ISBN 978-92-64-03945-2
How Regions Grow: Trends and Analysis

Chapter 3

Assessing the Impact of the Main Determinants of Regional Growth: A Parametric Analysis

Introduction

Chapter 1 described how, between 1995 and 2005, there was a significantly greater disparity in growth (three times larger) across OECD regions than across countries. Furthermore, we showed that growth does not occur uniformly within similar types of regions (*i.e.* predominately urban and rural regions). Not only was there a significant number of urban regions growing faster than rural regions, but also a significant number of rural regions out-performed urban regions in terms of GDP per capita growth rates over that decade.

The wide variation in economic performance among types of OECD regions reflects the regions' great heterogeneity in levels of income, rates of employment, mixes of high and low productivity activities, endogenous and exogenous assets, comparative advantages, stages of development and public policies. Therefore growth at the regional level results from a complex set of interconnected factors. Chapter 2 broke GDP growth down into a series of components (both endogenous and exogenous) so as to compare the performance of each region in terms of GDP growth associating common patterns in the decomposed components among successful and unsuccessful regions.

This chapter supplements that analysis. It investigates the impact of key structural endogenous factors on regional GDP growth while controlling for exogenous and national factors. The aim is to determine which factors are most relevant to generate growth at the regional level and which factors are needed if regions are to reap the benefits of globalisation. Are regions only required to improve their innovation capacity or do they also need to attract people, improve their infrastructure, and have an adequate labour market and business environment? Do regions need to improve only one of these factors or a bundle of them to remain competitive? We also distil the main drivers of growth and connect them with spatial aspects of agglomeration[1] to explore why some regions grow faster than others.

Main findings

We used four sets of analyses to explore trends and regional drivers of growth:

- A series of cross-section econometric models using elements stemming from neo-classical and endogenous growth theories along with the new economic geography (NEG, see Annex D).

 HOW REGIONS GROW: TRENDS AND ANALYSIS – ISBN 978-92-64-03945-2 – © OECD 2009

- Dynamic econometric modelling through panel data analysis which allows for spatial analysis to interact with time.
- Analysis based on a knowledge production function that relates human capital and research and development (R&D) to innovation outcomes such as patenting activity.
- Spatial econometrics, to add a geographical element to the classical econometric methods.

These models show that regional dynamism depends on endogenous factors such as infrastructure, education, innovation, economies of agglomeration and geographic characteristics:

- *Human capital*: regions with insufficient human capital will not grow, while those with increased levels will reap the benefits of endogenous elements of growth. For example, regions with a low rate of tertiary education are less economically vibrant than those with a high rate. The long-term effects of tertiary education on regional growth are also positive. Human capital has a strong impact on regional growth both directly and indirectly by increasing the rate of patenting.
- *Innovation, research and development*: research and development (R&D) is an indirect determinant of growth through its impact on patenting activity. Investments in R&D have a positive effect on patent activity in all categories considered; expenditures by businesses, the public sector, higher education institutions and the private non-profit sector. However, innovation is a longer-term process. When measured as the number of patent applications, it only appears to have a positive influence on regional growth after five years. Our results suggest that as capital and talent agglomerate they tend to positively influence growth in neighbouring regions. However, innovation remains a highly local element that does not necessarily influence growth in neighbouring regions.
- *Distance from markets*: endogenous sources of growth such as human capital and innovation are more important than a region's physical distance from markets. Although a region with good accessibility to markets has an added advantage for its growth prospects, these also depend on the presence of human capital, innovation, infrastructure and economies of agglomeration. While distance from markets is not relevant for innovation, proximity among the diverse local actors in a regional innovation system may well remain a key ingredient for innovation.
- *Infrastructure*: infrastructure is a necessary, but not sufficient, condition for growth. It is only relevant if human capital and innovation are also present in a region. Infrastructure and human capital require three years to positively influence growth.

- *Spatial effects:* neighbouring regions and presence of agglomeration. Geographic space plays a role in determining innovation in these models as agglomeration economies emerge as a relevant determinant of growth rates. Our results go a step beyond what NEG theories would predict (see Annex D), by showing that agglomeration economies are partly responsible for regional growth. The performance of neighbouring regions is strongly correlated with the performance of any given region in the OECD, suggesting that inter-regional trade and inter-regional linkages play an important role in a region's performance.

The findings of the chapter are useful for policy applications and policy-makers by providing them with a better understanding of the impact of key determinants of regional growth, the length of time needed for these factors to generate growth and which combinations of factors are most successful.

These results suggest that in order to promote regional growth, policy-makers should develop a comprehensive regional policy that not only links regions through infrastructure investments, but that also fosters human capital formation and facilitates the process of innovation. The risk of piecemeal visions for regional policy, such as only promoting human capital or only providing infrastructure, is that a "leaking" (*i.e.* leaking of jobs, talent, etc.) instead of a linking process will be created.

Review of the literature: neo-classical, endogenous and new economic geography

Growth has been viewed by some as a process determined by the accumulation of physical and human capital (neo-classical theory); others see it also as a process linked to a place's characteristics, such as innovation, knowledge and human capital (endogenous growth). Neo-classical theories rely entirely on capital accumulation (Solow, 1956; Swan, 1956),[2] and although technology is considered to be important, modelling difficulties have meant that technology is considered to be exogenous (Barro, 1997) and therefore excluded from the models. More recently, however, technology has been brought into these models through the inclusion of R&D theories (Romer, 1990; Grossman and Helpman, 1994; Barro and Sala-i-Martin, 1995).

Together, these growth theories tell us that economic growth can be explained by the stock of physical capital, human capital and innovation. While these factors have been largely analysed at the national level, there is a strong regional and even local dimension to all three. Most notably, in the process of innovation, the interaction of economic agents and the exchange of ideas demand social capital, urban spaces and face-to-face interaction. This latter is necessary – despite the reduced telecommunication costs with the emergence of the Internet – for ideas, patents, R&D, or production-line

 HOW REGIONS GROW: TRENDS AND ANALYSIS – ISBN 978-92-64-03945-2 – © OECD 2009

improvements to become new or improved products or upgraded processes. In addition, the emergence of a new body of literature, the NEG (see Annex D) has given us fresh insights into the concentration and dispersion of economic activity. Thus, increasing returns to scale external to firms are the main incentive for workers and firms to agglomerate, but dispersion of economic activity is possible depending on the interaction of two sets of opposing forces under varying levels of transport costs.

Neo-classical and endogenous growth models

Neo-classical growth theory was originally based on the proposition that long-run growth is the result of continuous technological progress in the form of new goods, markets or processes (Aghion and Howitt, 1998). Otherwise, the lack of technological change in the long run would cease growth by the effects of diminishing returns (Solow, 1956; Swan, 1956). Thus, the model can be expressed as a function of capital accumulation only, assuming perfect competition and decreasing returns to capital leading to equilibrium (Ramsey, 1928; Solow, 1956). Technological progress is recognised as an important growth determinant, but is regarded as exogenous mainly due to the implicit difficulties in modelling increasing returns. What is more, the original model considers that people save a fraction of their income, whereas a proportion of it is lost through depreciation (Solow, 1956; Swan, 1956). Economic growth is, under these circumstances, temporary. In fact, "any attempt to boost growth by encouraging people to save more will ultimately fail" (Aghion and Howitt, 1998: 13). Even if population expansion is included, growth stagnation is the result. Population growth will reduce capital per person, not by destroying it as depreciation does, but by diluting it since the number of people that must use it has increased. Therefore, long-run per capita growth rates can only be explained by technological progress.

The way in which the original neo-classical model includes technological change is by considering that an exogenously determined constant rate reflects the progress made in technology (Solow, 1956; Swan, 1956). Thus, the model implies conditional convergence; that is, if a country starts from a lower level of per capita output relative to other economies, the former is expected to attain a higher growth rate. Hence, the countries' output levels will tend to converge. Indeed, economies with less capital per worker are likely to attain higher rates of return and growth (Barro, 1997). Such a convergence is based on the assumption of diminishing returns to capital.

The inclusion of human capital as another form of capital which determines growth was one of the improvements made to the original model. The first attempts to internalise technology faced the technical difficulty of modelling increasing returns to scale. One solution was to consider that technological progress is the result of learning by doing (Arrow, 1962). Another similar school of thought was that growth rates are related to investment rates and the underlying rate of new ideas (Kaldor, 1957). However, neither approach could avoid regarding part of the technological progress as exogenous. Thereafter, the models tried to use diminishing returns in the struggle to internalise technology (Aghion and Howitt, 1998).

In the absence of technological improvements, neo-classical approaches were incapable of explaining long-run growth (Barro, 1997). For endogenous growth theorists, long-run growth was contemplated by considering that returns to capital did not diminish, since human capital entailed knowledge spillovers and external benefits (Romer, 1986; Lucas, 1988; Rebelo, 1991).

There are two distinct views on the role of human capital in the endogenous growth models. One approach (Nelson and Phelps, 1966) views growth as primarily driven by the *stock* of human capital that in turn affects a country's ability to innovate to catch up with more advanced countries. Differences in growth rates across countries are then attributable to differences in human capital stocks and thus in those countries' abilities to generate new ideas and technical progress. This allows for a one-off increase in the stock of human capital to have an indefinite impact on growth. A different approach (Lucas, 1988 based on the contributions of Becker, 1964 and Uzawa, 1965) views the *accumulation* of human capital as the key determinant of growth. In this view, countries can only grow in the long run as long as human capital keeps accumulating over time.

In addition, R&D theories were introduced and imperfect competition was factored into the model (Romer, 1990; Grossman and Helpman, 1994; Barro and Sala-i-Martin, 1995). The pursuit of long-run growth determinants represents the major contribution of the endogenous growth approach (Pack, 1994).

However, there is common ground between the two theories (see Table in Annex C). The neo-classical approach regards growth as being determined by capital intensities and human capital, and recognises the role played by technology in determining long-run growth but fails to include it in the model. The endogenous growth theory agrees on all three elements, but instead of regarding technology as exogenous it has tried to include it in the analysis. Theoretically, technology-treatment differences are crucial for determining long-run growth; empirically however, it is difficult to test. Particularly, in cases where data are limited, including technological progress in the model is remarkably difficult.

Models of the new economic geography

Perhaps the biggest difference between the NEG and the neo-classical and endogenous approaches is the relevance of scale. Neo-classical and endogenous economic theory is only concerned with relative terms: consumers' choices, firms' decisions, and wage-setting are all determined at the margin. The outcome of this process is unaltered in an economy with one individual, 1 000 individuals or 1 000 000 individuals.

Scale effects, on the other hand, do matter in the NEG. The process of agglomeration is precisely concerned with scale effects, where small initial differences can cause large effects over time through a self-feeding mechanism.

The main idea behind the NEG is to explain why consumers and firms tend to agglomerate together in geographic areas where other firms and consumers are already located. Studies of this phenomenon include Perroux's notion of "growth poles" (1955), Myrdal's analysis of "circular and cumulative causation" (1957), and Hirshman's concept of "forward and backward linkages" (1958).

The NEG formalises these kinds of cumulative causation mechanisms. Krugman (1991) provided the theoretical foundations by showing how regions that are similar or even identical in underlying structure can endogenously differentiate into either rich "core" regions or poor "peripheral" regions through a self-feeding mechanism of circular causation. Since the publication of Krugman's 1991 paper, the literature has considerably evolved. NEG models have now been applied to a variety of topics (Table 3.1), and a more precise description of each model is given in Annex D.

Table 3.1. **Summary table: The new economic geography**

Model	Assumptions	Agglomeration forces	Prediction
Krugman (1991)	• Two regions • Agriculture and manufacturing (IRS) prod. • Labour mobility • Transportation costs	• Internal scale economies • Cost of transportation • Proportion of mobile population in response to wage differentials (demand linkage)	Low transportation costs and economies of scale will agglomerate production and labour migration in the region with a higher initial production.
Krugman and Venables (1995)	• Two regions • Agriculture and manufacturing prod (intermediate and final goods with IRS) • Transportation costs	• Internal scale economies • Cost of transportation • Forward (cost) linkage • Backward (demand) linkage	As transportation costs fall below a critical value the region with the larger manufacturing share attracts more firms due to forward and backward linkages increasing the real income of the core region relative to the periphery. If costs continue to fall wage differential induces firms to relocate back to peripheral regions (convergence).

Table 3.1. **Summary table: The new economic geography** *(cont.)*

Model	Assumptions	Agglomeration forces	Prediction
Venables (1996)	• Two locations (regions) • One sector producing competitive goods, and two monopolistic sectors vertically linked • Transportation costs	• Internal scale economies • Transportation cost • Forward (cost) linkage • Backward (demand) linkage	For high and low transportation costs firms locate in both locations (convergence). For intermediate transportation costs some firms agglomerate in a single location while others may spread out in response to factor price differences.
Krugman and Venables (1996)	• Two countries (regions) • Two industries prod. intermediate and final goods (IRS) • Transp. costs	• Internal scale economies • Cost of transportation • Forward (cost) and backward (demand) linkage	For high transportation costs each country maintains a full range of industries. Low transportation costs lead to agglomeration of each industry in the country with a stronger initial position. For intermediate costs agglomeration occurs only when industries are initially very unequally distributed.
Englmann and Walz (1995)	• Two countries • Same technology • Labour mobility for skilled immobility for non-skilled • Local goods and services, R&D goods, industrial goods	• Immobility of one factor of production • Nontradeability of local inputs • Local limitation of knowledge spillover (case 1)	Case 1 assumes spillovers occur only locally. Agglomeration always occurs in the region with an initial advantage in the number of intermediate goods yielding a core-periphery pattern. Case 2 allows for interregional spillovers – the solutions comprise a stable steady state equilibrium with equal growth rates in both regions.
Puga and Venables (1996,1997)	• N countries • Manufacturing IRS and agriculture CRS sector • Transportation/trade costs	• Internal scale economies • Trade/transportation cost • Immobility of labour • Forward (cost) linkage • Backward (demand) linkage	Industrialisation will only occur in a few countries. When forward and backward linkages are strong enough, agglomeration occurs in one country raising the level of wages until reaching a critical mass. Industries relocate to another country creating agglomeration. Thus industry will spill over in a series of waves from country to country.

Table 3.1. **Summary table: The new economic geography** *(cont.)*

Model	Assumptions	Agglomeration forces	Prediction
Puga (1998)	• Two regions each of them can provide a location for a city and an agricultural hinterland • Manufacturing IRS and agriculture CRS sector • Transportation costs • Mobility between regions and sectors	• Internal scale economies • Cost of spatial interaction causes firms and workers to locate close to good market access • Elasticity of labour supply	A balanced system of cities emerges under high transportation costs. When transportation costs are low and elasticity of labour supply is high, the model predicts a primate urban pattern. Thus the greater emergence of metropolises in the less developed countries and their scarcity in Europe are due to lower costs of spatial interaction, stronger economies of scale, and more elastic supply of labour to the urban centre.
Puga (1999)	• Two regions • Manufacturing IRS and agriculture CRS sector • Transportation costs • Mobility between sectors • Regional mobility (case 1) • No regional mobility (case 2)	• Internal scale economies • Trade/transportation cost • Forward (cost) linkage • Backward (demand) linkage • Elasticity of labour supply	Case 1: under regional mobility high trade costs yield convergence, and lower trade costs (beyond a threshold) yield agglomeration. Case 2: under no regional mobility there is convergence at high trade costs, agglomeration at intermediate costs, and convergence at low costs. Thus European integration brings agglomeration only if labour is mobile. If labour is not mobile there is agglomeration but this fades at lower costs.
Martin and Ottaviano (2001)	• Two countries (regions) • Immobility of labour • Composite (IRS) and homogenous (CRS) good • Innov. by patents	• Internal scale economies • Cost of transaction (cost linkage) • Immobility of labour (demand linkage)	If equilibrium is present initially there is no incentive to relocate production of the increasing returns sector. If there are more firms initially producing differentiated goods in one region, agglomeration occurs as the cost for innovation in that region will be lower. All innovation occurs in that region.

Note: IRS = increasing returns to scale; CRS = constant return to scale.

Generally speaking all NEG models share the following characteristics:

- Assumptions of imperfect competition through increasing returns to scale in an economic sector – the monopolistic Dixit-Stiglitz model (Krugman, 1991) is the preferred choice.
- Costs associated with trade or transportation.

- Forces enhancing (centripetal forces) or discouraging (centrifugal forces) agglomeration.

The first two items are embraced by all models, while the third item, centripetal and centrifugal forces, varies. The three proposed centripetal forces include:

- Migration of labour (labour mobility between regions).
- Forward and backward linkages.
- Elasticity of labour supply (labour mobility between sectors).

All three forces positively enhance the formation of clusters. Workers tend to migrate towards the region with a higher initial industrial production since more goods and services are produced there than in regions with lower industrial production. The arrival of people increases local demand and local profits which in turn attract even more firms offering more goods and services.

Producers of final goods will find greater industrial concentration more attractive because a larger base of intermediate producers gives rise to forward (cost) linkages, while producers of intermediate goods will find it advantageous to produce near the large final good industry giving rise to backward (demand) linkages. The elasticity of labour supply operates very similarly to labour migration between regions. A high elasticity attracts non-industrial workers from the same region, increasing local demand and local profits, further attracting more firms. Centrifugal forces develop through lower competition in peripheral regions. Lower competition raises profits, thus attracting more firms.

Agglomeration economies occur when a firm enjoys increasing returns to scale (IRS) in a particular place. This could either be because of the presence of natural advantages (*i.e.* natural resources, location, etc.), monopolistic protection, political reasons (*e.g.* the decision to create a capital city) or any other reason. The presence of IRS also induces other firms to locate there as people come in search of higher wages, job opportunities and cultural values.

There are three main mechanisms that work to produce agglomeration economies (Duranton and Puga, 2004):

1. Mechanisms that deal with sharing of:
   - Indivisible facilities such as local public goods or facilities that serve several individuals or firms. Some examples, other than public goods, are facilities such as laboratories, universities and other large goods that cannot belong to one particular agent but where some exclusion is implicit in providing them.
   - The gains from the wider variety of input suppliers that can be sustained by a larger final-goods industry. In other words, the presence of IRS along

 HOW REGIONS GROW: TRENDS AND ANALYSIS – ISBN 978-92-64-03945-2 – © OECD 2009

with forward and backward linkages allows firms to purchase intermediate inputs at lower costs.

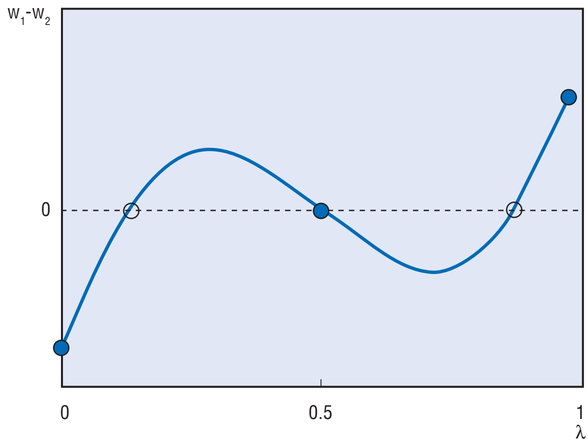
- ❖ The gains from the narrower specialisation that can be sustained with higher production levels. Several firms specialise in producing complementary products, reducing overall production costs.
- ❖ Risks. This refers to Marshall's idea that an industry gains from having a constant market for skills; in Krugman's words, a pooled labour market. If there are market shocks, firms can adjust to changes in demand accordingly as they have access to a deep and broad labour market that allows them to expand or contract their demand for labour.

2. Matching mechanisms by which:
   - ❖ Agglomeration improves the expected quality of matches between firms and workers, so both are better able to find a better match for their needs.
   - ❖ An increase in the number of agents trying to match in the labour market also improves the probability of matching.
   - ❖ Delays are alleviated. There is a possibility that contractual problems arising from renegotiation among buyers and suppliers result in one of the parties losing out by being held up by the other party in a renegotiation. This discourages investment. However, if the agglomeration is extensive enough, agents can change to an alternative partner.
3. Learning mechanisms based on:
   - ❖ The generation, diffusion, and accumulation of knowledge. This refers not only to the learning of technologies, but also the acquisition of skills.

The main findings stemming from the basic formulation of the core-periphery model are summarised next:

- At high transport costs, there is only one possible outcome: production will be divided equally among the two locations (black centre dot Figure 3.1).
- If transport costs are reduced, for instance due to a new motorway, to an intermediate level, five possible equilibria are possible, three of which are also stable (black dots in Figure 3.1). At that level of transport costs, the equal division of production among the two regions continue to be possible (black centre dot Figure 3.1). However, any changes in labour force in either region triggers a process of partial concentration in favour of the region with the largest labour market while some production is still retained in the other region; these two possible equilibria are unstable (white dots Figure 3.1) and can eventually lead to a full – catastrophic – concentration in either of the two regions (northeast and southwest black dots Figure 3.1).

Figure 3.1. **Possible equilibria with intermediate transport costs (new economic geography)**



1. λ refers to the proportion of labour force, whereas w refers to wages.
2. Black dots refer to stable equilibria and white dots to unstable ones.

Source: Based on Fujita, Krugman and Venables (1999).

- If transport costs continue to fall to lower levels, firms find it easier to concentrate in either of the two regions and benefit from agglomeration economies, pooled labour markets and increasing returns to scale, and still ship part of the production to the other region.

In sum, NEG models explain why economic activity tends to concentrate in particular geographic spaces. They also reveal that the benefits of agglomeration economies are sometimes offset by the costs that arise with concentrations. It is no surprise, therefore, that the theory has not established a clear understanding – or framework – of the links between economic concentration and growth.

Growth at the regional level

Analytical framework and selection of variables

The analysis in this chapter uses a series of econometric models to investigate the impact of the main determinants of regional economic growth. The modelling techniques are *i*) cross-sectional ordinary least square (OLS); *ii*) panel data analysis; *iii*) analysis based on a knowledge production function; and *iv*) spatial econometrics.

We selected the main determinants of regional growth from the most relevant theories summarised above (the neoclassical and endogenous growth theory and the more recent NEG literature). The models do not only include traditional variables related to neo-classical growth theories, but also endogenous determinants. One of the most salient elements of neo-classical models is convergence, either absolute or conditional depending on the model taken into account. The implication of convergence is that poorer regions further away from their steady-state level will tend to grow faster and thus converge. One of the ways in which economic growth models try to test this hypothesis is by including initial levels of income. A negative sign in the estimated coefficient therefore would denote that lagging regions are catching up and convergence is taking place. Conversely, a positive sign would imply higher growth rates in already richer regions and thus divergence would be occurring.

In addition to the convergence hypothesis, neo-classical growth theories rely heavily on capital as the main determinant of economic growth. Lack of data at the regional level prevents us from using a measure of physical capital such as private investment or gross fixed capital formation. However, we use infrastructure in roads as a proxy for physical capital.

Endogenous growth theories stress the role of human capital and R&D as the sources of boundless growth opening up the possibility for non-convergence. In that sense, our model uses a measure of human capital stock. However lack of data at the regional level restricts our ability to capture the effects of quality in human capital.

Technological progress is accounted for in the model by using input and output innovation measures. The former refers to R&D expenditure, whereas the latter refers to patenting activity. Since the effects of innovation inputs might be indirectly related to regional growth, we explore the links between innovation inputs to innovation outputs through a knowledge production function. This function measures the impact on patenting of R&D expenditures, personnel and employment in knowledge-based sectors.

As explained earlier, the NEG discusses the centre-periphery model: increasing returns to scale, external economies and transportation costs. The mechanics of the model rely on the action of two countervailing forces: *i*) centripetal, inducing agglomeration; and *ii*) centrifugal, favouring dispersion. Agglomeration forces include: a pooled labour market, backward and forward linkages driven by the interaction of increasing returns and transportation costs, and technological spillovers. In contrast, centrifugal forces include factor immobility, land rents, congestion costs or pure diseconomies of scale. Thus, our model includes variables for labour market pooling, measures of agglomeration economies – captured thorough sectoral specialisation indices multiplied by their size – and geographical measures (both distance to markets and accessibility to markets) as proxies for transportations costs.

In order to measure the effects of these determinants of growth we selected the following indicators as structural variables in our models:

- Initial level of GDP per capita (convergence hypothesis/neo-classical):
  - Logarithm of initial GDP per capital expressed in 2000 USD purchasing power parity.
- A measure of physical capital (neo-classical):
  - Physical infrastructure – motorway density (total motorway kilometres in a region to its population).
- A measure of human capital (endogenous growth):
  - Stock of labour – attainment rates in primary and tertiary education (percent of the working age population with primary, secondary and tertiary education rates).
- A measure of intellectual capital (endogenous growth):
  - GDP expenditures on R&D by sector of performance (*i.e.* business, government, private non-profit, and higher education).
  - R&D personnel by sector of performance (*i.e.* business, government, private non-profit, and higher education).
  - Patent applications.
  - Employment in knowledge intensive sectors.
  - Employment in high and medium-high technology manufacturing.
- A measure of labour market performance:
  - Employment rates.
- A measure of agglomeration economies (NEG):
  - Specialisation in sector *j* multiplied by the size of sector *j*. The sectors of interest are financial intermediation, agriculture and manufacturing, and specialisation is measured using the specialisation index.[3]
- A geographical measure (NEG):
  - Distance from markets (by blocks, see Annex E).
  - Market accessibility (by blocks, see Annex E).

The impact of these structural variables on regional growth is estimated in the four econometric models. Coverage includes the period 1995-2005, and the main source of data is the OECD (2008) Regional Database (RDB) for most indicators, with the exceptions of infrastructure and geographic measures.

Cross-section model

Model specifications

The first econometric technique is a simple cross-section regression model in which regional GDP per capita growth during the period 1995-2005 is regressed on a number of key structural variables at the beginning of the period. The specification of the model is quite simple and it assumes a linear relationship as a starting point.

Based on the availability of data the static model takes the following functional form:

$$\begin{aligned}\frac{1}{T}\ln\left(\frac{GPD_{t+T}}{GDP_t}\right) &= \alpha + \beta_1 \ln(Initial Y_t) + \beta_2 \ln(Infrast_t) + \beta_3 \ln(Prim\,Edu_t) + \beta_4 \ln(Tert\,Edu_t) + \\ &+ \beta_5 (Empl\,Rate_t) + \beta_6 \ln(Patents_t) + \beta_7 \ln(R\&D\,Total_t) + \beta_8 \ln(R\&D\,BUS_t) + \beta_9 \ln(R\&D\,GOV_t) \\ &+ \beta_{10} \ln(R\&D\,HE_t) + \beta_{11} \ln(Agg\,Ag_t) + \beta_{12} \ln(Agg\,Man_t) + \beta_{13} \ln(Agg\,Fin_t) + \beta_{14} \ln(Mkt\,Access_t) \\ &+ \beta_{15} \ln(Dist\,Mkts_t) + u_t \qquad (1)\end{aligned}$$

where t = 1995 and T = 10, and average growth of GDP is regressed on:

- *Initial* Y_t = initial GDP per capita
- $Infrast_t$ = motorway density defined by kilometres of motorway to population
- $Prim\ Edu_t$ = primary educational attainments
- $Tert\ Edu_t$ = tertiary educational attainments
- $Empl\ Rate_t$ = initial year employment rates
- $Patents_t$ = patent applications
- $R\&D\ Total_t$ = total research and development expenditures
- $R\&D\ BUS_t$ = research and development expenditures carried out by firms
- $R\&D\ GOV_t$ = research and development expenditures carried out by the government
- $R\&D\ HE_t$ = research and development expenditures carried out by higher education institutions
- $AGG\ Ag_t$ = agglomeration economies in agriculture defined by the size of the sector (*i.e.* employment in agriculture) times the index of specialisation (see endnote 3) in agriculture
- $AGG\ Man_t$ = agglomeration economies in manufacturing defined by the size of the sector (*i.e.* employment in manufacturing) times the index of specialisation (see endnote 3) in manufacturing
- $AGG\ Fin_t$ = agglomeration economies in financial informediation defined by the size of the sector (*i.e.* employment in financial intermediation) times the index of specialisation (see endnote 3) in financial intermediation

- *Mkt Access*$_t$ = access to markets (see appendix E for full explanation)
- *Dist Mkts*$_t$ = distance to markets (see appendix E for full explanation)

The first explanatory variable (*i.e.* initial GDP per capita) is included to account for convergence or divergence of regional income. A negative sign in this variable would signal that relatively poorer regions are growing faster and therefore a process of convergence is under way. Conversely, a positive sign would indicate that richer regions are growing faster, and thus that regional incomes are diverging. This convergence or divergence trend will be conditional on a series of variables that determine growth, listed above. As economic growth theories argue that the forces behind long-run growth are physical capital, human capital and innovation, a number of variables have been introduced to model them. First, as capital stock data at the regional level are not available, a measure of infrastructure (motorways) was included. Second, human capital is included in the form of educational attainment for primary schooling and for tertiary education. Third, innovation enters into the model using patents and research and development (R&D) expenditures. Several variables that reflect expenditure in R&D were included, such as those carried out by the government, the private sector, higher education institutions, and non-profit organisations. In addition to economic-growth theory variables, a proxy for the proper functioning of labour markets was included in the form of employment rates.

The model attempts to explain regional growth not only by the usual determinants of growth, but also by using – as much as possible – variables that describe the dynamics of concentration and dispersion which can be argued to be at the heart of growth and inequality. One of the reasons for firms' agglomeration lies in backward and forward linkages and other agglomeration economies. The model incorporates these types of external economies into the firm by introducing sectoral specialisation indicators. Similarly, the NEG also argues that a second reason for agglomeration is thick markets. The model explores the impact of distance from or access to markets on economic growth.

Interpreting the results

The benefits of co-ordinating infrastructure, human capital formation and innovation

The results of the model during the period 1995-2005 imply that convergence is taking place across OECD TL2 regions, but that this is conditional on a series of factors (Table 3.2). Infrastructure does not affect regional growth by itself, except when education and innovation are considered (Models 6-7). Although infrastructure was included as a proxy for physical capital, relying on this measure presents two caveats. First, motorways are only one part of infrastructure (public capital) and other types of investment with direct impact on productive activities – either by enabling them or by reducing costs – such as energy, telecommunications, rail or airports, are not accounted for. Second, public investment in any case does not take into account private stocks of capital simply because data for that sector are not available at the regional level. Thus, not surprisingly the results are not significant on their own. Having said that, there might be a policy-related interpretation as infrastructure is only significant in the presence of human capital and innovation. Put simply, motorways may open up markets, but may also provoke fierce competition that may lead to either local firm mortality or migration of production to core regions. NEG models describe how goods may be shipped from the core to seize increasing returns to scale external to the firm. However, as human capital and innovation are present in the region, capital finds it attractive to stay in their regions and benefit from a pooled and well-matched labour market. Therefore, there is some suggestion that policies for infrastructure, human capital formation and innovation should be co-ordinated to boost economic growth in any region.

In terms of human capital, it is interesting to note that while primary schooling is negatively associated with growth, tertiary education positively affects regional performance. This is in line with what we would expect the model to show. Regions with insufficient human capital will not grow while those with increased levels will reap the benefits of endogenous elements of growth.

It is also important to note that employment rates do not significantly affect growth, although they do of course affect per capita income levels. One possible explanation is the position of the economy relative to its steady-state level. That is, the results for employment may reflect the mechanics of convergence. Regions with lower employment rates are not fully exploiting their labour resources and therefore are located far away from their ideal production possibilities. As with lower incomes, the reorganisation of the regional economy to seize unused labour potential results in higher growth rates.

Furthermore, innovation does positively affect growth just as endogenous growth theory suggests. However, this positive relationship between innovation and growth only holds for patenting, and not for total R&D expenditure. Although this initially may seem puzzling, it might be related to the very process of innovation. R&D expenditure is in fact one of the many inputs used in the process to produce innovation. Patenting is only one possible outcome and patents alone often fail to affect economic growth as many patents end up not being used by industry; they are often more an outcome of a broader process. In fact, if we think in terms of a knowledge production function, the result is not puzzling at all. R&D expenditures should be related to patenting, not directly to growth, so this indirect relationship with economic expansion may explain these results. This indirect relationship is supported by the results in Model 11, when R&D expenditure by source of funding is taken into account. The fact that business-based and higher-education-based expenditure on R&D are not significant – and even negative – is very revealing as these sectors are usually where the bulk of patenting takes place.

Agglomeration economies drive regional growth

A crucial result in our models relates to agglomeration economies. Just as NEG models suggest, agglomeration economies are an important element of concentration. Our results go a step beyond what NEG theories would predict, by showing that agglomeration economies are partly responsible for regional growth (Table 3.2). However, it is possible that there might be a strong correlation between tertiary education and innovation indicators as the significance of the variable drops in the presence of endogenous growth variables.

Table 3.2. **OLS cross section results for regional economic growth in OECD TL2 regions, 1995-2005**

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11
Constant	0.0268	0.0013	0.1695	0.1553	0.1582	0.1934	0.193	0.3014	0.2972	0.104	–0.0126
	(2.65)**	(0.11)	(11.54)**	(9.06)**	(9.33)**	(6.44)**	(5.08)**	(6.27)**	(9.62)**	(1.46)	(–0.32)
Initial Y	–0.0006	0.0012	–0.0122	–0.0097	–0.0094	–0.015	–0.0152	–0.0261	–0.026	–0.0214	–0.0047
	(–0.59)	(0.95)	(–9.45)**	(–6.21)**	(–5.95)**	(–5.39)**	(–4.14)**	(–6.18)**	(–8.6)**	(–5.04)**	(–1.2)
Infrast	–	0.0075	–	0.0093	0.0132	0.0156	0.02	0.0155	0.0172	0.0148	0.0284
		(0.86)		(1.36)	(1.92)	(1.99)*	(2.31)*	(1.89)	(2.21)*	(1.89)	(3.23)**
Prim Edu	–	–	–0.0096	–0.0126	–0.0129	–0.0035	–0.004	–0.0075	–0.0079	–0.0091	–
			(–9.72)**	(–11.03)**	(–11.46)**	(–3.55)**	(–2.93)**	(–5.06)**	(–5.42)**	(–6.36)**	
Tert Edu	–	–	0.0076	0.0091	0.0097	–	–	0.0089	0.0087	0.0096	0.0067
			(8.79)**	(9.31)**	(9.81)**			(6.42)**	(6.83)**	(7.13)**	(4.58)**
Empl Rate	–	–	–	–	–0.0205	–	–	–	–	–	–
					(–2.37)**						
Patents	–	–	–	–	–	0.0015	–	–	–	–	–
						(2.5)**					
R&D Total	–	–	–	–	–	–	0.0019	–0.0007	–	–0.0009	–
							(1.71)	(–0.47)		(–0.6)	
R&D BUS	–	–	–	–	–	–	–	–	–	–	–0.0026
											(–2.3)*
R&D GOV	–	–	–	–	–	–	–	–	–	–	0.0028
											(2.98)**
R&D HE	–	–	–	–	–	–	–	–	–	–	–0.0078
											(–5.81)**
Agg Ag	–	–	–	–	–	–	–	–0.0014	–0.0009	–0.001	–
								(–2.04)*	(–1.41)	(–1.65)	

HOW REGIONS GROW: TRENDS AND ANALYSIS – ISBN 978-92-64-03945-2 – © OECD 2009

Table 3.2. **OLS cross section results for regional economic growth in OECD TL2 regions, 1995-2005** *(cont.)*

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11
Agg Man	–	–	–	–	–	–	–	–0.0047	–0.0052	–0.0028	–
								(–2.89)**	(–3.62)**	(–1.77)	
Agg Fin	–	–	–	–	–	–	–	0.0029	0.0031	0.0014	0.0015
								(2.02)*	(2.32)*	(0.96)	(1.03)
Mkt Access	–	–	–	–	–	–	–	0.0002	0.0009	–	0.0013
								(0.39)	(1.75)		(2.19)*
Dist Mkts	–	–	–	–	–	–	–	–	–	0.0333	–
										(3.67)**	
R	0.0011	0.0082	0.2916	0.3235	0.3451	0.1652	0.1515	0.4712	0.4728	0.5111	0.4014
Adj R	–0.002	0.0019	0.2989	0.3134	0.3329	0.1505	0.1338	0.442	0.4493	0.4841	0.3717
F	0.35	1.29	40.93**	32.16**	28.24**	11.23**	8.57**	16.14**	20.17**	18.94**	13.5**
N	333	315	292	274	274	232	197	173	189	173	170

Note: */ Significant at the 95% confidence level; **/ Significant at the 99% confidence level; BUS = business sector; GOV = government sector; HE = higher education institutions

Countries missing as the model grows in variables due to lack of complete data mainly in R&D expenditure: Model 1: Iceland; Model 2: Australia, Norway and New Zealand; Model 3: Iceland, Denmark, Japan and Turkey; Models 4 and 5: Iceland, Denmark, Japan, Turkey, Australia, Norway and New Zealand; Model 6: Iceland, Denmark, Japan, Turkey, Australia, Norway, New Zealand and Switzerland; Model 7: Iceland, Denmark, Japan, Turkey, Australia, Norway, New Zealand, Switzerland, Belgium, Ireland, Sweden and Mexico; Model 8, 9 and 10: Iceland, Denmark, Japan, Turkey, Australia, Norway, New Zealand, Switzerland, Belgium, Ireland, Sweden, Mexico and Germany.

Source: Calculations based on *OECD (2008) Regional Database*.

Transport costs and markets: a complex relationship

Our most problematic results lie in the variables that try to reflect the NEG idea that transport costs and distance to relevant markets determine concentration. Our measure of distance to markets shows the opposite sign we would expect in Model 10, suggesting that being far away from markets has a *positive* influence on growth. Our interpretation is that the catch-up process evinced by the conditional convergence pattern shown in Models 3 to 10 implies that regions in the periphery are growing faster despite being relatively further away from the main markets at the core. In addition there might be a measurement bias given that our measures do not account for travel time and transportation networks and also that they are heavily affected by the size of TL2 regions.

A more promising result is obtained in our measure of accessibility to markets. Although this is statistically insignificant in Models 8 and 9, Model 11 suggests that a region with good accessibility has an added advantage to its growth prospects, though this depends on the presence of human capital, innovation, infrastructure and economies of agglomeration.

In any case, endogenous sources of growth such as human capital and innovation are more important than the physical distance to markets.

Similarly, agglomeration economies at play even in the periphery seem to be even more relevant than distance. In contrast, insignificant results for distance to markets and the fact that access to markets is only related to growth in the presence of human capital, innovation and infrastructure, seem to suggest that access to markets is a necessary but not sufficient condition for growth.

We should underline the fact that as more variables are included, the number of available observations falls, because data on all variables are not available for all countries. Finally, there is a strong change in the values and significance of coefficients of all explanatory variables when controls for country-effects are taken into account.

Panel data model

Model specifications

We used a panel data model to look at the effects of structural variables on regional growth over time. The structural variables and the time period are the same as in the cross-sectional model. A panel specification offers some advantages over the cross-sectional specification by permitting us to factor out the *time effects* and the *cross-sectional* components of the data. While the cross-sectional model measures the impact of initial values on regional growth over a longer time period (*i.e.* ten years), the panel model measures the yearly impact of the independent variables on growth, controlling for country effects (cross-section) and time effects. In addition, panel data approaches allow for lagged effects on the phenomenon to be explained, so if a particular variable, say infrastructure, takes time to have an impact because it needs to be built and used, these models allow us to pinpoint the time needed for that impact to emerge.

Other advantages of the panel specification are:

- It allows for a significant gain in data observations by using regional data in all years.
- It captures and controls for national and time effects on regional economic growth.
- It has the ability to measure the impact of independent variables over time. As already mentioned, this can be achieved by lagging the independent variables and measuring their effects over time.

In the panel specification the unit of analysis starts at the regional level. The *cross-sectional effects* then capture the effects of countries and the *time* *effects* capture the effects of time on regional growth. The model allows us to measure forces affecting regional growth at three distinct levels:

- Forces at the regional level are captured through the coefficients of the independent variables.
- The *cross-sectional effects* capture the variation common to all regions of a country after the regional effects are accounted for by the independent variable coefficients. These identify the national factors influencing regional growth.
- The *time effects* measure the variation common to all regions in a given year after the regional and country effects are controlled for.

The panel model is specified as:

$$\ln\left(\frac{GDP_{i,t}}{GDP_{i,t-1}}\right) = \alpha + \beta_1 \ln\left(Initial Y_{i,t-1}\right) + \beta_2 \ln\left(Infrast_{i,t-1}\right) + \beta_3 \ln\left(Prim\,Ed_{i,t-1}\right) + \beta_4 \ln\left(Tert\,Ed_{i,t-1}\right) + \\ + \beta_5 (Emp\,Rate_{i,t-1}) - 1 + \beta_6 \ln\left(Patents_{i,t-1}\right) + \beta_7 \ln\left(R\&D\,Total_{i,t-1}\right) + \beta_8 \ln\left(R\&D\,BUS_{i,t-1}\right) \\ + \beta_9 \ln\left(R\&D\,GOV_{i,t-1}\right) + \beta_{10} \ln\left(R\&D\,HE_{i,t-1}\right) + \beta_{11} \ln\left(Agg\,Ag_{i,t-1}\right) + \beta_{12} \ln\left(Agg\,Man_{i,t-1}\right) \\ + \beta_{13} \ln\left(Agg\,Fin_{i,t-1}\right) + \beta_{14} \ln(Mkt\,Access_{i,t-1}) + \beta_{15} \ln\left(Dist\,Mkts_{i,t-1}\right) + u_i + e_{i,t} \qquad (2)$$

where the dependent and independent variables have already been specified in the previous section.

The panel model can be specified with fixed effects and random effects. One potential consequence of the fixed effect panel is that disturbances may be correlated within groups (*i.e.* countries). The random effects account for this correlation and therefore the random effects estimator should be selected, when possible, over the fixed effects estimator if it is statistically justifiable to do so since it offers more efficient estimates. A Hausman test can determine whether it is statistically justifiable to use random effects. For this reason Table 3.3 displays the results of the Hausman test for each model, whether it is statically justifiable, and whether it uses random effects (re) instead of fixed effects (fe).

Interpreting the results

- **Infrastructure:** As in the cross-section models, infrastructure does not influence regional growth by itself, but only in conjunction with human capital and innovation (Models 6-7), or with human capital, economics of agglomeration and accessibility (Models 8-9). Thus infrastructure is a necessary, but not sufficient, condition for growth.
- **Human capital:** The results for human capital variables confirm the findings of our cross-section models. Human capital influences growth: regional growth declines when there are insufficient levels of human capital (*i.e.* primary educational attainment rates) and it increases when

Table 3.3. **Panel results for regional economic growth in OECD TL2 regions, 1995-2005**

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11
Constant	0.104	0.105	0.145	0.108	0.086	0.008	0.092	0.166	0.125	–0.082	0.052
	(4.99)**	(4.40)**	(5.55)**	(3.81)**	(2.95)**	(0.2)	(2.1)	(2.91)**	(2.70)**	(–0.88)	(0.9)
Initial Y	–0.008	–0.009	–0.008	–0.005	–0.001	0.002	–0.003	–0.010	–0.008	–0.009	0.0004
	(–3.97)**	(–3.7)**	(–3.37)**	(–1.82)*	(–0.18)	(0.6)	(–0.75)	(–1.98)*	(–1.76)*	(–1.75)*	(0.1)
Infrast	–	0.001	–	0.008	0.015	0.020	0.014	0.014	0.015	0.013	0.018
		(0.2)		(1.1)	(2.05)*	(2.50)*	(1.83)*	(1.6)	(1.84)*	(1.4)	(2.27)**
Prim Ed	–	–	–0.008	–0.008	–0.008	–0.002	–0.005	–0.009	–0.006	–0.008	–0.004
			(–5.15)**	(–4.94)**	(–5.00)**	(–1.32)	(–3.19)**	(–3.71)	(–2.58)**	(–3.42)**	(–2.31)**
Tert Ed	–	–	0.005	0.005	0.005	–	–	0.003	0.004	0.002	–
			(3.38)**	(2.97)**	(3.32)**			(1.1)	(1.5)	(0.8)	
Empl Rate	–	–	–	–	–0.049	–	–	–	–	–	–
					(–4.17)**						
Patents	–	–	–	–	–	–0.001	–	–	–	–	–
						(–0.86)					
R&D Total	–	–	–	–	–	–	0.001	–0.001	–	0.0002	–
							(0.8)	(–0.06)		(0.1)	
R&D BUS	–	–	–	–	–	–	–	–	–	–	–0.003
											(–2.96)**
R&D GOV	–	–	–	–	–	–	–	–	–	–	0.002
											(2.89)**
R&D HE	–	–	–	–	–	–	–	–	–	–	0.001
											(0.51)
Agg Ag	–	–	–	–	–	–	–	0.0003	0.001	0.0001	–
								(0.49)	(1.16)	(0.17)	
Agg Man	–	–	–	–	–	–	–	–0.002	–0.003	–0.001	–
								(–1.52)	(–2.28)*	(–0.80)	
Agg Fin	–	–	–	–	–	–	–	0.004	0.003	0.004	0.001
								(2.62)**	(2.31)*	(2.35)*	(0.72)
Mkt Access	–	–	–	–	–	–	–	–0.001	–0.0002	–	–0.001
								(–0.83)	(–0.35)		(–1.38)
Dist Mkts	–	–	–	–	–	–	–	–	–	0.052	–
										(3.32)**	
Fixed (fe) or Random (re)	re	re	re	re	re	fe	re	re	fe	re	fe
Hausman test (Prob>chi2)	0.26	0.15	0.80	0.37	0.35	0.00	0.19	0.10	0.04	0.54	0.00
R^2 within	0.005	0.006	0.017	0.022	0.035	0.022	0.027	0.037	0.030	0.047	0.070
R^2 between	0.016	0.001	0.334	0.372	0.326	0.000	0.296	0.505	0.456	0.359	0.041
R^2 overall	0.002	0.000	0.082	0.090	0.083	0.012	0.054	0.107	0.108	0.136	0.050
Wald chi (re), F (fe)	15.72	13.75	41.53	45.07	63.52	6.52	33.70	46.37	5.05	54.04	7.37
n	3 166	2 850	1 650	1 529	1 494	1 165	1 062	942	1 320	936	813

Note: */ Significant at the 95% confidence level; **/ Significant at the 99% confidence level; BUS = business sector; GOV = government sector; HE = higher education institutions.
Source: Calculations based on *OECD Regional Database (2008)*.

 HOW REGIONS GROW: TRENDS AND ANALYSIS – ISBN 978-92-64-03945-2 – © OECD 2009

sufficient flows of human capital (*i.e.* tertiary educational attainment rates) are present, although the effect in the latter case fades away in Models 8-10.

- **Employment:** Again the results for this variable support our cross-section models. Employment rates – when used as an indicator of a region's distance to its production possibility frontier – have an adverse effect on regional growth. This indicator captures a region's ability to mobilise labour resources: the further away the region is from the production possibility frontier the higher the region's growth potential.
- **Initial income:** These results are less stable and significant than in our cross-section models. The results of initial income (GDP per capita) on regional growth in the cross-section models are interpreted as convergence or divergence given the dynamics occur over the medium and long term. In the panel specification they show convergence (relative to the previous period) in Models 1-4 and 8-10 and no effect in the rest of models.
- **Innovation:** Patents have no effect on regional growth on a yearly basis, and expenditures in research and development only influence regional growth in Model 11 (positive effects for government expenditure and negative effects for business expenditure). Since patents do influence growth over a longer time period, a subsequent analysis measures the time period over which patents positively affect growth. Our interpretation of these results is that the effect of patents is relevant over the long-run, when patents can become new products, change or create new processes. Thus, there is a need to look more deeply into the inter-temporal relationship of patenting and growth (see below). In contrast, our results for R&D are quite similar to those of the cross-section models.
- **Agglomeration economies:** The results for this variable are consistent with those obtained through the cross-section analysis. External economies positively affect growth, mainly through financial intermediation. In contrast a lack of agglomerations – captured though specialisation in the agricultural sector multiplied by the size of the sector – does not hurt regional growth over a one year period.
- Finally **distance to markets** has a positive relationship with growth as in the cross-sectional model. Again this result might be due to a catching up-process taking place in regions distant from markets, such as those in the Eastern European countries. Or it might be due to limitations in our measure of distance to market, which is highly affected by the large size of several TL2 regions.

An important result obtained in the panel specification is that after controlling for national factors through the country dummies, regional factors are quite important in determining a region's growth path. This reveals that national factors are not sufficient at the regional level to mobilise the available assets.

The different effects of human capital (*i.e.* tertiary education), infrastructure and patents on regional growth obtained in the cross-section and panel specification models might reflect differences in the time dimension. It is often argued that some of these variables (human capital and innovation) are determinants of growth in the medium and long term. For these reasons the next section explores the effects of human capital, patents and infrastructure on regional growth over different time periods.

Effects of infrastructure, human capital and innovation on regional growth over time

Table 3.4 reports the effects of infrastructure, human capital and patents on regional growth over a three-year and a five-year period using four different models. The results for Models 3-4 (Table 3.4) should be interpreted with caution because they account for a relative small number of observations; however, Models 1-2 include sufficient observations. The results for all models show that all three variables appear to influence regional growth positively over the medium term. In the short term they do not influence growth, and in some cases they have an adverse effect on growth:

- **Infrastructure takes three years to contribute to growth** when innovation is present (Model 2), and five years when human capital is present (Model 1). In both models the effects of infrastructure reverse in the short run. A possible explanation is that it not only takes time for infrastructure to yield some benefits in terms of growth, but also that endogenous growth variables are present to avoid a leaking of economic activity instead of the desired link to markets.
- **Tertiary education only has a positive effect on growth after three years** (Models 3-4), while in the short run (Model 4) it has a negative effect. These results should be interpreted with caution given the small number of observations; however, the long-term effects of tertiary education on regional growth are positive and appear very robust in the cross-sectional specification.
- **Patent applications take five years to have a positive effect on regional growth** (Models 2-4), while over a shorter time frame they have a negative effect. Again the effects of patents on regional growth over the long term are positive (according to the cross-sectional models). As can be expected, the process of patenting – and of innovation more broadly – is long-term and influences growth only in the long run. However, the relationship between input and outcome variables in the process of innovation is not clear (such as the relationship between R&D or human capital and patenting). The next section attempts to shed some light on these associations.

Table 3.4. **Panel results, lagging human capital, infrastructure and patents in OECD TL2 regions, 1995-2005**

	Model 1	Model 2	Model 3	Model 4
Constant	–0.053	–0.043	0.324	0.258
	(–0.64)	(–1.52)	(–1.39)	(0.72)
Initial Y	0.013	0.004	–0.033	–0.045
	(–1.71)*	(1.60)	(–1.71)*	(–1.66)*
Infrast	0.023	–0.093	–	–0.233
	(0.22)	(–2.61)**		(–0.56)
Lag 3 Infrast	–0.239	0.175	–	–0.439
	(–2.00)*	(3.84)**		(–0.64)
Lag 5 Infrast	0.221	–0.047	–	0.760
	(3.57)**	(–1.19)		(1.46)
Primary Education	0.024	–	0.012	–0.088
	(1.52)		(0.30)	(–1.31)
Lag3 Primary Education	–0.008	–	0.023	0.098
	(–0.53)		(0.6)	(1.4)
Lag 5 Primary Education	–0.023	–	–0.062	–0.030
	(–1.34)		(–1.32)	(–0.045)
Tert Ed	–0.001	–	–0.031	–0.244
	(–0.07)		(–0.67)	(–2.74)**
Lag 3 Tert Ed	–0.009	–	0.108	0.319
	(–0.48)		(2.40)**	(2.74)**
Lag 5 Tert Ed	0.013	–	–0.046	–0.037
	(0.75)		(–1.20)	(–0.64)
Patents	–	–0.010	–0.010	–0.010
		(–5.83)**	(–1.00)	(–1.11)
Lag3 Patents	–	0.000	–0.019	–0.030
		(–0.14)	(–2.67)**	(–3.09)**
Lag5 Patents	–	0.006	0.022	0.032
		(3.83)*	(2.99)**	(–3.68)**
Fixed (fe) or Random (re)	fe	fe	re	re
Hausman test (Prob>chi2)	0.01	0.02	1.00	1.00
R^2 within	0.108	0.105	0.384	0.671
R^2 between	0.011	0.028	1.000	1.000
R^2 overall	0.022	0.020	0.501	0.749
Wald chi (re), F (fe)	3.08	15.63	29.15	53.63
n	283	958	40	32

Note: */ Significant at the 95% confidence level; **/ Significant at the 99% confidence level.
Source: Calculations based on *OECD Regional Database* (2008).

Knowledge production function

The analysis finds that patent applications, albeit an imperfect output measure of regional innovation, appear to have a positive influence on regional growth over a five-year (Table 3.4) and a ten-year period (Table 3.2).

This section estimates the coefficients of a knowledge production function at the regional level for determining the effects of input indicators on innovation. As the effects of innovation inputs on regional growth might be indirect, this section should shed some light on the mixed results obtained in the previous models on the impact of innovation inputs on regional growth (*i.e.* the cross-sectional model finds no impact of total R&D on regional growth, a negative effect in business and higher education R&D expenditures and a positive effect on government expenditures in R&D, Table 3.2).

Model specifications

Research on regional innovation generally falls into three main approaches (Box 3.1).

Box 3.1. **Three main models for regional innovation research**

The linear model

In this view, research leads to inventions which then become innovations and produce greater levels of productivity and ultimately output:

The empirical studies first determine a link between R&D and patents and then they estimate the link between patents and growth.

This view sees differences in innovation capacity arising from an endogenous growth perspective creating persistent differences in wealth and economic performance.

The higher the investment in R&D, the higher the innovative capacity and the higher the economic growth.

The linear model overlooks key factors about how innovation is actually generated.

Systems of innovation or learning region approaches

These approaches regard innovation as part of a territory-embedded process where institutional networks can favour (or deter) innovation generation:

The capacity of these networks to act as catalysts depends on the combination of social and structural conditions in every territory; these are often referred to as a *social filter*.

The proximity and interaction of local synergies are very relevant; importance is assigned to inter-organisational networks, financial and legal institutions, technical agencies and research infrastructures, education and training systems, governance structures and innovation policies.

Box 3.1. **Three main models for regional innovation research** *(cont.)*

These embedded networks in regions (*i.e.* social economic structures and institutions) are very difficult to measure and to compare.

Diffusion and assimilation of innovation:

The knowledge spillovers approach looks at the micro level in innovative units (*i.e.* R&D departments within firms, universities and research centres) as well as local institutions and individuals. The interaction – with each other and with their external environment through networks – produces the transmission of knowledge in the form of knowledge spillovers. However, this approach is also not easy to carry out as it is difficult to capture spillovers.

Source: Crescenzi and Rodriguez Pose, 2006.

Our knowledge production function approach uses both the linear model and the systems of innovation approaches described in Box 3.1. We have not been able to capture knowledge spillovers, given the inherent difficulties of measuring them. Table 3.5 reports the coefficient of the knowledge production function, which measures the impact of the initial value of the independent variables on the final value of the dependent variable. The model is formally defined as:

$$\ln\left(Patents_{i,t+T}\right) = \alpha + \beta_1 \ln\left(Tert\,Ed_{i,t}\right) + \gamma_{j=1-4} \ln\left(R\,\&\,D\,Exp\,GDP_{j,i,t}\right) + \phi_{k=1-4} \ln\left(R\,\&\,D\,Personnel_{k,i,t}\right)$$
$$+ \beta_{10} \ln\left(KIS\,Emp_{i,t}\right) + \beta_{11} \ln\left(HTM\,Emp_{i,t}\right) + \beta_{12} \ln\left(Agg\,Ag_{i,t}\right) + \beta_{13} \ln\left(Agg\,Man_{i,t}\right)$$
$$+ \beta_{14} \ln\left(Agg\,Fin_{i,t}\right) + \beta_{15} \ln\left(Mkt\,Access_{i,t}\right) + \beta_{16} \ln\left(Dist\,Mkts_{i,t}\right) + e_{i,t} \qquad (3)$$

where t = 1995 and T = 10 R&D expenditures and R&D personnel include the (1) business and (2) government sectors, (3) higher education institutions and (4) the private non-profit sectors (*i.e.* j = 1-4 and k = 1-4 in Equation 3).

Interpreting the results

The results of the model are summarised below:

- Investments in R&D have a positive effect on patent activity in all categories considered; these are R&D expenditures by businesses, the public sector, higher education institutions and by the private non-profit sector. The fact that R&D expenditures in higher education institutions (HE) influences patenting activity negatively in Model 6 is quite puzzling and should be explored further. These results are also consistent with the previous models.
- R&D personnel only enhance patent applications in the business category, although the effects on patents are smaller than expenditures in R&D by the

Table 3.5. **Knowledge production function in OECD TL2 regions, 1995-2005**

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12	Model 13	Model 14	Model 15
Constant	–9.9	–9.4	–9.4	–10.2	–16.0	–14.3	–9.3	–13.6	–18.7	–16.9	–9.6	–7.5	–16.4	–5.0	–57.8
	(–11.68)**	(–8.73)**	(–8.70)**	(–8.88)**	(–9.56)**	(–8.64)**	(–5.04)**	(–7.08)**	(–8.12)**	(–6.93)**	(–10.06)**	(–6.49)**	(–6.17)**	(–0.55)	(–6.24)**
Tert Ed	1.2	1.1	1.1	1.2	1.6	1.5	1.0	1.6	2.3	1.6	1.1	0.8	1.2	1.1	1.1
	(16.30)**	(12.29)**	(12.12)**	(12.08)**	(11.39)**	(9.93)**	(6.41)**	(6.99)**	(7.09)**	(7.48)**	(12.86)**	(5.87)**	(16.20)**	(14.20)**	(3.77)**
R&D exp BUS to GDP	–	0.2	–	–	–	0.6	–	–	–	–	–	–	–	–	0.9
		(2.86)**				(1.96)*									(4.24)**
R&D exp GOV to GDP	–	–	0.9	–	–	0.8	–	–	–	–	–	–	–	–	–
			(2.22)**			(1.2)									
R&D exp HE to GDP	–	–	–	0.8	–	–2.4	–	–	–	–	–	–	–	–	–
				(1.69)*		(–2.87)**									
R&D exp in PNP to GDP	–	–	–	–	8.0	4.7	–	–	–	–	–	–	–	–	–
					(4.84)**	(5.48)**									
R&D personnel BUS	–	–	–	–	–	–	0.0001	–	–	–	–	–	–	–	0.00001
							(4.45)**								(1.0)
R&D personnel GOV	–	–	–	–	–	–	–	–0.2	–	–	–	–	–	–	–0.13
								(–1.2)							(–0.86)
R&D personnel HE	–	–	–	–	–	–	–	–	–0.7	–	–	–	–	–	–
									(–2.52)*						
R&D personnel PNP	–	–	–	–	–	–	–	–	–	0.2	–	–	–	–	–
										(1.5)					
KIS emp.	–	–	–	–	–	–	–	–	–	–	0.02	–	–	–	0.03
											(2.26)*				(1.17)
HTM emp.	–	–	–	–	–	–	–	–	–	–	0.1	–	–	–	0.02
											(1.66)*				(0.54)
Agg Ag	–	–	–	–	–	–	–	–	–	–	–	–0.2	–	–	0.03
												(–3.79)**			(0.32)

Table 3.5. **Knowledge production function in OECD TL2 regions, 1995-2005** *(cont.)*

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12	Model 13	Model 14	Model 15
Agg Man	–	–	–	–	–	–	–	–	–	–	–	0.1	–	–	0.02
												(0.6)			(0.09)
Agg Fin	–	–	–	–	–	–	–	–	–	–	–	0.4	–	–	0.06
												(2.72)**			(0.25)
Mkt Access	–	–	–	–	–	–	–	–	–	–	–	–	0.05	0.002	0.01
													(0.91)	(0.04)	(0.14)
Dist Mkts	–	–	–	–	–	–	–	–	–	–	–	–	1.4	–0.9	10.1
													(2.54)**	(–0.49)	(5.40)**
F-value	265.6	86.7	83.0	77.3	100.0	50.5	67.2	39.2	46.0	42.6	77.8	82.9	92.8	75.9	55.8
R^2	0.52	0.50	0.49	0.49	0.77	0.83	0.63	0.51	0.56	0.70	0.54	0.61	0.53	0.50	0.94
n	251	177	174	167	62	58	81	79	75	40	206	213	250	234	53

Note: */ Significant at the 95% confidence level; **/ Significant at the 99% confidence level; BUS = business sector; GOV = government sector; HE = higher education institutions; PNP = private non-profit sector; KIS = knowledge intensity services; HTM = high-and medium high-tech manufacturing.
Model 4 excludes 16 outlier regions, mostly from Australia, Canada and Turkey with a value of ln (distance to market) greater than 4.8.

Source: Calculations based on *OECD Regional Database* (2008).

HOW REGIONS GROW: TRENDS AND ANALYSIS – ISBN 978-92-64-03945-2 – © OECD 2009

business sector. The effects of personnel in higher education also appear negative; this is surprising but it might reflect the lack of R&D commercialization in higher education institutions. The small effects of R&D personnel on innovation reflect that fact that the marginal contribution of each person to innovation is not homogeneous across individuals; instead it varies significantly.

- As expected, the presence of knowledge intensive services and high technological manufacturing enhances regional innovation activity in terms of patent applications.
- The presence of economies of agglomeration only has a positive influence on innovation in the case of financial intermediation, while a lack of agglomeration economies reduces patenting activity.
- Finally, being distant from markets seems have a positive effect on innovation (Models 13 and 15), contrary to expectations, although this result is mostly driven by outlying regions.[4] When these regions are taken out of the sample (Model 14), the positive effect vanishes. This means that regional accessibility (and lack of it) does not influence patenting activities. These results suggest that a region's distance from markets or from other regions does not necessarily hinder its capacity to innovate, mainly because communication costs are falling. However, it is important to underline that while distance from markets is not relevant for innovation, proximity among the diverse local actors in a regional innovation system may well remain a key ingredient for innovation. However, our model does not include proximity among actors simply because it is difficult to measure.

In essence the evidence in this section supports the linear view (Box 3.1) in which human capital and R&D expenditure lead to innovation. It also shows that their influence is greater than R&D personnel. Within regional innovation systems, the presence of a specialised workforce – in high-tech manufacturing and knowledge intensive sectors – enhances innovation, as does the presence of economies of agglomeration.

Spatial econometric model

Model specifications

Spatial econometric techniques improve classical econometric methods when there is spatial dependence in the observations. Traditional econometrics have largely ignored these issues.[5] When spatial dependence is present in the data the coefficients estimated by classical econometric methods might be biased and inconsistent.

Spatial econometrics are generally characterised by: *i*) spatial dependence (or spatial autocorrelation) between sample data observations at

 HOW REGIONS GROW: TRENDS AND ANALYSIS – ISBN 978-92-64-03945-2 – © OECD 2009

various points in space (*i.e.* lack of independence which is often present among observations); and *ii*) spatial heterogeneity that arises from relationships or model parameters that vary with sample data as we move through space.

There are different types of spatial data, depending on whether the unit of analysis is an individual data point (*e.g.* geo-referenced or point pattern data, such as a firm or a household) or a geographical region (or areal data, such as administrative divisions). For each type of data, different techniques and models are used. For our area of interest (*i.e.* regions in OECD countries) we have areal (administrative) data.

Spatial dependence means that observations at location i depend on other observations at location $j \neq i$, *i.e.* $y_i = f(y_j)$, for $i = 1,...,n$ and $j \neq i$. We can allow the dependence to be among several observations by letting the index i take any value. To detect spatial patterns (association and autocorrelation), some standard global and local spatial statistics have been developed. These include Moran's *I*, Geary's *C*, *G* statistics, LISA and GLISA (see Annex F).

Two main reasons are commonly given for expecting to find dependence between data and spatial areas:

i) Data collection associated with spatial units might reflect measurement error. This would occur if the administrative boundaries do not accurately reflect the nature of the underlying process generating the sample data.

ii) The spatial dimension of socio-demographic, economic or regional activity may truly be an important aspect.

We therefore need to model the functional spatial dependence. Turning to spatial heterogeneity (the variation in the relationship over space), we might expect a different relationship to hold for every point in space. We cannot however hope to estimate a set of n parameter vectors given a sample of n observations: we simply do not have enough observations. We therefore need to provide a specification for the variation over space (*i.e.* we need to impose restrictions). Specifically, we need to formulate a parsimonious model which reflects the spatial structure in the data.

Because spatial relationships can be defined in an infinite number of ways, we impose a spatial structure on the data by constructing a *spatial weight matrix*, which portrays the neighbourhood structure among spatial units (see Annex F for more details). Spatial dependence is postulated to decrease with distance. We have two sources of information on distance: *i*) the location in Cartesian space represented by longitude and latitude (when we have geo-referenced data, where the distance between two points can be calculated); and *ii*) contiguity (for non-geo-referenced data), reflecting the relative position of one regional (spatial) unit to another such unit, *e.g.* two units are neighbours if they share a common border or edge, and for which we can sometimes calculate the centroid (geographical centre) coordinates and thus the distance between (the centre of) administrative units.

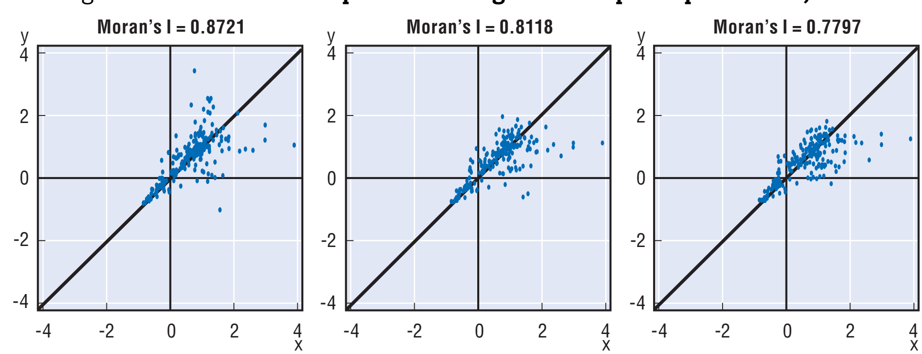
Therefore, the spatial structure (interconnectedness) of administrative units such as TL2 OECD regions can be represented by a weight matrix based either on contiguity criteria and/or on the geographical distance between centroids (or other spatial units such as main cities).

Once we have specified a spatial structure, we can use the spatial weight matrix in a spatial model, which will provide unbiased and consistent estimates in presence of spatial dependence (see Annex F for a brief description of the main spatial models).

Interpreting the results

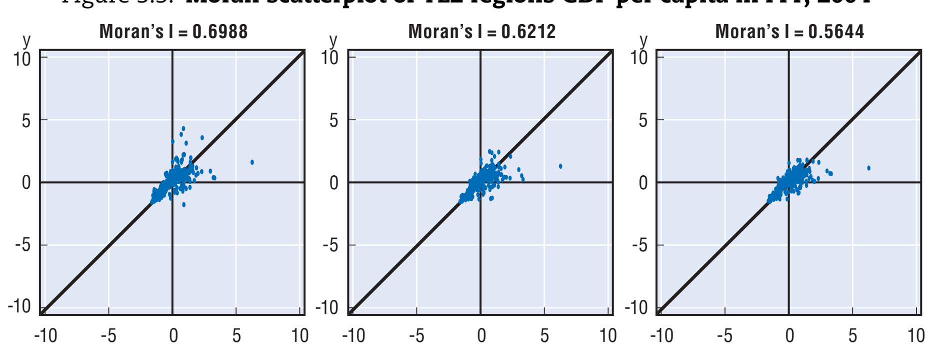
Figures 3.2-3.4 plot the value of GDP PPP per capita (for 1995 and 2004), and GDP PPP per capita growth (1995-2005 period) for all OECD TL2 regions as well as their respective value in neighbouring regions (their spatial lags). The

Figure 3.2. **Moran scatterplot of TL2 regions GDP per capita in PPP, 1995**



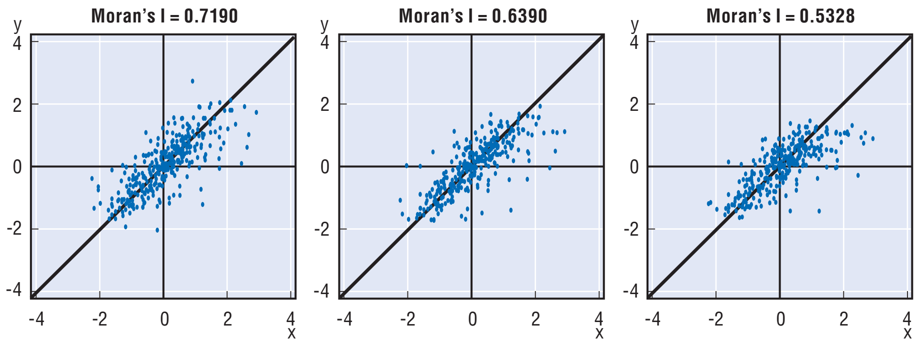
Source: Calculations based on *OECD Regional Database* (2008).

Figure 3.3. **Moran scatterplot of TL2 regions GDP per capita in PPP, 2004**



Source: Calculations based on *OECD Regional Database* (2008).

Figure 3.4. **Moran scatterplot of TL2 regions growth in GDP per capita PPP, 1995-2005**



Source: Calculations based on *OECD Regional Database* (2008).

graphs are divided according to the definition of k-nearest neighbour weight matrices used (*k* = 2, 5 and 10, see Annex F for a definition of k-nearest neighbour weight matrices). The figures display the Moran's *I* statistic, which shows the degree of spatial correlation. Its value ranges from 1 (strong positive spatial autocorrelation) to –1 (strong negative spatial autocorrelation). The datasets are divided into the following four quadrants:

- The first quadrant (top left) associates low values of GDP (or GDP growth) with high values of GDP (or GDP growth) in neighbouring regions (LH).
- The second quadrant (top right) associates high values of GDP (or GDP growth) with high values of GDP (or GDP growth) (HH).
- The third quadrant (bottom right) associates high values of GDP (or GDP growth) with low values (HL) of GDP (or GDP growth).
- The fourth quadrant (bottom right) associates low values of GDP (or GDP growth) with low values (LL).

Figure 3.2-3.8 display a strong positive spatial correlation, and the correlation lessens when more neighbours are added (*i.e.* when k increases). The fact that values are more clustered in 2005 than in 1995 suggests there has been some convergence between neighbouring regions.

Spatial regressions

The spatial econometric regression uses the same model specifications as the cross-sectional and the panel models, with the addition of spatial components. The first regressor is the lagged dependent variable. If it is significant and positive it means that there is significant positive spatial correlation of the dependent variable. The spatial lag model is used here, that is the mixed regressive-spatial autoregressive (SAR) model (see Annex F for a detailed explanation):

$$y = \rho W_1 y + X\beta + \varepsilon$$
$$\varepsilon \sim N(0, \sigma^2 I_n) \quad (4)$$

The weight matrices used in equation 4 are k = 2, 5 and 10 (see Annex F for definitions and a deeper explanation).

Table 3.6. **Spatial results regional economic growth in OECD TL2 regions, 1995-2005**

	Model 1			Model 2			Model 3		
	k-2	k-5	k-10	k-2	k-5	k-10	k-2	k-5	k-10
Lagged Dependent	0.337	0.549	0.645	0.229	0.43	0.52	0.277	0.666	0.628
Variable	(7.226)**	(10.112)**	(10.77)**	(4.384)**	(6.733)**	(7.243)**	(4.891)**	(5.157)**	(8.89)**
Constant	0.017	0.016	0.015	0.126	0.103	0.1	0.134	0.143	0.084
	(1.936)	(1.933)	(1.766)	(7.427)**	(6.143)**	(5.802)**	(4.777)**	(4.78)**	(3.18)**
Initial Y	–0.0003	–0.0007	–0.0008	–0.008	–0.007	–0.007	–0.01	–0.01	–0.006
	(–0.38)	(–0.795)	(–0.911)	(–5.455)**	(–4.672)**	(–4.544)**	(–3.94)**	(–4.45)**	(–2.45)*
Infrast	–	–	–	0.012	0.013	0.014	0.015	0.02	0.015
				(1.937)	(2.225)*	(2.396)*	(2.216)*	(2.474)*	(2.24)*
Prim Ed	–	–	–	–0.01	–0.008	–0.008	–0.002	–0.003	–0.003
				(–8.565)**	(–7.114)**	(–6.7)**	(–3.183)**	(–3.116)**	(–3.144)**
Tert Ed	–	–	–	0.007	0.006	0.005	–	–	–
				(7.152)**	(5.808)**	(5.36)**			
Patents	–	–	–	–	–	–	0.001	0.001	0.0005
							(1.62)	(1.104)	(1.128)
R^2	0.262	0.362	0.336	0.407	0.464	0.46	0.309	0.232	0.411
Breusch-Pagan test	7.43	4.4	3.185	1.642	2.332	1.501	35.18	54	32.34
(p-value test)	0.006	0.036	0.074	0.801	0.675	0.826	0.00	0.00	0.00
n	333	333	333	274	274	274	232	232	232

*/ Significant at the 95% confidence level; **/ Significant at the 99% confidence level.
z-value in parentheses.
Source: Calculations based on *OECD Regional Database* (2008).

This coefficient finds that the performances of neighbouring regions strongly influences the performance of any given region in the OECD (Table 3.9), suggesting that inter-regional trade and inter-regional linkages play an important role in a region's performance.

This spatial correlation with growth also confirms that infrastructure and human capital are drivers of economic expansion, but it does not confirm previous results for innovation (Model 3). These results suggest that as capital and talent agglomerate they tend to positively influence growth in neighbouring regions, but innovation remains a highly local element that does not necessarily influence growth in neighbouring regions. It is also possible

 HOW REGIONS GROW: TRENDS AND ANALYSIS – ISBN 978-92-64-03945-2 – © OECD 2009

that our models should attempt to incorporate lagged values, as in our panel data analysis, at the same time that spatial econometrics is carried out.

Conclusions

In sum, this chapter reveals that regional growth depends on endogenous growth factors such as education and innovation, but also on infrastructure and on forces described by the NEG such as economies of agglomeration and geographic characteristics. The results show that policies can benefit from an integrated approach: polices aiming at providing infrastructure only are bound to be unsuccessful as endogenous growth factors such as human capital and innovation need also to be taken into account.

The dynamic panel model found that infrastructure and human capital require three years to positively influence regional growth, while innovation is a longer-term process, having a positive effect on regional growth only after a five-year period.

Our analysis based on a knowledge production function related innovation input variables such as human capital and R&D to innovation outcomes such as patenting activity. We found that *i*) human capital has a strong impact on regional growth both directly (from previous analysis), and indirectly, through patenting; *ii*) R&D is an indirect determinant of growth through its impact on patenting activity; and *iii*) geographical space plays a role in determining innovation in these models as agglomeration economies emerge as a determinant.

Our final analysis, spatial econometrics, found that the performance of neighbouring regions influences the performance of any given region in the OECD.

The main lessons from these results are that in order to promote regional growth, policy-makers should develop a comprehensive regional policy that not only links regions through infrastructure investments, but that also fosters human capital formation and facilitates the process of innovation. The risk of piecemeal visions of regional policy or of sectoral polices, such as only promoting human capital or only providing infrastructure, is that a "leaking" instead of a linking process will be created.

Notes

1. The term *economies of agglomeration* is used in *urban economics* to describe the benefits that firms obtain when locating near each other.
2. The basic proposition is that holding constant population expansion, and in the absence of technological progress, diminishing returns to scale will bring about convergence (Aghion and Howitt, 1998).

3. We defined the specialisation index as $Sp_i = \frac{Y_{ij}/Y_j}{Y_i/Y}$ where Y_{ij} is total employment of industry *i* in region *j*, Y_j is total employment in region *j* of all industries, Y_i is the national employment in industry *i*, and *Y* is the total national employment of all industries. A value of the index above 1 shows specialisation in an industry and a value below 1 shows non-specialisation.
4. We observe 16 regions as outliers in the data, mostly from Australia, Canada and Turkey, with a value of ln (distance to market) greater than 4.8.
5. Perhaps because they violate the Gauss-Markov assumptions used in regression modelling, *i.e.* that the distribution of the sample data exhibits a constant mean and variance as we move across observations.