an interest here, to get the facts right. It should be an incentive to cooperate closely.

The Need for a Common Results Framework

Poverty Reduction Support Plans (PRSPs) that many developing countries have developed during the past eight years or so set out a clear framework for achieving results. These plans were accompanied by a monitoring system that would generate information to be used in annual progress reports to donors. And donors would use this to report progress to their own constituencies back home. It would also provide governments with a source of policy learning. Substantial progress has been made, especially in monitoring poverty reduction and other outcomes. But progress has been much less substantial at the sectoral level.

In recent years, donors have increasingly delivered their aid in the form of various types of budget support. Aid is transferred directly to a country's budget, provided the country has made satisfactory progress along a set of earlier agreed policy reform measures and outcome indicators. Such a set of reform measures and progress indicators together constitute a Performance Assessment Framework (PAF). As budget support enables a significant scaling up of aid with only a limited increase in transaction costs for receiving countries and donors, Performance Assessment Frameworks might play an even more central role in future aid discussions.

The idea of linking results, resources and the nature of the development relation between partners has been picked up in the Results, Resources and Partnership approach. This approach was started by World Bank and OECD/DAC, but later on Sweden and oureselves provided new impetus to help and bring the planning, monitoring and results assesments at the core of donor-recipient aid relations. Knowing the numbers on

resources, on required inputs and on results to be achieved can help to deliver much better on MDG's.

But can we use country progress reporting systems as a framework to assess progress? The Paris Agenda requires us to use country systems as much as possible. And we are ready to do so. We rather have a good-enough country progress report that is produced by the government and used for domestic debate, than a perfect glossy report that is used only for discussion by donors. But again can we use the country's domestic progress reporting systems? Given their current weakness we can not. Countries' local progress reports are often judged to be of insufficient quality for donors to base their disbursement decisions on. Donors often have questions about the relevance, quality and timeliness of the monitoring data presented in these progress reports. And information on policy actions undertaken was weak. As a consequence we have had to develop donor oriented Performance Assessment Frameworks in parallel to national performance assessment systems. Of course these frameworks are well linked to targets in PRSP's or other policy frameworks. But the existing data sources are often not linked to providing information that is required to monitor and asses progress in a meaningful manner.

In most cases donors cannot use local performance assessment systems because of the lack of a clear monitoring framework and its link to a reliable data collection system to measure progress. This is even more the case in sector plans which are often the most crucial ones for donor support. While progress has been made in monitoring the PRSP's and national development plans, progress in developing sector plans has remained behind. There are probably three reasons for this:

- Firstly, in our partner countries, there is no tradition and expertise within sector ministries to clearly define objectives and link them to a set of acivity/ input -output-outcome indicators with an indication of how and when these will be measured and what theirn current baseline value is;

- Secondly, donors are often too confused and in disagreement on indicators and benchmarks, on policies and on desired outcomes to provideany meaningful help in designing outcome-based sector plans and accompanying monitoring systems and making them work. The same counts for sub-national plans;

- Thirdly, the statistical data to measure these indicators are often lacking and there is no clear and harmonised donor approach to strengthen country data collection systems and country statistical capacity.

Strengthening Statistical Systems

Available donor support to statistics is often insufficient and provided in a fragmented and often poorly coordinated manner. In Mozambique for example there are more than 20 donors active in supporting statistics, but coordination is minimal. In many developing countries, various joint support mechanism exist for strengthening planning systems and public financial management and these have proven very effective. But donors have not yet been able to work together to jointly support the strengthening of statistical capacity in developing countries. The Paris Declaration on Aid Effectiveness has not been applied to statistics.

In practice, donors help partner countries to provide data which are more often than not of use to themselves. To fill the hunger for data, donors refer to survey data, the quality and collection of which can be controlled. At the same time, the ever so important routine data for regular monitoring and timely adjustment of planning and implementation is often neglected. More collaboration is needed, including division of labour. Structural improvements for routine data systems should be at the core of capacity building in statistics, whereas survey data is an area in which donors could be usefully supportive if the framework in which those data should fit are well established. This requires training, institutional support, technical assistance. But most importantly, it requires a well formulated demand for data, politically backed. For data will reveal the reality, will thus provide strong messages and will require positive responses.

Summing Up

In summary, the MDG's have set the agenda. Objectives have been quantified, indicators developed and expected resuts formulated. Partner countries want to deliver, donor countries want to account for their efforts. Current donor-recipient relations, defined within the Paris framework, depart from mutual accountability, joint learning, evidence based dialogue and management for development results. The Monterrey conference calls for more support, scaling up of efforts. For all this to happen facts and figures are required. Better country systems to monitor progress along objectives as defined in national, sub-national and sector plans will be essential. Such monitoring systems then need to be linked to a demand-led and country-led data producing system. This requires three things: (1) help countries strengthen outcome-based sector plans and thematic monitoring frameworks through longterm advisory services and build trust; (2) donors need to scale up their technical and financial support to strengthening statistical data collection; and (3) support to statistical development needs to be provided in a much more coordinated fashion to make it more effective. Strengthening a well balanced national country statistical strategy should be central in this.

This all is a hugh list of requirements. You can imagine why I referred to MDG's as my nightmare. But they should not be a nightmare, they should be a dream. For me as a manager to reduce the workload of providing this biennual Results Reports. For us as a donor to understand much better the intricacies and difficulties which are the real world of development. For our partners to help planning and managing for results and be proudly accountable for what is achieved. For all of us to become a bit wiser on what works and what doesn't.

Chapter 40
Case Study: Introducing Standard Indicators in European Commission Funded Projects for Aggregation of Aid Effects[i]

Androulla Kaminara
Director, Operations Quality Support, DG EuropeAid, European Commission

Abstract

The purpose of this paper is to present a pilot phase project which aims to gradually introduce a selective list of standard indicators to improve the accountability of European Commission (EC) financed external development cooperation operations. This aims to illustrate how the use of indicators can be promoted and how the risk of misuse can be avoided.

The paper presents the objective of the initiative, the chosen approach, the selection criteria for the indicators, the method of introducing the indicators in the project cycle and the evaluation that will follow after one year.

The potential risks linked to the approach are highlighted and possible mitigating actions are suggested. Finally lessons learnt from the exercise are presented.

Objective

The objective of this paper is threefold:

1. To give an example of the process of promoting more systematic use of indicators for external aid projects and programmes financed by the EC. This is necessary in order to allow the assessment of the quantifiable effects of the EC's operations in order to improve accountability and visibility of external aid operations.
2. To highlight the risk of the misuse of these indicators and to report on the discussions that have taken place internally in the EC as well as with the Member States and other donors.
3. Provide some lessons learnt for others that have embarked on a similar process.

i Donors, partner countries and others that would like to contribute or get more information on this issue should send an e-mail to: EuropeAid-E-INDICATORS@ec.europa.eu

The Need of Standard Indicators for Accountability

The EC is accountable to European citizens for its operations of external aid. The EC does this (like other donors) by providing detailed annual financial accounts of commitments and payments, representing inputs, in the field of external aid. However, for years the EC has had difficulties in providing, in an easy and systematic way, quantified data on the results of interventions (such as, for instance: the number of newborn babies vaccinated, the percentage decrease in the repetition rate in schools in a region).

This difficulty in establishing global assessments of results is essentially due to the fact that each project or programme keeps only project specific data in a non-standardised way. For each project or programme supported by the European Commission, there exists a large variety of indicators. These indicators are used to ensure follow-up and evaluation and are aligned to the reform process in the partner country. They do not, however, allow for the consolidation of results across countries or regions because they are not easily comparable and they rarely allow for the aggregation of data.

In order to improve accountability for EC financed aid operations, we are preparing to gradually introduce a mechanism envisaging the systematic use of a selective set of standard indicators for all project/programmes in a particular sector. These would allow for the aggregation of data from different programmes or projects in order to obtain a quantified general overview of the effects of aid in a given sector, country or region, and to appraise trends over a given period of time. It should be underlined that this aim is separate from the aim of assessing progress at individual country level or the individual project level.

General Context

The concept for accountability reporting has been developed in a general context characterised by:

- The increase of European Union (EU) & EC development aid funds;
- The commitment under the Paris Declaration to enhance donors' and partner countries' respective accountability to their citizens and parliaments for their development policies, strategies and performance,[1]
- Under the European Consensus on Development: the importance of progress indicators and regular evaluation of assistance to better focus EU assistance[2;]
- Under Development Cooperation Instrument (DCI) regulation; the requirement of annual reports detailing the implementation and results and, as far as possible, the main outcomes and impacts of the assistance[3];
- The new possibilities of co-financing between Member States and the Commission under the revised Financial Regulation.[4]

Our Approach

Our aim was to go beyond mere input (financial data: commitments, payments) in the reporting on EC financed development cooperation action. To achieve this, we analysed a number of alternative options since the issue is complex and controversial:

1. Drop the idea of having this type of indicators because it is too complex
2. Wait until there is a donor consensus on the topic
3. Adopt a "Learning by doing" approach through a pilot phase that focuses on a limited number of sectors that are found to be most developed in this respect and, at the same time, launch a dialogue on the issue with the Member States and other donors.

The last pragmatic option was chosen and we have also decided to have an evaluation after one year in operation.

Selection Criteria

In order to minimise the creation of any additional burden on the partner countries, in line with the Paris Declaration, and to minimise the risk of delaying the preparation of programmes, the selective lists of standard indicators per sector were selected on the basis of the following criteria:

- Limited number of indicators/sector;
- Relevance with respect to the development objective;
- Practices of the key donors in the concerned sectors;
- Consistency with the indicators used for programming and with the mdg indicators (millennium development goals)[5];
- Allowing aggregation;
- Possibility of calculating the values of the indicators with and without the project;
- Mandatory consideration of the indicators for all projects in a given sector;
- Easily available data sources for calculating their values.

A Pilot Phase and a Progressive Approach

This system will be introduced progressively with a step-by-step approach. In the inception phase, it will only concern those sectors for which indicators meeting the criteria have been identified and defined:

- Human Development (Health, Education and Vocational training);
- Good governance (Civil Society, Decentralisation and Elections);
- Security (Fight against Organised Crime, Police Reform and Antipersonnel Mines);

- Migration and Asylum including Trafficking in Human Beings;
- Food Security, Rural Development and Environment;
- Infrastructure (Transport, Water, Urban Development, Energy and ICT).

Each selected indicator has one or more associated DAC code.[6] For all new projects with DAC codes that have an associated indicator, it is mandatory to take this indicator into account during project/programme preparation.

The introduction of indicators concerns the identification and formulation phases of the project cycle.

Starting from the envisaged launch date of the operation on 1st September 2007, the identification of an eligible project will have to include a list of the indicators to be taken into consideration for the subsequent formulation phase.

For these projects, the formulation phase will have to systematically provide the values with and without the project.

The collected data will therefore initially cover only the expected effects of projects (*ex-ante* level, at the time of the decision). The introduction, collection and use of the data measuring the real effects actually obtained after implementation of the projects (*ex-post* level) will have to be subject to further analysis.

The action will be reviewed annually on the basis of the acquired experience and the exchanges of information with other donors - in particular interested Member States. Such an evaluation will particularly look into the following issues:

- Additional sectors to include;
- The list and the definition of the indicators (corrections, removals, additions);
- The procedures of introduction, data collection and of analysis of the results.

The EC attaches great importance to the transparency of this operation for the Member States and other donors. In this respect, the EC is ready to play a catalyst role as it is clear that it is important to have the maximum number of summary indicators common to bilateral and multilateral agencies, which would help us all to cope with the demand for information regarding results of aid provided. As a consequence, we have had a number of discussions with *e.g.* the World Bank, United Nations, the Member States and their development agencies.

Some Examples of Indicators

Here are some examples of the indicators envisaged for the pilot phase. Where relevant, the indicators will be disaggregated by gender:

Road Transport

- Rural access;
- Percentage of road network in good or fair condition;

- Number of Km maintained, rehabilitated, constructed;
- Water and Sanitation;
- % of Population with access to improved water source;
- % of Population with access to improved sanitation;
- Health;
- Proportion of 1-year-old-children immunised against measles;
- Proportion of births attended by skilled health personnel;
- Tuberculosis case detection rate;
- Tuberculosis cure rate;
- Education;
- Net enrolment ratio in primary schools;
- Gender parity index in primary, secondary and tertiary education;
- Elections;
- Percentage of errors (omissions or incorrect information) in voter registry.

Risks of Misuse of These Indicators

The initiative to introduce the use of standard indicators has been discussed internally and externally. There are a number of risks that have been highlighted in these discussions.

Firstly, the indicators are likely to give an incomplete picture of reality. In order to determine the effect of the assistance from the EC, only some standard indicators were retained: on average the list includes 4 indicators per sector. Thus, for certain topics or fields of intervention the selected indicators will obviously not cover all the facets of the various projects/programmes and will reflect only in a very partial manner the contribution of the EC. This limitation is inherent to any attempt aimed at establishing a global synthesis or general tendencies. In this context, it is recommended that each report of quantified data is accompanied by text clarifying the range of the results obtained.

In addition, it must be clear that this exercise, aimed at meeting the needs of EC reporting, is not intended to replace the strategies developed by the partner countries, with the support of the donor community. Nor is the exercise aimed at measuring in its totality, the result and the impact of the policies of reform which the partner countries carry out with the support of external financing on a program, sector or country scale.

It will also be important to take the country/regional context into account, especially for fragile states. Indicators will and should move in a different direction in different contexts. For example, in a post-conflict context, one might expect to see an increase in

the role of the police for internal security functions but once the situation stabilises this role could be decreased. Hence, each indicator should be complemented by definitions and explanations of their use, scope and ideal value.

The selected indicators are mainly at output and outcome level. Output indicators generally are not able to show the effects of the interventions for the target groups while with outcome indicators, it is difficult to attribute the effects of the assistance to the specific projects or programmes. These limitations of the different types of indicators were often presented as a main problem in the preparatory discussions. Therefore, when using the indicators it is important to highlight these intrinsic risks.

Secondly, there is a potential risk of distorting the priorities of the government in the partner country. One of the main concerns expressed by Members States is the risk that the existence of the set of standard indicators would potentially distort country priorities and encourage the development of programmes to fit the indicators. As indicated above, the pilot phase includes only a limited number of sectors and an evaluation has been planned to assess the impact of the indicator set with an examination of the benefits, costs and risks. This is particularly important since the programming documents, Country Strategy Papers and National Indicative Programmes are developed in line with the partner countries' strategies. The indicators should not in any way affect this process.

Thirdly, there is the risk of obliging the partner countries to set up new monitoring systems. The EC believes the majority of selected indicators to be standard, even basic. It is planned to introduce an evaluation in the identification phase of the compatibility of the standard indicators with partner countries' monitoring system and statistical data collection.

Lessons Learnt

From the process of developing the standard indicators a number of lessons can be learnt:

- Even though people agree that there is a need for standard indicators, there still is resistance to change and scepticism because of the complexities of the subject and the inherent inability to find "perfect" solutions;

- It is important to never underestimate the time it takes to convince people internally and externally of the benefits of introducing a new scheme;

- Prior to the start of such an endeavour it is important to carry out a full analysis of the needs and implications and to ensure that the selection of indicators is the result of an open debate with all concerned;

- Overlaps and complementarities with other data collection instruments need to be considered;

- It is too complex to solve everything before the beginning of the pilot phase. As a result there is a need to have an evolving approach. The evaluation of the pilot phase will be able to highlight issues and any negative effects.

Notes

1 Paris Declaration (2005), §3, iii: http://www.oecd.org/dataoecd/11/41/34428351.pdf

2 European Consensus on Development §15, Official Journal of the European Union 2006/C 46/01: http://eur-lex.europa.eu/LexUriServ/site/en/oj/2006/c_046/c_04620060224en00010019.pdf

3 Regulation (EC) No 1905/2006 of the European Parliament and of the Council of 18 December 2006 establishing a financing instrument for development cooperation, §34, Official Journal of the European Union 2006/L 378/41: http://eur-lex.europa.eu/LexUriServ/site/en/oj/2006/l_378/l_37820061227en00410071.pdf

4 Revised Council Regulation 1995/2006 of 13 December 2006, specifically new article 53: Official Journal of the European Union 2006/L 390/1: http://eur-lex.europa.eu/LexUriServ/site/en/oj/2006/l_390/l_39020061230en00010026.pdf
The EC will now have the possibility to manage funds from the Member States, third countries and International Organisations

5 MDG indicators: http://mdgs.un.org/unsd/mdg/Host.aspx?Content=Indicators/OfficialList.htm

6 Explanation and list of DAC codes:
http://www.oecd.org/document/21/0,2340,en_2649_33721_1914325_1_1_1_1,00.html

Corrigendum 1

Chapter 20
Policy and Statistical Issues Underpinning Financial Stability: The IMF Perspective

Robert W. Edwards[i]
Director, Statistics Department, International Monetary Fund

Notice of correction

This chapter replaces Chapter 20, pages 317-326 of
OECD (2008) *Statistics, Knowledge and Policy 2007: Measuring and Fostering the Progress of Societies*, OECD, Paris

Abstract

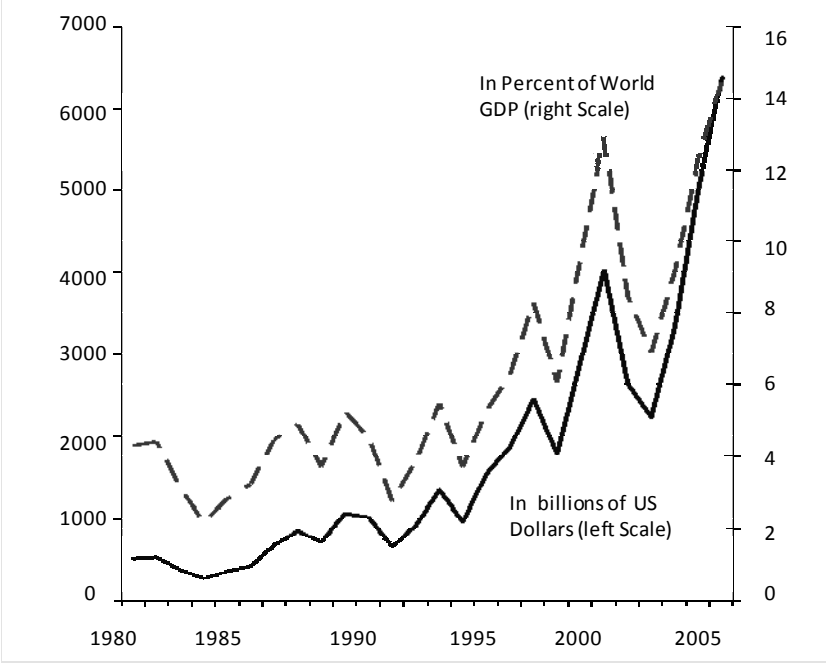
In the past decade, financial globalisation has led to an unprecedented deepening in financial markets and internationalisation in capital allocation. While it has brought benefits, financial globalisation has also brought new risks and challenges for policy makers, particularly through its linking of national economies into a vast network of closely interconnected balance sheets, and consequent potential for severe cross-border spillover effects. In this environment, financial instability in one country may be transmitted to other countries or affect developments in regional or global markets. To meet the new challenges, the IMF has been sharpening the focus of its financial surveillance. While continuing to emphasise its financial sector assessment programme (a joint effort with the World Bank), the IMF has been further integrating finance into its Article IV surveillance; strengthening its global surveillance through multilateral discussions on common issues of systemic importance; and expanding its framework for vulnerability assessments to include the analysis of balance sheet vulnerabilities and use of new data tools such as financial soundness indicators as well as market-based, forward-looking indicators. In the statistical area, the IMF has launched several initiatives in order to gather timely, comprehensive, and internationally comparable data to help meet the new data needs of financial surveillance, better track cross-border flows and positions, and support rigorous, diagnostic, analytical work. In meeting its challenges, the IMF will continue to leverage the use of its comparative advantage - its universal reach and expertise on macroeconomic and financial issues and on statistics.

i The author wishes to thank the following for their inputs into the drafting of this paper: Marina Moretti and Andrea Maechler of the IMF's Monetary and Capital Markets Department, and Armida San Jose and Ewe-Ghee Lim of the IMF's Statistics Department. The views expressed herein are those of the author and should not be attributed to the IMF, its Executive Board, or its management.

Financial stability in a Globalised World

The last decade has witnessed an unprecedented deepening in financial markets and internationalisation in capital allocation. The rise to $6.4 trillion in total cross-border inflows has been impressive – it now represents over 14% of world GDP (Figure 1). These flows have also been associated with the globalisation of financial institutions. Cross-border financial sector mergers and acquisitions, mostly concentrated in the banking sector, are now around 40% of all financial sector acquisitions, compared to less than one% a decade ago. The increase in foreign bank ownership has been particularly rapid in Eastern Europe and Latin America

Figure 20.1 Total Cross-Border flows, 1980-2005
(in % of world GDP and in billions of U.S. dollars)



Source: IMF staff calculations based on International Financial Statistics and World Economic Outlook.

The globalisation of financial institutions has complex implications for financial stability.[1] From the perspective of individual institutions, globalisation helps diversify risks and may well have improved financial stability, particularly in the face of relatively small shocks. But as national economies become part of a vast network of balance sheets, often closely interconnected through the financial sector, severe crises in the future may be more easily transmitted across borders and therefore may have become more broad-ranging and difficult to deal with. For instance, financial systems with substantial foreign bank presence may be more resilient to traditional domestic banking crises, but more vulnerable to foreign shocks that affect parent banks. More broadly, financial instability in one country may easily spill over to other countries or affect developments in regional or global markets, as was the case, for example, during the Asian and Russian crises in the late 1990s.

The global community is thus presented with new challenges. To maximise the benefits of globalisation while containing the potential risks, efforts need to continue to reduce the likelihood of systemic crises and limit their severity if they occur. This requires measures conducive to deep, broad, and resilient financial systems, which can support an optimal allocation of resources, an effective monetary policy, and a sound fiscal stance. In addition, policy makers need to ensure effectively coordinated oversight of internationalised financial institutions and to put in place cross-border crisis management and resolution arrangements that are sufficiently robust to handle a severe shock and minimise spillovers. Progress is being made in these areas, but more needs to be done, in a wider range of countries.

The International Monetary Fund (IMF) has an important role to play in helping countries meet the challenges of financial globalisation. Interactions between economies and financial sectors are becoming increasingly important, and addressing financial sector vulnerabilities and risks increasingly needs a global approach. Through the continued use of its comparative advantage – that is, its universal reach and expertise on macroeconomic and financial issues, the Fund can further enhance its monitoring of the global economy and help individual countries strengthen the resilience of their economies. In addition, it can use the same comparative advantage in the statistical area to help provide timely, comprehensive, and internationally comparable data to meet the new data needs of financial surveillance, better track cross-border flows and positions, and support rigorous, diagnostic, and analytical work.

The IMF Financial Surveillance Framework

The IMF is sharpening the focus of its financial surveillance to remain in step with a rapidly changing world. Financial surveillance is ultimately rooted in concerns about financial crises that are changing in nature. The late 1990s crises in emerging markets were largely financial account crises caused by abrupt shifts in global asset allocations (that is, crises that affect stocks or balance-sheets) rather than budgetary or current account crises caused by fiscal or terms-of-trade shocks (that is, crises affecting flows). What distinguishes these events from past episodes is the speed and amplitude at which they propagate, both domestically and cross-border. Efforts at preventing these phenomena require new directions in financial surveillance that recognise the complex networks of interconnected balance sheets and emphasise the spillovers of financial market disturbances, both between countries and between real and financial sectors.

New Directions in Financial Surveillance

Given the complexity of balance sheet linkages, there is no single, widely-accepted methodology for assessing financial sector stability. Financial stability assessments need to cover a potentially wide range of topics – depending on country circumstances – and to take a holistic view of the financial system. For this, financial surveillance needs to address two key questions:

- What are the critical channels of interaction between the macroeconomy, financial markets, and the financial institutions?

- What is the role of the financial sector per se in initiating, amplifying, or muting disturbances in the economy and in transmitting such effects internationally?

In addressing the first question, it is important to highlight that the linkage between financial institutions and markets, on the one hand, and macroeconomic performance and policies, on the other, runs in both directions. Financial soundness affects macroeconomic performance and policies, and macroeconomic performance and policies have consequences for the financial sector. In order to understand this two-way relationship, better use must be made of a broader set of indicators, including data on financial flows, the balance sheets of the financial and other sectors, and market-based price data. A great deal of progress has been made in this area in the last decade but more needs to be done.

In addressing the second question, it is important to understand the role of the financial sector as a risk transmission mechanism. Potential **risks** to the financial system – that is, events that might trigger a crisis – must be distinguished from the underlying **vulnerabilities** that will determine the impact of such events should they occur. Vulnerabilities in the financial sector are not merely exposures to risk, including balance sheet vulnerabilities across the various sectors, but also the system's ability to absorb or withstand a shock (*i.e.* its resilience). To assess those risks and vulnerabilities, possible channels of contagion also need to be analysed explicitly, across borders, and in both directions.

Integrating Finance into the IMF's Work

Continued Emphasis on the FSAP

The Financial Sector Assessment Program (FSAP), a joint effort by the IMF and the World Bank launched in 1999, is a key element of the Fund's engagement on financial issues with its member countries. Supported by experts from a range of national agencies and standard-setting bodies, work under the programme seeks to identify the strengths and vulnerabilities of a country's financial system; to determine how key sources of risk are being managed; to ascertain the sector's developmental and technical assistance needs; and to help prioritise policy responses. The programme is designed to assess the stability of financial systems as a whole, rather than individual institutions, and to emphasise prevention and mitigation rather than crisis resolution. As part of the process, the FSAP provides assessments of observance of various internationally-accepted financial sector standards and codes, set within the broader institutional and macroprudential context.

With a total of 120 initial assessments under the FSAP now completed or underway, and an additional 20 or so in the pipeline, the programme is increasingly moving into a phase of FSAP updates (27 updates have been completed or are underway, and 30 or so more are being planned). FSAP updates provide an opportunity to refresh the initial assessment, albeit with possible differences in scope. With the initial assessment as a reference point, updates can choose focus over comprehensiveness and concentrate on key stability and development issues.

Coverage of Financial Issues in Article IV Surveillance

Annual consultations with member countries under Article IV of the IMF's Articles of Agreement provide another vehicle for financial surveillance. The coverage of financial sector issues and balance sheet vulnerabilities in Article IV reports is being enhanced with a view to better reflect the immediacy and extent to which financial

market developments can affect people's lives. This would strengthen the overall effectiveness of Article IV surveillance, in addition to ensuring continuity in the coverage of financial sector issues in-between FSAP assessments, which occur at relatively low periodicity (on average, every five to six years).

Efforts are also underway to integrate more fully financial analysis into the traditional macroeconomic framework examined as part of Article IV work. This requires more attention to macro-financial links, better use of available data and information – both quantitative and qualitative – and a strengthened focus on spillovers and cross-country links. A key challenge in this area is the absence of a clear intellectual framework for integrating financial sector analysis, much of which is microeconomic, with macroeconomic analysis. Incorporating these two different dimensions of surveillance is one of the most compelling aspects of the Fund's work.

A More Global Perspective

Integrating multilateral and regional aspects in financial surveillance represents a further challenge. Strengthening global surveillance can be achieved by facilitating discussions within groups of countries on common issues of systemic importance and strengthening the cross-country analysis of macroeconomic and financial risks and their interactions, as well as formulating regional work plans that focus on particular policy issues facing a region.

Some work on cross-border financial sector issues has already been undertaken and more is in train. Regional FSAPs are being conducted for currency unions, and are particularly appropriate where significant regulatory and supervisory structures are at the regional level. The IMF is also engaging in financial sector regional projects that examine special issues relevant to a particular region or group of countries – recent examples include studies on Central America, the Maghreb, and the Nordic-Baltic region. Efforts are equally underway to incorporate in the FSAP elements that transcend the borders of the country under consideration – such as heavy dependence on cross-border financing or the challenges of home-host supervision.

Analytical Tools and Indicators for Financial Surveillance

Surveillance must strive to extract diagnostically useful information about risks and vulnerabilities from all of the available financial data. This requires an in-depth understanding of existing data sources and their appropriate use through a wide range of financial analysis tools.

Use of Indicators in Surveillance

The ability to conduct financial surveillance presupposes the existence of indicators that can be used as a basis for analysing the health and stability of the financial system. These indicators comprise financial soundness indicators (FSIs) – that is, aggregated data on individual banking institutions and their non-bank clients, and indicators that are representative of the markets in which these institutions operate – as well as market-based data drawn from price and volatility measures of various capital market instruments.

The monitoring of FSIs is a key element of surveillance. They provide a measure of banks and non-banks' exposure to different types of risk and their capacity to handle

shocks (their resilience) that affect solvency or liquidity. FSIs consist mainly of aggregate balance sheet or income statement measures and their concept originates from prudential and commercial measurement frameworks that were developed to monitor individual entities. This type of aggregation of individual institution-level indicators (microprudential indicators) into financial soundness indicators (macroprudential indicators) necessarily involves a loss of information because the distribution of prudential indicators of individual institutions is also a crucial dimension of financial stability. Although aggregation is required for facilitating macroprudential analysis and international comparison, the stability assessments are strengthened by allowing some disaggregation through peer groups or the monitoring of the distributional characteristics of various indicators. In addition, FSIs themselves are concurrent indicators of financial soundness, available often with a lag. Therefore, proper interpretation of FSIs requires a range of forward-looking analytical tools, notably stress testing of individual institutions as well as analyses of the determinants of FSIs and forecasts of their future course.

FSIs can be complemented by market-based indicators, which are forward-looking indicators of soundness and are available with higher frequency. Asset prices, for example, reflect market perceptions and can influence economic developments. In a world of highly integrated financial markets, expectations themselves are subject to a slew of heterogeneous influences. Information from high-frequency financial data (about expectations, arbitrage incentives volatility, risk premia, and the like) is useful to identify policy inconsistencies and warning signals. FSIs are part of the broader framework for vulnerability assessment, which includes a wide range of analytical tools, including scenario analysis, different methods of debt sustainability analysis, and a variety of balance-sheet-type methods of risk analysis.

The Balance Sheet Approach (BSA)

The analysis of vulnerabilities across various balance sheets and sectors is an increasingly important part of a country's risk assessment. Balance sheet weaknesses in the public and private non-financial sectors are now recognised as a key factor that exacerbated the financial account crises of the last dozen years.

The Balance Sheet Approach (BSA) is based on stock variables in countries' sectoral balance sheets (assets and liabilities of financial firms, nonfinancial firms, households, government, and their sub-sectors, as appropriate) and the consolidated aggregate balance sheet (for the country). The balance sheet analysis focuses on the determinants and evolution of stocks of assets and liabilities, and the likely shocks to the stock variables, both of which can trigger large adjustments in flows (including cross-border capital flows, shifts in holdings of domestic or foreign currency assets, etc.). An approach of this type can, therefore, be a useful complement to the traditional flow analysis that is based on data related to fiscal, balance of payments, and financial programming.

Where markets are deep enough to provide relatively reliable information, the Balance-Sheet Risk Approach can also be used to derive a set of risk indicators that can serve as barometers of financial sector vulnerability. This approach combines balance sheet data with high frequency financial market prices to impute the market value and volatility of assets, both of which are needed to understand changes in the overall level of risk facing each sector but are not directly observable. An important advantage is the ability to translate continuously adjusting financial market price information into current market value estimates of asset values, which is particularly important given the speed with which economic conditions change relative to the time span between releases of

consolidated accounting balance-sheet information. This framework allows to compute a forward-looking measure of credit quality (probability of default) and could be used as a metric for comparing the risk profiles of a sector/country under different policy assumptions.

Statistical Initiatives in Support of Financial Surveillance

Collecting and presenting high quality statistical data to support the analytical work needed for financial surveillance is a complementary challenge for the IMF. Indeed, it has been noted that the surprise element at the onset of a financial crisis is often the lack of high quality, comprehensive data. The Asian crisis, for example, revealed major gaps in statistical coverage of the domestic financial sector and the external sector that permitted serious vulnerabilities to remain undetected.

In general, two types of data are needed for effective financial surveillance. One, as already noted, is high frequency market-based price data, which are forward-looking and needed to assess changes in market views and expectations, and impending changes in vulnerabilities. The other is timely and internationally comparable economic and financial data. While market-based data are available from commercial vendors, the IMF has in recent years undertaken several initiatives aimed at meeting the economic and financial data needs stemming from these various new features of the financially globalised world:

- The intensification of cross-border financial flows and positions, greater potential for spillover effects, and need for a global, multilateral perspective on surveillance;
- The lack of good quality cross-country comparable financial soundness indicators; and
- The balance sheet nature of the major crises of the past dozen years.

Data on Cross-Border Financial Flows and Positions

To meet the data needs of global, multilateral surveillance, the IMF in recent years launched five major statistical initiatives to improve as well as increase the available data on the cross-border linkages among economies: the Coordinated Portfolio Investment Survey; data for the International Investment Position; the Data Template on International Reserves and Foreign Currency Liquidity; dissemination of quarterly data on the currency composition of official foreign exchange reserves; and the Joint External Debt Hub. In addition, it plans to launch a Coordinated Direct Investment Survey (CDIS).

The first four initiatives and the planned CDIS all involve external sector data which were deemed by the IMF's Executive Board in 2002 as important data for the assessment of an economy's vulnerabilities. The fifth initiative involves creating a platform that brings different databases together for worldwide dissemination.

External Sector Data – Portfolio Investment, IIP, and Direct Investment

In 1997, the IMF launched the first **Coordinated Portfolio Investment Survey (CPIS)**, in response to global asymmetries in reported balance of payments data, particularly those in portfolio investment flows of equities and debt securities. Starting

with 29 economies, the CPIS's distinguishing feature is that it provides data by partner countries by requiring all participants to provide a breakdown of their stock of portfolio investment assets by the country of residency of the nonresident issuer. This feature allows the derivation of a country's foreign portfolio investment liabilities from creditor sources, facilitating cross-checking of data and improving data quality. From 2001, the IMF began to undertake the survey on an annual basis and presently, about 70 countries participate in the survey.

After 1998, the IMF placed a greater emphasis on collecting data on countries' **International Investment Position (IIP)**. The IIP is important for vulnerability assessment because it provides a balance sheet snapshot of the levels, sectoral distribution, and maturity of a country's external liabilities (*e.g.*, external debt), and the size and composition of its external claims (*e.g.*, banks' foreign claims), that may be available to meet its external obligations. The IIP also complements the IMF's existing collection of balance of payments data. In 2001, the IMF Board further promoted IIP data by including it as a prescribed category of the IMF's Special Data Dissemination Standard (SDDS). As a result, the number of countries reporting IIP data increased from 37 in 1998 to around 110 countries presently, albeit with varying degrees of component detail.

Plans are underway to launch a **Coordinated Direct Investment Survey (CDIS)** in order to improve understanding of the pattern and direction of foreign direct investments worldwide. The CDIS, which will be modeled after the CPIS in terms of collecting partner data, has a targeted reference date of end-2009. The survey will be a collaborative effort with several international partners, namely, the European Central Bank, Eurostat, the Bank for International Settlements (BIS), the OECD, UNCTAD, and the World Bank.

External Sector Data – International Reserves

Two of the initiatives involved data on international reserves, another critical input for vulnerability assessments. In 1999, the IMF launched the **Data Template on International Reserves and Foreign Currency Liquidity (Reserve Template**) with the objective of improving the assessment of a country's official foreign currency liquidity position. The Reserve Template disseminates data on international reserve assets and potential short-term foreign currency obligations (and claims) including on off-balance sheet activities (such as those arising from derivative operations). This initiative inter alia helps to rectify the problem indicated earlier regarding the lack of data on off-balance sheet derivative activities during the Asian crisis. All SDDS subscribers are required to provide data for the Reserve Template.

In December 2005, the IMF published for the first time quarterly data on the **currency composition of official foreign exchange reserves (COFER)** on its website. COFER data, which distinguish reserves denominated in U.S. dollars, euros, pounds sterling, Japanese yen, Swiss francs, and other currencies had previously been published only on an annual basis in the IMF Annual Report. In response to heightened public interest, the IMF decided to make the data publicly available on a quarterly basis. Presently, COFER data are reported on a voluntary basis by 119 member countries of the IMF, comprising all 24 industrial countries and 95 out of the 160 developing countries.

External Sector Data – common Platform for Dissemination

The fifth external sector data initiative involved the creation of a platform for worldwide dissemination of different databases on external debt. In May 2006, the IMF, BIS, OECD, and the World Bank jointly launched the Joint External Debt Hub (JEDH) to bring together the data that they each compile on the external debt of countries. This database complements external debt statistics based on national sources, filling important coverage gaps, particularly in the area of private sector external liabilities. It disseminates data provided by over 60 SDDS subscribers, for whom it is a reporting requirement; data from creditor and market sources; and comprehensive metadata. Altogether, the database facilitates cross-country comparisons of external debt flows and positions for 214 economies.

Financial Soundness Indicators

To encourage compilation of internationally comparable FSIs for financial sector surveillance, the IMF undertook the following initiatives.

Table 20.1 Financial Soundness Indicators: The Core and Encouraged Sets

Core Set	
Deposit takers	
Capital adequacy	Regulatory capital to risk-weighted assets
	Regulatory Tier 1 capital to risk-weighted assets
	Nonperforming loans net of provisions to capital
Asset quality	Nonperforming loans to total gross loans
	Sectoral distribution of loans to total loans
Earnings and profitability	Return on assets
	Return on equity
	Interest margin to gross income
	Noninterest expenses to gross income
Liquidity	Liquid assets to total assets (liquid asset ratio)
	Liquid assets to short-term liabilities
Sensitivity to market risk	Net open position in foreign exchange to capital
Encouraged Set	
Deposit takers	Capital to assets
	Large exposures to capital
	Geographical distribution of loans to total loans
	Gross asset position in financial derivatives to capital
	Gross liability position in financial derivatives to capital
	Trading income to total income
	Personnel expenses to noninterest expenses
	Spread between reference lending and deposit rates
	Spread between highest and lowest interbank rate
	Customer deposits to total (noninterbank) loans
	Foreign-currency-denominated loans to total loans
	Foreign-currency-denominated liabilities to total liabilities
	Net open position in equities to capital
Other financial corporations	Assets to total financial system assets
	Assets to gross domestic product (GDP)
Nonfinancial corporations sector	Total debt to equity
	Return on equity
	Earnings to interest and principal expenses
	Net foreign exchange exposure to equity
	Number of applications for protection from creditors
Households	Household debt to GDP
	Household debt service and principal payments to income
Market liquidity	Average bid-ask spread in the securities market 1/

Real estate markets	Average daily turnover ratio in the securities market 1/ Real estate prices Residential real estate loans to total loans Commercial real estate loans to total loans

1/ Or in other markets that are most relevant to bank liquidity, such as foreign exchange markets.

After several rounds of extensive consultation with experts from international agencies, standard setting bodies, and member countries, the IMF completed and published the Compilation Guide: Financial Soundness Indicators (Guide),[2] which provides guidance on the concepts, definitions, sources, and techniques used in compiling FSIs so as to encourage compilation of internationally comparable FSIs. The Guide provides guidance for compiling the 12 core and 27 encouraged FSIs (Table 20.1). To pilot the implementation of the Guide, the IMF in 2004 launched a Coordinated Compilation Exercise (CCE). In particular, the objective of the CCE is to develop member countries' capacities to compile FSIs; promote cross-country comparability of FSIs; and disseminate methodologically sound FSIs. Sixty two countries participated in the CCE, whose terms of reference required the compilation of the 12 core FSIs while encouraging compilation of as many of the encouraged FSIs as possible. As a result of the CCE, metadata and end-2005 FSI data for 57 of the participating countries have been posted on the IMF's website. Posting for the remaining countries is expected in the coming months.

Relevant international and regional organizations as well as compilers/coordinators from countries participating in the CCE met on May 29–31, 2007 to discuss their views and experience in compiling FSIs. A report to the IMF's Board later in 2007 will draw lessons from the CCE and make proposals for future IMF work on FSIs.

Meeting the data needs of the Balance Sheet Approach (BSA)

The BSA is a data intensive framework that requires significant amounts of stock/position data on economic sectors, and on currency denomination and maturities of sectoral assets and liabilities.

The development of the BSA has advanced with the introduction in 2005 of the standardized report forms (SRFs) for reporting monetary and financial statistics. The SRFs are based on sectoral balance sheets for the central bank, other depository corporations, and other financial corporations (as defined in the *1993 SNA*) and provide breakdowns by sectors and by domestic and foreign currency for both domestic and external assets and liabilities. SRF data can provide most of the information needed for the intersectoral framework of the BSA.

The external sector initiatives discussed above provide the other major source of data for the BSA – particularly CPIS, data in the IIP, and JEDH. These databases can be used to identify positions of the domestic sectors with the nonresident sector, often with breakdowns by maturity and instruments, and in some instances, currency.

The IMF recently studied the feasibility of organizing the available monetary and external sector data in the form of BSA matrices. It found that construction of BSA matrices was feasible for countries reporting SRF monetary data and data for the CPIS and the IIP. However, given differences in frequency and timeliness, monthly updates of BSA matrices must rely exclusively on SRF data at this stage. The work on the BSA

matrices is ongoing. Presently, about 80 countries report SRF data to the IMF, which are published in the *IFS Supplement on Monetary and Financial Statistics*.

Conclusion

Looking ahead, much work remains to be done. On the policy and analytical front, efforts need to continue to enhance the analytical framework of financial surveillance with a view to providing timely insights about the financial risks and vulnerabilities that individual countries face in a global economy. For this, more attention must be paid to macro-financial linkages, better use must be made of available data and information – both quantitative and qualitative – and there needs to be more focus on spillovers and cross-country links.

In the statistical area, wider production and reporting of more timely data is essential to further elucidate the patterns of global financial flows and understanding of the direction of direct investment worldwide. Continued efforts are also needed to expand the coverage of monetary statistics essential for the analysis of balance sheet linkages, particularly in the area of non-bank financial institutions and the non-financial private sector. Possible directions of future work on FSIs might include the regular collection of time series data as well as an expansion of the set of indicators to include additional FSIs for the non-bank sector.

Notes

1 For a detailed discussion on financial stability implications of globalisation of financial institutions, see Chapter 3 of the 2007 Global Financial Stability Report (Washington, International Monetary Fund).

2 The Guide was released in electronic format in 2004 and issued as an official IMF publication in 2006.

Corrigendum 2

Chapter 21
International Integration and Societal Progress: A Critical Review of Globalisation Indicators

Philippe de Lombaerde[i]
Research Fellow, United Nations University (UNU-CRIS), Belgium

P. Lelio Iapadre[ii]
Associate professor of international economics, University of L'Aquila, Italy
Professorial lecturer of international economics, Johns Hopkins University, SAIS Bologna Center, Italy
Associate Research Fellow, United Nations University (UNU-CRIS), Belgium

Notice of correction

This chapter replaces Chapter 21, pages 327-331 of
OECD (2008) *Statistics, Knowledge and Policy 2007: Measuring and Fostering the Progress of Societies*, OECD, Paris

Abstract

This paper presents a critical survey of different proposals to measure globalisation, from the perspective of a more general view of the relationship between international integration and societal progress.

A higher degree of international integration can be seen in itself as an indicator of societal progress, inasmuch as it reveals that human societies more and more acknowledge their common destiny. In addition, international integration fosters the provision of essential ingredients of societal progress, such as trans-national public goods and economic growth.

The available composite indicators of globalisation, although going beyond the limits of a purely economic definition of international integration, fail to perform adequately their task for a variety of conceptual and methodological reasons.

A promising alternative is based on the recognition that the scope of international integration is not necessarily global, as cross-border interactions among human societies are often limited in their geographic reach. A new generation of statistical indicators is therefore being developed, in order to clearly distinguish between regional and global integration.

i E-mail: pdelombaerde@cris.unu.edu

ii E-mail: iapadre@ec.univaq.it

Introduction

The quality of social life depends on a wide range of environmental, cultural, economic and political factors, whose measurement represents a fundamental challenge for understanding their role and for devising proper policies. The task is made even more complex by the process of international integration, which raises the degree of interdependence among human societies across the globe, so that it is increasingly meaningless to approach the study of any community without explicitly considering its linkages with the rest of the world. 'Globalisation' is the word most widely used to describe this process, even if the scope of international integration is not necessarily global.

The purpose of this short paper is to critically review different proposals to construct globalisation indicators, from the perspective of a more general view of the relationship between international integration and societal progress. Special attention will be given to the underlying conceptual frameworks and to the question whether indicators are built on broader or narrower concepts of international integration (society versus economy, regional versus global integration, etc.), and to the coherence between the measurement needs, on the one hand, and the selection of the variables and indicators, on the other. A thorough discussion of the different technical solutions adopted in the construction of the different composite globalisation indicators is outside the scope of this paper.[1]

After critically revising the existing supply of globalisation indicators, we address the nexus between international integration and societal progress in last two sections concludes.

Genealogy of Globalisation Indicators

The Kearney/FP globalisation index (Kearney/FP-GI) is generally considered as the first proposal to construct a composite multi-dimensional globalisation index, supported by a database (Kearney 2001-2006). The index covers the economic, technological, political and personal aspects of globalisation, taking inspiration from the approach used to build the Human Development Index (UNDP 1998).

The economic dimension of this and other globalisation indicators benefited from previous work on international openness and competitiveness, including: the World Economic Forum's indicator of competitiveness, since 1979 (Lopes-Claros *et al.* 2006), Gwartney and Lawson's work on economic freedom, since 1996 (1996, 2006), and the World Market Research Center globalisation index (G-index) (Randolph 2001). More recently the OECD has taken the lead as a facilitator of new work on economic globalisation indicators, which has materialised in a Handbook (OECD 2005a) and a set of indicators (OECD 2005b), but has not involved the construction of a composite index of globalisation.[2]

Several proposals followed the Kearney/FP-GI, all trying to improve it on some aspect(s). Lockwood and Redoano (2005), consistent with Lockwood's critique of the Kearney/FP-GI (Lockwood 2001, 2004), designed the CSGR globalisation index (CSGR-GI). Whereas they partly present a different set of variables, the index mainly differs from the Kearney/FP-GI on the operational aspects (adjustment, normalisation, and weighting

of specific sub-indicators). Martens and Zywietz (2004, 2006), based on Zywietz (2003), proposed a Modified Globalisation Index (MGI). The authors take the Kearney/FP-GI also as their point of reference but start from a broader definition of globalisation, including environmental and military dimensions, and consequently, reduce the weight of the economic dimension. In addition, some technical improvements to the construction of the indicator are introduced. Heshmati's indicator (Kearney/FP/H) (Heshmati 2006), does not alter anything to the choice of variables and structure of the Kearney/FP-GI but a sophisticated statistical weighting procedure is added.

As far as the different dimensions of international integration are concerned, Dreher's globalisation index (DGI) is a more significant departure from the Kearney/FP-GI than the previous ones (Dreher 2005). Dreher expands the variables concerning personal contact and information flows, includes a cultural convergence variable, and re-introduces economic policy measures, which had been used before the Kearney/FP-GI to assess the degree of international economic integration.

Finally, an alternative for the traditional approaches to the measurement of the economic dimension of globalisation is represented by the recent work of Riezman, Whalley and Zhang (2004), who construct different measures of globalisation by comparing actual data to a counterfactual full integration equilibrium. Although pointing to a promising new strategy of research, they admittedly fail to obtain robust and reliable ordinal measures of globalisation, even if their indicators offer some information about the relative ranking of countries. A further problem of their approach is their reliance upon simple general equilibrium models based on the assumption of perfect competition, which do not appear apt to represent the actual features of global markets, characterised by various degrees of monopoly power.

Conceptual Frameworks

As there is no unique definition for globalisation the conceptual frameworks behind the globalisation indicators are diverse.[3] Until the late 1990s globalisation was still often considered as a synonym of 'global economic integration'. In his review article, for example, Brahmbhatt (1998, p.2) proposes as a definition of globalisation: "the increasing freedom and ability of individuals and firms to undertake voluntary economic transactions with residents of other countries, a process entailing a growing contestability of national markets by foreign suppliers". The definition used by the World Markets Research Center, developers of the G-index refers to "the ever closer knitting together of a one-world economy" (Randolph 2001, p.5). More recently, the OECD in its Handbook also still affirmed that "[g]lobalization refers above all to a dynamic and multidimensional process of economic integration whereby national resources become more and more internationally mobile while national economies become increasingly interdependent" (OECD 2005a, p.11).

However, inspired by the conceptual work of Held *et al*. (1999), Scholte (2000) and several others, a shift towards a multi-dimensional concept of globalisation has been observed. The Kearney/FP index, for example, measures the globalisation of a country in four dimensions (components): (i) the degree of integration of its economy in the world economy, (ii) the internationalisation of the personal contacts of its citizens, (iii) the use of internet technology, and (iv) the extent of its international political engagement. Martens and Zywietz (2004, 2006) add two additional dimensions in their MGI: (i) the involvement of a country's military-industrial complex with the rest of the world, and (ii)

the intensity of globalisation in the ecological domain. The spread of ideas, information, images, and people have been grouped together by Keohane and Nye (2000, p.4) and labeled 'social globalisation', as contrasted with 'economic' and 'political' globalisation.

Martens and Zywietz define globalisation then as: "the intensification of cross-national cultural, economic, political, social and technological interactions that lead to the establishment of transnational structures and the global integration of cultural, economic, environmental, political and social processes on global, supranational, national, regional and local levels" (Rennen & Martens 2003).

Dreher (2005) builds on the definitions proposed by Clark (2000, p.86) and Norris (2000, p.155) and refers to a process of "creating networks of connections among actors at multi-continental distances, mediated through a variety of flows including people, information and ideas, capital, and goods", a process "that erodes national boundaries, integrates national economies, cultures, technologies and governance, and produces complex relations of mutual interdependence".

Table 21.1 shows which dimensions have been used in the construction of the different globalisation indicators, and how the variables are distributed across dimensions.

Table 21.1 Globalisation Indicators - Number of Variables and Dimensions

Globalisation index	Number of variables	Number of categories	Dimensions
G-Index (Randolph 2001)	6	2	'old' economy (3 variables), 'new' economy (3)
Kearney/FP-GI (2001)	11	4	globalisation in goods and services (2), financial globalisation (3), globalisation of personal contact (3), internet connectivity (3)
Kearney/FP-GI (2003)	13	4	economic integration (4), personal contact (3), technology (3), political engagement (3)
Kearney/FP-GI (2004)	14	4	economic integration (4), personal contact (3), technological connectivity (3), political engagement (4)
Kearney/FP-GI (2005, 2006)	12	4	economic integration (2), personal contact (3), technological connectivity (3), political engagement (4)
CSGR-GI (Lockwood & Redoano 2005)	16	3	economic globalisation (4), social globalisation (9), political globalisation (3)
MGI (Martens & Zywietz 2004, 2006)	11	7	global trade (1), global finance (2), global politics (2), organised violence (1), people on the move (2), technology (2), environment (1)
DGI (Dreher 2005)	23	3	economic integration (8), political engagement (3), social globalisation (12)
Kearney/FP/H (Heshmati 2006)	13	4	economic integration (4), personal contacts (3), technology (3), political engagement (3)

The design of an adequate conceptual framework for the construction of globalisation indicators is thus clearly not an easy task, not the least because the very purpose of the whole effort can vary according to the needs and aims of the researcher or policy-maker. The following points can be taken into account:

First, some reflection is needed on the actors involved in the globalisation process. The builders of globalisation indicators usually do not address explicitly the fact that different actors are taking part in the process: states, regions, individual citizens, organised civil society, companies, etc. although by selecting certain specific indicators

they obviously implicitly privilege the behaviour of certain actors over others. Paying attention to the actors is probably a good shortcut to check that one is not neglecting important aspects of globalisation. Indicators that reflect the activity of global actors could add value to the usual measures based on transactions among national actors. The work done by the OECD on the activities of multinational enterprises (MNE) is an interesting reference point (OECD 2005a, p.75-136).

Second, there is a need to further differentiate concepts; globalisation indicators should be more clearly distinguished from indicators of international economic integration, openness, universalisation, westernisation, etc. (Scholte 2002; Martens & Zywietz 2004). Otherwise, they risk not to add much to existing knowledge and understanding. In practice, conceptual clarity seems to be missing and often globalisation indicators (or one or more of their sub-indicators) rather reflect different (albeit related) phenomena.

Third, one should be aware of the fact that 'adding dimensions' to a globalisation indicator, in practice, boils often down to a disaggregation of existing flows and may result in double-counting. For example, if military or cultural dimensions are added via the incorporation of measures of cross-border arms transactions or exchange of cultural products, this implies using some elements of goods and services trade flows. If, in addition, the aggregate variables are also kept (because they supposedly reflect the economic dimension of globalisation), certain flows are counted twice, which questions the validity of the whole weighting procedure. Unless there are good (theoretical, normative) reasons to give a double weight to certain flows, a possible procedure would be to subtract cross-border arms transactions or exchange of cultural products, in these examples, from total trade in goods and services.

Fourth, as Scholte (2002) also points out, 'globality' as a state should be distinguished from the 'globalisation process'.[4] Indeed, any state is the result of the corresponding process; what makes the difference is the perspective of the observer; when we want to see the process, we have to observe it across time; if we want to see the state, a snapshot at a given instant is required. From a conceptual point of view, the fact that globalisation refers to a long term and complex process, is generally acknowledged. One of the implications thereof for the construction of indicators is that these can theoretically refer to different logical components of this process, such as 'inputs' in the process, 'features' of the process, or 'outputs' (results, effects) of the process. According to Heshmati (2006, p.2), for example, the purpose of the construction of an index of globalisation is "to be able to quantify its sources and impacts", which would suggest that the purpose is not to measure the characteristics and progress of the process itself. Brahmbatt (1998, p.2-3) shares the view that globalisation indicators show both prerequisites and outcomes. Prerequisites or 'drivers' of the process include e.g. the progressive reduction in official obstacles to the conduct of cross-border economic transactions and the fall of business transaction costs, whereas the outcomes of the process refer to increased cross-border transactions (international trade, FDI, financial integration, labour migration) or international price convergence. An alternative organisation of the variables (and related indicators) is followed by Held *et al*. (1999) and distinguishes between the extensity, intensity, velocity and impact of global interactions. Some confusion seems to exist concerning these logical categories of variables; apparently various authors use 'outcome indicators' (cfr. supra) to measure the 'intensity/extensity' of the process, and not as synonyms of 'impact indicators' in Held's terminology. One way out could be to define 'inputs' in the process following Brahmbatt (see above), to define 'features of the process', following Held, in terms of the velocity, extensity and intensity of global

interactions, and to define 'outputs' of the process as effects of the increased global interactions on variables (growth, employment, income inequality, cultural convergence, …) different from those directly used in the definition of globalisation in terms of global interactions.

Fifth, and related to the previous point, a better understanding of the different logical components of the globalisation process should make it easier to detect blank spots in the construction of indicators. The current practice, as already mentioned, tends to focus on 'outcomes' or 'extensity/intensity' indicators (Martens and Zywietz (2004, p.12), and has thus moved away from policy indicators (input indicators) that were covered in the earlier work on economic globalisation (World Economic Forum, Heritage Foundation), with Dreher (2005) as a noticeable exception. Paradoxically, the gains related to incorporating more dimensions in the globalisation indicators, might have come at the cost of coverage in terms of logical categories. The practical difficulty to directly quantify policy measures in non-economic areas partially explains this (Lockwood 2001, p.6). From the perspective of the construction of globalisation indicators two options lay open to bring policy in. The first one is to explicitly include policy variables, as in the earlier economic globalisation indicators, and also in the DGI. These policy variables can refer to trade barriers, FDI policies, competitiveness enhancing policies, connectivity policies, migration policies, etc. Another option is to adjust globalisation measures for structural characteristics in order to filter out net policy effects.[5] An obvious problem with the first option is to weigh the policy variables with respect to other categories of variables. A problem with the second option is that structure-adjustment is usually not done in a systematic way. In the case of the CSGR-GI, for example, one can ask why structure-adjustment is applied to economic variables and not to other variables (Caselli 2006, p.17). However, in our opinion the most important question is related to the very purpose of the globalisation indicators. In most proposals, the designers do not choose clearly between an indicator showing the de facto degree of globalisation of a country (or a region, a group of citizens, a group of companies, …), which would not call for structure-adjustment, or an indicator assessing the scope and quality of globalisation policies of a country (or a region, a group of citizens, a group of companies, …). We would recommend to clearer distinguish between the two concepts, for example, systematically using a pair of indicators: one referring to the de facto degree of globalisation, and one referring to the quality of globalisation policies.

Sixth, when computing composite indices of globalisation, the problem of variable selection is intimately linked to the problem of weighting the different sub-indices. Although we do not address the operational aspects of the construction of composite indicators in detail here, it can be said that no single weighting procedure is a priori superior to any other procedure. Different theoretical and methodological considerations can be put forward to prefer one procedure over another. The choice is not straightforward and an element of arbitrariness will always remain. However, from an empirical point of view, robustness tests of the original Kearney/FP index have shown that changing the normalisation procedure and using statistical weights does only have small effects on country rankings (Lockwood 2001; Martens & Zywietz 2004, 2006). However, structure-adjustment has a major impact (Lockwood 2001, p.12-14). Heshmati (2006) performed a sensitivity analysis of Kearney/FP-GI via Principal Component Analysis. The author also finds that the Kearney/FP-GI performs relatively well and that the value added of statistical weighting is limited. The decisions on the choice of variables and on structure-adjustment appear therefore to be more important than the decisions about weighting procedures.

Finally, all the efforts to include more and better information in globalisation indicators should be balanced against the requirements of parsimony, efficiency and transparency. Analysts like Caselli (2006, p.15-16, 25-26), for example, argue already in favour of including less variables in the construction of globalisation indicators than usually is the case. The Human Development Index is thereby referred to as good practice.

Defining the Scope of International Integration

Starting from the simplifying assumption that measuring globalisation amounts to measuring the degree of international integration, defined as the extent to which the space dimension of a given social process tends not to be restricted by national borders, the precise definition of international integration and related indicators changes according to the perspective of the observer.

The first and most common option is to assume the standpoint of a single country or territory with respect to the rest of the world. In this case, measuring international integration amounts to assessing to what extent that particular country is open to relationships with the rest of the world, treated as a single partner country. The typical example, in the economic domain, is the trade-to-GDP ratio, universally considered as the most intuitive measure of international openness.[6]

In this class of indicators, no attention is normally paid to the geographical distribution of foreign relationships. A country with very intense linkages with only one neighbouring partner can in principle be considered as open as another country with moderate linkages with every possible partner.

An alternative approach would be to combine traditional measures of international openness with indicators of geographical diversification of bilateral relationships.[7] The simplest way to do so is by computing the ratio between the number of actual partners and the total number of potential partners (the total number of countries in the world). However this index would not account for any difference across partners in the intensity of the relationship, so that, for any given level of aggregate foreign openness and number of partners, a country having intense links with only one of them and marginal interactions with the others would be treated in the same way as a country interacting with all of them at the same level of intensity. In order to solve this problem, more precise measures of diversification are available, such as the inverse of the Herfindahl concentration index, sometimes called the 'number of equivalent partners'. The Herfindahl indexes of total exports and imports were included by the OECD in its list of supplemental indicators for measuring the extent of trade globalisation in their Handbook (OECD 2005a, p.185).

Although improving with respect to the previous option, indicators of geographical diversification, even when computed as the inverse of concentration indices, fail to inform properly on the geographical reach of the integration process, because they treat every partner in the same way, independently of its distance, so that a country linked exclusively with a certain group of neighbouring partners would not be distinguished from a country interacting with an equal number of partners scattered all over the world. The severity of this problem is obviously negatively related to the total number of partners, but still it cannot be neglected, also because of its interaction with the problem of concentration, in the sense that bilateral relationships tend to be relatively less intense with distant partners.[8] A possible solution lies in giving higher weights to more distant

partners when computing the diversification indices, which however raises the problem of finding proper measures of distance.[9]

However, even assuming that all these technical problems can be solved satisfactorily, openness indicators are anyway limited to a national perspective. Indicators at a higher level, regional or global, could only be obtained as appropriate averages of national measures.

A more radically different approach would be to take the perspective of a group of countries, be it a region, or an arbitrarily defined set of countries belonging to different regions. In this case, the central issue becomes that of distinguishing between intra- and extra-group integration, but again, this can be done in two different ways, either by treating both the group and the rest of the world as two single partners, or by exploiting the available information at bilateral level, and building appropriate measures of distance-weighted geographic diversification for intra- and extra-group relationships.

Traditional measures of regional integration tend to follow the first and simplest approach, but the second option is clearly superior, particularly for large regions with many member countries, such as the European Union.

At the world level, an appropriate average of national indicators could be enough to meet the need for a simple measure of globalisation, but additional information about the distribution of the indicators across countries would give important insights.[10] However, as pointed out by Caselli (2006), measures of global integration based exclusively on averages of national indicators fail to capture the specific nature of some aspects of globalisation, which can be defined only at the global level, without any reference to specific locations.[11] In other words, certain processes, the classical example being climate change, although being the result of activities with a specific territorial location, do not entail bilateral cross-border interactions, and can be better defined with reference to the entire planet. On the other hand, these intrinsically global processes, although very important, do not seem to represent exhaustively the nature of globalisation. Interactions across national borders are still fundamental not only for pragmatic reasons (data availability), but also in terms of policy relevance, given their role in forging global interdependence. A reference to cross-border interactions seems essential even when the unit of analysis is defined at local level, such as sub-national regions, cities, and individual agents.

A proper specification of regional and global integration indicators would be particularly useful to shed light on the empirical basis of the debate about regionalism and multilateralism.[12] At the policy level, with particular reference to the trade domain, there is increasing concern about what is customarily named the 'proliferation' of preferential integration agreements. The traditional debate is centred around the alleged negative effects of regional integration on economic welfare and on the functioning of the multilateral trading system. More recently, the 'new regionalism' literature has shifted the attention to the issue of the optimal allocation of competences across a multi-level architecture of international relations, where a proper application of the subsidiarity principle clearly reveals that regional integration performs important functions in the production of trans-national public goods (see below, section on "Measuring the Provision of Trans-national Public Goods). At the same time, concerns are now focussed on the rapid development of bilateral integration agreements, which can create problems not only for the multilateral system, but also for regional integration processes. The debate about these normative issues could receive useful inputs by a correct measurement of the actual intensity of regional vs. global integration processes. It is sometimes argued

that the word 'globalisation' is used improperly to describe integration processes which tend to be mostly regional.[13] It is therefore particularly important to build appropriate measures of the intensity and the extension of international transactions, as suggested above.[14]

Linking International Integration to Societal Progress

The debate about globalisation and its effects on human societies is so intense and widespread that it would be almost impossible to summarise its arguments in the limited space of this note. We will here focus on only two of the possible linkages between international integration and societal progress:

In the next section we will consider the way in which the concepts of societal well-being and progress can or should influence the building of international integration indicators. The section "The Globalisation-Societal Progress Nexus: Empirical Evidence" we will briefly discuss some problems related to assessing the empirical evidence about the effects of international integration on societal progress, with particular reference to its economic dimension and more precisely to the linkages between international integration, competition, innovation and growth.

Measuring the Provision of Trans-national Public Goods

The concepts of societal well-being and progress have been developed in order to overcome the well-known limitations of per-capita GDP in measuring adequately a country's standard of living. So, a wide range of environmental, social, cultural, and political variables has been combined with per-capita GDP in order to build more comprehensive composite indicators of the quality of life at the national level, such as the above quoted Human Development Index.

In principle, any of such indicators could be averaged across countries, so as to obtain global and/or regional measures of societal well-being. These aggregation exercises, although useful, fail to capture properly the specific contribution that international integration in itself can give to societal progress, through its effects on the provision of trans-national public goods.

Societal progress depends strongly on an adequate supply of global public goods (Kaul, Grunberg and Stern 1999), which is currently hampered by the mismatch between the trans-national scope of the problems and the national level of policy-making (jurisdictional gap), by the lack of mechanisms ensuring the involvement of trans-national social organisations in the decision-making process at the supra-national level (participation gap), and by the flaws of the institutional means designed to ensure international co-operation (incentive gap). The economic theory of institutions and fiscal federalism theory can be applied to understand under what circumstances the production of trans-national public goods requires the establishment of specific institutions and at what level (local, national, regional, or global). However, this economic approach must be integrated into a wider political vision of global governance, taking into account the problems of legitimacy, fairness and accountability of international institutions (Higgott 2006).

Since the reach of spill-overs generated by trans-national public goods is often limited to regional groupings of countries, it has been argued that their production should be fostered by regional institutions (Sandler 2004).

From this perspective, building and developing regional institutions can be seen as a direct contribution to societal progress, inasmuch as it ensures an improvement in the availability and quality of regional public goods. Indeed, the establishment of such institutions in itself can be seen as a sign of progress, since it reveals that participating societies are starting to share a common understanding of their regional identity, which is the necessary pre-condition to effectively tackle their joint problems. Any comprehensive indicator of global or regional integration should take this institutional dimension into account, by including measures not only of the number of institutions, but also of their importance in terms of membership, resources, and depth of the integration achieved.

Available indicators of globalisation fail to fully meet this need. Most of them (CSGR-GI, MGI and DGI) include only numerical variables, measuring the extensity of international political integration without any attempt at gauging its intensity, in terms of financial resources or other possible criteria. Only the Kearney/FP-GI combines numerical indicators of institutional integration (memberships in international organisations and ratification of multilateral treaties) with measures of the intensity of 'political engagement' (personnel and financial contributions to UN peacekeeping missions and governmental international transfers). However, these four variables are equally weighted, which means that purely numerical extensity indicators are given the same importance as intensity indicators. Moreover, as argued in section "Defining the Scope of International Integration", globalisation indicators actually measure the degree of international integration at national level, regardless of the geographic reach of foreign interactions. As a consequence, the challenge of distinguishing between global and regional integration remains open, even in the political domain.

An alternative way to take global (and regional) public goods into account would be by looking at the consumption/use of these goods, instead of looking at their financing. An example of how this can be done is provided by the MGI, which includes a measure of the environmental impact of trade flows, based on the balance between their 'ecological footprint' and bio-capacity. One could think of similar measures related to other trans-national public goods, such as peace, economic stability, etc.

The Globalisation-Societal Progress Nexus: Empirical Evidence

Turning now to the empirical evidence, even if no firm conclusion has yet been reached, it seems fair to say that most studies tend to show a positive effect of international integration on societal well-being.

For example, Dreher (2005), using his composite globalisation indicator, finds strong evidence of a positive effect of international integration on growth. Less convincingly, Kearney/FP use simple correlation coefficients to claim that there are positive effects of globalisation on environmental performance (2003) (although not maintained in 2006), life expectancy, women's well-being (2004), and freedom (2005). Negative correlations are found between globalisation and corruption (2005). There is no attempt at controlling for the risk of spurious correlation.

In the economic domain, although controversies about the trade and growth nexus are still very intense, there is a large consensus over the idea that growth tends to be more rapid in open than in inward-looking countries. Measuring the dynamic effects of international integration, that go beyond its once-for-all impact on the allocation of productive resources, is therefore one of the most important challenges to be met, in order to gauge its contribution to societal progress. At the same time, it is a very difficult task,

given the complexity of the growth phenomenon and the difficulty of tracing it back to each of its underlying factors.

International integration is expected to foster growth by promoting investment in physical, human, and knowledge capital. This is both a consequence of its positive static impact on production efficiency, and of its direct spur to investment return. The more intense competition generated by international integration translates into a selection effect, picking up the most productive and innovative enterprises.

Accounting for the dynamic effects of international integration does not amount only to providing evidence about the final result of this complex chain of interactions, that could be done simply by comparing growth rates of per-capita GDP. It also calls for finding adequate indicators of all the intermediate steps, that could be used as inputs in a properly specified econometric exercise. For example, it is necessary to devise reliable measures of the degree of competition, in order to check to what extent the removal of trade and investment barriers actually weakens monopoly power in imperfectly competitive markets. Another important aspect to be considered is the process of innovation, and more generally knowledge creation and diffusion, which is crucial to the understanding of growth, but elusive of any simple quantitative indicator.

These problems become even more severe and important when trying to distinguish between the effects of global and regional international integration. In the traditional literature about trade creation and diversion, the static welfare effects of integration offer a relatively easy criterion to compare preferential and multilateral liberalisation strategies. This comparison becomes more difficult in the most recent studies of regional integration, where the emphasis is put on its growth effects. It is normally very problematic to ascertain if preferential trade liberalisation is able to foster growth more or less than integration into the multilateral trading system. Proper indicators of the dynamic effects of global and regional integration could be very useful to address these problems.

The same applies to other aspects of societal well-being, such as equality of income distribution, health, environment, education, culture, and social cohesion, for which problems related to selecting the most appropriate theoretical models and statistical techniques to detect possible effects of international integration are augmented by the limitations of available data and indicators.

In all dimensions of international integration there is a strong need for new and better quality data, as well as for more precise statistical indicators, distinguishing clearly between regional and global interdependence.

Conclusions

The measurement of international integration, in all its aspects, is increasingly important to a complete understanding of societal well-being and its dynamics. A higher degree of international integration can be seen in itself as an indicator of societal progress, inasmuch as it reveals that human societies more and more acknowledge their common destiny. In addition, international integration fosters the provision of essential ingredients of societal progress, such as trans-national public goods and economic growth.

These issues are often approached in the context of the debate about globalisation. In this paper, we have tried to show that the available composite indicators of globalisation, although going beyond the limits of a purely economic definition of international

integration, fail to perform adequately their task for a variety of conceptual and methodological reasons.

A promising alternative is based on the recognition that the scope of international integration is not necessarily global, as cross-border interactions among human societies are often limited in their geographic reach. A new generation of statistical indicators is therefore being developed, in order to clearly distinguish between regional and global integration. Nevertheless, many challenges remain open. The transmission channels from international integration to societal progress are complex and still not fully understood, even in the economic domain, as it is clearly shown by the debate about the growth effects of international trade. Many fundamental concepts, such as competition and innovation, defy precise definition and measurement. Even more importantly, the need of taking all the multidimensional aspects of societal progress into a comprehensive and simple indicator of international integration is still to be met.

Notes

1 See, for example, Nardo et al. (2005) for an excellent general treatment of these operational issues.

2 The OECD indicators aim at measuring the magnitude and intensity of economic globalisation, in four areas: international trade, foreign direct investment (FDI), the activity of multinational firms, and the production and international diffusion of technology.

3 There is now already a rich literature on globalisation. It is beyond the scope of this article to review the evolution of the concept of globalisation in depth. We refer to Scholte (2002) for an excellent overview. He convincingly argues for an understanding of globalisation "as the spread of transplanetary – and in recent times more particularly supraterritorial – connections between people […] globalisation involves reductions in barriers to transworld contacts. People become more able –physically, legally, culturally, and psychologically – to engage with each other in 'one world' […] globalisation refers to a shift in the nature of social space". The author further questions the 'methodological territorialism' which is still dominating the social sciences (and hence the construction of globalisation indicators). See also, Caselli (2006).

4 In the literature about regional integration, which will be considered below, a similar point has been raised by De Lombaerde and Van Langenhove (2006).

5 'Structure-adjustment' refers to the adjustment of outcome measures for structural characteristics, regressing the measures on a set of structural variables and using the residuals to construct the index, so that they better reflect the effect of policies on a country's integration in the world economy. Lockwood (2001:6-9), for example, applied Pritchett's approach (1996) to adjust the Kearney/FP variables. The cost to pay is a considerable loss of transparency, readability and user-friendliness of the index. One should also be aware of the fact that after such adjustment the indicator might reflect less what could be called de facto globalisation. For example, Martens and Zywietz (2006) filter out the effect of countries being landlocked, but one could ask whether relatively low levels of international integration and connectedness is not exactly what one would like the globalisation indicator to reveal for landlocked countries.

6 It should be reminded that, for a variety of reasons, trade openness indicators tend to be negatively correlated with country size, so that the resulting ranking is not reliable.

7 The recently flourishing literature about extensive and intensive margins of trade refers to a similar problem, i.e. the decomposition of world trade growth into the increase in the number of bilateral relationships (extensive margins) and the growth in the volume of trade per relationship (intensive margins). See Helpman, Melitz and Rubinstein, 2007.

8 In the case of trade, this observation is nicely captured by gravity models of bilateral flows.

9 A more pragmatic solution would consist of measuring the relative importance of 'extra-regional' interactions (trade, FDI, migration, tourism, etc.) as a proxy for 'global' interactions, rather than using the usual 'international' interactions, provided that one is able to define the relevant 'region'. However, as will be argued below, extra-regional integration measures do not solve the problem, even leaving aside the issue of defining the region, because the 'extra-regional world' is still seen as a unique partner, independently of the country distribution of transactions.

10 The recent literature on socio-economic networks can be very useful to this purpose. See, for example, Rauch (1999, 2001) and Kali and Reyes (2007).

11 See also, Scholte's (2002) critique of 'methodological territorialism'.

12 See, for example, Woolcock (2006) and Fiorentino et al. (2007).

13 In the case of multinational enterprises, see Rugman and Verbeke (2004) and Rugman (2005).

14 See, for example, the work that has been undertaken by UNU-CRIS on regionalization indicators in the context of the Regional Integration Knowledge System (RIKS) (www.cris.unu.edu).

References

Brahmbhatt, M. (1998), "Measuring Global Economic Integration: A Review of the Literature and Recent Evidence", Working Paper, Washington: The World Bank.

Caselli, M. (2006), "On the nature of globalisation and its measurement. Some notes on the A.T. Kearney/Foreign Policy Magazine Globalisation Index and the CSGR Globalisation Index", UNU-CRIS Occasional Papers, (2006/03).

Clark, W.C. (2000), "Environmental Globalisation", in: J.S. Nye and J.D. Donahue (eds), Governance in a Globalizing World, Washington: Brookings Institution, pp. 86-108.

De Lombaerde, P. and L. Van Langenhove (2006), "Indicators of Regional Integration: Conceptual and Methodological Aspects", in: P. De Lombaerde (ed), Assessment and Measurement of Regional Integration, London: Routledge, pp. 9-41.

Dreher, A. (2005), "Does Globalisation Affect Growth? Evidence from a New Index of Globalisation", Research Paper Series, Thurgauer Wirtschaftsinstitut, (6).

Fiorentino, R., L. Verdeja and C. Toqueboeuf (2007), "The Changing Landscape of Regional Trade Agreements: 2006 Update", WTO Discussion Papers, (12).

Gwartney, J., R. Lawson and W. Block (1996), Economic Freedom of the World 1975-1995, Vancouver: Fraser Institute.

Gwartney, J. and R. Lawson (2006), Economic Freedom of the World 2006 Annual Report, Vancouver: Fraser Institute.

Held, D., A. McGrew, D. Goldblatt and J. Perraton (1999), Global Transformations: Politics, Economics and Culture, Stanford: Stanford University Press.

Helpman, E., M. Melitz and Y. Rubinstein (2007), "Estimating Trade Flows: Trading Partners and Trading Volumes", mimeo, Harvard University, available at: http://www.economics.harvard.edu/faculty/helpman/papers/TradingPartners.pdf

Heshmati, Almas (2006), "Measurement of a Multidimensional Index of Globalisation", Global Economy Journal, 6(2):1-28.

Higgott, R. (2006) "Global Governance and Global Public Goods", International Conference on Governance and the Politics of Development, Den Haag, Institute of Social Studies: 1-16. Keynote Address.

Kali, R. and J. Reyes (2007), "The Architecture of Globalisation: a Network Approach to International Economic Integration", Journal of International Business Studies, (online publication 17 May 2007; doi: 10.1057/palgrave.jibs.8400286).

Kaul, I., I. Grunberg and M. A. Stern (1999), Global Public Goods: International Cooperation in the 21st Century, UNDP, New York, Oxford: Oxford University Press.

Kearney, A.T. (2001), "Measuring Globalisation", Foreign Policy, 122:56-65.

Kearney, A.T. (2002), "Measuring Globalisation", Foreign Policy.

Kearney, A.T. (2003), "Measuring Globalisation: Who's Up, Who's down?", Foreign Policy.

Kearney, A.T. (2004), "Measuring Globalisation: Economic Reversals, Forward Momentum", Foreign Policy, 2:54-69.

Kearney, A.T., (2005), "Measuring Globalisation", Foreign Policy, 3:52-60.

Kearney, A.T., (2006), "The Globalisation Index", Foreign Policy, 6:74-81.

Keohane, R.O. and J. Nye (2000), "Introduction", in: J.S. Nye and J.D. Donahue (eds), Governance in a Globalizing World, Washington: Brookings Institution, pp. 1-44.

Lockwood, B. (2001), "How Robust is the Foreign Policy/Kearney Index of Globalisation?", CSGR Working Paper, (79/01), Warwick: University of Warwick.

Lockwood, B. (2004), "How Robust is the Foreign Policy-Kearney Globalisation Index?", The World Economy, 27:507-523.

Lockwood, B. and M. Redoano (2005), "The CSGR Globalisation Index: an Introductory Guide", CSGR Working Paper, (155).

Lopes-Claros, A., M. Porter, X. Sala-i-Martín and K. Schwab (2006), Global Competitiveness Report 2006-2007, London: Palgrave McMillan.

Martens, P. and D. Zywietz (2004), "Rethinking Globalisation. A Modified Globalisation Index", e-Readers, University College Maastricht, (LS212).

Martens, P. and D. Zywietz (2006), "Rethinking Globalisation. A Modified Globalisation Index", Journal of International Development, 18(3):331-350.

Nardo, M., M. Saisana, A. Saltelli, S. Tarantola, A. Hoffman and E. Giovannini (2005), "Handbook on Constructing Composite Indicators: Methodology and User Guide", OECD Statistics Working Paper, Paris: OECD, [STD/DOC(2005)3].

Norris, P. (2000), "Global Governance and Cosmopolitan Citizens", in: J.S. Nye and J.D. Donahue (eds), Governance in a Globalizing World, Washington: Brookings Institution, pp. 155-177.

OECD (2005a), Measuring Globalisation: OECD Handbook on Economic Globalisation Indicators, Paris: OECD.

OECD (2005b), Measuring Globalisation: OECD Economic Globalisation Indicators, Paris: OECD.

Pritchett, L. (1996), "Measuring Outward Orientation in LDCs: Can It Be Done?", Journal of Development Economics, 49(2):307-335.

Randolph, J. (2001), "G-index: Globalisation Measured", World Markets Research Center, <www.globalinsight.com>.

Rauch, J.E. (1999), "Networks versus Markets in International Trade", Journal of International Economics, 48:7-35.

Rauch, J.E. (2001), "Business and Social Networks in International Trade", Journal of Economic Literature, 39:1177-1203.

Rennen, W. And P. Martens (2003), "The Globalisation Timeline", Integrated Assessment, 4(3):1-8.

Riezman, R., J. Whalley and S. Zhang (2005), "Metrics Capturing the Degree to which Individual Economies are Globalized", CESIFO Working Paper, (1450).

Rugman, A.M. and A. Verbeke (2004), 'A Perspective on Regional and Global Strategies of Multinational Enterprises', Journal of International Business Studies, 35(1): 1-15.

Rugman, A.M. (2005), The Regional Multinationals, Cambridge: Cambridge University Press.

Saltelli, A., M. Nardo, M. Saisana and S. Tarantola (2004), "Composite Indicators – The Controversy and the Way Forward", paper presented at Statistics, Knowledge and Policy. OECD World Forum on Key Indicators, Palermo, 10-13 Nov.

Sandler, T. (2004), "Demand and Institutions for Regional Public Goods", in A. Estevadeordal, B. Frantz and T.R. Nguyen (eds), Regional Public Goods. From Theory to Practice, Washington: IDB-ADB, pp. 11-30.

Scholte, J.A. (2000) Globalisation. A Critical Introduction. Basingstoke: Palgrave.

Scholte, J.A. (2002), "What is Globalisation? The Definitional Issue – Again", CSGR Working Paper, (109/02), Warwick: CSGR, University of Warwick.

UNDP (1998), Human Development Report, Oxford: Oxford University Press.

Woolcock, S. (2006), Trade and Investment Rulemaking: The Role of Regional and Bilateral Agreements, Tokyo: UNU Press.

Zywietz, D. (2003), Measuring Globalisation. Constructing a Modified Indicator, Maastricht: Faculty of Economics and Business Administration, Maastricht University.

OECD PUBLICATIONS, 2, rue André-Pascal, 75775 PARIS CEDEX 16
PRINTED IN FRANCE
(30 2008 08 1P) ISBN 978-92-64-04323-7 – No. 56153 2008

Statistics, Knowledge and Policy 2007

MEASURING AND FOSTERING THE PROGRESS OF SOCIETIES

Is life getting better? Are our societies making progress? Indeed, what does "progress" mean to the world's citizens? There can be few questions of greater importance in today's rapidly changing world. And yet how many of us have the evidence to answer these questions?

For a good portion of the 20th century there was an implicit assumption that economic growth was synonymous with progress: an assumption that a growing GDP meant life must be getting better. But we now recognise that it isn't quite as simple as that. Despite high levels of economic growth in many countries many experts believe we are no happier than we were 50 years ago; that people trust one another – and their governments – less than they used to; and that increased income has come at the expense of increased insecurity, longer working hours and greater complexity in our lives. Much of the world is healthier and people live longer than they did just a few years ago, but environmental problems like climate change cast a shadow over an uncertain future. Indeed, it sometimes seems that for every action to demonstrate societal progress, an equal but opposite reaction demonstrates precisely the opposite. And when the experts disagree, what hope do the citizens have to engage in democratic debate about their future and make the right choices at the ballot box? Access to accurate information is vital when we come to judge our politicians and hold them accountable. But access to a comprehensive and intelligible portrait of that most important of questions – is life getting better – is lacking in many societies.

Concerns about this have been growing. And over the past 10 years or so there has been an explosion of interest in produce measures of societal progress. Measures that go beyond GDP to represent a broader view of the ways in which societies are progressing and regressing. Measures which are based on the values of a society, not those of a single political party or an elite few. Such sets of progress measures can help governments focus in a more joined up way on what really matters: they can foster a more informed debate on where a society is, where it wants to head, and – crucially – the choices it needs to make if it is to get there. By measuring progress we can foster progress.

The OECD's 2nd World Forum on *Statistics, Knowledge and Policy "Measuring and Fostering the Progress of Societies"* held in Istanbul in June 2007 brought together a diverse group of leaders from more than 130 countries to debate these issues. These proceedings contain the papers presented at the Forum.

The full text of this book is available on line via this link:
www.sourceoecd.org/generaleconomics/9789264043237

Those with access to all OECD books on line should use this link:
www.sourceoecd.org/9789264043237

SourceOECD is the OECD online library of books, periodicals and statistical databases.
For more information about this award-winning service and free trials ask your librarian, or write to us at **SourceOECD@oecd.org**.

OECD*publishing*
www.oecd.org/publishing

ISBN 978-92-64-04323-7
30 2008 08 1 P

9 789264 043237