OECD Statistics Working Papers 2011/08

Linking Policies to Well-Being Outcomes Through Micro-Simulation

Michael Wolfson

https://dx.doi.org/10.1787/5kg1zq0qwmf1-en

OECD

Unclassified **STD/DOC(2011)8**

Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development **21-Nov-2011**

English - Or. English

STATISTICS DIRECTORATE

STD/DOC(2011)8
Unclassified

LINKING POLICIES TO WELL-BEING OUTCOMES THROUGH MICRO-SIMULATION

STATISTICS DIRECTORATE

WORKING PAPER NO. 43

This paper has been prepared by Michael Wolfson, Canada Research Chair, Population Health Modeling / Populomics, Faculty of Medicine and Institute of Population Health, University of Ottawa; former Assistant Chief Statistician, Statistics Canada.

For further information please contact:
Michael Wolfson
E-mail: Michael.Wolfson@uOttawa.ca

JT03311814

English - Or. English

OECD STATISTICS WORKING PAPER SERIES

The OECD Statistics Working Paper Series - managed by the OECD Statistics Directorate – is designed to make available in a timely fashion and to a wider readership selected studies prepared by OECD staff or by outside consultants working on OECD projects. The papers included are of a technical, methodological or statistical policy nature and relate to statistical work relevant to the Organisation.

The Working Papers are generally available only in their original language - English or French - with a summary in the other. Comments on the papers are welcome and should be communicated to the authors or to the OECD Statistics Directorate, 2 rue André Pascal, 75775 Paris Cedex 16, France.

The opinions expressed in these papers are the sole responsibility of the authors and do not necessarily reflect those of the OECD or of the governments of its Member countries.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

---

http://www.oecd.org/std/research

ABSTRACT

A major challenge in the measurement of well-being and progress is to link indicators of high-level societal outcomes with specific policy interventions. This is important not only for better informing the public, but also to provide the means for policy makers and advisors to assess the impacts of their policies and programmes and to increase their effectiveness and cost-efficiency. This paper looks at four major areas of social policies– health status, literacy and learning, economic security, and economic inequality– with the aim of understanding how to link broad outcome measures of progress in these areas, on the one hand, and the policies bearing on them, on the other. Emphasis is given to the powerful benefits to be derived from coupling longitudinal, multivariate data and powerful statistical methods with recently developed analytical tools such as micro-simulation. The paper also emphasises the need for "principled" summary indicators, i.e. indicators embedded within coherent data systems, and the importance of internationally comparable data based on common concepts and definitions.

RESUMÉ

Lorsqu'on mesure le bien-être et le progrès, l'une des principales difficultés consiste à relier les indicateurs de résultats sociaux à des actions spécifiques. Cela est crucial, non seulement pour mieux informer le public, mais aussi pour permettre aux acteurs politiques d'évaluer l'incidence de leurs actions et de leurs programmes, afin de leur permettre d'accroître l'efficacité et l'efficience des mesures en place. Ce rapport porte sur cinq domaines majeurs, pour lesquels le suivi des moteurs du progrès social ou du bien-être est essentiel : i) la santé, ii) l'alphabétisation et l'apprentissage iii) la sécurité économique, iv) les inégalités économiques et v) le manque de temps – l'objectif étant de comprendre comment améliorer les liens entre, d'une part, les indicateurs généraux du progrès dans ces domaines, et d'autre part, les outils permettant d'influer sur les résultats. L'accent est mis sur les grands avantages que présente l'association entre des données longitudinales multi variées, de puissantes méthodes statistiques et des outils d'analyse récents tels que la micro-simulation. Sont également soulignés : la nécessité d'utiliser des indicateurs synthétiques structurés sur des principes définis et s'inscrivant dans des systèmes de données cohérents, et l'importance de disposer de données comparables à l'échelle internationale, fondées sur des définitions et des concepts communs.

TABLE OF CONTENTS

Figures

Boxes

LINKING POLICIES TO WELL-BEING OUTCOMES THROUGH MICRO-SIMULATION: EXAMPLES FOR SELECTED AREAS OF SOCIAL POLICY

M.C. Wolfson[1]

Introduction

1. For a number of years, the OECD has been leading a broad effort to measure the progress of societies in an innovative and more comprehensive manner[2]. A major challenge in this context is connecting new and better measures of well-being and progress to government policies, and to their real world implementations as programmes and institutional structures, as well as to their constituent processes. One basic reason for this challenge is that some fundamental measures of societal progress – ideally outcome measures – often appear too general to be amenable to policy intervention. By articulating a coherent set of measures of both outcomes and of processes reflecting the operations of government policies and programmes, it should be possible to make the connections between broad measures of progress on the one hand, and the means to influence results on the other. This is not only important to make the public better informed, but also to provide the means for policy makers and advisors to assess better the impacts of their policies and programmes, and thereby to amend or adjust them in order to be more effective and more cost-effective.

2. This paper is structured around four major areas where measuring, monitoring, and understanding the drivers of societies' progress is key: health status, literacy/learning, economic security, and economic inequality.[3] The analysis then highlights common threads in these areas, particularly from the viewpoint of statistics and analytical methods. These common threads include:

- The need for "principled" summary indicators for overall monitoring of progress;
- The need for these summary indicators to be embedded within coherent data systems;
- The importance of international comparisons, hence of measures and data systems based on common concepts and definitions;

---

1 Canada Research Chair, Population Health Modeling / Populomics, Faculty of Medicine and Institute of Population Health, University of Ottawa; former Assistant Chief Statistician, Statistics Canada. I am indebted for comments in an earlier shorter version to Andrew Sharpe, Centre for the Study of Living Standards and Christian Dea, Human Resources Development Canada, and to members of the Coordination Group for the OECD-hosted Global Project on Measuring the Progress of Societies, who discussed the paper at the meeting in Busan, Korea, October 2009. I also wish to thanks Conal Smith, Kate Scrivens and Marco Mira d'Ercole for comments on this longer version of the paper.

2 www.oecd.org/progress

3 This particular set of domains is illustrative. It is not intended to be comprehensive nor to reflect any specific proposal. Still, the domains in this listing are commonly included in virtually all discussions of social indicators and measures of progress.

- The centrality of longitudinal and multivariate micro-data – *i.e.* data sets with representative samples of individuals' trajectories of socio-economic status, program participation, and program impacts on their lives – in order to be able to unravel and understand the connections among outcomes, policies and programmes, and processes;

- The powerful benefits to be derived by coupling rich data and powerful statistical methods with recently feasible computer-intensive analytical tools such as micro-simulation.

"Principled" summary indicators and public policy

3. There is a wide range of proposals for summary indicators and measures, in addition to the two most widely cited international indicators, GDP per capita and life expectancy. At the same time, there is sometimes scepticism as to the usefulness of new indicators, especially to the extent there are no obvious "policy levers" that can affect them.

4. In general, a summary indicator combines information or data from a variety of sub-indicators or sub-domains, often by means of a formula taking the form of some sort of weighted average. We use the term "principled" to describe those summary indicators where the formula for aggregating the information from the sub-domains or sub-indicators, usually a set of weights, is based on some clear and reasonable principle(s), and is not completely arbitrary.[4] Not all summary measures in the literature or under discussion are principled in this sense. However, those in wide-use for public policy do meet this criterion. Life expectancy is effectively an aggregation of age-specific mortality rates, where the aggregation is based on the stable population age structure consistent with this series of mortality rates. GDP is based on an aggregation of transactions using market prices as de facto weights. Measures of inflation aggregate myriad price changes of specific commodities using weights based on the observed basket of goods and services purchased on average in the economy. In all these cases, the summary index is "principled" as we define it (though not completely free of arbitrariness.)

5. Summary measures are important because they can become the focus of public and policy interest. The two summary socio-economic measures that over decades have generally received the greatest attention domestically among OECD countries are the unemployment rate and the inflation rate. While neither of these measures relates in detail to the circumstances of any specific individual in a society, they have enough validity for the range of individuals' experiences that they are widely accepted as important. In turn, these indicators have for decades been the objects of economic policy. Politicians and their Ministry advisors, by means of fiscal, monetary, and labour market policy, continually endeavour to stabilise their economies at low levels of unemployment and inflation, while central banks are charged with maintaining price stability. They further have created structures of social insurance to mitigate the most severe impacts of unemployment, often with indexing provisions to assure that benefits are not eroded by inflation, and policies oriented toward innovation and longer term economic growth in order to ensure adequate employment over the longer term.[5]

6. These summary indicators have validity for citizens as well as for government ministries with policies and programmes (*e.g.* monetary policy, automatic fiscal stabilisers, unemployment insurance, support for industrial R&D). Moreover, these summary indicators have long-standing associations with causal theories, statistical systems and simulation models. The principal theory underpinning GDP and employment rates derives from Keynesian macro-economics; the main statistical framework is the System of National Accounts; and the simulation models are macro-econometric. In other words, there is an

[4] I am indebted to an oral comment from Dan Usher for this usage of the term "principled".

[5] In the current fiscal climate of unsustainable public deficits, though, many countries are cutting back on these programmes.

elaborate theoretical, professional and statistical infrastructure that connects specific policy levers to the broad summary indicators. Of course, these connections are nowhere near perfect, as clearly evidenced by the recent global financial collapse. And there are competing underlying causal theories. But the response to events that are unanticipated by the data, theory and models, as in science generally, is to go back, re-examine the data, revise theories and models, perhaps collect new kinds of data, and develop and deploy new policy instruments.

7. This example of macro-economic summary indicators, data and analytical tools, and policy is relevant in the context of the OECD's initiative for measuring well-being and societal progress. This latter field is quite new in comparison, and policy makers may well wonder about the usefulness of broad summary measures if there is nothing they can do to move them. The analogy with the macro-economy suggests that new summary measures, in isolation, may at best become an exercise in frustration for policy makers and their advisors. To be useful, as in the macro-economic policy domain, such new measures require complementary infrastructure in the form of causal theories, statistical data systems, simulation models and appropriately skilled professional staff. We illustrate this point in the following sections.

Application to selected policy areas

Health

8. We begin in the domain of health, one of the most important and extensive areas of government policy. The fundamental outcome measure in the health area should be obvious and uncontroversial – the health status of the population. However, the only widely used internationally comparable population "health status" measure continues to be life expectancy, which is in fact no more than a weighted average of mortality rates.[6] There are no widely used internationally comparable measures of health outcomes that use data on the health conditions of the living.

9. One approach to develop such health outcome measures would be to use data on chronic disease prevalence. These data, conceptualised in bio-medical terms, are becoming increasingly available, either from administrative data on health care encounters, or from self-reports in household interview surveys. Another approach is to use data on functioning – basic aspects of health such as mobility, pain, vision, hearing, and memory. This concept of functioning is distinct from bio-medical or clinical notions of disease, and is explained in the WHO's International Classification of Functioning, Disability and Health (ICF, WHO 2010). A short form of the ICF version is operationalised in the survey question set of the Budapest Initiative on Measuring Health Status (UNECE). While both disease and functioning approaches are desirable, health outcome measures based on functioning have the advantages of being more understandable to a broader public and to policy makers; they are more readily collected in household surveys; and can be collected directly from patients in clinical settings.

10. Moreover, there is a well-established tradition and methodology for deriving a summary index of population health from self-reported data on functioning in the form of the widely used concept of QALYs or quality adjusted life years (*e.g.* Torrance and Feeny, 1989; Gold *et al*., 1996). Given survey data on the distribution of QALYs in the population, the simplest population health outcome measure is HALE – *health-adjusted* life-expectancy. HALE is a generalisation of the widely used concept of life expectancy that can be simply computed using life table methods (Sullivan, 1971; Wolfson, 1996; Wolfson and Lievesley, 2007).

[6] The weights are the population counts by year of age for the hypothetical steady-state population from the life table based on the same mortality rates.

11. From a policy perspective, though, taking increases in HALE as the principal policy objective can be very daunting, since quality adjusted life years reflects myriad factors, many of which are not the object of, and often not even amenable to, government policy.[7] Still, it is fundamentally important to construct an analytical framework whereby the connections to areas of government policy can be made. Figure 1 below, from Statistics Canada and CIHI (2009) shows these kinds of influences and pathways.

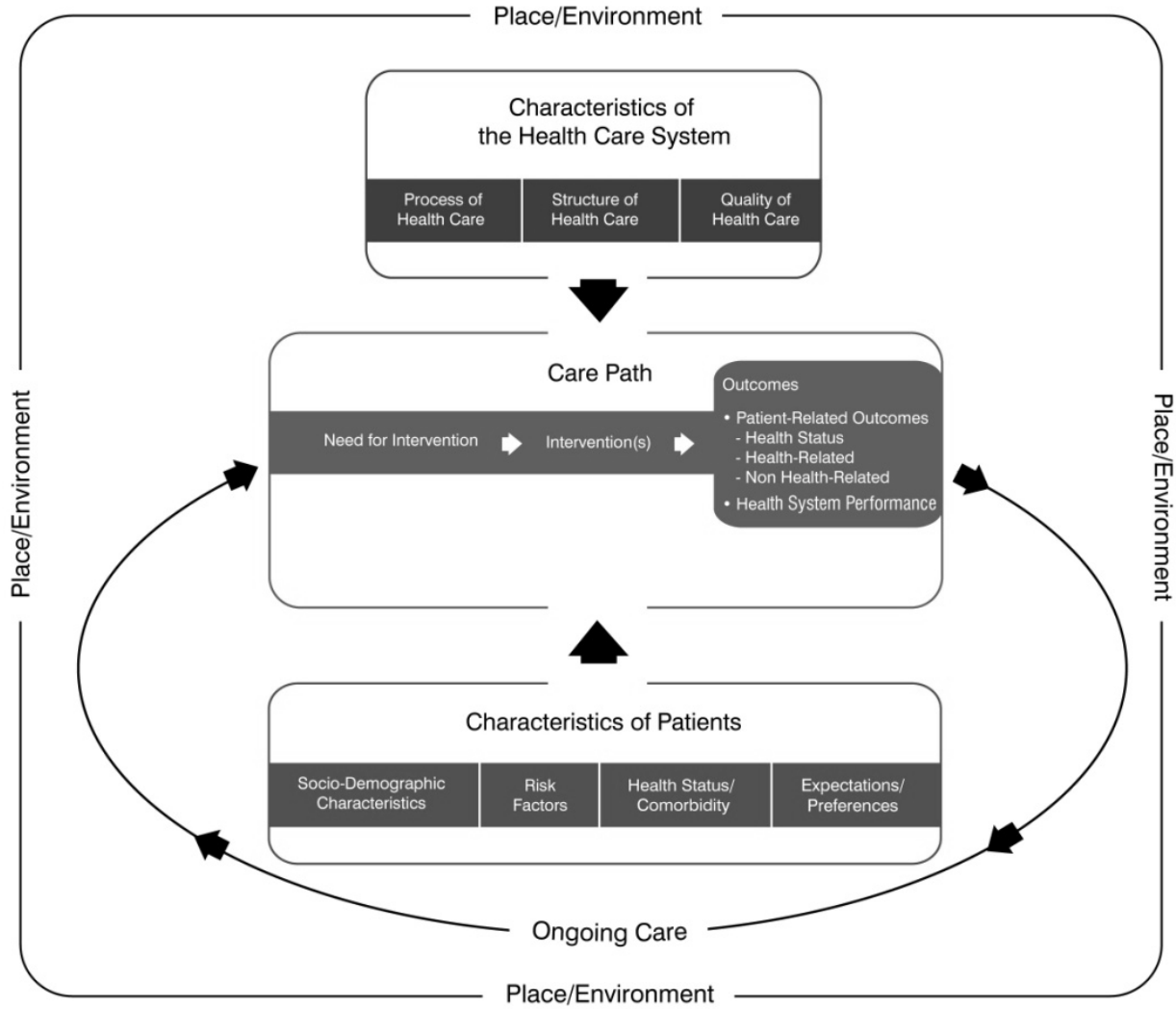
12. This diagram shows health status outcomes in the box at the middle-right, along with other outcomes that are not central to health policy, but nevertheless important to the population – such as employment status, or are important to government as manager of the health programmes and policies – such as system performance measures like costs per case (these are grouped together as "Non-Health Related outcomes" in the diagram). Figure 1 also shows as central the path of care followed by individuals – first determining who gets care, and then what kind of care they receive. These are the core health care processes, hence reflective of health care policy.

13. At the same time, the figure makes clear that other factors are at play, not least a range of individual patient characteristics, noted in the lower part of the diagram, and health care system characteristics noted in the upper part. For example, the success or failure of heart surgery cannot be ascribed only to the hospital and surgical team without also taking account of whether the patient is a smoker or seriously obese. On the other hand, it is clearly insufficient to use only measures at the top of the diagram, such as doctors and hospital beds per capita. While these are obviously important, they are only intermediate to the core information requirement – the central care path drawn from individual-level data on patient trajectories through their various health care encounters. The fundamental outcome of a health care encounter can only be assessed by comparing the individual's health status (using a standard metric like the question set underlying QALYs) before and after these healthcare interventions, where hopefully there has been an improvement.[8]

[7] Still, in this regard, HALE is no different from main macroeconomic indicators such as the growth rate of GDP and the unemployment rate. These too, while clearly the object of macroeconomic policy, are driven to a substantial degree by factors beyond the control of governments.

[8] This 'before' and 'after' health status comparison is essential for discrete interventions like a cataract or hip replacement procedure. In other situations, like chronic disease management, a longer series of data points on each individual's health status and encounters is needed.

Figure 1. Pathways of influence on health-care outcomes



14. These kinds of data would serve to connect health outcomes to health care policies – such as the kinds of surgical and pharmaceutical interventions funded, how hospitals and health care delivery are organised, and the remuneration of healthcare providers. But the diagram also indicates a broader range of factors, such as the socio-demographic status of patients, and more generally the "place / environment" within which these events unfold. For example, there is overwhelming evidence that even after controlling for well-known risk factors like smoking and physical inactivity, health status varies systematically with individuals' income (de Looper and Lafortune, 2009). And, while it is not as strong, there is also evidence that living in a poor or segregated neighbourhood also has adverse health effects. Health care providers should not be held accountable for poorer health outcomes if the individuals they are treating are themselves at higher risk for poor outcomes due to non-medical factors such as these. Unfortunately, data for taking account of these broader non-medical determinants of health rarely exist in complete form. Most often, they are at best piecemeal; so a priority for connecting outcomes to policies is the gathering of more complete and coherent covariate data.

15. While such data are fundamental for connecting health outcomes to "health affecting" policies (both intended and inadvertent); they are not sufficient. Data like these must be used to assess the strength of connections – between and among various processes reflecting the implementation of policies, and to

outcome measures. The first step in doing so is typically statistical analysis – where the outcome of interest is "explained" (*e.g.* as in a regression equation) by a mix of process measures and other exogenous factors. This step is very helpful in sorting out which factors co-vary importantly with the outcome of interest, and hence are candidates for some sort of causal role in the outcome.

16. In the health area, this is the focus of epidemiology and health services research. Indeed, there is a huge industry of randomised clinical trials whose objective is precisely to connect specific policies, most often approval for prescribing a drug, to health outcomes. Unfortunately, nowhere near a corresponding effort is devoted to connecting the largest portion of health-care activity – ranging from acute in-patient hospital care, such as surgical procedures, to rehabilitation – to health outcomes in an analogous manner.[9] This is the largest gap in the statistical and methodological tool kit that policy makers need in order to connect policies and their attendant processes to outcomes in the health area.

17. Many of the non-health care factors which are clearly determinative of health outcomes as summarised in HALE are amenable to public policy. These range from diet and physical (in)activity to poverty. It is therefore important to have means for disentangling the effects of a number of co-evolving factors. The standard approach to this problem in a wide range of fields – from cosmology to global climate change – is computer simulation modelling. In economic policy areas such as fiscal and industrial policy, this includes both time-series macro-econometric and input-output table based models. For income tax and unemployment insurance policy, there is a decade long practice of using detailed micro-simulation models.

18. Correspondingly, in the health area, micro-analytic approaches are also necessary, given the important heterogeneity of both the population and the range of health-care interventions.[10] It is for these reasons that Statistics Canada, on behalf of a range of partners, has developed the POpulation HEalth Microsimulation (POHEM) model. This model has been used, for example, to support policy decisions on colorectal cancer screening (Public Health Agency of Canada, 2002), to show the difference between clinical trial results and the prospective impacts of a drug in real world settings (Will *et al*. 2001), and to explore the relative impacts of both conventional (smoking and obesity) and broader (education, psycho-social) risk factors (Wolfson and Rowe, 2009).

19. In sum, better connecting health outcomes to policies in the health area requires:

- agreement and adoption of a standard measure of health outcomes, where HALE is proposed;
- coherent and wide-ranging data collection, with an explicit longitudinal micro-data foundation, and;
- extension of standard analytical methods beyond aggregate trends, correlations and regressions to include micro-simulation modelling of disease onset, progression, treatment, and functional health outcomes, these simulation models should be established at the individual level and embedded in a framework that encompasses the full range of causal factors.

---

9 Nor is anywhere close to the same effort devoted to monitoring the health impacts of pharmaceuticals, once they have passed their initial regulatory approval. The incentives facing firms, patients and governments play a role in the asymmetric treatment of data collection and analysis of drugs pre- and post-marketing.

10 In fact, micro-analytic approaches would be highly beneficial for macro-economic modelling, as argued over half a century ago by Orcutt (1957) and for their underlying statistical data, *i.e.* the System of National Accounts (*e.g.* Ruggles and Ruggles, 1967).

Literacy and Learning

20. Perhaps the best examples of summary outcome measures in the area of literacy and learning are the overall literacy rates estimated from the OECD's International Adult Literacy Surveys (IALS, and its successor Programme of International Adult Assessment Competency PIAAC) and the standardised achievement scores from the OECD Programme of International Student Assessment (PISA). These surveys have been internationally harmonised, and use, as their conceptual foundation, a graded set of tasks that are typical of and necessary to living in a modern complex society. For example, IALS builds on questions about being able to read and understand a public transit schedule and a drug prescription.

21. Literacy, when understood in this sense, is correlated with, but not at all identical to educational attainment (or the number of years of schooling completed). Still, education is the most important factor determining literacy. Thus, from the perspective of connecting learning outcomes to policy, data on the educational process are fundamental to understanding and quantifying the relationship between public policies and main outcomes.[11]

22. Obvious areas of education policy include numbers of years of mandatory school attendance, class size, and curriculum. As with health care, however, general data on average numbers of years of schooling completed, and average class size, are only a beginning for understanding literacy outcomes. For example, it is important to distinguish at least three kinds of curriculum. The "delivered curriculum" is the content and material actually covered in the pupil's classroom over the school year, while the "prescribed curriculum" is the content teachers are told to cover by central authorities. The two are often not identical. The third is "home curriculum", a shorthand used by education researchers to describe the home environment as it pertains to children's learning.

23. Data on the home curriculum can be gathered via household surveys, and include such items as household structure, parental educational attainment (*e.g.* as a marker of the extent of parental vocabulary), numbers of books in the home, whether the child participates with the parent in reading, and parental attitudes toward education. Parents' socio-economic status, including their income, occupation and migrant status, has consistently being reported as among the most important determinants of students' literacy. Ideally, data on the delivered curriculum are drawn from sampling actual classrooms and the teachers' lesson plans, rather than relying only on prescribed curriculum documents.

24. One of the most controversial but essential items in measuring the educational process is standardised testing. Teachers often complain strongly that such testing distorts their teaching practices. And it is certainly possible to overdo standardised testing. However, this testing as it is most widely practiced is seriously inadequate from the perspective of understanding and connecting educational policies and processes to literacy and learning outcomes. As currently practiced, standardised testing typically occurs once per school year, and is structured at best as a repeated cross sectional data collection.[12] This approach allows monitoring general patterns of achievement from year to year, or comparing across schools or school districts. It may, for example, allow some analysis of the extent to which class size affects learning outcomes. But it is still seriously limited for judging what factors are most important in determining pupils' literacy and learning. As with health status, several factors are at

---

11 Of course, the education system has many objectives, with literacy being only one. Others include socialisation, citizenship, problem-solving skills, and a range of specific competencies. This section starts with one widely accepted outcome, literacy, and works back to the policies and processes which are most relevant.

12 In some jurisdictions, there are many standardised tests conducted over the school year. But this does not change the fact that they are still almost always repeated cross-section data collections, and are not longitudinally linked to enable tracking individual pupils' progress in learning.

work – ranging from the home curriculum, teacher characteristics, classroom characteristics (*e.g.* the social and behavioural characteristics of the other students in the class), and school characteristics (such as "school climate", *e.g.* approaches to bullying, emphasis on teamwork). The only way to disentangle this range of factors is with a longitudinal survey design, as in the National Longitudinal Survey of Children and Youth sponsored by Human Resources and Skills Development Canada and run by Statistics Canada.[13]

25. Moreover, it is far better to assess "learning gain" for each pupil during a school year than at a discrete point in time. This requires administering the same standardised test both at the beginning and at the end of the school year, with the key outcome variable being the change ("gain", hopefully) in test score over the year. (Since children often forget quite a bit over summer holidays, it is also important to have the end of year results from the previous school year for comparison.) Such longitudinally linked standardised testing however is rare. Understandably, teachers are resistant, because it is the only rigorous method to ascribe learning outcomes at the level of classrooms and hence individual teachers. This is clearly evident with the recent controversy in the Los Angeles school system, where freedom of information laws were used to gain access to detailed data on students' learning gains. These data were then associated with named teachers via a sophisticated statistical analysis and the results were published by the Los Angeles Times, albeit violating the teachers' privacy rights according to some (Buddin, 2010; LA Times, 2010).

26. Of course, from a public policy (and parent) perspective, these data would be highly valuable. To be fair to teachers and their school systems, learning gain as a dependent variable in a regression analysis should be examined not only as a function of classroom and teacher characteristics, but also in terms of parental characteristics including the home curriculum. Teachers or the school board should be held accountable for their contributions to the learning process, but not for the strengths or weaknesses of their students attributable to aspects outside the school – such as their pupils having more highly educated parents on average, or income inequalities rising over time

27. Beyond the age of compulsory schooling, it is also important to understand why some children pursue further education, and of what kind. This entails collecting longitudinal data on young people's educational trajectories through trade school, community college, and other forms of post-secondary education (PSE). For example, studies in Canada using the Youth in Transition Survey combined with the OECD's PISA standardised test results have shown that there is a clear gradient by parental income and education in the likelihood of pursuing post-secondary education – a factor not easily amenable to public policy. But at the same time, the geographic distance between students' residences and post-secondary institutions is also significant – a factor with is clearly under the control of governments (Frenette, 2007). And of course, post-secondary educational financial programmes, such as subsidised tuition, and student loan and grant programmes, are major components of education policy.

28. Beyond the formal educational system, literacy is also influenced by formal and informal on-the-job training, as well as by individuals' own behaviours, particularly the extent to which they continue reading. Since there are so many strands of influence on adult literacy, longitudinal microdata covering all of the main driving factors are needed to understand the main determinants.

29. Moreover, in areas such as post-secondary student finance, it is highly desirable to be able to link policies such as the structure of student loans to outcomes such as post-secondary education attendance, completion and subsequent labour market experiences.

---

[13] This was also a "snowball" or network survey in that the starting point of the survey was an individual child, but data were also gathered from the child's network of "significant others" – parents, teacher and school principal.

30. In comparison to health, where at least substantial longitudinal data sets from both household surveys and administrative records on health care encounters have been assembled, the situation in the area of literacy and learning is generally not as strong. On the one hand, summary and internationally comparable measures of literacy exist, unlike internationally comparable summary measures of functional health status. On the other, longitudinal data, with its capacity to support analysis of the factors driving transitions dynamics – *e.g.* dropping out of school, learning more while at school, attending post-secondary education, using skills and knowledge gained in schooling in subsequent labour market activity, retaining or losing literacy skills – are unfortunately rare.

31. Given this weak empirical base, there is a corresponding absence of simulation models which, as in health, provide the best methodological approach for connecting policies to summary measures. We are unaware, for example, of a quantitative analysis that examines within the same overall framework the comparative costs, benefits and distributional impacts of greater early childhood development programmes versus different integration strategies for special needs students, or changes in high school class sizes.

32. In sum, better connecting literacy and learning outcomes to policies in the education and training area requires:

- continued use and expansion of internationally comparable measures of literacy and competencies (*e.g.* cross curricular competencies, OECD, 2004)

- collecting data on these measures longitudinally, and in combination with their main correlates, in order to construct quantitative measures of learning gain (literacy loss) and of the impacts of various causal factors; and

- as better longitudinal data are developed and analysed, constructing micro-simulation models of learning dynamics embedded in a framework that encompasses the full range of causal factors in order to provide the basis for evaluating trade-offs in areas of educational investment.

Economic Security

33. The economic security of a society's population can be characterized in many ways; and complete economic security is unattainable (and arguably is not even desirable). However, there are two broad concepts of economic security on which there is wide agreement. One is the avoidance of very low income – an anti-poverty objective. The other is general maintenance of living standards through economic vicissitudes – *e.g.* recessions and consequent unemployment, or illness and consequent health care costs, or disability followed by inability to work, or over the life cycle, especially during retirement.

34. Providing economic security of these sorts is central to many government policies, including welfare, unemployment and disability insurance, publicly funded or supported health insurance, and public pensions. Summary outcomes related to these policies and programmes can be measured.

35. First, when looking at anti-poverty outcomes, data from cross-sectional household income surveys, or from routinely collected administrative data such as income tax returns (provided those with low and very low incomes still provide data to the administrative authorities) can be used to provide head counts of the numbers of individuals and households falling below a low income line.[14] In addition to overall measures of income poverty, it is also necessary to be able to "drill down" and provide

[14] More sophisticated poverty indices also take account of the "depth of poverty", *i.e.* not only whether a family's income is below the poverty line but by how much it falls below that threshold. (Foster, Greer, Thorbecke, 1984)

distributional tables – *e.g.* the numbers of individuals with incomes below the "poverty line" broken down by age, family type, geographic area, precipitating circumstance, and how far below the line they are. It is also helpful to examine those who are just above the line, but vulnerable to falling into poverty if there is a change in the contingencies they face (*e.g.* due to heavy indebtedness and a possible increase in interest rates).

36. The second objective, maintenance of living standards, requires longitudinal data, though the time scales are different for the economic security issues most common in working age (*e.g.* unemployment, illness) and in older age (*e.g.* retirement, widowhood). For policy domains like unemployment, longitudinal data over a few years are sufficient, though a sub-annual frequency, (ideally weekly or monthly, but at least quarterly), would be more useful. In contrast, for pension and disability policy, data spanning decades are needed in order to measure properly the key outcome indicator, the "replacement rate" – the ratio of disposable or "consumable" income after the 'event' (*e.g.* becoming disabled, retiring from paid work) to the corresponding income before the event. Also, there are often a number of government tax and transfer programmes that interact; and family circumstances matter, so that ideally, replacement rates should be estimated empirically on a "net" basis (*i.e.* after all taxes and transfers as well as retirement saving and dis-saving) and taking account of family or household circumstances.[15]

37. Such net replacement rates require careful data collection and most often cannot be derived from programme data alone. There is also considerable sensitivity in the results to rather subtle choices in the construction of the replacement rate measure, such as the discount rate, treatment of varying family circumstances over the period of years (*i.e.* the equivalence scale used), and the specific range of years used for the numerator and denominator (Wolfson, 2011).

38. And as for low income, a summary index, such the proportion failing to achieve "adequate" replacement income, is insufficient. It should be complemented by a range of distributional statistics such as replacement rates by income group (Larochelle-Cote, Myles and Picot, 2010), and results showing the sensitivity to various analytical assumptions (*e.g.* discount rate, threshold used to define "adequate living standards").

39. The most obvious policies and programmes affecting these two broad outcomes are government funded income transfer programmes such as unemployment insurance and welfare for those of working age, and public pensions for the elderly. The role of publicly funded or supported health insurance is also fundamental. However, there are many other programmes and policies that have important effects on these outcomes. These may include labour market regulations (*e.g.* re layoffs), tax incentives (*e.g.* the tax expenditures in respect of workplace pension plans and individual retirement saving, refundable income tax credits in some countries), housing and childcare subsidies, disability policies, and fiscal and monetary policy.

40. Since economic security is only meaningful in the context of risks of various sorts, it is also essential to assess the benefits of various social insurance programmes designed to mitigate risks in the context of various scenarios. For example, how well does the public pension system do in supporting maintenance of living standards if inflation rates are higher than expected, or if per capita economic growth is slower, or if life expectancy were to increase more rapidly than the baseline scenario, or individuals were to continue working to later ages than currently anticipated? This kind of policy assessment consists

[15] The OECD has been using micro-simulation models to calculate replacement rates for different income groups and household types; these micro-simulation models refer to both unemployment and income-replacement benefits (see OECD, 2007, *Benefits and Wages*) and pensions (see OECD, 2011, *Pensions at a Glance*).

essentially of posing and then answering a series of "what if?" questions. Doing so requires some sort of simulation modelling capacity, as illustrated for the last two kinds of risks in Wolfson and Rowe (2007).

41. A major challenge in this area is assessing behavioural responses. It is not the case that the effects of a given program can be simply measured by the amount of transfer income (say) that it provides. This amount can be considered the "first round impact" of the program, but it is likely that individuals would change their behaviour in the absence of a given program or policy. For example, in the absence of public pension programmes, a great many individuals would certainly have much lower incomes post retirement. But it is also likely that there would be additional private saving for retirement. Similarly, in the absence of unemployment insurance, individuals would likely be less inclined to take seasonal jobs. However, the state of knowledge for quantifying behavioural responses to these kinds of policy changes remains unsatisfactory, notwithstanding decades of published econometric studies. A key reason is the propensity of mainstream economic analysis to start from psychologically unrealistic assumptions of individualistic utility maximisation and unduly simplified assumptions about the range of real world individual choice. The best option in this unfortunate circumstance is more pragmatic and heuristic sensitivity analysis, since improvements in the state of knowledge in this area do not appear likely in the near term.

42. In order to connect programmes and policies to outcomes in this case as well, micro-simulation models are essential. In some areas, like individual income tax policy, this has been the norm in Finance ministries in many OECD countries since the 1960s. Every potential income tax change is run through the tax simulation model to determine its overall fiscal impact as well as its distributional impacts, for example by income level, family composition, and geographic region. In these situations, cross-sectional tax/transfer models are typically be used. Good examples include the Social Policy Simulation Database and Model in Canada (Statistics Canada), TRIM in the United States. (Urban Institute), STINMOD in Australia (NATSEM), and EuroMod in Europe (ISER). Behavioural changes can be assessed "out of model" by, for example, examining not only the estimated first round impacts on individuals in various situations, but also the changes in their effective marginal tax rates, which these kinds of models can also estimate.

43. Micro-simulation modelling provides the same potential for connecting policies to outcomes in other areas, such as public pension policy. But in these cases, more complex longitudinal micro-simulation is essential. These models, in turn, require more extensive data, especially longitudinal data on individuals' demographic and income trajectories. Quite a few OECD countries have such models (*e.g.* Flood, 2008; Blanchet *et al*.) and some cross country comparative analysis has been developed in Europe with the EU funded Euromod project (ISER). But capacity for producing broad summary measures of economic security, especially in a longitudinal or life cycle context, and then connecting these outcomes to policies and programmes, remains weak in most OECD countries.

As for the policy areas already discussed stronger development of longitudinal data and simulation models – as well as the professional staff to design and collect the data, undertake the analysis, develop the simulation models, and then provide the policy advice – is needed.

Economic Inequality

44. The Stiglitz-Sen-Fitoussi Commission report (2009) has highlighted economic inequality as a major area where greater prominence should be given to routine and standardised measurement as the basis for assessing societies' progress. Major investments since World War II have resulted in high quality, standardised and regular measures of the "size of the pie" in each country, *i.e.* GDP via the System of National Accounts. But nowhere close to as much international effort has been devoted to regular, standardised, and high quality measurements of "how the pie is divided".

45. The general approaches to measuring outcomes in the previous three areas (health, literacy and learning, economic security) are reasonably clear and well-accepted. But there are greater challenges and less consensus in defining and measuring the "basic outcome" for economic inequality. These challenges can be structured under several headings:

- Degree of inequality. While there is broad agreement that extreme inequality is undesirable, there is also agreement that perfect equality (in outcomes, rather than opportunities) is not only unattainable but also likely undesirable. Thus, unlike average measures of health status, literacy and economic security , where almost always more is unequivocally better, defining a norm for inequality is a major challenge

- Income, wealth or consumption. There are several reasonable, but conceptually distinct measures of the "economic position" for individuals and households, so it is an open question which should form the basis for a summary measure of a country's degree of economic inequality. The main options are household income, wealth or net worth, and consumption. Further, there are gray areas, such as owner-occupied housing, where some account should be taken not only of its asset value as a component of net worth, but also of the flow of benefits derived from this asset – *i.e.* imputed rental income in SNA terms. Perhaps the focus should be on joint distributions such as that for income and wealth, possibly using an adjustment like the annuity equivalent income flow that could be generated by the stock of wealth (Projector and Weiss, 1969; Wolfson, 1979).

- Treatment of pension wealth. While many pensions are explicitly funded, the connection between funds and future benefits is often partial and often obscure at the individual level. Moreover, there are typically decades long periods between the "now" when economic inequality is being measured, and the future when the pension benefits will be paid, which in turn begs the thorny question of how to deal with time discounting.

- Accounting period. Most often, individual and household income data are based on an annual accounting period. But the economic notion of permanent income entails a multi-year accounting period, while for other purposes a lifetime perspective may be most appropriate.

- Household economies of scale. How should the fact that two people can live more inexpensively together than apart be recognized when assessing economic inequality? The usual method is to use an equivalence scale, but the current OECD equivalence scale is quite different from the one most commonly used by the research community (OECD).

- "Boundary" of the household. Household survey techniques usually dictate that the reporting unit is either all the residents in a dwelling unit, or a subset based on relationships (*e.g.* parent, spouse). Data from administrative sources are usually even more restricted in terms of the family members included (*e.g.* excluding adult children still living at home). However an individual's economic circumstances may be substantially influenced by extended family members and "significant others" who live in another dwelling, and hence are not included in the usual statistical data.

- Government-provided benefits in kind. The usual income data exclude the value of government benefits provided "in kind" such as health care, educational services, and housing subsidies. But variations in policies across jurisdiction in this "income packaging" can have a major impact on measured economic inequality (Smeeding *et al*., 1993)

- "Hidden" income and/or wealth. Especially for the wealthy, much of their income accrues in forms that are "not theirs" per the definitions used in official statistics, even though they

represent very real sources of "command over resources" – for example income received in closely held companies, trust funds, stock options, deferred income, tax havens, and in the form of company provided perks.

46. Decades of accumulated research evidence have shown that the choices for each of these factors can have a substantial impact on measured inequality. As a result, these challenges in measuring economic inequality are best addressed by first achieving international consensus on the relevant concepts, definitions, and measures. An important component is the recently updated manual from the Canberra Group (2011).

47. The second (much more expensive) step is a careful approach primary data collection. For example, to capture the economic positions of those at the top of the economic ladder, both a stratified sample with an added and specially designed high income stratum is needed, as well as a snowball sampling approach to gather information on pension entitlements and closely held company income and assets. In Statistics Canada's experience, many higher income respondents do not really know their economic position, and refer the interviewer to their accountant. Proper measurement of economic inequality requires that the survey budget include funds for specially trained individuals who can follow up in cases like these.

48. Once higher quality data are available on the outcome of interest, *e.g.* inequality in equivalised household disposable income, the related policy information can be gathered in a manner analogous to that for economic insecurity – for example, data on income tax liabilities and the key provisions affecting individuals at various points along the income spectrum. In turn, the first-round fiscal and distributional impacts of government policies and programmes, including various tax provisions, can be assessed using tax/transfer micro-simulation models like Statistics Canada's SPSD/M (Statistics Canada, no date), which are standard in policy departments.

49. Wealth inequality raises further challenges from the perspective of connecting government policies to outcomes, since wealth at any point in time essentially represents the cumulative effects of savings and wealth transfers (*i.e.* bequests and gifts inter vivos). In this case, longitudinal data on savings, owner-occupied housing, private pension accumulations, wealth transfers, and the structures of wealth taxation (*e.g.* estate, inheritance and gift taxes) must also be integrated.

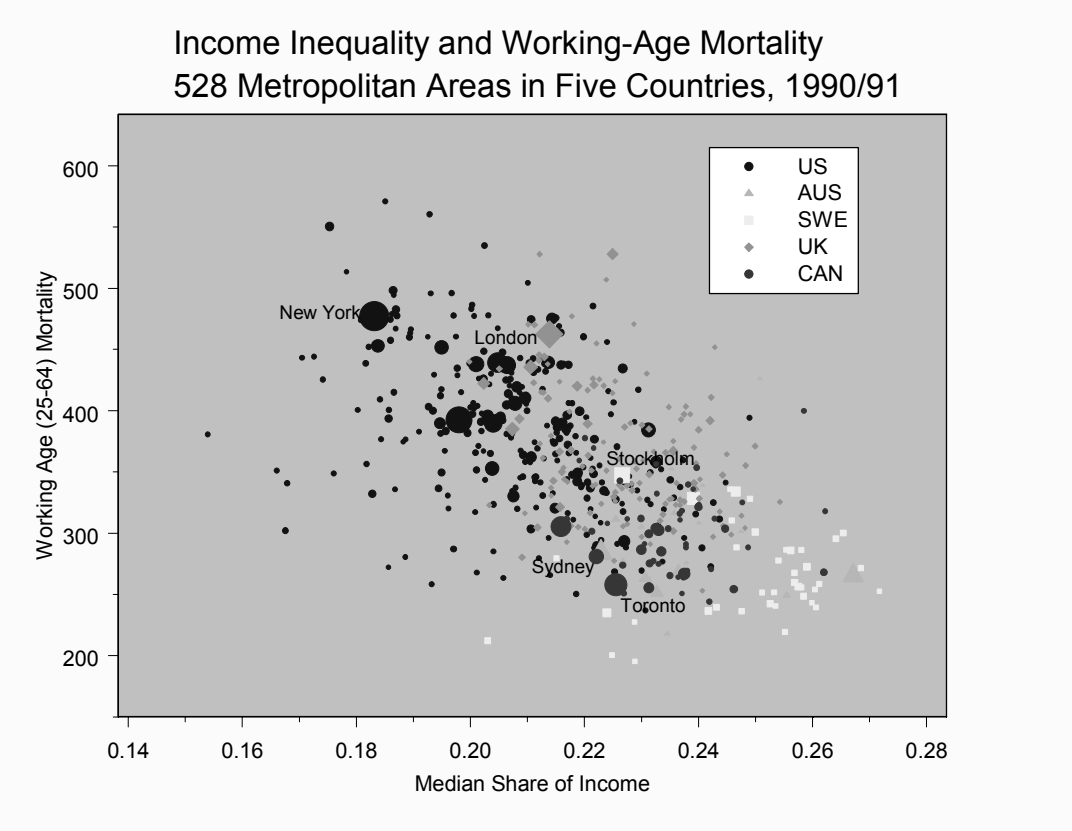
50. Most OECD member countries have reasonable data on the extent of income inequality (though not wealth or consumption inequality). And in many countries, recent data show a significant increase in inequality, particularly in the upper tail of the distribution. Whether this is considered a problem to which public policy should pay attention is often a matter of political inclination, and is often controversial.

51. To some degree, the connections between economic inequality and policy are clear. For example, if more income equality is desired, then increased cash redistribution via existing programmes like welfare, public pensions, publicly funded health insurance, and income taxes are obvious policy levers. But there tends to be more political support for equality of opportunity (*ex ante*) than for equality of outcome (*ex post*). In this case, the policy levers are less clear – though full employment and universal education are obvious contributors.

52. Perhaps more interesting are the connections between economic inequality and other broad social outcomes, like economic growth and health status. There have been lively academic debates in both cases, for example whether more economic equality is good for economic growth (*e.g.* Persson and Tabellini, 1994; Ravallion, 2001) or for population health (*e.g.* Wilkinson and Pickett, 2009). In these cases, the only empirical approach is by observation of natural experiments, either international comparisons or comparisons among sub-national entities like states or cities. For example, Ross *et al*. (2004) in Figure 2

show that there is a clear correlation between average life expectancy and income inequality across cities in some countries (United States and United Kingdom), but not in others (Canada, Australia, Sweden). While this kind of association was challenged as a statistical artefact due to an ecological fallacy (Gravelle, 1998), Wolfson *et al.*, (1999) showed this could not be the main explanation, and Backlund *et al.* (2007) reinforced this result in an analysis of the 50 U.S. states. Still, one lesson from this recent exchange on the relationship between economic inequality and average health outcomes is the importance not only of natural experiment data but also of multi-level data – both on individuals and on their socio-economic contexts. Of course, a fundamental pre-requisite is that the data across (natural) experimental units, be they countries or states or cities, be comparable.

Figure 2. Income inequality and working-age mortality



53. One final point worth noting is the current disjunction between the data sources for the "size of the pie" and "how the pie is divided". The size of the economic pie is based most often on the SNA, with GDP as the most often used measure, while the division of the pie is typically based on household survey data. This incoherence was emphasized by the Stiglitz-Sen-Fitoussi (2009) commission, who recommended that both be derived from household data – not least because the GDP per capita data lack fundamental face validity for many individuals. As the seminal Atlantic Monthly title that introduced the Genuine Progress Indicator put it, "if the economy is up, why is everyone down?" (Cobb *et al.*, 1995). The Stiglitz-Sen-Fitoussi recommendation was to shift to median household income as the main summary measure of the economic situation of the typical person or household. The main summary measures of economic inequality can then be defined coherently in terms of the same underlying household distribution of income. Furthermore, it would be most beneficial if the SNA and household income distribution statistics were brought into coherence with one another by creating a proper household sub-sector of the

personal sector, and aligning the household survey and SNA data. The feasibility of such sectoring and explicit microdata foundations for the SNA (also in line with the Ruggles' (1967) recommendations noted earlier) was shown over 20 years ago by Wolfson and Adler (1988).

54. In sum, internationally comparable data for summary measures of economic inequality are reasonably available for income, but not for wealth or consumption. Moreover, there is a fundamental lack of coherence between SNA income and the income measured for inequality purposes. These two sources should be made coherent, for example by following the Stiglitz-Sen-Fitoussi (2009) proposals to use median household income as the starting point. Extensions of the SNA in this direction would also be desirable. For other measures of inequality, further work on internationally agreed concepts and definitions, along the lines of the Canberra Group (20011), is needed.

55. Beyond the measurement of economic inequality, perhaps the most important challenge is to exploit natural experiments across jurisdictions, and hence across policy regimes, to determine whether and the extent to which economic inequality has other major socio-economic impacts – for example on economic growth and population health. This will require much greater attention to comparable multi-level data.

56. The four areas explored in this paper are central to people's well-being and social progress but they are not the only ones that matter for people's life. Box 1 briefly looks at time crunch, a well-being outcome whose conceptualization is not as developed as for the previous four areas. While data in this field are more limited, the type of approach outlined in the previous sections could also be applied here.

Box 1. Time Crunch

Patterns of time use, and the widespread feelings of "time crunch", while recognized in principle, have received far less attention than the previous four areas in the context of broad measures of progress and well-being. Typical hours of work dropped substantially up to the middle of the 20th century in leading OECD countries, though as recently as 2004, Korea had annual work hours almost twice those in France. While some 1950s futurists predicted that by now, the work week would be down to 20 hours, this has clearly not occurred. Instead of a shorter working week, the benefits of increased productivity in terms of patterns of time use have been primarily realised in the form of longer paid vacations each year and fewer years spent in paid work over people's lifetimes – with more time being spent at younger ages in education, and even more so at later ages in retirement. Still, there is no question that societal progress is impaired if there is a general sense across the population of being stressed because of insufficient time.

The pervasiveness of these feelings is clearly revealed in surveys when the question of "time crunch" is asked (*e.g.* Allen *et al*.). Yet another indication of concerns about available time is the broadly negative reactions of most ordinary people to questions about policies to "improve productivity" in public opinion polling. In Canada, people often associate (sometimes mistakenly) the notion of increasing productivity with working longer hours.

As in other areas, the first step to improve policies is to improve measures of outcomes. The straightforward method is to include questions on the prevalence of "time crunch" on household surveys on a regular basis. Additionally, given the growing literature on subjective well-being (SWB), it is important that questions on a range of covariates – both standard socio-economic variables, and various measures of SWB and stress, are included. To connect such feelings with public policy and programmes, questions on time crunch should be embedded in time use surveys, *e.g.* through 24 hour diary recall surveys.

One analytical challenge is to separate sources of reported time crunch between those which reflect the (thoughtful, unconstrained) choices of the individual from those where social actions (or their absence) play a major role. For example, a person may have chosen to pursue too many hobbies, in which case their time crunch is "self-inflicted". But the person may also be a "tweener" with both a rambunctious child and a frail elderly parent to look after, in which case absence of adequate childcare and/or homecare for the frail elderly, *i.e.* government programmes, may be the main factors accounting for their stress and time-crunch. The analysis is further complicated by the fact that different life decisions play out over different time scales. For example, the decision on what hobbies to pursue is one that can be changed on almost a daily basis. In contrast, the decisions to have a professional career and to have children, especially for women, are ones whose implications play out over decades.

Common threads and opportunities

57. The examples in the areas above share several common threads, and these can be used to indicate important opportunities for providing the empirical foundations for stronger connections between policy and program processes on the one hand, and well-being outcomes on the other.

1. The first common thread in the substantive areas reviewed is that while indicators – whether of outcomes (*e.g.* health status) or describing policies (doctors per capita) are interesting, they are not nearly sufficient. It is essential to embed summary measures of outcomes and related indicators of processes associated with programmes and policies in a coherent data system. Indicators do not exist in splendid isolation; rather they are the summary distillation of a mass of more detailed data. Moreover, these underlying data should be designed and constructed to serve not only the derivation of a handful of summary indicators, but also to support more detailed analysis – a "drill down" capacity.

2. The specification of an appropriate set of indicators, and the supporting foundation of a coherent data system, should be based on a sensible and explicit set of concepts and definitions. Indeed, these concepts should hang together coherently, *i.e.* in a conceptual framework. For the kinds of policy processes being discussed here, international comparisons are potentially highly valuable. In effect, the variations in policies and processes across jurisdictions provide a series of natural experiments. In turn, to be able to learn from these experiments, it is fundamental that the data collected within each jurisdiction can be meaningfully compared across jurisdictions. Thus, international collaboration in specifying concepts, definitions and conceptual frameworks is essential. One pertinent example is that for income distribution developed by the Canberra Group (2011).

3. Another common thread is the fundamental importance of micro data. Only simple analyses can be supported by aggregated data like doctors per capita, average class size, average earnings replacement rates, Gini coefficient, and the proportion of individuals expressing that they are time crunched. It is very important to be able to drill down into the supporting data, for example to examine distributions and correlations. People in all jurisdictions are highly heterogeneous, and these differences almost always make a difference in program and policy impacts.

4. The processes of human development, and hence the impacts of programmes, unfold over time, so that cross-sectional snapshots, even repeated snapshots, while useful, are still inadequate. A life course perspective is often required; hence longitudinal microdata are needed, for representative samples of the population. Additionally, there is growing evidence of the very wide range of factors that typically have to be taken into account in various outcome analyses, in order to avoid, "omitted variable bias". Hence, the data need to be suitably multivariate, and in some cases – for example in the case of education, where data on all of pupils, their classrooms, their schools, and their neighbourhoods are likely salient – the data need to be multi-level. In this regard, international sharing of experience with regard to powerful and cost-effective data collection options would be beneficial – for example the judicious use of record linkage of both administrative and survey data.

5. The kind of data ideally suited to connecting programmes and policies to wellbeing outcomes is likely to be controversial with the providers of services, such as doctors and teachers, because it (almost uniquely) provides a strong empirical basis for systematically holding them accountable for the quality of the work they are doing. However, if outcomes are to be successfully connected to the relevant processes, it is inevitable that the resulting data and analyses will have the potential to be embarrassing to some of the providers of these public services, or to the policy

makers who define the structures of these and other major social programmes. The resultant resistance to data collection, while often subtle, is nonetheless important to understand, and to confront in a reasoned and persistent manner.

6. Finally, the relationships between wellbeing outcomes and programmes and policies are typically complex and multi-factorial. As a result, in order to disentangle and estimate the "effect sizes" (to borrow a term from epidemiology) of specific underlying processes on key outcomes, it is essential to build simulation models based on the detailed microdata. Then, using the model, it is possible to pose and rigorously answer carefully constructed "what if" questions. By comparing these results with the baseline situation, it is then possible to ascribe effects on outcomes to specific processes or components of processes. As noted above, Statistics Canada has developed a suite of micro-simulation models that illustrate these possibilities.

Conclusions

58. One often hears discussions conducted as if the design of a public policy, especially in areas like those reviewed above, were like civil engineering and designing a bridge. But the differences are fundamental. In the case of a bridge, knowledge of girder strength, and sophisticated methods like finite element analysis, mean that a structure that can withstand the anticipated loads can be easily designed. The bridge is then built (usually with a "safety factor" of multiples of the required girder cross sections), and it stands for a century or more.

59. Socio-economic policy is not at all like building a bridge. The state of knowledge is not nearly as strong, and the objects of the analysis – individuals, groups and institutions – all have dynamics of their own, unlike inanimate steel structures. Thus, the paradigm for effective socio-economic policy is much more an iterative loop than in case of bride building: such loop involves designing the policy; implementing it along with an appropriate statistical system for continuous monitoring and evaluation; engaging in routine monitoring and evaluation; and repeatedly adjust the policy (as well as the statistical information flows) based on accumulating experience; and repeating this sequence frequently. In other words, socio-economic policy is not a once and for all design exercise, but one of continuous adaptation. The summary measures and analytical tools and processes mentioned above are all a reflection if this fundamental approach to public policy.

60. The work spearheaded by the OECD over the past decade on "measuring the progress of societies", in combination with the recent report of the Stiglitz-Sen-Fitoussi Commission (Stiglitz *et al*., 2009) has greatly raised the profile and legitimacy of widespread and longstanding concerns with the conventional statistical measures of living standards. The two most widely used internationally comparable measures, GDP per capita and life expectancy, are certainly useful but nowhere near adequate for assessing social progress. A new and broader set of routinely collected and standardised measures is needed.

61. But the kinds of measures needed to monitor social progress are often at such a high level that government Ministers and other policy makers may feel frustrated since the policy tools available to them do not connect with the problems revealed by these summary measures. We have sketched, in the discussion above, how these essential connections can be made. They require investments in infrastructure: – better data, more extensive analysis, computer simulation models allowing rigorous answers to the various "what if?" questions entailed by careful policy analysis, and trained professional staff to undertake all these tasks. There are clearly many challenges, but the tasks are certainly feasible.

REFERENCES

Allen C, H. Dryburgh, D Horlor (2011), "Canada's General Social Survey on Time Use: Challenges and Potential", http://www.statcan.gc.ca/pub/89-622-x/89-622-x2006003-eng.htm (accessed Sept 5, 2011).

Buddin R. (2010), "How Effective Are Los Angeles Elementary Teachers and Schools?", http://www.latimes.com/media/acrobat/2010-08/55538493.pdf (accessed Sept 27, 2011)

Backlund E, G Rowe, J Lynch, MC Wolfson GA Kaplan and, P Sorlie (2007) "Income inequality and mortality: a multi-level prospective study of 521, 248 individuals in 50 US States", *International Journal of Epidemiology*.

Blanchet D, E Crenner and S Le Minez (2011e), "The New Version Of The Destinie Microsimulation Model, Increased Flexibility And Adaptation To Users' Needs", http://jms.insee.fr/files/documents/2009/93_2-JMS2009_S13-2_BLANCHET-ACTE.PDF (accessed Sept 5, 2011)

Canberra Group (2011), Handbook on Household Income Statistics, 2nd Edition (Final Draft), http://www.unece.org/fileadmin/DAM/stats/groups/cgh/FinaldraftCanberra.pdf; (accessed Oct 5, 2011)

Cobb C, T Halstead, J Rowe (1995) "If the Economy is Up, Why is America Down?", *Atlantic Monthly*, October 1995, vol 276, no. 4 also cover page.

De Looper, M. and Lafortune, G. (2009), "Measuring Disparities in Health Status and in Access and Use of Health Care in OECD Countries", OECD Health Working Papers No. 43.

Flood L (2008), "A Swedish Micro-Simulation Model" in A. Klevmarken and B. Lindgren (eds), *Simulating an Aging Population: A Microsimulation Approach Applied to Sweden*, Emerald Group Publishing ltd.

Foster J, J Greer and E. Thorbecke (1984), "A Class of Decomposable Poverty Measures", *Econometrica*, Vol. 52, No. 3, May.

Frenette M. (2007), "Do Universities Benefit Local Youth? Evidence from University and College Participation, and Graduate Earnings Following in the Creation of a New University", Analytical Studies Branch Research Paper Series, No. 283, Statistics Canada http://www.statcan.gc.ca/pub/11-015-x/edu/4154101-eng.htm

Gold MR, JE Siegel, LB Russell, MC and Weinstein, eds. (1996), *Cost-Effectiveness in Health and Medicine*, New York, NY, Oxford University Press.

Gravelle H. (1998), "How much of the relation between population mortality and unequal distribution of income is a statistical artefact?", *British Medical Journal*, 316: 382–385.

ISER (2011), Euromod http://www.iser.essex.ac.uk/euromod, (accessed Sept 5, 2011).

Larochelle-Cote S. J. Myles and G. Picot (2010), "Replacing Family Income During the Retirement Years: How Are Canadians Doing?" http://www.statcan.gc.ca/pub/11f0019m/11f0019m2010328-eng.pdf (accessed Sept 27, 2011)

Los Angeles Times, "Grading the Teachers: Value-Added Analysis", http://www.latimes.com/news/local/teachers-investigation/ (accessed Sept 27, 2011)

National Committee on Colorectal Cancer Screening (2010), Health Canada Recommendations for Population-based Colorectal Cancer Screening, http://www.phac-aspc.gc.ca/publicat/ncccs-cndcc/ccsrec_e.html (accessed Sept 27, 2011)

NATSEM (2011e), Static Incomes Model, http://www.canberra.edu.au/centres/natsem/research-models/projects_and_models/stinmod (accessed Sept 5, 2011).

OECD (2011), "What Are Equivalence Scales", http://www.oecd.org/dataoecd/61/52/35411111.pdf (accessed Sept 5, 2011)

OECD (2011), *Pensions at a Glance: Retirement Income Systems in OECD and G20 Countries*, OECD: Paris.

OECD (2007), *Benefits and Wages 2007: OECD Indicators*, OECD: Paris.

OECD (2004), *Problem Solving for Tomorrow's World – First Measures of Cross Curricular Competiencies from PISA 2003*, http://www.oecd.org/dataoecd/25/12/34009000.pdf.

Orcutt G (1957), "A New Kind of Socio-Economic System", *Review of Economics and Statistics*, Vol. 39, No. 2, May.

Persson T and G. Tabellini (1994), "Is Inequality Harmful for Growth?", *American Economic Review*, Vol. 84, No. 3 http://didattica.unibocconi.it/mypage/upload/48805_20081009_055128_IS_INEQUALITY_HARMFUL_FOR_GROWTH.PDF .

Projector D S and G. S. Weiss (1969), "Income – Net Worth Measures of Economic Welfare", *Social Security Bulletin*.

Public Health Agency of Canada (2002), Technical Report for the National Committee on Colorectal Cancer Screening, http://www.phac-aspc.gc.ca/publicat/ncccs-cndcc/techrep-eng.php (accessed Oct 13, 2011)

Ravallion M. (2001), "Growth, Inequality and Poverty: Looking Beyond Averages", *World Development*, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.17.8110&rep=rep1&type=pdf (accessed Sept 5, 2011).

Ross N. A, D. Dorling, J. R. Dunn, G. Henriksson, J. Glover, J. Lynch, and G. Ringbäck Weitoft (2004), "Metropolitan Income Inequality and Working-Age Mortality: A Cross-Sectional Analysis Using Comparable Data from Five Countries", *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, Vol. 82, No. 1.

Rowe G. and M. C. Wolfson (2009), "Simulating Life Course Dynamics of Functional Health" paper presented to the International Microsimulation Association 2nd biennial meeting, Ottawa.

Ruggles, R. and N. Ruggles (1967), "Data Files for a Generalised Economic Information System", *Social Science Information*, 6: 187-196.

Smeeding T., P. Saunders, J. Coder, S. Jenkins, J. Fritzell, A. Hagenaars, R. Hauser, M.C. Wolfson (1993), "Poverty, Inequality, and Family Living Standards; Impacts Across Seven Nations: the Effect of Non-Cash Subsidies for Health, Education and Housing", *Review of Income and Wealth*, Series 39, No. 3, Sept http://www.roiw.org/1993/229.pdf .

Statistics Canada (2011), Social Policy Simulation Database and Model – Overview, http://www.statcan.gc.ca/spsd/OVERVIEWE.pdf (accessed Sept 5, 2011).

Statistics Canada and CIHI (2008) "A Framework for Health Outcomes Analysis" http://secure.cihi.ca/cihiweb/products/health_outcomes_web_e.pdf_.

Stiglitz J., A. Sen, J-P. Fitoussi (2009), Report by the commission on the Measurement of Economic Performance and Social Progress, http://www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf (accessed Sept 7, 2010).

Sullivan DF, (1971), "A Single Index of Mortality and Morbidity", HSHMA Health Reports Vol. 86, No. 4

Torrance G.W, and Feeny D., " (1989), "Utilities and quality-adjusted life years", *International Journal of Technological Assessment in Health Care*; 5:559-575.

UNECE (2010) (United Nations Economic Commission for Europe) papers describing the Budapest Initiative, http://www.unece.org/stats/documents/2010.01.health.htm (accessed Sept 5, 2011).

Urban Institute (2010), "Transfer Income Model", http://trim3.urban.org/T3Welcome.php (accessed Sept 5, 2011).

Wilkinson R. and K. Pickett (2009), *The Spirit Level: Why More Equal Societies Almost Always Do Better*, Penguin.

Will B.P, K. M. Nobrega, J.M. Berthelot, W.Flanagan, M.C Wolfson, D.M. Logan and W.K. Evans (2001), "First do no harm: extending the debate on the provision of preventive tamoxifen", *British Journal of Cancer*, 85(9), 1280–1288

Wolfson, M. C. (1979), "Wealth and the Distribution of Income, Canada 1969-70", *The Review of Income and Wealth*, June 1979.

Wolfson, M. C. and H. Adler (1988), "A Prototype Micro-Macro Link for the Canadian Household Sector", *The Review of Income and Wealth*, December.

Wolfson M.C. (1996), "Health-Adjusted Life Expectancy", *Health Reports*, Statistics Canada, Vol 8, No. 1, Summer.

Wolfson M. C, G.A. Kaplan, JW Lynch, NA Ross, E Backlund (1999) "The relationship between income inequality and mortality is not a statistical artefact: an empirical assessment", *British Medical Journal*, Vol 319, 9 October.

Wolfson M.C. and G. Rowe (2007), "Aging and Intergenerational Fairness – A Canadian Analysis", in Peter Lambert (Ed.), *Research on Inequality*.

Wolfson M.C. and D. Lievesley (2007) "Making Progress in Health and Health Care", 2nd OECD World Forum on Measuring the Progress of Societies, Istanbul, June 2007 http://www.oecd.org/document/13/0,3343,en_21571361_31938349_38086797_1_1_1_1,00.html (see agenda at Friday 11am) (accessed Sept 27, 2011)

Wolfson M.C. and G. Rowe (2009), "Using Functional Health Trajectories to Quantify the Relative Importance of Selected Health Determinants", paper presented to the International Microsimulation Association 2nd biennial meeting, Ottawa.

Wolfson M. (2011), "On the Replacement Adequacy of Canada's Retirement Income System Estimates Using Statistics Canada's LifePaths Microsimulation Model", Institute for Research In Public Policy, Montreal http://www.irpp.org/pubs/IRPPstudy/IRPP_Study_no17.pdf **(**accessed Sept 27, 2011**).**

World Health Organisation (2010), *International Classification of Functioning, Disability and Health* (ICF), http://www.who.int/classifications/icf/en/ (accessed Sept 27, 2011).