Volume 5 No 2 2000

Cahiers Papers

Volume 5 No 2 2000

BEI EIB

Cahiers Papers

Regional convergence in Europe: Theory and empirical evidence

BEI EIB

European Investment Bank

Editorial policy

The *EIB Papers* are published twice a year by the Chief Economist's Department of the European Investment Bank. The journal is aimed at encouraging high-quality economic research and debate on matters of a European interest. As such the *Papers* are intended to be accessible to non-specialist readers and emphasise policy dimensions rather than technical issues. They present the results of research carried out by Bank staff together with contributions from external scholars and specialists. The winning essays of the biennial *EIB Prize* are also published in the winter edition during competition years.

Articles will only be accepted for publication on the condition that they have not already been published elsewhere. All articles in the *EIB Papers* may be freely reproduced and quoted; however, the Editor would appreciate an appropriate acknowledgement and a copy of the publication in question.

The views expressed in the articles are those of the individual authors and do not necessarily reflect the position of the EIB.

Les *Cahiers de la BEI* sont publiés deux fois par an par le département de l'économiste en chef de la Banque Européenne d'Investissement. Le journal a pour but d'encourager la recherche de haute qualité en économie et le débat sur des sujets d'intérêt européen. En tant que tels, il est attendu des *Cahiers* qu'ils soient accessibles aux lecteurs non spécialistes et qu'ils mettent l'accent sur les dimensions politiques plutôt que les problèmes techniques. Ils présentent les résultats de la recherche menée par le staff de la Banque ainsi que des contributions de chercheurs et de spécialistes extérieurs. Les essais gagnants du *Prix de la BEI* proposé tous les deux ans sont aussi publiés dans l'édition d'hiver les années de compétition.

Les articles ne seront acceptés pour publication qu'à condition qu'ils n'aient été publiés ailleurs. Tous les articles des *Cahiers de la BEI* peuvent être librement reproduits et cités; cependant, l'Editeur apprécierait qu'un crédit approprié leur soit donné et qu'une copie de la publication lui soit envoyé.

Les vues exprimées dans les articles sont personnelles aux auteurs et ne reflètent pas nécessairement la position de la BEI.

Cahiers Papers

BEI EIB

Regional convergence in Europe: Theory and empirical evidence

European Investment Bank

Contents

Preface

Ewald Nowotny

Vice-President

The main raison d'être of the EIB was set out in the Treaty of Rome in 1957: it is to support the balanced development of the European Union, and in particular to facilitate the development of "less-developed regions". As with the EU budgetary funds, this has largely been achieved by helping to finance investments located in lagging regions. Over the last decade there have been major steps in the integration of Europe. The Single Market has been established, and the euro has become a reality. This has much reduced the economic and financial barriers between countries that could have been one justification for policy intervention of this kind.

The last decade has also seen a revolution in the thinking about geography by the economics profession. In traditional neo-classical models, regional differences come about because local resources are different. However, the seminal paper by Paul Krugman in 1991 made the link between imperfect competition and the spatial distribution of economic activity. He showed that economic agglomeration could arise because firms choose to be located in certain areas given economies of scale in production process.

Over the last year, work at the Chief Economist's Department of the EIB has tried to pull these diverse strands together. The purpose of this particular edition of the EIB Papers *is to try to provide an intellectual framework for thinking about reducing regional disparities. Another edition (also of Volume 5) takes this framework to discuss the possible implications for EU policy.*

The questions addressed in this edition are: what do economic theory and the empirical evidence tell us about the forces that lead to convergence or divergence? What is the motivation for government intervention: is it to increase equity or efficiency? What issues does government intervention raise? For example, should state support for lagging areas come via a priority for public infrastructure or through a priority for support for the private sector?

I believe the analysis presented here does provide a convincing economic logic for regional development policy - though designing the correct vehicle for support may be very much more complex than we had thought hitherto. Having a clear economic framework to address these issues is essential when we look forward to the new challenge of enlarging the EU to countries with standards of living well below the EU average.

Editors' Introduction

Both editions of this year's EIB Papers are devoted to regional development in Europe. In this edition, we discuss the theoretical framework for understanding the existence of regional disparities. A second edition ("Regional Development in Europe: An assessment of policy strategies", EIB Papers, Volume 5, Number 1) uses this background to analyse a number of case studies. A summary of both publications is given in the overview paper by **Christopher Hurst, Jacques-François Thisse** *and* **Patrick Vanhoudt** *available in that edition. Here we give a brief introduction to the justifications for policy intervention.*

It is well known that over the past two centuries of rapid global growth, the gap in per capita incomes between rich and poor countries in the world has widened dramatically. In more recent decades, this gap has seemed to stabilise somewhat as many once-poor countries made faster progress, but in the aggregate, income inequality among nations has failed to diminish. It seems that the rich get richer, and so do the poor, but without ever catching-up. In broad terms, Europe does not seem to have reacted much differently.

Daniel Moucque *(Directorate General for Regional Policies, European Commission, Brussels) draws on the European Commission's Sixth Periodic Report to emphasise the relative better performance of the poorest regions in the EU. His paper shows that, over the decade from 1986 to 1996, the 25 poorest regions have been able to increase their per capita income from 52 percent of the EU average to 59 percent. In the Cohesion countries (Ireland, Greece, Portugal, Spain), income per head went up from 65 percent to 77 percent of the Union's average.*

This apparent shrinking of the lowest tail of the income distribution certainly deserves to be highlighted. However, looking at the average income of a sub-sample of regions gives only a limited picture of the evolution of the full distribution. For example, taking the whole sample, the coefficient of variation, a measure of dispersion, indicates that the distribution of per capita income among the EU regions has been rather stable. To be more precise, while the standard deviation was 27 percent of the average in 1986, it was still 26 percent in 1996 - an almost identical figure (1). In fact, using the same indicator (the sub-sample average with respect to the total sample average) there is little evidence that the richest regions are growing less quickly than the average. For instance, for the 20 percent richest regions this indicator increased from 131 percent of the EU average in 1986 to 135 percent in 1996, and went up from 144 percent to 150 percent for the top decile.

---

1) We look at the sample of 142 NUTS-2 regions for which the Regio database of Eurostat provides data on PPP adjusted per capita incomes in 1986 and 1996, and ignore the Dutch province of Groningen. This latter region has known an enormous decline due to the evolution in its local gas production and is an outlier in the sample. Clearly, the New Länder of Germany are also excluded since they only entered the sample mid-way through the period.

Perhaps even more striking is the fact that the composition of any part of the distribution may change. Let us look again at the members of the lowest decile of the distribution. Some regions that were not in the lowest 10 percent of the distribution in 1986 had entered there by 1996 (e.g. Peloponnisos and Dytiki Makedonia in Greece, and Calabria in Italy had replaced Galicia and Castilla-la Mancha in Spain, and Algarve in Portugal). Thus, in spite of the improvement of the average income in this part of the distribution, some regions had dropped into the sample from above. They had such slow growth rates over the period that they experienced a relative decline in their standard of living. This illustrates that the evolution of the shape of the distribution is a complex phenomenon. There are winners and losers, and yet other regions that maintain a relatively static position. Taking this into account, our interpretation of the broad picture of European regional GDP is that not much convergence has taken place.

Moucque also notes that the situation regarding unemployment reflects no convergence whatsoever. Indeed, disparities have rather shown a tendency to increase. The catching-up of the poorest regions with the EU average has been predominantly due to the productivity growth of those that remained employed, rather than to job creation. This may be less controversial than it seems at first sight. Moucque argues that the most critical factors for lagging behind are an unfavourable sectoral structure together with low levels of education among the work force and a lack of innovative capacity. Technical change in such an environment increases the potential for economic growth, but it may also increase unemployment. As discussed in an earlier edition of the EIB Papers ("Employment in Europe", Volume 3, Number 1, 1998) there are a number of potential reasons for this, though the root causes are rigidities in EU labour markets.

While one can debate the extent to which there has - or has not - been convergence in the EU in the past, it is clear that the enlargement of the Union eastwards, will increase the disparity among EU regions once more. The challenge to obtain social and economic cohesion among regions will go on, but what, in fact, does economic theory teach us about determinants of regional inequalities?

Angel de la Fuente *(Instituto Análisis Económico, Barcelona) makes clear that this question is rather easy to answer when three conditions are satisfied. To be precise, if all markets are highly competitive, market imperfections are essentially non-existent, and if there are no external benefits or spillovers, economies will converge towards an equilibrium per capita income that is dictated by fundamentals. These include each economy's propensity to invest, and the rates of growth of technology and population. In other words, even in a perfect world, some degree of regional imbalances would exist, though in this case disparities are driven purely by preferences. Policies that alter economic fundamentals, such as increasing investment, would consequently have an impact on the level of inequality. Interestingly, when the predictions from this neo-classical theory are tested against aggregated data (i.e. mainly at the country level), they seem to survive rather well. The overwhelming bulk of the evidence is for convergence towards such country specific equilibria, particularly if human capital is included as a type of investment.*

However, levels of inequality among European regions are often said to be excessively high and, as documented earlier, they do not show much tendency to decline. In this context, policy makers sometimes mention studies indicating that regional imbalances in the EU are currently twice as high as in the US. Is the EU economy somehow operating in a non-optimal way?

In his paper, **Jacques-François Thisse** *(UC Louvain) argues that widening disparities might not be bad in some cases, yet need monitoring and correction in others. Thisse starts from the fundamental questions of why local clusters emerge in different places and why economic activity is located in the most urbanised areas of the Union. He explains that there are a number of economic forces that drive agglomeration. These are technological and pecuniary externalities, and a co-ordination problem.*

For example, technological externalities may arise when companies learn from each other how to do things better. In this case, Thisse argues that the formation and the size of clusters that emerge depends on the relative strength of three distinct forces: the magnitude of localisation economies (i.e. externalities affecting all the firms belonging to the same sector in an area), the intensity of competition, and the level of transportation costs. Although the joint effect is not clear, some partial relations are. Firstly, low transportation costs are likely to drive the economy towards more agglomeration, because firms do not fear losing business in distant markets. Secondly, more product differentiation, which relaxes price competition, induces more firms to locate together. In that way, they can exploit the benefits from being in large clusters without being punished by competition. Thirdly, the impact of localisation economies becomes larger when market size increases, and increases in the demand for differentiated products induce greater asymmetry between clusters. Consequently, the process of European integration could have been a source of growing disparities, rather than one of overall convergence!

Should this be a cause for concern? What if per capita income over the last decade increased in a poor region by, say, 10 percent, and by 20 percent in a rich one? Both regions clearly improved their situation, although the poorer region was not able to catch-up. Is this a bad thing? To answer this, we need to compare the market outcome with the decisions of a hypothetical planner who allocates firms to maximise the standard of living in the society. This solution may display a similar pattern to that arising when firms are free to choose locations. Surprisingly, in certain circumstances the planner could also opt for a larger disparity among regions. This occurs because firms may try to relax competition through keeping other firms at arms length, even though it would be more efficient for them to be located at the same spot.

Thisse also argues that over- or under-concentration at the interregional level also depends on pecuniary externalities. By this we mean the following: when some workers choose to move away from their region, they are likely to affect both the labour and product markets in various ways. However, workers do not take into account the impact of their migration decisions on the well being of those who stay put, nor on those living in the region of destination. Still their relocation will change the level of demand inside the regions, possibly making the region of destination

even more attractive for firms. The relocation will at the same time depress the labour market in this region so that, all else being equal, the wage is affected negatively. Taking these pecuniary externalities into account, Thisse shows that there is a range of transport costs for which there is too much agglomeration and dispersion compared to the planner's optimum. However, the market outcome remains the most efficient either when transport costs are very low or very high. This means that we cannot say much in general about whether the natural level of agglomeration is optimal or not from an efficiency point of view.

The statistical evidence regarding convergence referred to by de la Fuente, and mentioned earlier, gives little insight into this issue. Most studies use data from countries or relatively large regions, and the scale is such that local agglomeration effects are lost in the averages. In any case, Thisse doubts the value of much empirical work at the regional level since the administrative regions considered are historical accidents rather than well-defined economic units.

A third possibility to explain why imbalances can prevail - a co-ordination failure. In some cases a region does not take-off because a minimum threshold of economic activity has not been reached. No-one knows how a new business would perform in such a region. Indeed, even the prices for some goods and services may not be known in advance. Since many economic agents must work together to launch a new market, the absence of adequate information is a recipe for a Catch-22 situation, and the region remains permanently underdeveloped. The consequence of such a market failure on economic efficiency is clearly negative since optimal investment decisions are simply not taken.

Policy may also be motivated by equity considerations. If people are, for whatever reason, stuck in a region, they might unreasonably suffer from industrial relocation away from their region. Cultural and language barriers explain labour immobility at the EU level, but people also tend not to move within countries. The reasons for this include regulations in housing markets, and a lack of information on job opportunities elsewhere.

However, **Philippe Martin** *(Ecole Nationale des Ponts et Chaussées, Paris) clarifies that the equity motivation behind regional policies is not as straightforward as it seems. Following a similar logic as Thisse, he shows that public policies aimed at altering economic geography may sometimes have a contradictory impact. For example, a new highway linking a lagging region may simply expose that region to increased competition from imports. The net result of public aid may simply be a transfer of income from rich to poor without improving the recipient's productivity. In that case, intervention could lead to lower overall prosperity if the implemented policies drain resources from the most innovative regions.*

Martin also notes that these long-term impacts may be hidden in the short-term. This is because spending on public infrastructure has both demand and supply effects. Whereas the short-run demand effects are relatively easy to understand in a Keynesian framework through the impact of

the multiplier, the long-run supply effect is much more difficult to assess. The European Commission has employed several Keynesian-type of models to show that the Structural Funds have contributed to the Cohesion countries' growth performance. Martin argues, however, that these national estimates are very difficult to interpret at the local level. Even worse, the long-term supply effects may be in exactly the opposite direction if the policy intervention - take the better highway example again - leads to a relocation of business away from the region in question. As a result, Martin suggests that the relation between growth and public spending remains fragile, at best.

This conclusion is reinforced in the study by **Patrick Vanhoudt**, **Thomas Mathä** *and* **Bert Smid** *(EIB). They develop a long-term growth model in order to investigate the differential impact of private and public investment on the economic performances of regions and countries within the EU. The authors find evidence of reverse causality between public investment and economic growth. What seems to be the case is that richer countries have been able to invest more in public capital, thereby willing to forego a higher pace of growth. At the regional level, the results indicate that spending on public capital seems to have mainly been used as an instrument for income redistribution. It has not, however, closed productivity gaps. On the other hand, private capital investments and increases in the level of education have been effective in stimulating regional growth and reducing disparities. The authors also find that the speed of conditional convergence (i.e. the concept developed by de la Fuente) in Europe is only half the size of the one reported for the US. This may reflect a low degree of factor mobility - especially of labour - in Europe.*

In sum, we can say that the possibility of a co-ordination failure in the economic development of lagging regions does give a general justification for policy intervention from an efficiency point of view. However, the other economic forces leading to agglomeration mean that the long-run outcome of particular policies may be counter-productive. Any policy recommendation must therefore rest on a detailed understanding of the agglomeration forces at work. Unfortunately, this is a formidable exercise. To quote Martin: "the policy mistakes are going to be numerous because the information requirement is too severe."

Perhaps for this reason, much of what has been done can be seen in terms of income redistribution, but appears rather less effective in terms of productivity catch-up. This is particularly the case for public investment. The question is what happens to local institutions when a region gets use to a steady stream of transfers from the outside? What happens when public grants and loans are not continued? These broader issues are elaborated further in the companion edition of the EIB Papers on regional development policies.

Christopher Hurst and Patrick Vanhoudt

A survey of socio-economic disparities between the regions of the EU

Daniel Moucque

1. Introduction

The European Union is facing challenges that have important implications for its economy as a whole and for regional cohesion in particular. The transition to the euro had already started, as had the process of enlargement towards Central and Eastern European countries. This occurs against a backdrop of increasing globalisation and a 'second industrial revolution' based on information technology.

The current essay consequently surveys the resulting social and economic trends in the regions. Section 2 finds that there is a clear distinction between a successful, prosperous and competitive EU core and regions, often in the periphery, which perform less well. Section 3 briefly surveys the factors that underlie competitiveness, and suggests that the main reasons for poor performance can be found in unfavourable sectoral structures, a lack of innovative capacity, poor accessibility and low education. The situation in Eastern Europe is discussed in Section 4. Finally, Section 5 summarises and concludes.

2. The situation in the regions

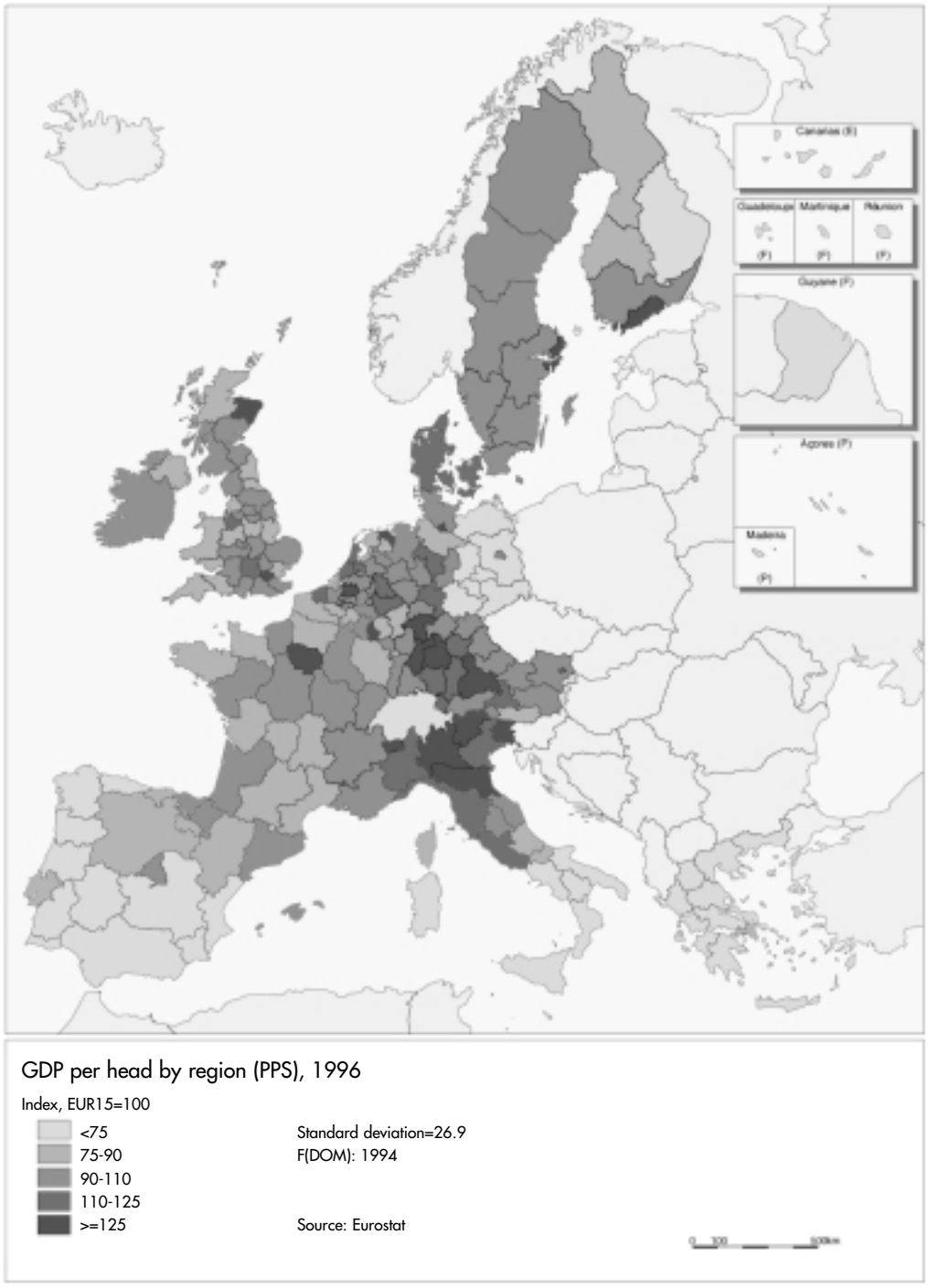
In previous Periodic (1981, 1984, 1987, 1991, 1994) and annual Cohesion (1995a,b to1999) Reports, the first signs of real convergence of lagging regions were detected, but the message was mixed, with some indicators showing convergence while others were unclear. The evidence now shows that the output per head of the poorest regions has converged towards the EU average. For example, over the 10 years from 1986 to 1996:

- GDP per head in the 10 regions where this was lowest, increased from 41% of the EU average to 50%. In the 25 poorest regions, it rose from 52% to 59%.
- GDP per head in the four Cohesion countries went up from 65% of the EU average to 76½%, and, according to estimates, to 78% in 1999.

This pace of convergence has been driven largely by closer European economic integration. As an example, exports and imports between the Cohesion Four and other EU Member States have doubled in real terms over the past decade and now amount in each case to around 120 billion ECU. In addition, the Structural Funds have contributed to the reduction in regional disparities across the Union. Although the precise impact is difficult to measure, the four main macroeconomic models used to assess the Funds suggest that one third of the reduction in disparities is due to the Funds. One model suggests that around ½ percentage point or more has been added to the growth of Objective 1 regions (cf. infra). Another suggests that, the cumulative effect of the Funds so far has increased the GDP of Greece, Ireland and Portugal by nearly 10% and that of Spain (much of which is not covered by Objective 1) by over 4%. Around half of this represents a supply side improvement and not just a boost in aggregate demand.

Daniel Moucque is with the Directorate General for Regional Policies at the European Commission, Brussels. He was editor of the 6th Periodic Report, and one of the co-authors of the 1st Cohesion Report. This paper summarises the main findings of the 6th Periodic Report.

Figure 1. The regional distribution of GDP per head, 1996



However, significant disparities remain; even where catching up is occurring relatively fast, the full process can take a generation or more. In addition, although most regions may experience at least some convergence, their performance varies widely, as can be observed from Figure 1. The more favoured lagging regions, particularly those hosting capital cities such as Dublin or Lisbon, are catching up much more rapidly than their rural hinterlands.

The situation regarding unemployment is less positive. Despite the recent cyclical recovery, unemployment in the EU still stood at just under 10% in late 1998, equivalent to 16½ million people. The overall increase is concentrated in some regions, while others are hardly affected. The 25 regions with the lowest unemployment have hardly changed over the last decade, with rates steady at around 4%. Rates for the most affected regions have climbed from 20% to nearly 24% and are continuing to increase despite decreasing rates elsewhere in Europe. Figure 2 illustrates the regional distribution of unemployment in 1996.

One particular concern here is the high proportions of long term unemployment; 48% of the unemployed have been so for more than one year. A closely related problem is the exclusion of certain individuals and social groups – such as women and young people – from the labour market. These forms of unemployment are particularly worrying, since they are relatively resistant to general improvements in the economy. The high overall rates in the 25 most affected regions are in fact largely driven by such problems. In such regions, the long term unemployed account for 60% of total unemployment (as against 30% in the 25 regions least affected). In addition, only 30% of women of working age there have a job, and youth unemployment rates average 47%.

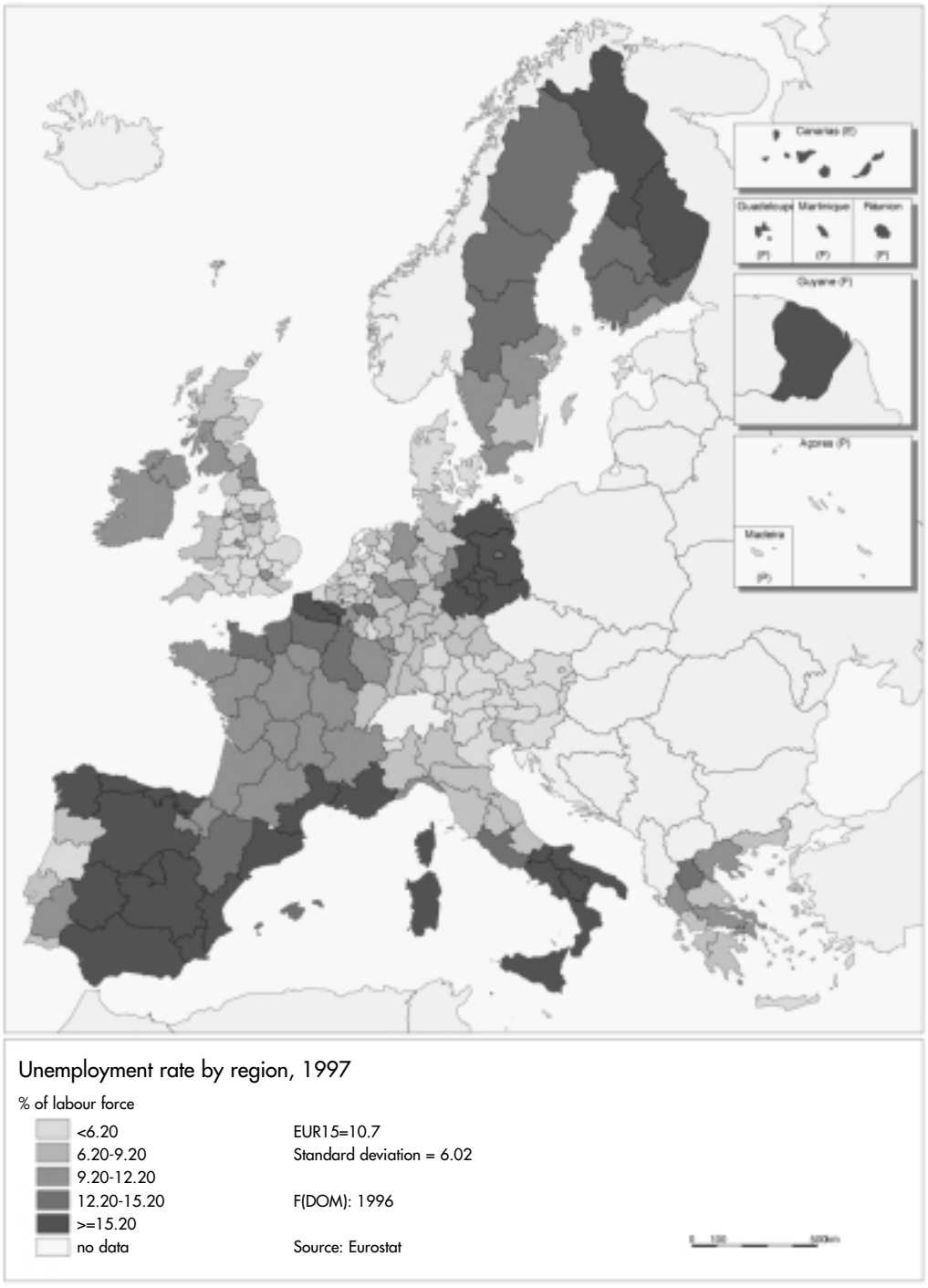
The resumption of growth alone will not tackle such problems. What is needed is an integrated approach combining improvements to the economic base in these regions with training measures aimed at reskilling those at a disadvantage in the labour market and re-inserting them into the world of work. In addition, where female and young workers are so underused, mainstreaming of policies aimed at them is not an option but a necessity.

The Commission has designated the most disadvantaged regions as "Objective 1 areas". They are characterised by a GDP per capita that is lower than 75% of the Union's average for the last three consecutive years. In terms of population, Objective 1 areas account for 22% of the EU population, or around 83 million people. Lower productivity and lower employment rates than in other parts of the Union have been the main causes for the differences vis-à-vis the EU-average in the standard of living in these areas. For some regions, notably Ireland and regions in Spain and Southern Italy, productivity is close to (or in the case of Ireland, above) the EU average, and the main challenge is the generation of employment. Conversely, Portugal and Eastern Germany have relatively high employment rates, but both would need to boost productivity by 50% to converge on the EU average. For Greece, significant improvements would need to be made both in productivity (by 40%) and employment (20%).

The closing of the gap in lagging regions is predominantly due to increases in productivity rather than higher employment.

However, the closing of the gap that has occurred since 1989 in Objective 1 regions is predominantly due to increases in productivity growth, rather than higher employment. Indeed, unemployment is a major problem in many such regions. Just over one in six of the labour force in Objective 1 regions are unemployed, compared with one in ten in the EU as whole.

Figure 2. The distribution of unemployment.



3. Drivers of competitiveness and catch-up

In fact, the regions of the EU can be roughly divided into three types (though some regions do not fit neatly into a single category):

- Large urban service centres. These regions typically perform well in terms of both GDP and employment. The 25 regions most concentrated in services have an output per head that is 27% above the EU average. Since the service sector is the main source of employment in the EU – jobs in market services in particular increasing by 12 million over the past decade – service centres generate significant employment opportunities, often extending well beyond the region concerned. Nevertheless, there can still be serious unemployment blackspots within the cities themselves.

- Industrial regions, the economy of which tends to be centred on medium-sized cities, which are often part of a network. The fortunes of these regions depend strongly on the health of the particular industries located there. Since much of the sector is performing well, manufacturing regions are often successful; the 25 regions in which employment is most concentrated in manufacturing have an output per head 8% above the EU average and unemployment of over 1½ percentage points below the average. However, a minority of industrial regions particularly affected by restructuring have high rates of unemployment, sometimes (but not always) combined with moderately low GDP per head.

- Rural regions, with relatively high employment in agriculture. These regions generally perform reasonably well in terms of unemployment, although problems may show up in other ways, e.g. in terms of high outward migration. However, some agricultural subsectors are low value-added and face significant restructuring pressures. The 25 regions with the very highest dependence on agriculture (and this can be extreme, covering anything up to 40% of the labour force) are particularly affected and have an average unemployment rate of 14.7%. This underlines the importance of facilitating diversification.

An unfavourable sectoral structure together with a lack of innovative capacity are the most important factors underlying lagging competitiveness.

Studies conducted for the Sixth Periodic Report found that an unfavourable sectoral structure together with a lack of innovative capacity seems to be among the most important factors underlying lagging competitiveness, suggesting that the key development challenge in the regions affected is to improve the productive base and their potential for growth.

Poor accessibility and low levels of education are often contributing factors to reduced competitiveness. Even though disparities in the education levels of the work force are tending to narrow, significant differences remain in the relative number of young people in education and initial vocational training beyond compulsory schooling. The weight of the past is reflected in the high proportion of people of working age with only a basic level of education. Three-quarters of those aged 25 to 59 in Portugal and two-thirds in Spain have no qualifications beyond basic schooling. These figures are substantially lower, however, for the 25 to 39 age group, reflecting the progress being made to raise levels.

In addition, the technology gap (measured by such indicators as patent applications and spending on research, see figure 3) between the Cohesion countries and the other Member States far exceeds the gap in GDP per head (except for Ireland, which has more or less caught up in both respects). The disparities are most significant in terms of output indicators, i.e. in terms of the innovations which stem from research and development, underlining the need to improve the efficiency of the process

by which research effort is translated into new products or more efficient ways of doing things in lagging regions. In this respect, it is important to bear in mind that companies can innovate and become more competitive through the transfer of technology, possibly by means of direct investment, without necessarily having to do their own research and development and applying for patents.

Small and medium enterprises (SMEs) play a major role in employment creation and the development of lagging regions. The number of SMEs is highest in the Southern Member States, although this is partly due to their different pattern of sectoral specialisation. In addition, SMEs tend to be concentrated in more favoured regions of these countries, particularly capital cities, while in the poorest regions there are comparatively few. Tackling such imbalances must be part of an integrated approach to regional development which also takes account of the sectoral distribution of SMEs and the extent of their presence in the more dynamic sectors. Recent research suggests that the potential contribution of SMEs to development depends on other conditions, such as the availability of support services and on their links with large firms and/or the networks between them.

Foreign direct investment (FDI) contributes to regional development, not just by increasing the capital stock but also by introducing new products and techniques. In order for lagging regions to derive the full benefits of FDI, however, it is important that the firms making the investment become integrated into the local economy. Over the past 10 years, the EU has been the world's major investor abroad, but it has also received large inflows of FDI. In relation to GDP, Ireland especially but also Portugal and Spain have benefited from above average inflows of investment from countries outside the EU as well as from other Member States.

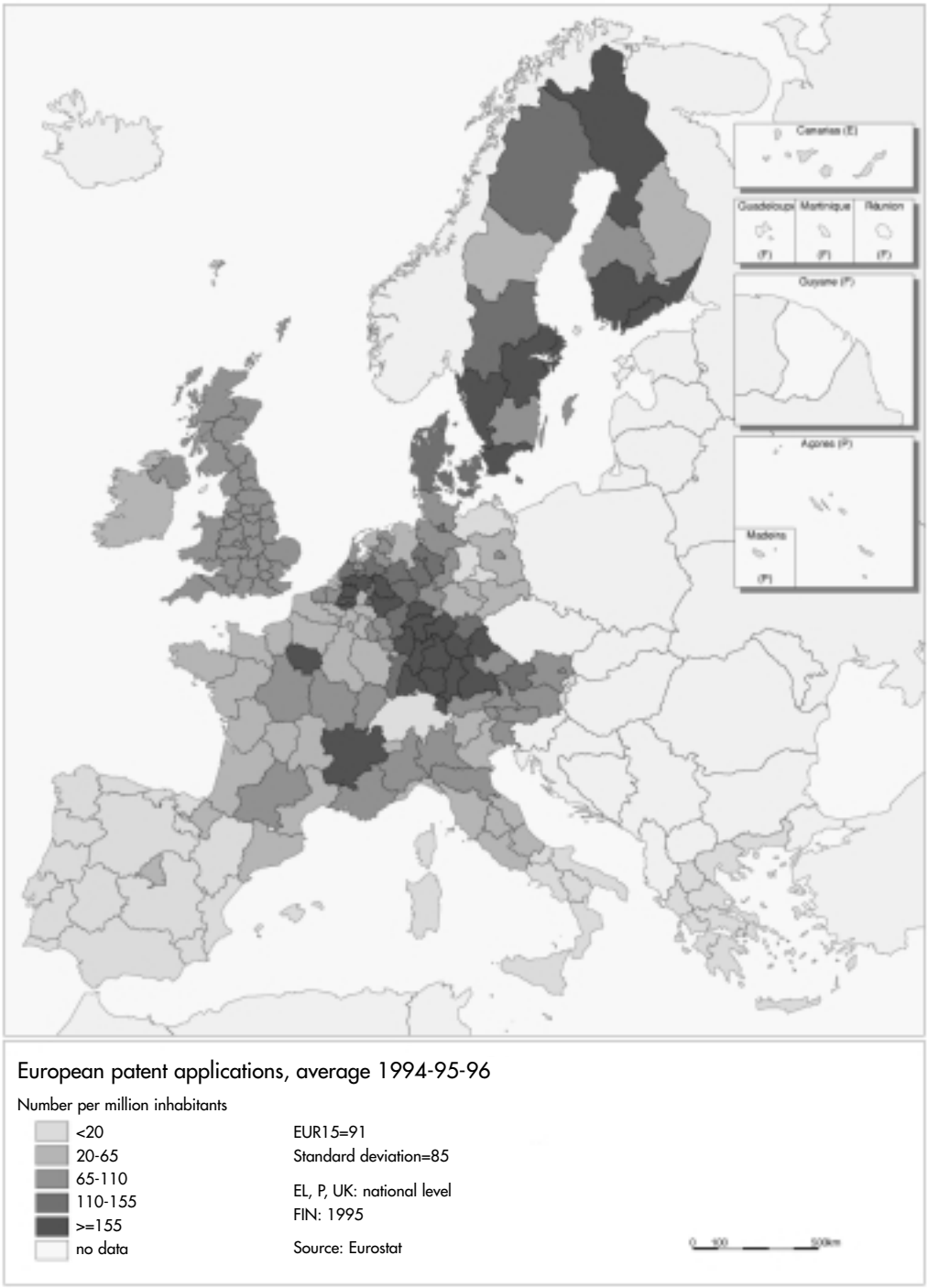
Despite progress in recent years, significant disparities in transport infrastructure remain between regions, and the four Cohesion countries still lag behind other parts of the Union, particularly in terms of the standard of provision. More progress has been made in reducing disparities in telecommunications infrastructure. The Cohesion countries still have somewhat less extensive networks, as measured by the number of telephone lines per 100 inhabitants. However, with the notable exception of Greece, the gap in the quality of networks, as measured by the extent of digitalisation, has largely been eliminated.

The availability of reliable sources of energy at reasonable cost is closely linked to economic growth and development. Investment in energy infrastructure is necessary to close the remaining disparities in provision between different regions. In particular, the market in natural gas is still very segmented, and certain regions continue to be at a disadvantage in terms both of market structure and of infrastructure.

Institutional factors are increasingly seen as key elements in competitiveness.

Institutional factors are increasingly seen as key elements in competitiveness. Such factors include the endowment of social capital, in the form of the business culture and shared social norms of behaviour which facilitate co-operation and enterprise, which is of particular importance for regional development. Networks between firms are both a product of social capital and an element of it. These combine the economies of scale normally open only to large firms with the dynamism and flexibility of small units and, as such, are especially important for innovation. The success of Northern Italy, for example, or the lagging development of many parts of the South, cannot be explained simply in terms of the structure of economic activity, accessibility and education levels.

Figure 3. Number of patent applications as an indicator of innovative capacity.



The efficiency of public administration is another institutional factor of importance. In recent years, there have been significant changes in the principles governing public sector management, a key feature being emphasis on performance evaluation, so that lessons from the past can be systematically fed into decision-making to improve policy in the future (to create a 'learning organisation'). Another feature is a shift towards decentralisation and partnership, enabling different levels of government as well as the private sector to participate in the policy process and to bring their different kinds of expertise and experience to bear.

A long-term strategy is needed, which builds up social capital parallel to physical capital, and improves the skills of the work force.

This, therefore, argues strongly for an integrated approach to regional development that explicitly acknowledges the complexity of the process and takes due account of the interaction between factors, intangible as well as tangible. The need, in sum, is for a long-term strategy which addresses simultaneously the many aspects of the problem of a lack of competitiveness and attempts to build up the social capital of a region – its business culture, administrative structure, institutional relationships and so on – in parallel with its physical infrastructure, the skills of its work force and its productive base.

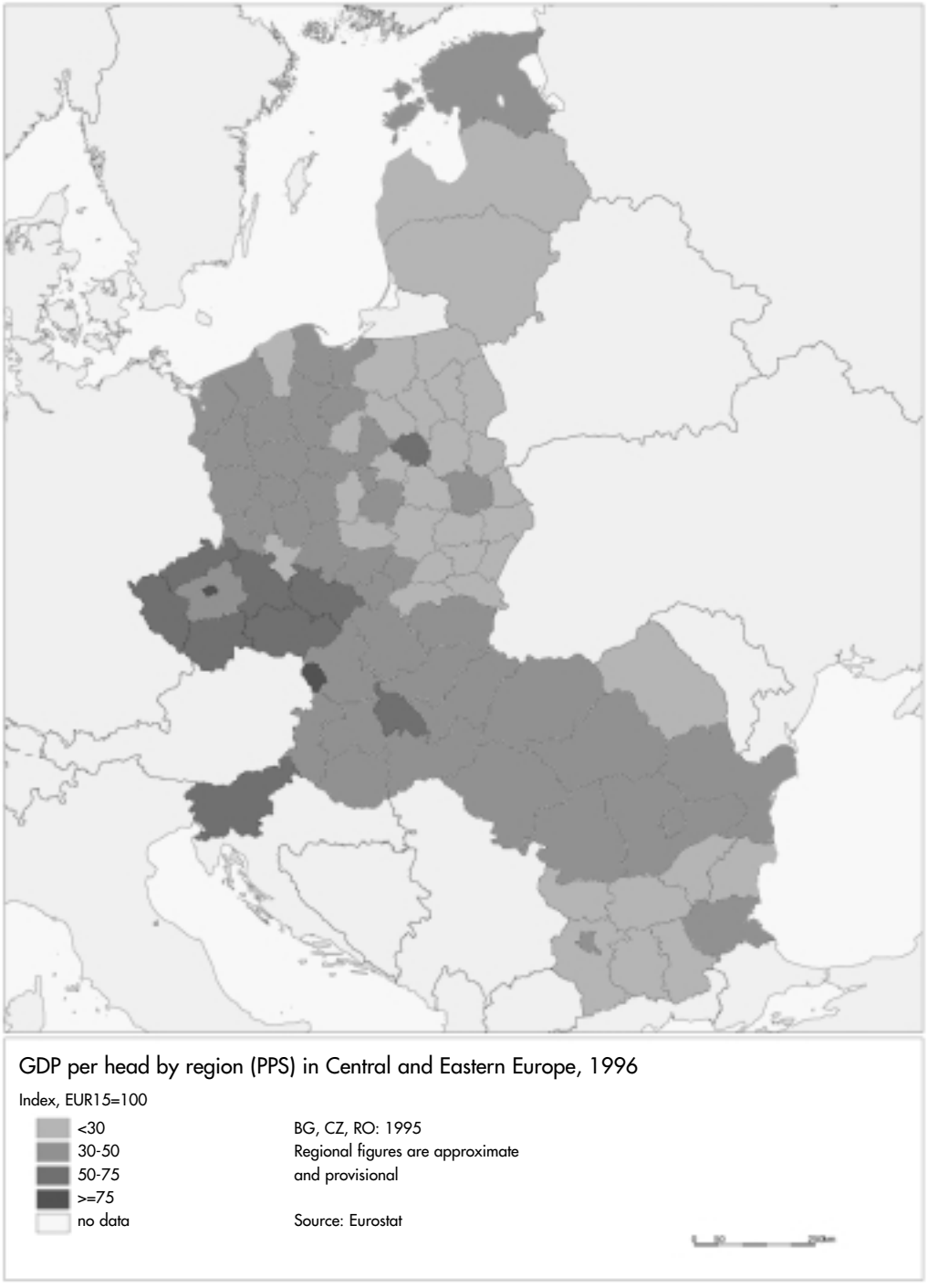
4. Enlargement to Eastern Europe

The situation in the Central and Eastern European (CEE) countries has evolved rapidly since the collapse of the previous regime about one decade ago. After initial sharp falls in income and output, most of the CEE countries have experienced growth since 1993 or 1994. The recovery has, in general, been most marked in the countries that have made the most progress in moving towards a market economy, underlining the gains to be achieved from reform. On the basis of the recovery and closer economic integration with the EU, many CEE countries have made large strides towards preparing for EU membership.

European-wide economic integration is reflected in growing trade flows. By 1995, the EU was the main trade partner of all CEE countries, and the share of the latter in total EU trade is now superior to that of Japan. This has given rise to a significant EU trade surplus with the countries and EU-CEE exchanges are increasingly dominated by intra-industry trade. The CEE countries as a group are also experiencing a significant inflow of foreign direct investment, though flows are concentrated in a few countries with well-advanced reform programmes. EU Member States are by far the main source of investment, further confirming the increasing degree of economic integration.

The economic recovery from 1993 onwards has allowed some CEE countries to narrow the gap in output per head with the Union. Figure 4 shows that in 1996, GDP per head in the countries was nonetheless barely around 40% of the EU average, thus far less than the economic performance in the Cohesion countries. In addition, this masks significant imbalances, such as Latvia, whose GDP per head is only 25% of the EU average, and Slovenia, for which this figure is closer to 67%. Only two regions, Prague and Bratislava, have a GDP per head above 75% of the EU average. Consequently, much remains to be achieved. Even if the Eastern applicants maintain a 2% point growth differential with respect to the EU, it will take about half a century before most of them approach the EU average.

Figure 4. Standards of living in Eastern Europe.



Regional imbalances within CEE countries are characterised by the relative prosperity of urban centres and certain Western regions bordering the EU, which have benefited from the expansion of the service sector. Conversely, employment has plummeted in other regions as a result of large-scale job losses in traditional industries and reductions in agriculture. Nevertheless, employment in agriculture and industry remains high in some regions, reflecting delayed restructuring.

Even if the Eastern applicants keep a 2 percent point growth differential with respect to the EU, it will take about half a century before most of them approach the EU average.

Unemployment has risen significantly in most countries, but with considerable variation in rates, ranging from 5% in the Czech Republic to 14% in Bulgaria, Latvia and Lithuania. There are also significant regional disparities with, again, large urban centres and most Western regions having lower unemployment. The labour force has declined as the availability of jobs has diminished and people have withdrawn from the work force and, in many CEE countries, participation rates are now close to the EU average. Participation is regionally differentiated, often with high rates in areas where restructuring is still incomplete.

Most governments have begun to introduce development policies in recognition of the need to address regional disparities. This has been facilitated by decentralisation of government and encouraged by the prospect of EU membership. Accordingly the legal, institutional and budgetary structure for regional policy which will be necessary to participate in EU structural policy has begun to be established. However, completing these structures and procedures is likely to be a long process. CEE regional policies are still weak, lacking a comprehensive strategy and a programming approach. Measures tend to take the form of limited projects, implemented through sectoral policies that are only loosely co-ordinated. There remains a need to strengthen the Ministries responsible for regional policy and to develop their operational capacity, as well as to formulate national strategies for regional policy on the basis of which sectoral policies can be co-ordinated. Financial procedures also need to be improved so if the support from the EU Structural Funds is to be used efficiently.

5. Conclusion

Evidence now shows that GDP, or output, per head of the poorest regions has converged towards the European Union's (EU) average. To be more precise, these regions have typically closed the gap with the EU average by around 10 percentage points over the last decade, although even at this rate it will be several decades before the process is complete. The Structural Funds are making a significant contribution to closing the gap – the average result of a number of macro-economic models ascribes a third of the convergence to the Funds.

However, the situation with regard to unemployment is less positive. Disparities are high and tending to increase, driven in the worst affected regions by social exclusion, especially of young workers, and long term unemployment. The implication of output convergence and labour market divergence is a trend in Western Europe towards compact pockets of problems of economic and social restructuring (including urban problems) and away from the larger scale regional problems typified by Objective 1 (i.e. lagging development).

In addition there is the challenge of enlargement, where although regional unemployment rates fit well within Western norms, there are severe problems of lagging development, low GDP, poor quality infrastructure and environmental damage. In fact, GDP is so low that, even if the Eastern

applicants maintained a 2% point growth differential with the West, it would be of the order of half a century before most of them approached the EU average.

In sum, much progress has been made with regard to disparities in output and many of the factors, such as education and infrastructure, which underlie this. However, much work remains to be done; the economic – and particularly the technology – gap between core and periphery in Western Europe remains a challenge. This, along with labour market problems (including social exclusion) and accession, constitutes the three main challenges to cohesion over the next few years.

References

European Commission. (1981). The European regions: First periodic report on the economic and social situation in the regions of the Community, COM/80/816, Brussels.

European Commission. (1984). The European regions: Second periodic report on the economic and social situation in the regions of the Community, Official Publication Office, Luxembourg.

European Commission (1987). The European regions: Third periodic report on the socio-economic situation in the regions of the Community, Official Publication Office, Luxembourg.

European Commission. (1991). The European regions: fourth periodic report on the socio-economic situation in the regions of the Community, Official Publication Office, Luxembourg.

European Commission. (1994). The European regions: fifth periodic report on the socio-economic situation in the regions of the Community, Official Publication Office, Luxembourg.

European Commission. (1999). The European regions: sixth periodic report on the socio-economic situation in the regions of the European Union, Official Publication Office, Luxembourg.

European Commission. (1995a). Annual Cohesion Report 1993/94. COM/95/1/fin, Brussels.

European Commission. (1995b). Annual Cohesion Report 1994. Annex. COM/95/222/fin, Brussels.

European Commission. (1996). Annual Cohesion Report 1995. COM/96/388/fin, Brussels.

European Commission. (1997). Annual Cohesion Report 1996. COM/97/302/fin, Brussels.

European Commission. (1998). Annual Cohesion Report 1997. COM/98/543/fin, Brussels.

European Commission. (1999). Annual Cohesion Report 1998. COM/99/483/fin, Brussels.

Convergence across countries and regions: Theory and empirics

Angel de la Fuente

1. Introduction

In the last decade or so, growth has come to occupy an increasingly important place among the interests of macroeconomists, displacing to some extent their previous preoccupation with the business cycle. This change is largely due to two factors. The first one is the realisation that, in terms of medium and long-term welfare, the trend is more important than the cycle - provided the volatility of income remains as low as it has been during the last few decades (Lucas, 1987). The second factor is the increasing dissatisfaction with the traditional neo-classical models that summarised the pre-existing consensus on the determinants of growth - essentially because of their perceived inability to account for such key features of the data as the observed increase in international inequality or the absence of capital flows toward less developed countries.

Dissatisfaction with the received theory has motivated the search for alternatives to the traditional neo-classical model that has driven the recent literature on endogenous growth. At the theoretical level, numerous authors have developed models in which departures from traditional assumptions about the properties of the production technology or the determinants of technical progress generate predictions about the evolution of the international income distribution that stand in sharp contrast with those of neo-classical theory. Some of these models emphasise the role of growth factors that were ignored by previous theories and generate policy implications that are considerably more activist than those derived from the traditional models. At the empirical level, there is also a rich literature that attempts to test the validity of the different theoretical models that have been proposed, and to quantify the impact of various factors of interest on growth and on the evolution of international or interregional income disparities.

This essay provides an introduction to the theoretical and empirical literature on growth and convergence across countries and regions. It is organised as follows. Section 2 contains some general considerations on the convergence and divergence mechanisms identified in the growth literature. In Section 3 develops a simple descriptive model that attempts to capture the main immediate determinants of the growth of output and illustrates how some key properties of technology determine the evolution of the international or interregional income distribution. Section 4 focuses on the empirical implementation of growth models through convergence equations and illustrates their theoretical implications. Finally, Section 5 discusses some loose ends as well as recent empirical developments, and concludes with a brief summary.

2. Convergence and divergence in growth theory

As the reader will soon discover, the concept of convergence plays a crucial role in the literature we will survey. Although we will eventually provide a more precise definition of this term, we can provisionally interpret it as shorthand for the tendency towards the reduction over

Angel de la Fuente is with Instituto de Análisis Económico, Barcelona.

time of income disparities across countries or regions. Hence, we will say that there is convergence in a given sample when the poorer economies in it tend to grow faster than their richer neighbours, thereby reducing the income differential between them. When we observe the opposite pattern (i.e. when the rich grow faster and increase their lead) we will say that there is divergence in the sample.

Economic theory does not provide unambiguous predictions about the convergence of income levels across regions, but it does identify factors that may generate this.

Economic theory does not provide unambiguous predictions about the convergence or divergence of per capita income levels across countries or regions. It does, however, identify a series of factors or mechanisms that are capable in principle of generating either convergence or divergence. At some risk of oversimplifying, we can classify growth models into two families according to their convergence predictions.

According to those in the first group, being poor is, to some extent, an advantage. In these models, other things equal, poor countries grow faster than rich ones. This does not necessarily imply the eventual elimination of inequality (other things may not be equal), but it does mean that the distribution of relative income per capita across territories will tend to stabilise in the long run, provided some key "structural" characteristics of the different economies remain unchanged over time. In the second set of models, in contrast, rich countries grow faster and inequality increases without any bound.

The source of these contrasting predictions must be sought in very basic assumptions about the properties of the production technology at a given point in time and about the dynamics of technological progress. A first necessary condition for convergence is the existence of decreasing returns to scale in capital (or, more generally, in the various types of capital considered in the model). This assumption means that output grows less than proportionally with the stock of capital. This implies that marginal productivity will decrease with capital accumulation, reducing both the incentive to save and the contribution to growth of a given volume of investment, and creating a tendency for growth to slow down over time. The same mechanism generates a convergence prediction in the cross-section: poor countries (in which capital is scarcer) will grow faster than rich ones. Under the opposite assumption (of increasing returns in capital), the preceding neo-classical logic is inverted and we obtain a divergence prediction. In this case, the return on investment increases with the stock of capital per worker, favouring rich countries that tend to grow faster than poor ones, thereby increasing inequality further.

The second factor to consider has to do with the determinants of technological progress. If countries differ in the intensity of their efforts to generate or adopt new technologies, their long-term growth rates will be different. One possible objection is that the persistence of such differences is not plausible. For instance, it may be argued that the return on technological capital should decrease with its accumulation, just as we would expect to find for other types of capital. In this case, large differences across countries in rates of technological investment would not be sustainable, and there would be a tendency towards the gradual equalisation of technical efficiency levels. It is far from clear, however, that the accumulation of knowledge should be subject to the law of diminishing returns. If the cost of additional innovations falls with scientific or production experience, for instance, the return on technological investment may not be a decreasing function of the stock of

accumulated knowledge, and cross-country differences in levels of technological effort could persist indefinitely.

Technical progress could be an important divergence factor, but there are also forces that point in the opposite direction.

Hence, technical progress could be an important divergence factor. But there are also forces that point in the opposite direction. As Abramovitz (1979, 1986) and other authors have pointed out, the public good properties of technical knowledge have an international dimension that tends to favour less advanced countries, provided they have the capability to absorb foreign technologies and adapt them to their own needs. The idea is simple: not having to reinvent each wheel, followers will be in a better position to grow quickly than the technological leader, who will have to assume the costs and lags associated with the development of new leading-edge technologies (1). The resulting process of technological catch up could contribute significantly to convergence, particularly within the group of industrialised countries that are in a position to exploit the advantages derived from technological imitation.

In addition to decreasing returns and technological diffusion, the literature identifies a third convergence mechanism that, although featured less prominently in theoretical models, is likely to be of great practical importance. This mechanism works through structural change, or the reallocation of productive factors across sectors. Poorer countries and regions tend to have relatively large agricultural sectors. Given that output per worker is typically much lower in agriculture than in manufacturing or in the service sector, the flow of resources out of agriculture and into these other activities tends to increase average productivity. Since this process, moreover, has generally been more intense in poor economies than in rich ones in the last few decades, it may have contributed significantly to the observed reduction in productivity differentials across territories.

In conclusion, economic theory identifies forces with contrasting implications for income dynamics. Convergence mechanisms feature prominently in the neo-classical and catch-up models that dominated the literature until recently. The perceived failure of the optimistic convergence predictions of these models, however, has motivated the search for alternatives and contributed to the development of new theories that incorporate various divergence factors. Some of the pioneers of the "endogenous growth" literature (especially Romer 1986, 1987a, b) focused on the possibility of non-decreasing returns to scale in capital alone, while other authors, such as Lucas (1988), Romer (1990) and Grossman and Helpman (1991), developed models in which the rate of technical progress was determined endogenously and could differ permanently across countries, reflecting differences in structural characteristics. In both cases, the theory allows for the possibility of a sustained increase in the level of international or interregional inequality. The Box, which draws on de la Fuente (1995), provides a formal framework that illustrates these two classes of growth theories.

1) The idea seems to be due originally to Gerschenkron (1952) and has been developed among others by Abramovitz (1979, 1986), Baumol (1986), Dowrick and Nguyen (1989), Nelson and Wright (1992) and Wolff (1991).

Box 1: A formal model

Immediate determinants of the rate of growth

Let us consider a world in which there are only two factors of production and one final good. Capital (K) and labour (L) are combined to produce a homogeneous output (Y) that can be consumed directly or used as capital in the production process. We will assume that the production technology can be adequately described by an aggregate production function of the form

$$(1) \quad Y = \Phi K^{a}(AL)^{1-a} = \Phi A L Z^{a}$$

where A is an index of labour-augmenting technical efficiency and K denotes a broad capital aggregate that includes both human and physical capital. The variable $Z = K/AL$ denotes the capital/labour ratio in efficiency units and the coefficients a and $1\text{-}a$ measure the elasticity of output with respect to factor stocks.

To allow the possibility of increasing returns in the simplest possible way, we will assume that the term Φ, although perceived as an exogenous constant by individual agents, is in fact a function of the form $\Phi = Z^{b}$ that captures the external effects associated with investment. This specification is basically the one proposed by Romer (1986) building on Arrow (1962) to capture the possibility that capital accumulation may generate positive spillovers. Under these assumptions, output per worker, Q, is given by

$$(2) \quad Q = AZ^{\alpha}$$

where $\alpha = a + b$ measures the degree of returns to scale in capital taking into account this factor's indirect contribution to productivity through possible externalities.

Given equation (2), the growth of output per worker must be the result of the accumulation of productive factors or the outcome of technical progress. Taking logarithms of (2) and differentiating with respect to time, we see that the rate of growth of output per capita $\dot{Q}/Q = g_{Q}$, where $\dot{Q} = dQ/dt$, can be written as the sum of two terms that reflect, respectively, the rate of technical progress and the accumulation of productive factors:

$$(3) \quad g_{Q} = g_{a} + \alpha g_{Z}.$$

Let us explore the immediate determinants of g_{a} and g_{Z}, starting with the second factor. Denoting by s the share of investment in GDP and by δ the rate of depreciation, the increase in the aggregate capital stock is given by the difference between investment and depreciation, i.e.

$$(4) \quad \dot{K} = sY - \delta K = sLQ - \delta K$$

where $\dot{K} = dK/dt$ is the instantaneous increase in the capital stock. Since $Z = K/AL$, the growth rate of the stock of capital per efficiency unit of labour, g_{Z}, is the difference between $g_{k} = \dot{K}/K$ and the sum of the rates of technical progress (g_{a}) and labour force growth (n). Using (2) and (4), it is easy to see that

$$(5) \quad g_{Z} = g_{k} - g_{a} - n = sZ^{\alpha-1} - (n + g_{a} + \delta),$$

where the term $Z^{\alpha-1}$ $(=Q/(K/L))$ is the average product of capital. Substituting this expression into (3), we have:

(6) $g_Q = (1-\alpha)g_a + \alpha s Z^{\alpha-1} - \alpha(n+\delta).$

Finally, we have to specify the determinants of the rate of technical progress, g_a. We will assume that g_a is an increasing function of the fraction of GDP invested in R&D, θ, and of the opportunities for technological catch-up, measured by the log difference $(b = \ln X - \ln A)$ between a "technological frontier" denoted by X and the country's own technological index, A, or:

(7) $g_a = \gamma\theta + \varepsilon b.$

The parameters ε and γ measure, respectively, the speed of diffusion of new technologies across countries and the productivity of R&D. We will also assume that best-practice technology improves at a rate g_x which we will take as exogenous from the perspective of each given country and assume constant for simplicity.

Substituting (7) into (6) we finally arrive at an expression,

(8) $g_Q = (1-\alpha)(\gamma\theta + \varepsilon b) + \alpha s Z^{\alpha-1} - \alpha(n+\delta),$

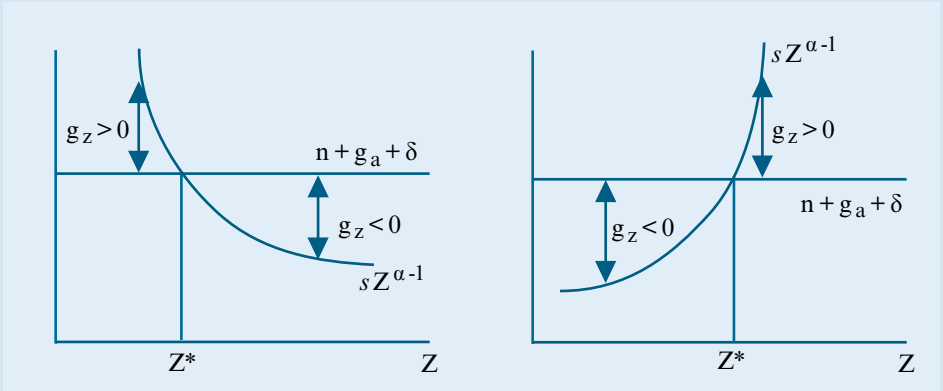
that gives the rate of growth of output per worker, g_Q, as a weighted sum of two terms that capture the immediate determinants of the rates of technical progress and capital accumulation.

Dynamics

To study the dynamics of the system, given by equation (5), it will be convenient to organise the analysis according to the impact on growth of two separate processes - capital accumulation and technical progress - for two countries (the "leader" and the "follower").

Assuming that the rate of technical progress, g_a, is an exogenous constant, we can draw both terms on the right-hand side of (5) as functions of Z. As shown in Figure 1, the rate of factor accumulation, g_z, is the difference between the product of the investment rate and the average product of capital, $sZ^{\alpha-1}$, and the constant $(n + g_a + \delta)$, and corresponds, therefore, to the vertical distance between the two lines, as shown in the figure.

Figure 1. Dynamics of capital accumulation



a.- Decreasing returns in capital ($\alpha < 1$) **b. - Increasing returns** ($\alpha > 1$)

The two panels of Figure 1 show that the behaviour of the dynamical system depends crucially on the value of α. When $\alpha < 1$, that is, when the neo-classical assumption of decreasing returns holds, the return on investment decreases with the stock of capital. Hence, the term $Z^{\alpha-1}$, is a decreasing function of Z and cuts the horizontal line given by the constant $(n + g_a + \delta)$ at the point Z^* characterised by $g_z = 0$. This implies:

$$(9) \quad Z^* = \left[\frac{s}{n + g_a + \delta}\right]^{\frac{1}{1-\alpha}}$$

From a dynamic point of view, the key finding is that under the assumption of decreasing returns the curve $sZ^{\alpha-1}$ cuts the horizontal line from above, making the growth rate of Z a decreasing function of its level. This implies that the steady state or long-term equilibrium described by Z^* is stable. Notice that g_z is positive (that is, Z increases over time) when the stock of capital per worker is small (and therefore the return on investment is high), and negative (Z decreases over time) when Z is "large" (larger than Z^*), for in this case the volume of investment is not enough to cover depreciation and equip newborn workers with the average stock of capital. Hence, we can interpret the steady-state value of the stock of capital per unit of labour, Z^*, as the one corresponding to a long-term equilibrium to which the economy gradually converges for any given initial value of Z .

When the external effects associated with the accumulation of capital are sufficiently strong that $\alpha > 1$, the situation is very different, as shown in panel b of Figure 1. Since the return on investment, measured by $Z^{\alpha-1}$, is now an increasing function of the stock of capital per efficiency unit of labour, the rate of accumulation increases with Z instead of falling. Hence, Z grows when it is larger than Z^* and falls when it is smaller, moving farther and farther away from the steady state, which must now be interpreted as a threshold for growth rather than as a long-run equilibrium.

To analyse the impact of technical progress on growth and convergence it will be convenient to work explicitly with two countries, f and l, (follower and leader). Let us define the technological distance between leader and follower by:

$$b_{lf} = a_l - a_f = (a_l - x) - (a_f - x) = b_f - b_l$$

where b_l and b_f denote the technological distance between each of these countries and the best-practice frontier. Observe that the evolution of the technological gap between leader and follower, b_{lf}, satisfies the following equation:

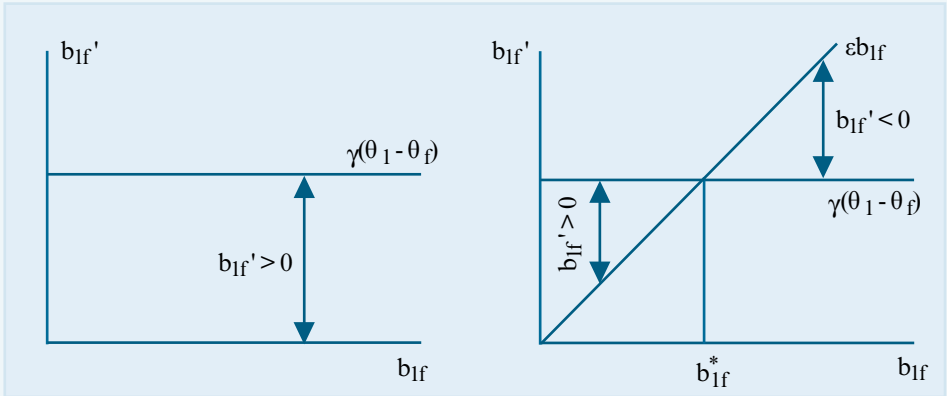
$$(10) \quad \dot{b}_{lf} = \dot{a}_l - \dot{a}_f = \gamma(\theta_l - \theta_f) - \varepsilon(b_l - b_f) = \gamma(\theta_l - \theta_f) - \varepsilon b_{lf}$$

Figure 2 displays the dynamics of this equation under two assumptions on the value of ε. When there is no technological diffusion ($\varepsilon = 0$), the leading country (which by assumption invests more in R&D) always has a higher rate of productivity growth. As a result, $\dot{b}_{lf}$ is always positive and the technological distance between leader and follower, b_{lf} , grows without bound as shown in Figure 2a.

When $\varepsilon > 0$, on the other hand, the line εb_{lf} is positively-sloped and cuts the horizontal line $\gamma(\theta_l - \theta_f)$ at a finite value of b_{lf} we will denote by b_{lf}^*. Under this assumption, the model is stable: $\dot{b}_{lf}$ is positive (that is, the technological gap increases over time) when b_{lf} is below its stationary value, b_{lf}^*, and negative (b_{lf} decreases) otherwise (see Figure 2b). Hence, the technological gap converges to a finite value, b_{lf}^*, defined by $\dot{b}_{lf} = 0$, which reads:

$$(11) \quad b_{lf}^* = \frac{\gamma(\theta_l - \theta_f)}{\varepsilon} .$$

Figure 2. Evolution of the technological distance between leader and follower



a.- No technological diffusion ($\varepsilon = 0$)

b. - Catch-up with technological diffusion ($\varepsilon > 0$)

In the long run, the (logarithm of the) ratio of the technical efficiency indices of the two countries converges to a constant value that is directly proportional to the difference between their rates of investment in R&D, and inversely proportional to the speed of technological diffusion.

Combining the results of the partial analyses undertaken so far, we can distinguish between two cases. When the technology exhibits increasing returns in capital ($\alpha > 1$) or there is no technological diffusion ($\varepsilon = 0$), the model is unstable and the growth paths of the two countries diverge. If there are decreasing returns and technological diffusion ($\alpha < 1$ and $\varepsilon > 0$), however, the model is stable. In the long run, the rates of growth of the two countries converge to the world rate of technical progress, g_x, and the ratio of their per capita incomes approaches a strictly positive constant value. Thus, long-run income disparities can be attributed to differences in levels of investment in physical and technological capital and in rates of population growth. Notice, however, that the extent to which such differences in "fundamentals" translate into long-term productivity differentials depends on the strength of the two convergence mechanisms present in the model. For given values of θ, s and n, the income differential across countries will be a decreasing function of the rate of technological diffusion (ε) and the degree of returns to scale in capital (α). Hence, both convergence mechanisms tend to mitigate the level of international inequality induced by cross-country differences in fundamentals, but do not eliminate it.

3. From theory to empirics: A framework for empirical analysis

When it comes to trying to distinguish empirically between these two families of models, set out above, a starting point is the observation that the main testable difference between them is the sign of the partial correlation between the growth rate and the initial level of income per capita. While this correlation should be negative according to standard neo-classical models (that is, other things equal poorer countries should grow faster), in some models of endogenous growth the expected sign would be the opposite one. This suggests that a natural way to try to determine which group of models provides a better explanation of the growth experience involves estimating a *convergence equation*, that is, a regression model in which the dependent variable is the growth rate of income per capita or output per worker and the explanatory variable is the initial value of the same income indicator.

The correct formulation of the empirical model, however, requires that we control for other variables that may affect the growth rate of the economies in the sample. As we have seen in a previous section, neo-classical and catch-up models predict that poor countries will grow faster than rich ones only under certain conditions. In Solow's (1956) neo-classical model, for instance, the long-term level of income is a function of the rates of investment and population growth and can, therefore, differ across countries. In a similar vein, Abramovitz (1979, 1986) emphasises that the process of technological catch-up is far from automatic. Although relative backwardness carries with it the potential for rapid growth, the degree to which this potential is realised in a given country depends on its "social capability" to adopt advanced foreign technologies (i.e. on factors such as the level of schooling of its population and the availability of qualified scientific and technical personnel) and on the existence of a political and macroeconomic environment conducive to investment and structural change.

Even in models where convergence forces prevail, long-term income levels can vary across territories, reflecting underlying differences in "fundamentals".

In short, even in models where convergence forces prevail, long-term income levels can vary across territories, reflecting underlying differences in "fundamentals". If we do not control for such differences, the estimated relationship between growth and initial income could be very misleading. Imagine, for instance, that the Solow model (with decreasing returns and access by all economies to a common technology) is the correct one, and that richer countries display on average higher rates of investment and lower rates of population growth than poorer countries (which is why they are richer in the first place). According to the model, these two factors would have a positive effect on the growth rate (during the transition to the long-run equilibrium) that could conceivably dominate the convergence effect that makes growth a decreasing function of income with other things constant. It is clear that if we do not include the rates of investment and population growth in the equation, we could find that the estimated coefficient of initial income is positive and conclude, erroneously, from this fact that the predictions of the Solow model fail to hold. To put it in a slightly different way, the problem would be that when we do not control for the determinants of the steady state, we are actually testing the hypothesis that all economies converge to the same long-run equilibrium. The rejection of this hypothesis, however, has no implications for the validity of the Solow model, since this model makes no such prediction except when the economies in the sample are exactly alike.

On the basis of the preceding discussion, we can conclude that a "minimal" model for the empirical analysis of convergence would be an equation of the form

$$(1) \quad \Delta y_{i,t} = \gamma x_{it} - \beta y_{i,t} + \varepsilon_{it} ,$$

where $y_{i,t}$ is income per capita or per worker in territory i at the beginning of period t, $\Delta y_{i,t}$ the growth rate of the same variable over the period, ε_{it} a random disturbance, and x_{it} a variable or set of variables that captures the "fundamentals" of economy i, that is, all those characteristics of this territory that have a permanent effect on its growth rate.

3.1 Structural convergence equations

Many empirical studies of growth and convergence have proceeded by estimating some variant of equation (1). In early studies the empirical specification was frequently *ad hoc* and only loosely tied

with the theory (2). In recent years, however, researchers have increasingly focused on the estimation of "structural" convergence equations derived explicitly from formal models. One of the most popular specifications in the literature is the one derived by Mankiw, Romer and Weil (1992; henceforth MRW) from an extended neo-classical model à la Solow that would be equivalent to the one developed in the Box under the assumption that the rate of technical progress is an exogenous constant common to all countries (3). Working with a log-linear approximation to the model around its steady state, MRW show that the growth rate of output per worker in territory i during the period that starts at time t is given approximately by the following equation (4):

$$(2)\quad \Delta y_{i,t} = g + \beta(a_{io} + gt) + \beta\frac{\alpha}{1-\alpha}\ln\frac{s_{it}}{\delta+g+n_{it}} - \beta y_{i,t}$$

where

$$(3)\quad \beta = (1-\alpha)(\delta+g+n),$$

and where g is the rate of technical progress, δ the depreciation rate, α the coefficient of capital in the aggregate production function, t the time elapsed since the beginning of the sample period, a_{io} the logarithm of the index of technical efficiency at time zero, s the share of investment in GDP and n is the rate of growth of the labour force.

It is important to understand that the estimation of equation (2) does not imply that we are literally accepting the assumptions of the underlying Solow-type model (i.e. we do not need to assume that the investment rate is exogenous or constant over time). What we are doing is simply assigning to some of the parameters of the Solow model (in particular, to s and n) the observed average values of their empirical counterparts during a given period. During this period, the economy will behave approximately as if it were approaching the steady state of the Solow model that corresponds to the contemporaneous parameter values. In the next period, of course, we are likely to observe different values of the investment and population growth rates and therefore, a different steady state, but this poses no real difficulty. In essence, all we are doing is constructing a convenient approximation to the production function that allows us to recover its parameters using data on investment flows rather than factor stocks. This is very convenient because such data are easier to come by and can be expected to be both more reliable and more comparable over time and across countries than most existing estimates of factor stocks. It must be kept in mind, however, that the only information we can extract from the estimation of a convergence equation of the form (2)

2) See for instance Kormendi and McGuire (1985), Grier and Tullock (1989) and Barro (1991).

3) That is, the specification of the rate of technical progress as a function of $R\&D$ expenditure and the technological gap is abandoned, being replaced by the simple assumption that the rate of technological progress is an exogenous constant, g, equal for all countries. The part of the model that describes capital accumulation, on the other hand, would be exactly as developed in Section 2.

4) Barro and Sala-i-Martin (1990, 1992a) derive a similar expression from a variant of the optimal growth model of Cass (1965) and Koopmans (1965) with exogenous technical progress. The resulting equation is similar to (2) except that the investment rate (which is now endogenous) is replaced by the rate of time discount among the determinants of the steady state. The convergence coefficient, β, is now a more complicated function of the parameters of the model, but it still depends on the degree of decreasing returns to capital and on the rates of population growth, depreciation and technical progress. A second difference between the two models is that, whereas the MRW model can be easily extended to incorporate investment in human capital, Barro and Sala do not include this factor as an argument of the production function, although they do bring it into their empirical specification, in an ad hoc way, as a determinant of the steady state.

concerns the properties of the production technology. As Cohen (1992) emphasises, the estimated equation does not, in particular, tell us anything about the actual dynamics of the economy or the position of a hypothetical long-run equilibrium - although it does allow us to make predictions about long-term income levels conditional on assumptions about the future behaviour of investment and population growth rates.

The empirical implementation of equation (1) or (2) does not, in principle, raise special problems. Given time series data on income, population and investment for a sample of countries or regions, we can use (2) to recover estimates of the rate of convergence and the parameters of the production function. The convergence equation can be estimated using either cross-section or pooled data. Most of the earlier convergence studies took the first route, averaging the variables over the entire sample period and working with a single observation for each country or region. The second possibility, which has become increasingly popular, involves averaging over shorter sub-periods in order to obtain several observations per country.

In either case, one difficulty which immediately becomes apparent is that three of the variables on the right-hand side of the equation (g, δ and a_{io}) are not directly observable. In the first two cases, the problem is probably not very important. Although these coefficients can be estimated inside the equation (and this has been done occasionally), the usual procedure in the literature is to impose "reasonable" values of these parameters prior to estimation. The standard assumption is that $g = 0.02$ and $\delta = 0.03$, but researchers report that estimation results are not very sensitive to changes in these values.

The possibility that initial levels of technical efficiency may differ across countries raises a difficult problem that requires panel data techniques.

The possibility that initial levels of technical efficiency (a_{io}) may differ across countries does raise a more difficult problem. Although some authors have argued that it may be reasonable to assume a common value of a_{io} because most technical knowledge is in principle accessible from everywhere, casual observation suggests that levels of technological development differ widely across countries. If this is so, failure to control for such differences (or for any other omitted variables) will bias the estimates of the remaining parameters whenever the other regressors in the equation are correlated with the missing ones. In other words, we can only legitimately subsume technological differences across countries in the error term if they are uncorrelated with investment rates and population growth. This seems unlikely, however, as the level of total factor productivity is one of the key determinants of the rate of return on investment.

The standard solution for this problem is to turn to panel data techniques in order to control for unobserved national or regional fixed effects. The simplest procedure involves introducing country or regional dummies in order to estimate a different regression constant for each territory. It should be noted, however, that this is equivalent to estimating the equation with the dependent and independent variables measured in deviations from their average values (computed over time for each country or region in the sample). Hence, this procedure (as practically all panel techniques designed for removing fixed effects), ignores the information contained in observed cross-country differences and produces parameter estimates which are based only on the time variation of the data within each territory over relatively short periods. Since what we are trying to do is characterising the long-term dynamics of a sample of economies, this may be rather dangerous, particularly when the data contain an important cyclical component or other short-term noise.

The structural convergence equation methodology has some important advantages and limitations, both of which are derived from the close linkage between theory and empirics that characterise this approach. Its most attractive feature is that it allows us to use the relevant theory to explicitly guide the formulation of the empirical model - that is, the formal model is used to determine what variables must be included in the regression and how they must enter in order to obtain direct estimates of the structural parameters of the model. It is clear, however, that such guidance comes at a price, as our estimates will be, at best, only as good as the underlying theoretical model. Hence, an inadequate specification of this model can yield very misleading conclusions.

An inadequate specification of the model can yield very misleading conclusions.

Although this problem arises to some extent whenever we run a regression, there are reasons to think that it may be particularly important in the present context. In most of the recent empirical work on growth and convergence, the theoretical model of reference is some version of the one-sector neo-classical model with exogenous technical progress that underlies equation (2). Since the only convergence force present in this model is what we may call the neo-classical mechanism, the usual finding of a negative partial correlation between growth and initial income must be interpreted in this framework as evidence that the aggregate production function displays decreasing returns to scale in reproducible factors. In fact, this assumption is precisely what allows us to draw inferences about the degree of returns to scale from the estimated value of the convergence coefficient. The problem, of course, is that if there are any other operative convergence mechanisms, the inference will not be valid, as the estimated value of the convergence parameter will also capture their effects.

As we have seen, the literature identifies at least two factors other than decreasing returns that can generate a negative partial correlation between income levels and growth rates holding investment and population growth constant: technological diffusion and structural change. Although none of these convergence mechanisms is incompatible with the neo-classical story, the observation that this is not the only possible source of convergence suggests that it may be dangerous to accept without question an interpretation of the convergence coefficient based too literally on the preceding model. For instance, if income per capita is highly correlated with the level of technological development, the coefficient of initial income in a convergence regression could capture, at least in part, a technological catch-up effect. To avoid the danger of drawing the wrong conclusions about the properties of the technology, it may be preferable to interpret existing estimates of the convergence parameter, β, (particularly in the case of unconditional convergence equations) as summary measures of the joint effect of several possible convergence mechanisms. The value of this parameter (i.e. the partial correlation between the growth rate and initial income) will depend on the coefficient of capital in the production function, the speed of technological diffusion, the impact of sectoral change and on the response of investment rates to rising income), and will be positive (i.e. growth will be negatively correlated with initial income) whenever the forces making for convergence dominate those working in the opposite direction.

3.2 Some convergence concepts

Before we proceed to review the empirical evidence, it is convenient to introduce some concepts of convergence that will feature prominently in the discussion below. Perhaps the first question that arises concerning the evolution of the distribution of income per capita is whether the dispersion of this variable (measured for instance by the standard deviation of its logarithm) tends to decrease

over time. The concept of convergence implicit in this question, called σ-*convergence* by Barro and Sala-i-Martin (1990, 1992a,b), is probably the one closest to the intuitive notion of convergence. It is not, however, the only possible one. We may also ask, for instance, whether poorer countries tend to catch up with richer ones, or whether the relative position of each country within the income distribution tends to stabilise over time. The concepts of *absolute* and conditional β-*convergence* proposed by Barro and Sala-i-Martin correspond roughly to these two questions.

To make more precise these two notions of convergence, we can use a variant of equation (1) in which we assume that each economy's fundamentals remain constant over time (that is, that $x_{it} = x_i$ for all t) and we interpret the variable y_{it} as relative income per capita, that is, income per capita normalised by the contemporaneous sample average. Omitting the disturbance term, the evolution of relative income in territory i is described by

$$\Delta y_{i,t} = \gamma x_i - \beta y_{i,t}.$$

Setting $\Delta y_{i,t}$ equal to zero in this expression, we can solve for the steady-state value of relative income,

$$y_i^* = \frac{\gamma x_i}{\beta}$$

It is easy to check that if β lies between zero and one, the system described by the above equation is stable. This implies that the relative income of territory i converges in the long run to the equilibrium value given by y_i^*. Notice that the equilibrium can differ across countries as a function of the "fundamentals" described by x_i.

In terms of this simple model, we will say that there is conditional β-convergence when β lies between zero and one, and absolute β-convergence when this is true and, in addition, y_i is the same for all economies - i.e. when all countries or regions in the sample converge to the same income per capita.

Even though they are closely related, the three concepts of convergence are far from being equivalent. Some type of β-convergence is a necessary condition for sustained σ-convergence, for the level of inequality will grow without bound when β is negative (i.e. when the rich grow faster than the poor). It is not sufficient, however, because a positive value of β is compatible with a transitory increase of income dispersion due either to random shocks or to the fact that the initial level of inequality is below its steady-state value (as determined by the dispersion of fundamentals and the variance of the disturbance). The two types of β-convergence, moreover, have very different implications. Absolute β convergence implies a tendency towards the equalisation of per capita incomes within the sample. Initially poor economies tend to grow faster until they catch up with the richer ones. In the long run, expected per capita income is the same for all members of the group, independently of its initial value. As we know, this does not mean that inequality will disappear completely, for there will be random shocks with uneven effects on the different territories. Such disturbances, however, will have only transitory effects, implying that, in the long run, we should observe a fluid distribution in which the relative positions of the different regions change rapidly.

With conditional β-convergence, on the other hand, each territory converges only to its own steady state but these can be very different from each other. Hence, a high degree of inequality could persist, even in the long run, and we would also observe high persistence in the relative positions of the different economies. In other words, rich economies will generally remain rich while the poor continue to lag behind.

It is important to observe that, although the difference between absolute and conditional convergence is very sharp in principle, things are often much less clear in practice. In empirical studies we generally find that a number of variables other than initial income enter significantly in convergence equations. This finding suggests that steady states differ across countries or regions and, therefore, that convergence is only conditional. It is typically the case, however, that these conditioning variables change over time and often tend to converge themselves across countries or regions. Hence, income may still converge unconditionally in the long run, and this convergence may reflect in part the gradual equalisation of the underlying fundamentals. In this situation, a conditional and an unconditional convergence equation will yield different estimates of the convergence rate. There is, however, no contradiction between these estimates once we recognise that they are measuring different things: while the unconditional parameter measures the overall intensity of a process of income convergence which may work in part through changes over time in various structural characteristics, the conditional parameter captures the speed at which the economy would be approaching a "pseudo steady state" whose location is determined by the current values of the conditioning variables.

4. Convergence across countries and regions: Empirical evidence and theoretical implications

Having reviewed the theoretical and empirical framework used in the convergence literature, we are now in a position to examine the empirical evidence and discuss its implications. Contributions to the empirics of economic growth highlight three interesting empirical regularities. First, evidence of some sort of β-convergence is found in practically all available samples. While convergence is only conditional at the national level, in most regional samples a negative correlation between initial income and subsequent growth emerges without controlling for other variables. This second result is consistent with the existence of absolute convergence at the regional level - but most of the studies we have reviewed do not explicitly test this hypothesis (5). Secondly, the process of convergence seems to be extremely slow. Many of the existing estimates of the convergence parameter cluster around a value of 2 percent per year which implies that it takes around 35 years for a typical region to reduce its income gap with the national average by one half. Hence, the expected duration of the convergence process must be measured in decades. Finally, it is interesting to observe that the estimated convergence coefficient is remarkably stable across samples. This

The process of convergence seems to be extremely slow.

5) Those that do test it by including different sets of conditioning variables generally reject it, as the significance of many of these variables implies important cross-regional differences in steady states. See for instance Dolado et al., *(1994) and Mas* et al., *(1995) for the Spanish provinces, Herz and Röger (1996) for the German Raumordnungsregionen, Neven and Gouyette (1995) and Faberberg and Verspagen (1996) for various samples of European regions, Holtz-Eakin (1993) for the states of the US and Paci and Pigliaru (1995), and Fabiani and Pellegrini (1996). As we have noted in Section 3, however, this evidence does not conclusively reject the hypothesis of absolute convergence, as conditioning variables (and hence steady states) may themselves be converging over time. Note that the study by Vanhoudt* et al., *this volume, tackles this issue for Europe in more detail. The authors report strong support of conditional convergence within Europe - even at the regional level - with a speed of convergence of about half the size found for the United States.*

stability suggests that the mechanisms that drive convergence in income per capita across different economies seem to operate in a regular fashion. Hence, we can at least hope to provide a unified structural explanation of the convergence process in terms of a "general" theoretical model.

A good candidate for a "general" model is a simple extension of the neo-classical one. All that is needed is to extend the concept of capital to intangibles.

Perhaps the dominant view in the literature is that a good candidate for this "general" model is a simple extension of the one-sector neo-classical model with exogenous technical progress. Just about the only departure from the traditional assumptions required in order to explain the empirical evidence is a broadening of the relevant concept of capital in order to include investment in intangibles such as human and technological capital. This conclusion is reached essentially by interpreting the results we have just reviewed within the framework of the growth model given in (2). According to our previous discussion, the finding of (at least conditional) β-convergence in most national or regional samples can be interpreted as evidence in favour of the neo-classical assumption of decreasing returns to capital, as this result would not be consistent with increasing returns models that predict an explosive behaviour of income and its distribution. On the other hand, the apparent slowness of the convergence process does suggest that we are not that far from having constant returns in reproducible factors - a result that seems considerably more plausible if we think in terms of a broad capital aggregate, rather than the rather restrictive concept of capital we find in old-fashioned neo-classical models.

Since this broader concept of capital is probably one of the most significant contributions of the recent literature to our understanding of the mechanics of growth, the issue deserves a fairly detailed discussion. The reader will recall that within the framework of the Solow model the convergence coefficient (β) depends on the degree of returns to scale, measured by α (the coefficient of capital in the aggregate production function with, as discussed in the Box, $\alpha = a + b$, where a is the coefficient of capital in the "private" production function and b captures the possible externalities), and on the rates of technical progress (g), population growth (n) and depreciation (δ). More specifically, we have seen that the relationship among these variables is given by equation (3): $\beta = (1 - a - b)(\delta + g + n)$.

Using this expression and making reasonable guesses about the values of some of the parameters, we can extract information about key properties of the production technology from empirical estimates of the convergence rate. To start, let us consider the expected value of β under conventional assumptions about the values of the remaining parameters. Within the framework of a traditional neo-classical model (with constant returns to scale in capital and labour, perfect competition and no externalities) we would have $b = 0$ and the coefficient a would be equal to capital's share of national income, which is around one third. The average rate of population growth in the industrial countries during the post-WWII period is approximately 1%. Available estimates of the rate of technical progress are around 2% per year. Finally, estimates of the rate of depreciation vary considerably. In the convergence literature it is commonly assumed that $\delta = 0.03$, but a higher value (around 5 or 6% per year) may be more reasonable. Given these assumptions, the expected value of β lies between 0.04 and 0.06.

The empirical results of Barro and Sala-i-Martin (1990, 1992a,b), Mankiw, Romer and Weil (1992) and other authors point towards a much lower convergence rate. Since the estimated value

of the parameter is still positive, the evidence is consistent with decreasing returns to capital (i.e. $a + b < 1$). The low value of β, however, suggests that we are relatively close to having constant returns to capital. Maintaining our previous assumptions about the values of the remaining parameters, a convergence coefficient of 0.02 would imply a value of α between 0.67 and 0.78 - more than twice the share of capital in national income.

One possible explanation (Romer, 1987b) is that this result may reflect the existence of important externalities associated with the accumulation of physical capital. While these external effects would not be sufficiently strong to generate increasing returns in capital alone, they might still account for the apparent slowness of convergence. Other authors, however, argue that a more plausible explanation is that the omission of variables which are positively correlated with investment in physical capital may bias upward the coefficient of this variable. Barro and Sala-i-Martin (1990, 1992a) argue that a value of capital's coefficient around 0.7 only makes sense if we count accumulated educational investment as part of the stock of capital.

Mankiw, Romer and Weil (1992) advance the same hypothesis and test it explicitly by estimating a structural convergence equation similar to equation (2) that explicitly incorporates a proxy for the rate of investment in human capital as a regressor. Their results, and those obtained by other authors who estimate similar specifications, tend to confirm the hypothesis that investment in human (and technological) capital plays an important role in the growth process (6). As Mankiw (1995) points out, once human capital is included as an input in the production function, the resulting model is consistent with some of the key features of the data. Countries that invest more in physical capital and education tend to grow faster and therefore eventually attain high levels of relative income. Cross-country differences in rates of accumulation, moreover, are sufficiently high to explain the bulk of the observed dispersion of income levels and growth rates.

5. Loose ends and recent developments

It is probably fair to say that just a few years ago the extended neo-classical model we have just described summarised a consensus view on the mechanics of growth. In recent years, however, this consensus has been challenged by a series of papers that, relying on panel data techniques, obtain results that are difficult to reconcile with the prevailing theoretical framework. In this section we will summarise some of the key findings of these studies and discuss the theoretical difficulties they raise.

One of the key findings of the "classical" convergence studies is that convergence to the steady state is an extremely slow process. It has recently been argued, however, that this result may be due to a bias arising from the use of econometric specifications that do not adequately allow for unobserved differences across countries or regions. To get around this problem, a number of authors have proposed the use of panel techniques that allow for unobserved fixed effects. As we will see in this section, their results raise some puzzling questions.

For example, Marcet (1994), Raymond and García (1994), Canova and Marcet (1995), de la Fuente (1996a,b), Tondl (1997) and Gorostiaga (1998), estimate fixed-effects convergence models

6) See for instance Lichtenberg (1992), Holtz-Eakin (1993), Nonneman and Vanhoudt (1996) and de la Fuente, (1998b). de la Fuente (1997) provides a detailed review of this literature.

using panel data for a variety of regional samples. Their results suggest a view of the regional convergence process that stands in sharp contrast with the one advanced before: instead of slow convergence to a common income level, regional economies within a given country seem to be converging extremely fast (at rates of up to 20 percent per year) but to very different steady states. Cross-national studies provide a roughly similar picture: Knight *et al.*, (1993), Canova and Marcet (1995), Islam (1995) and Caselli *et al.*, (1996), among others, find evidence of rapid convergence across countries (at rates of up to 12% per annum) toward very different steady states whose dispersion can be explained only in part by observed cross-national differences in population growth and investment rates. In both cases, many of the standard conditioning variables (such as human capital indicators) lose their statistical significance, the estimated coefficient of physical capital adopts rather low values, and the size and significance of the regional or national fixed effects suggests that persistent differences in levels of technical efficiency play a crucial role in explaining the dispersion of income levels.

Should we take these results at face value? Before we do so and abandon the only workable models we have so far, it seems sensible to search for some way to reconcile these empirical findings with some kind of plausible theory. I believe that this can be done - at least to some extent. My argument is essentially that a more reasonable interpretation of the extremely high convergence rates obtained in recent studies is that, if we have correctly estimated the relevant parameter (and we may not), then convergence is much too fast to be simply the result of diminishing returns to scale. This observation points to two complementary lines of research. The first one asks whether panel specifications of growth equations do in fact yield estimates of the relevant parameter. The second proceeds by identifying plausible mechanisms that may help account for rapid convergence and incorporating them into theoretical and empirical models.

We can reject the assumption that increasing returns generate an explosive behaviour in the distribution of income across economies.

On the first issue, Shioji (1997a, b) and de la Fuente (1998a) provide some evidence that panel estimates of the convergence rate may tell us very little about the speed at which economies approach their steady states (and therefore about the degree of returns to scale in reproducible factors) - essentially because these estimates are likely to capture short-term adjustments around trend rather than the long-term growth dynamics we are really interested in. Both authors show that correcting for the resulting bias in various ways brings us back to convergence rates that are broadly compatible with sensible theoretical models.

On the second issue, de la Fuente (1995, 1996b) and de la Fuente and Doménech (2000) estimate a further extension of the neo-classical model that allows for cross-country differences in total factor productivity and for a process of technological catch-up and show that technological diffusion can go a long way towards explaining rapid convergence across countries and regions (7). It follows that fast convergence does not require us to abandon the broad concept of capital we have so laboriously developed over the last decades. In fact, the parameters of the aggregate production functions estimated by these authors at the regional and national level are not far from those

7) Dowrick and Nguyen (1989) also investigate the quantitative importance of technological catch-up as a convergence factor, but their empirical specification makes it difficult to disentangle this effect from the neo-classical convergence mechanism. Helliwell (1992), Coe and Helpman (1995) and Engelbrecht (1997) provide additional evidence on technological diffusion. There is also some evidence that a significant part of what appears to be TFP convergence at the aggregate level is in fact due to factor reallocation across sectors. See for instance Paci and Pigliaru (1995), de la Fuente (1996b), Caselli and Coleman (1999) and de la Fuente and Freire (1999).

obtained by Mankiw, Romer and Weil (1992) and suggest, in particular that educational investment plays a crucial role in growth that may not be apparent in previous studies in part because of data deficiencies. On the other hand, these studies do show that total factor productivity (TFP) differences across countries and regions are substantial and account for around half of observed productivity differences in the OECD in recent years, a result that is broadly consistent with the findings of Klenow and Rodríguez (1997) on the explanatory power of the extended neo-classical model. These results highlight the importance TFP dynamics as a crucial determinant of the evolution of productivity, a subject brought up recently by Prescott (1998), while retaining a significant role for differences in factor stocks as sources of income differentials across economies.

Recent results have still been useful in shaking-up an exaggerated confidence in our ability to explain why some countries are richer than others.

If we try to summarise, the key points that stand out in the theory and empirics on economic growth are as follows. In the current state of the literature, the conclusions we can draw must necessarily remain rather tentative. Practically all existing studies on the subject find clear evidence of some sort of long-run convergence both across countries and across regions over the post-war period. These findings allow us to reject with a fair degree of confidence a series of recent models in which the assumption of increasing returns generates an explosive behaviour of the distribution of income across economies that cannot be found in the data. Many of the results we have reviewed are consistent with an extended neo-classical model built around an aggregate production function that includes human capital as a productive input.

Recently, results using panel data techniques have suggested that educational investment was not productive and that the bulk of productivity differences across countries or regions has little to do with differences in stocks of productive factors. In my opinion, this has been largely a false alarm, but it has been useful in shaking up what was probably an exaggerated confidence in our ability to explain why some countries or regions are richer than others with an extremely simple model, and in directing researchers' attention to the determinants of technological progress and to some of the difficult econometric issues involved in the estimation of growth models.

References

Abramovitz, M. (1979). "Rapid Growth Potential and its Realization," in *Thinking About Growth and Other Essays on Economic Growth and Welfare*. Cambridge U. Press, 1989, pp. 187-219.

Abramovitz, M. (1986). "Catching Up, Forging Ahead and Falling Behind," in *Thinking About Growth and Other Essays on Economic Growth and Welfare*. Cambridge U. Press, 1989, pp. 220-44.

Arrow, K. J. (1962) "The Economic Implications of Learning by Doing," in *Collected Papers of Kenneth J. Arrow, vol. 5: Production and Capital*. Harvard U. Press, 1985, pp. 157-80.

Barro, R. (1991). "Economic Growth in a Cross Section of Countries. "*Quarterly Journal of Economics*, CVI (2), pp. 407-43.

Barro, R. and Sala-i-Martin, X. (1990). "Economic Growth and Convergence Across the United States." NBER Working Paper no. 3419, Aug. 1990.

Barro, R. and Sala-i-Martin, X. (1992a). "Convergence." *Journal of Political Economy*, 100(2), pp. 223-51.

Barro, R. and Sala-i-Martin, X. (1992b). "Regional Growth and Migration: A Japan-United States Comparison." *Journal of the Japanese and International Economies*, 6, pp. 312-46.

Baumol, W. (1986). "Productivity Growth, Convergence and Welfare: What the Long-Run Data Show." *American Economic Review*, 76(5), pp. 1072-85.

Canova, F. and Marcet, A. (1995). "The Poor Stay Poor: Non-convergence across Countries and Regions." CEPR Discussion Paper no. 1265.

Caselli, F. and Coleman, W. (1999). "How regions converge." CEPR Discussion Paper no. 2191.

Caselli, F., Esquivel, G. and Lefort, F. (1996). "Reopening the convergence debate: a new look at cross-country growth empirics." *Journal of Economic Growth*, 1(3), pp. 363-89.

Cass, D. (1965). "Optimum Growth in an Aggregative Model of Capital Accumulation." *Review of Economic Studies*, XXXII, July, pp. 223-40.

Coe, D. and Helpman, E. (1995). "International R&D spillovers." *European Economic Review*, 39, pp. 859-87.

Cohen, D. (1992). "Tests of the 'convergence hypothesis':a critical note." CEPR Discussion Paper no. 691.

de la Fuente, A. (1995). "Catch-up, Growth and Convergence in the OECD." CEPR Discussion Paper no. 1274.

de la Fuente, A. (1996a). "Convergencia y otras historias: economía regional desde una perspectiva neoclásica. "*Revista de Economía Aplicada*, IV(10), Primavera, pp. 5-64.

de la Fuente, A. (1996b). "On the sources of convergence: A close look at the Spanish regions." CEPR Discussion Paper No. 1543.

de la Fuente, A. (1997) "The empirics of growth and convergence: a selective review." *Journal of Economic Dynamics and Control*, 21(1), pp. 23-74.

de la Fuente, A. (1998a). "What kind of regional convergence?" CEPR Discussion Paper no. 1924.

de la Fuente, A. (1998b). "Convergence equations and income dynamics: The sources of OECD Convergence, 1970-95." CEPR Discussion Paper no. 1794.

de la Fuente, A. and Freire, M. J. (1999). "Estructura sectorial y convergencia regional." Mimeo, Instituto de Análisis Económico, Barcelona.

de la Fuente, A. and Doménech, R. (2000). "Human capital in growth regressions: how much difference does data quality make?" Mimeo, Instituto de Análisis Económico, Barcelona.

Dolado, J., González-Páramo, J. M. and Roldán, J. M. (1994). "Convergencia Económica entre las Provincias Españolas: Evidencia Empírica (1955-1989)." *Moneda y Crédito* 198, pp. 81-131.

Dowrick, S. and Nguyen, D. T. (1989). "OECD Comparative Economic Growth 1950-85: Catch-up and Convergence." *American Economic Review* 79(5), pp. 1010-1030.

Engelbrecht, H.-J. (1997). "International R&D spillovers, human capital and productivity in the OECD economies: an empirical investigation." *European Economic Review*, 41, pp. 1479-88.

Faberberg, J. and Verspagen, B. (1996). "Heading for divergence? Regional growth in Europe reconsidered. *Journal of Common Market Studies*, 34(3), pp. 431-48.

Gerschenkron, A. "Economic Backwardness in Historical Perspective," in B. Hosilitz, editor, *The Progress of Underdeveloped Area*s. Chicago University Press, 1952.

Grier, K. and Tullock, G. (1989). "An Empirical Analysis of Cross-National Economic Growth, 1951-80." *Journal of Monetary Economics*, 24, pp. 259-276.

Grossman, G. and Helpman, E. (1991). *Innovation and Growth in the Global Economy*. MIT Press, Cambridge, Mass.

Helliwell, J. (1992). "Trade and Technical Progress." NBER Working Paper no. 4226.

Herz, B. and Röger, W. (1996). "Economic growth and convergence in Germany," *Weltwirtschaftliches Archiv*,1/1996, pp. 133-43.

Holtz-Eakin, D. (1993). "Solow and the states: Capital accumulation, productivity, and economic growth." *National Tax Journal*, XLVI, pp. 425-39.

Islam, N. (1995). "Growth Empirics: A Panel Data Approach". *Quarterly Journal of Economics*, CX(4), pp. 1127-70.

Klenow, P. and Rodríguez-Clare, A. (1997). "The neo-classical revival in growth economics: has it gone too far?" Mimeo, U. of Chicago.

Knight, M., Loayza, N. and Villanueva, D. (1993). "Testing the neo-classical theory of economic growth: a panel data approach." *IMF Staff Papers*, 40, pp. 512-41.

Koopmans, T. (1965). "On the Concept of Optimal Economic Growth," in *The Econometric Approach to Development Planning*. North-Holland, Amsterdam.

Kormendi, R. and Meguire, P. G. (1985). "Macroeconomic Determinants of Growth: Cross-Country Evidence." *Journal of Monetary Economics*, 16, pp. 141-63.

Lichtenberg, F. (1992). "R&D Investment and International Productivity Differences." NBER Working Paper no. 4161.

Lucas, R. (1987). *Models of Business Cycles*. Basil Blackwell, Oxford.

Lucas, R. (1988). "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22(1), pp. 3-42.

Lucas, R. (1990). "Why Doesn't Capital Flow from Rich to Poor Nations?" *American Economic Review Papers and Proceedings*, 80, pp. 92-6.

Mankiw, G. (1995). "The Growth of Nations." *Brookings Papers on Economic Activity* I, pp. 275-326.

Mankiw, G., Romer, D. and Weil, D. (1992). "A Contribution to the Empirics of Economic Growth." *Quarterly Journal of Economics*, CVII(2), pp. 407-37.

Marcet, A. (1994). "Los pobres siguen siendo pobres: Convergencia entre regiones y países, un análisis bayesiano de datos de panel," en *Crecimiento y convergencia regional en España y Europa, Vol. II*. Instituto de Análisis Económico, Barcelona.

Nelson, R. and Wright, G. (1992). "The Rise and Fall of American Technological Leadership: the Post-war Era in Historical Perspective." *Journal of Economic Literature*, Vol. XXX, pp. 1931-64.

Neven, D and Gouyette, C. (1995). "Regional convergence in the European Community." *Journal of Common Market Studies*, 33, pp. 47-65.

Nonneman, W. and Vanhoudt, P. (1996). "A further augmentation of the Solow model and the empirics of economic growth for OECD countries," *Quarterly Journal of Economics,* CXI (3), pp. 943-53.

Paci, R. and Pigliaru, F. (1995). "β convergence and/or structural change? Evidence from the Italian regions." Nota di Lavoro 72.95, Fondazione Eni Enrico Mattei.

Prescott, E. (1998). "Needed: A theory of total factor productivity." *International Economic Review*, 39(3), pp. 525-51.

Romer, P. (1986). "Increasing Returns and Long-Run Growth." *Journal of Political Economy*, 94(5), pp. 1002-37.

Romer, P. (1987a). "Growth Based on Increasing Returns due to Specialization." *American Economic Review Papers and Proceedings* 77(2), pp. 56-62.

Romer, P. (1987b). "Crazy Explanations for the Productivity Slowdown." *NBER Macroeconomics Annual* 2, pp. 163-210.

Romer, P. (1990). "Endogenous Technical Change." *Journal of Political Economy*, 98, pp. S71-S102

Shioji, E. (1997a). "Convergence in panel data: evidence from the skipping estimation." Mimeo, Universitat Pompeu Fabra, Barcelona.

Shioji, E. (1997b). "It's still 2%. Evidence of convergence from 116 years of the US states panel data." Mimeo, Universitat Pompeu Fabra, Barcelona.

Solow, R. (1956). "A Contribution to the Theory of Economic Growth. "*Quarterly Journal of Economics*, LXX(1), pp. 65-94.

Wolff, E. (1991). "Capital Formation and Productivity Convergence Over the Long Term." *American Economic Review*, 81(3), pp. 565-79.

Agglomeration and regional imbalance: Why? And is it bad?

Jacques-François Thisse

1. Introduction

In a world of globalisation, it is tempting to foresee the "death of distance" and, once the impediments to mobility have declined sufficiently, to wait for the predictions of the neo-classical theory of factor mobility to materialise. According to this theory, production factors respond to market disequilibrium by moving from regions in which they are abundant toward regions in which they are scarce. In equilibrium the capital-labour ratio is equal across regions, thus implying that both factors receive the same return in each region. In other words, the mobility of production factors would guarantee the equalisation of their returns across regions. As a consequence, there would be no reasons anymore to worry about where activities locate.

Yet, in the *First Report on Economic and Social Cohesion*, the European Commission observes that "*economic activity is strongly concentrated in the most urbanised areas of the Community. Regions with more than 500 inhabitants per square kilometres account for only 4% of the land area of the Union but for more than half the population. This implies that between two-thirds and three-quarters of the EU's total wealth creation occurs in urban areas*" (p.24). For the European Commission (1996, p.13), this is clearly a very bad state of affairs: "*Imbalances do not just imply a poorer quality of life for the most disadvantaged regions and the lack of life-chances open to their citizens, but indicate an under-utilisation of human potential and the failure to take advantage of economic opportunities which could benefit the Union as a whole.*" And, indeed, most regional policy debates in industrialised countries implicitly assume that there is too much spatial concentration in economic activity. Regional planners and analysts point to the inability of the market to organise the space-economy in a rational way and forcefully argue that public intervention is needed, but they fail to explain the nature of this market failure. This is precisely what I want to investigate in Section 2.

Using a simple model of monopolistic competition, Section 3 then illustrates what seems to be the main spatial feature of modern economies, namely the emergence of a *putty-clay economic geography*. More precisely, the recent fall in trade costs seems to allow for a great deal of flexibility on where particular activities can locate, but once spatial differences develop, locations tend to become quite rigid. Hence, regions that were once similar may end up having very different production structures.

In the subsequent section, I discuss a fairly neglected fact: regions *per se* do not exist and their size and shape critically depend on the criteria used to determine their borders. When the drawing of regional borders endows some entities with large economic agglomerations (such as metropolises or urban networks), the corresponding regions are likely to grow faster than others, thus providing a possible explanation why contiguous regions may exhibit different patterns of development. In Section 5, I suggest a few guiding principles for what could be a better institutional regional system at the level of the European Union. Some remarks conclude in Section 6.

Jacques-François Thisse is with CORE, Université catholique de Louvain and CERAS, Ecole nationale des ponts et chaussées, Paris.

2. Why is there a regional question?

2.1 The breakdown of the price mechanism in a spatial economy

As an economist, I find it natural to ask the question: *why would the market be unable to cope with location?* In order to answer this question, I must turn to general equilibrium analysis because this setting provides us with the benchmark that economists use to evaluate the market. The most elegant and general model of a market economy is the one developed by Kenneth Arrow and Gérard Debreu. It can be briefly described as follows. The economy is formed by agents (firms and households) and by commodities (goods and services). A firm is characterised by a set of production plans, each production plan describing a possible input-output relation. A household is identified by a relation of preference, by a bundle of initial resources and by shares in the firms' profits. Essentially, when both consumers' preferences and firms' technologies are convex, there exist a price system (one price per commodity), a production plan for each firm and a consumption bundle for each household that satisfy the following conditions: at the prevailing prices (i) supply equals demand for each commodity; (ii) each firm maximises its profit subject to its production set; and (iii) each household maximises her utility under her budget constraint defined by the value of her initial endowment and her shares in firms' profits. In other words, all markets clear while each agent chooses her most preferred action at the equilibrium prices.

In the Arrow-Debreu world, a commodity is defined not only by its physical characteristics, but also by the place where it is available. The same good traded at different places is treated, therefore, as a different commodity. Within this framework, choosing a location is part of choosing commodities. Hence, the Arrow-Debreu model integrates spatial interdependence of markets into general equilibrium in the same way as other forms of interdependence, thus suggesting that there is no regional question.

The inability of the competitive paradigm to deal with agglomeration, shown more than 20 years ago, seems to be ignored by a majority of the economics profession.

Unfortunately, however, things are not that simple. Although the inability of the competitive paradigm to deal with the process of agglomeration has been shown more than 20 years, it still seems to be ignored by a vast majority of the economics profession. Yet, Starrett (1978) has obtained a very general, although intriguing, result.

In order to illustrate his result, I consider the extreme case of a *homogeneous* space with a finite number of locations. By a homogeneous space, I mean the following: (i) the production set of a firm is the same in all locations; (ii) consumers' preferences are the same at all locations; and (iii) the natural resources are equally distributed across locations. Clearly, these assumptions are highly unrealistic. However, they are made in order to control for the impact that "nature" may have on the distribution of economic activity. Indeed, since our purpose is to understand why there exist large economic agglomerations, we are interested in finding socio-economic mechanisms that explain such agglomerations without appealing to physical attributes of locations. For the rest, Starrett describes the economy following the lines of the competitive framework as described in the foregoing.

Without making any convexity assumptions regarding preferences or technologies, Starrett shows the following result:

The Spatial Impossibility Theorem. *If space is homogeneous and transport is costly, then there is no competitive equilibrium involving transportation.*

What does it mean? At least two things. First, if economic activities are perfectly divisible, then a competitive equilibrium exists and it is such that each location operates as an autarchy. Typically, each local economy is identical to the others in that they have the same relative prices and the same production structure (backyard capitalism). This is hardly a surprising outcome since, by assumption, there is no reason for the economic agents to distinguish among locations and since each activity can operate at an arbitrarily small level. Agents thus succeed in reducing transport costs at their absolute minimum.

Second, if economic activities are not perfectly divisible, the transport of some goods between some places is unavoidable and, in this case, the above result tells us that no competitive equilibrium exists. One must then wonder about the reasons for such a market breakdown. In order to develop some insights, it is convenient to consider a 2-firm-2 location example in which (i) transportation is costly, (ii) firms must trade with one another, (iii) each firm can be set up in one place only, and (iv) firms' demand for land is somewhat but not perfectly elastic.

Let me now discuss the meaning of each of these assumptions for the non-existence of a competitive equilibrium that involves transportation. Clearly, such an equilibrium exists if firms can ship goods at zero transport costs or if firms can split their activities between the two locations and trade with each other within each location. In either of these cases, transportation vanishes from the economy, which becomes dimensionless. This is not what we are interested in so that the first three assumptions may be regarded as being rather undemanding for our purpose.

Consider now assumption (iv). If firms can substitute between land and other inputs such as capital, for example by building high-density structures, they can economise on the use of land while locating at the same place. In the realistic case where firms' demand for land has a positive and finite elasticity, both firms can locate together at location A. Hence, the land rent there is positive ($R_A > 0$) while the rent at the vacant location B is zero. If the land rent R_A is not too high relative to transport costs, this no-transportation configuration is an equilibrium. On the other hand, if R_A is very high, then at least one firm (say 2) can increase its profit by setting up at location B where $R_B = 0$ (recall that, at a competitive equilibrium, prices are given to economic agents and uninfluenced by their action), even though this entails some positive transportation cost between locations A and B. However, this new configuration cannot be an equilibrium. Since the price of a good is always lower at the place it is produced, firm 1 at location A finds the price of its output higher and the price of its input (excluding land) lower at location B, while firm 2 at location B finds the price of its output higher and the price of its input (neglecting land) lower at location A. Therefore, if land rents were the same at two locations, both firms would like to move, while a differential in the two land rents can at best prevent only one firm from moving. In other words, there exists no competitive rent system (1).

Thus, what creates problem in decentralising decisions across different locations is the fact that *economic agents may want to be separate because each one must choose to use a positive amount of land that cannot be made arbitrarily small*. When agents trade together, this physical separation implies that there is a tension between the incentive to reduce transport costs and the need to consume land. This tension cannot be solved at a competitive equilibrium because, space being homogeneous, *an agent's most preferred location depends only upon the locations chosen by the*

1) The reader may find the competitive assumption unrealistic in the case of two firms. As shown by the Spatial Impossibility Theorem, the number of firms can be made large without changing the conclusions.

others (2). Such a choice context has a "strategic" flavour that makes it difficult to be handled by the market.

Either space is heterogeneous (as in the neo-classical theory of international trade), or production and consumption externalities exist (as in urban economics) or markets are imperfect (as in economic geography).

Consequently, if we want to understand something about the spatial distribution of economic activities and, in particular, the formation of major economic agglomerations, the Spatial Impossibility Theorem tells us that we must assume either that *space is heterogeneous* (as in the neo-classical theory of international trade), or that production and consumption *externalities* exist and are many (as in urban economics) or that *markets are imperfectly competitive* (as in economic geography). The implications of the research strategy selected are important. For example, if we choose to rely on the heterogeneity of space, the market outcome is socially optimal. By contrast, if we choose one of the other two approaches, the market outcome is likely to be socially undesirable, thus pointing to the need for some public policy. Although it is obvious that space is heterogeneous, this seems weak as the main explanation for the existence of large metropolises as well as for the persistence of substantial regional income inequalities. When, then, should we distinguish between the other two possible solutions?

2.2 Externalities in the space-economy

For many years, the concept of externality has been used to describe a great deal of situations and it is important to have a clear perception of what they are. Following Scitovsky (1954), I consider two types of externalities: "technological externalities" (also called spillovers) and "pecuniary externalities". The former deals with the effects of non-market interactions that are realised through processes directly affecting the utility of an individual or the production set of a firm. By contrast, the latter refers to the benefits of economic interactions that take place through usual market mechanisms via the mediation of prices. Technological externalities imply that prices do not reflect the social values of goods and services whereas, for the same reason, pecuniary externalities are relevant when markets are imperfectly competitive (even in the absence of technological externalities).

According to Anas *et al.*, (1998), cities are replete with technological externalities. The same would hold in local production systems (Pyke *et al.*, 1990). In particular, communication externalities remain very critical in various fields such as management, administration, research, and finance. Knowledge, ideas and, above all, tacit information, can be considered as impure public goods that generate spillover effects from one firm or institution to another. Consequently, if economic agents possess different pieces of information, pooling them through informal communication channels can benefit everyone, hence the importance of proximity (Glaeser, 1999). This may be explained by the fact that the transmission of ideas that are not yet formalised cannot take place in a totally standardised way. The initial steps in the development of a new technology, say, require repeated contacts between the agents involved in order to establish a common language, interpret individualised pieces of information, and bring them into the operational state. Such a process is facilitated by spatial proximity. Even in this age of telecommunications "knowledge crosses corridors and streets more easily than oceans and continents" (Feldman, 1994, p.2). Thus, at the local level, it seems reasonable to appeal to technological externalities to understand the formation of clusters.

However, at the interregional level, it makes sense to assume that market imperfections are more central to explain what is going on. This is not to deny the reality of spillovers and congestion effects,

2) It should be kept in mind that the essence of the theory of competitive markets is that agents make their production or consumption decisions on the basis of the price system alone. Thus, adding considerations such as firm 1 in A recognises that if it locates in B firm 2 will also change its decision, takes us away from that theory.

but I believe that they have little to do with the imbalance between large regions. At this spatial scale, the reasons for over- or under-concentration have more to do with linkages between firms and consumers-workers, through product and labour markets. In particular, pecuniary externalities become critical because firms and individuals do not account for the impact that their migration decisions have on the well-being of those who stay put as well as on those who live in the region of destination since prices do not reflect the exact impact of agents' decisions. Consequently, when migration flows are substantial, one may expect the spatial economy to be inefficiently organised.

According to the spatial scale retained, each type of externality may help understand the spatial agglomeration of economic activities. Both will be illustrated in Section 3.

2.3 Co-ordination failure in a spatial economy

Given the implications of the Spatial Impossibility Theorem, we discover another fundamental reason that could explain why there is a regional question: *the incompleteness of spatial markets*. Indeed, an underdeveloped region has typically a small number of markets, which deprives firms and workers of the opportunity to signal demand and supply for jobs and intermediate services (Matsuyama and Takahashi, 1998). For the first welfare theorem to hold, that is, for the market outcome to be efficient, firms and households must be able to make their production and consumption decisions on the basis of publicly posted prices not only for the active markets, but also for all potential markets. Since an economic good is differentiated by its location, this means that firms and households must know a very large number of prices, one price being associated with each good at each location (neglecting for simplicity land and transport prices). Stated differently, in order to figure out whether or not an agent wants to buy or sell a good in a given location, she must know the price of this good even though nobody chooses to exchange the good at this particular location. It seems hard to believe that something like that could happen since the price of a good is not quoted before a market is open.

Yet, we know from Samuelson (1952) that there exists a relationship between the equilibrium prices of the same good at different locations that allows one to obviate this difficulty. The equilibrium price at a location in which the good is not produced is equal to the minimum of the marginal production cost (over the places where the good is produced) plus the corresponding transport cost. Once firms and consumers know the matrix of transport costs between all location pairs, they are able to infer from the prices where the good is actually produced, the equilibrium prices at all the other places. It is precisely because these potential prices are either too high or too low that the corresponding good is not traded at such places.

On the other hand, once markets are imperfect, only the prices quoted at the active locations are available. Economic agents are no longer able to apply the procedure described above because prices now vary with the decisions they make. For an economic agent to be able to evaluate whether or not a void location is a desirable alternative, she must be able to anticipate what will be the actions taken there by other agents. This is already enough to make the informational requirements gigantic. This is not the end of the story, however. The price level of a good that prevails at a location varies with the agents transacting there. Hence, there are likely to exist different price levels that depend on the parties involved in opening a new marketplace. We thus encounter an additional difficulty: potentially acting agents must co-ordinate on a single profile of actions for these actions to

The underdevelopment of a region may result from a lack of co-ordination between the agents who could be potentially involved in the opening of new markets.

be consistent. To sum-up, when we account for the fact that markets are imperfectly competitive, it appears that the underdevelopment of a region may result from the lack of co-ordination between the agents who could be potentially involved in the opening of new markets in this region.

Such a co-ordination problem becomes especially acute once it is recognised that the activity level of a region in a developed economy depends, at least to a large extent, on the availability of a wide array of service-firms providing untradables. As the profitability of these firms depends, in turn, on the size of the final sector (Abdel-Rahman and Fujita, 1990), the lack of information regarding the prices of intermediate goods in a lagging region may well be sufficient to prevent firms operating in the final (exporting) sector from locating there.

If the agglomeration of economic activity within a few regions appears to be inefficient, then co-ordinating *all* the decisions needed to open the relevant new markets in some lagging regions is both profitable and socially desirable. However, the information that such an operation requires is tantamount, due to the very large number of agents involved, and seems out of reach from any decision maker.

3. Why is there agglomeration, and is it bad?

In this section, two processes of agglomeration are discussed. In the first, I focus on the formation of clusters of firms in an economy whose markets are supposed to be unaffected by clusters' size, presumably because they are small relative to the rest of the economy (Section 3.1). In the second, I will shift to general equilibrium and will assume that both workers and firms are mobile, thus generating market effects at the level of the whole economy (Section 3.2). In both settings, consumers and firms can locate in one region only, which stands for the fundamental indivisibility that appears at the level of the person or of the plant. Also common to both settings is the fact that the emerging locational configuration is the outcome of the interplay between centrifugal and centripetal forces.

The most typical feature of the analysis is that the two processes are self-reinforcing. In particular, we will see that, once transport costs (broadly defined in order to include all the impediments to trade) have sufficiently decreased, regions that were initially similar end up with very contrasted production patterns. Hence, divergence instead of convergence should be expected as integration develops. Yet, as will be discussed in the concluding section, further decreases in transport costs may well foster the dispersion of some activities due to factor price differentials.

3.1 The formation of clusters

In the urban economics literature, it is common to focus on what is called *localisation economies*, that is, externalities affecting all the firms belonging to the same sector and located in the area within which these externalities produce their effects (Henderson, 1988). As discussed in the foregoing, the key element in the transmission of knowledge and ideas is given by the various institutions that foster personal contacts. As observed by Saxenian (1994), the institutional and economic environment influencing the collective process of learning within a given area is probably as important as the microeconomic linkages between firms and other economic agents. They are probably the main local factors that are at the origin of localisation economies.

The formation and size of clusters depend on the relative strength of three distinct forces: the magnitude of localisation economies, the intensity of competition, and the level of transport costs.

It is well known from industrial organisation that geographical proximity renders competition on the product market fiercer, thus inducing firms to locate far apart (d'Aspremont *et al*., 1979). This implies that firms' decisions to congregate or to separate depends on the trade-off between localisation economies and price competition. Furthermore, even though localisation economies lead to a reduction in a firm's cost, the same holds for its co-located competitors, thus intensifying the process of competition within the cluster and making the final outcome *a priori* unclear.

Even if price competition is relaxed through product differentiation, it is still true that firms want to be separate when transport costs are high. Since the emergence of industrial clusters is generally confined to small geographical areas, it is reasonable to assume that the spatial distribution of demand is unaffected by firms' locational behaviour. Therefore, the cost reduction associated with the agglomeration may be more than offset by the fall in exports. Consequently, transport costs have to be sufficiently low for firms to gather. Collecting all these arguments together, we may conclude that *firms supplying differentiated products must be able to serve almost equally well all markets* (globalisation) *in order to enjoy the local advantages associated with the formation of a cluste*r (localisation).

This argument can be made more precise, using the simple model discussed in Box 1 (3). This shows that the equilibrium states of the economy depend upon a factor, X, that combines into one variable the effects of transportation costs and the intensity of localisation economies. X also depends upon the total size of the market, the slope of the demand curve, and the substitutability between varieties.

Whether the economy has unique stable equilibrium that involves identical clusters, asymmetric clusters, or a single cluster in which all firms are agglomerated depends upon the value of this aggregate factor. The possible equilibrium patterns are displayed in Figure 1 where X is represented by the horizontal axis. The vertical axis shows the difference between the number of firms in each of two regions (N_A firms in region A, and N_B firms in region B). A dispersed equilibrium is unique (the solid line) when $X \leq 0$ so that there is "convergence". However, once X becomes positive, two more equilibria emerge (again the solid lines) and the story changes. These two equilibria involve a large and a small cluster (up to a permutation); they are stable while the dispersed equilibrium represented by the dotted line is now unstable. Once the value $X = 1$ is reached, all firms are concentrated within the large cluster whereas the small one vanishes: there is "divergence". Such a pattern provides a very neat example of what I meant by a putty-clay geography: the large cluster may equally arise in A or B but, once this has occurred, the corresponding region accommodates more and more firms.

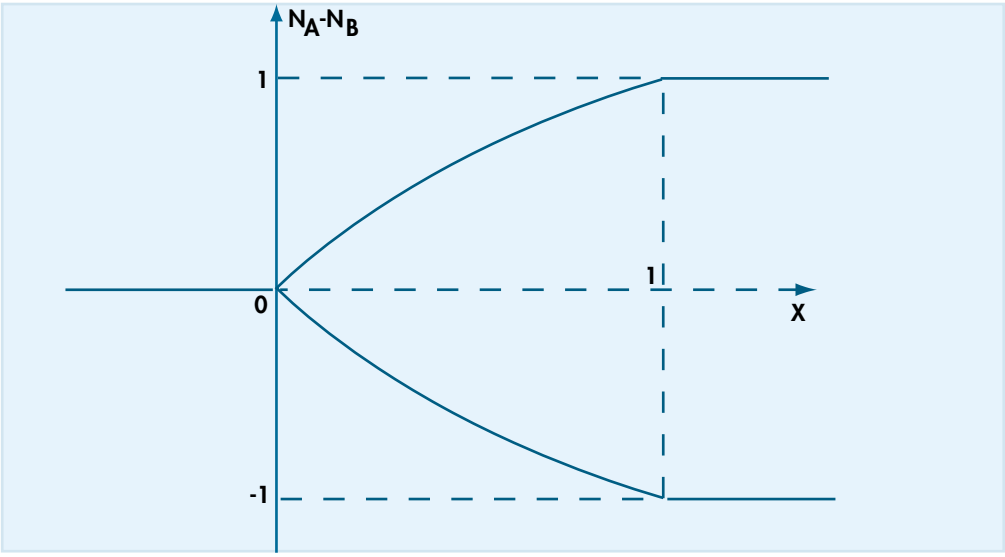
Low transport costs are likely to drive the economy towards more agglomeration because firms do not fear losing distant markets.

Despite the symmetry of the economy, *regions tend to nest clusters of different sizes* when there are localisation economies (4). Moreover, *a decrease in transport cost leads to more asymmetry between the clusters* (i.e. X increases). Low transport costs are likely to drive the economy towards more agglomeration in one region because firms do not fear the prospect of losing their business on distant markets. Second, we see that *more product differentiation induces more firms to locate within the large cluster* (again X increases). This is because product differentiation allows firms to relax price competition, thus leading them to exploit more the benefits associated with the presence of localisation economies. What makes these two properties relevant for us is the fact that transport costs keep decreasing whereas products become more and more differentiated. Therefore, we can safely conclude that these two effects combine to generate more agglomeration. In the limit, a single cluster may involve all firms.

3) For more details, see Belleflamme et al., *(2000).*

4) Indeed, when $\theta = 0$ the equilibrium always involves two identical clusters since X is always negative.

Figure 1. The emergence of asymmetric clusters



Finally, as the desirability of the differentiated product rises, the degree of asymmetry between the two clusters grows. This occurs because the relative impact of localisation economies rises with the market size. Consequently, economic growth would go hand in hand with a more agglomerated pattern of production.

Is such a concentration a waste of resources? In order to answer this question, I assume that a benevolent planner is able to impose marginal cost pricing as well as to choose firms' locations that maximise social welfare. It can then be shown that the efficient solution displays a pattern similar to that arising when firms are free to choose prices and locations at the market equilibrium. More surprising, perhaps, is the fact that the *large cluster never involves too many firms in equilibrium*: the planner wants to set up more asymmetric clusters than what arises at the market solution. This result requires some explanation. At the efficient solution, prices are set at the lowest admissible level and locations are chosen so as to maximise the social benefits of agglomeration net of transport costs. By contrast, at the market equilibrium, firms maximise their own profits. In doing so, they take advantage of their spatial separation in order to relax competition, thus making higher profits. These two effects combine to generate the above-mentioned discrepancy between the market and efficient outcomes. More generally, economic agents worry only about their role as "receivers", neglecting the fact that they are also "transmitters" to the others, in the collective process leading to the upsurge of localisation economies. This provides a strong argument to support the idea that clusters are not too big from the efficiency point of view.

Unless dispersion corresponds to both the equilibrium and the optimum, the difference between regional surpluses generates a conflict between regions about firms' locations. Indeed, the region with the larger cluster benefits from larger localisation economies, and thus lower prices, as well as from lower transportation costs on its imports. This occurs because the planner focuses only upon global efficiency and not on interregional equity. This makes sense when lump sum transfers compensating the consumers of the less industrialised region are available. However, *when such redistributive instruments are not available, a trade-off between global efficiency and interregional equity arises.*

Box 1. The formation of clusters

Consider an economy with a large number of firms producing each a differentiated variety.* Firms decide, first, to locate in either of two possible regions (say A and B) in which clusters are nested and, then, compete in prices. In order to focus on the impact of localisation economies, it is assumed that both regions A and B are characterised by the same market conditions. In each region, firms' demand functions are generated by a representative consumer who has a utility function quadratic in the varieties of a differentiated product and linear in a homogeneous product. This utility function exhibits love for variety, a feature that seems to characterise consumers in modern economies.

The demand function for firm i is linear and given by:

$$(1) \quad q(i) = a - b\,p(i) + c\int_0^1 [p(j) - p(i)]\,dj$$

where the variables $p(j)$ stand for the competitors' prices, whereas a expresses the intensity of preferences for the differentiated product and c the substitutability between varieties (the higher c, the closer substitutes the varieties). This demand system has the intuitively appealing property that a firm pricing above (below) the average market price has less (more) outlets than many of its competitors. However, the impact of such a price gap decreases as varieties become more differentiated because each firm increases its market power.

When firm i produces in region $r = A,B$, its marginal cost $k_r(N_r)$ decreases with the number of firms N_r located in r:

$$(2) \quad k_r(N_r) \equiv k\text{-}\theta.N_r$$

where k is a constant that stands for the marginal cost prevailing in the absence of agglomeration and θ is a parameter that measures the intensity of the localisation economies.

In selecting its price, each firm neglects its impact on the market but is aware that the market as a whole has a non-negligible impact on its behaviour through the average market price. This provides a setting in which individual firms are not competitive (in the classic economic sense of having infinite demand elasticity) but, at the same time, they have no strategic interactions with one another. In other words, there is monopolistic competition.

A spatial equilibrium is such that no firm can earn a higher profit by changing location. This arises at an interior point such as $0 < N_A < 1$ when the profit differential $\Delta\prod(N_A,N_B)$ between the two regions equals zero. However, an equilibrium may also arise when all firms agglomerate in region A (or in region B) provided that the profit differential evaluated at the corresponding agglomeration is nonpositive (nonnegative) so that no firm has an incentive to leave that cluster. Thus, we may observe two identical clusters (full dispersion), two asymmetric clusters, or a single cluster (agglomeration).

The profit differential can be shown to be given by:

$$(3) \quad \Delta\pi(N_A,N_B) = -C_1(N_A - N_B)\left[(N_A - N_B)^2 - X\right]$$

where C_1 is a positive constant whereas X depends on the basic parameters of the economy:

$$(4) \quad X \equiv \frac{\left[4a - 2b(2k + t - \theta)\,\theta\right] - ct^2}{c\theta^2}$$

t being the unit transport cost between the two regions.

This implies that full dispersion, in which the number of firms is the same in the two clusters, is always an equilibrium, as suggested by the neo-classical approach. However, this equilibrium becomes unstable when other equilibria emerge as solutions to the quadratic equation $(N_A - N_B)^2 - X = 0$, that is, when the parameters of the economy are such that X is positive. More precisely, it can be shown that the economy has a unique stable equilibrium that involves: (i) identical clusters if and only if $X \leq 0$; (ii) asymmetric clusters if and only if $0 < X < 1$; (iii) a single cluster in which all firms are agglomerated if and only if $1 \leq X$.

*Formally, the set of firms is described by the interval [0,1].

3.2 The core-periphery structure

The interregional economy is replete with *pecuniary externalities*. For example, when some workers choose to move away from their region, they are likely to affect both the labour and product markets in various ways. The result will be a change in the well-being of those who stay put. Moreover, the moving workers do not account either for the impact of their decision on the workers and firms located in the region of destination. Still, their moves will increase the level of demand inside this region, thus making the place more attractive to firms. They will also depress the local labour market so that, everything else being equal, the local wage is affected negatively. In sum, these various changes may increase or decrease the attractiveness of the destination region for outside workers and firms. Such pecuniary externalities are especially relevant in the context of imperfectly competitive markets because prices do not perfectly reflect the social values of individual decisions. They are also better studied within a general equilibrium framework in order to account for the interactions between the product and labour markets.

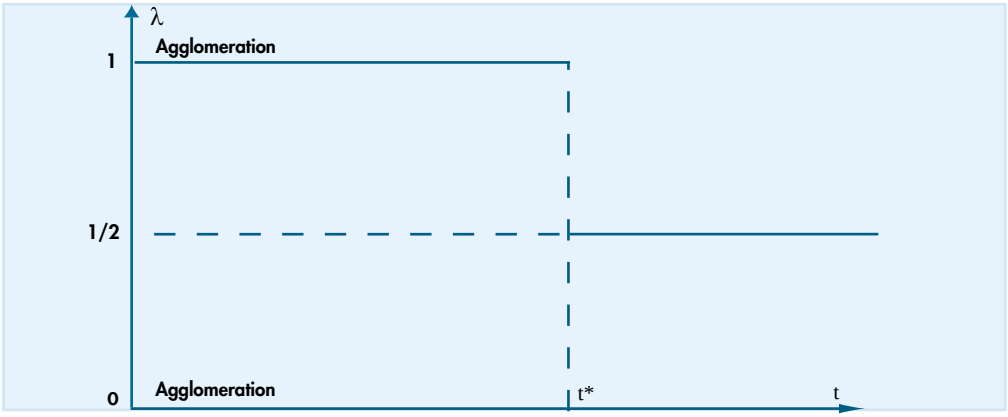
At first sight, this seems to be a formidable task. Yet, as shown by Krugman (1991), several of these various effects can be combined and studied within a simple enough general equilibrium model of monopolistic competition, which has come to be known as the *core-periphery model*. I present here some of the main conclusions of this model, using an alternative framework developed by Ottaviano *et al.*, (1999), which involves downward-sloping linear demands and a linear transport cost as in Section 3.1.

The intuition that underlies the agglomeration process is simple. The spatial immobility of a group of workers is a centrifugal force because they consume all types of goods. The centripetal force is less straightforward and finds its origin in a demand effect generated by love for variety. If a larger number of firms supplying a differentiated good are located in a region, the number of regional varieties is greater. In addition, the local price index is lower there than in the other region because local competition is fiercer. Everything else being equal, these two effects combine to make this region more attractive, thus inducing migration. The resulting increase in the number of consumers creates a larger local demand for the differentiated product, which in turn leads more firms to locate there. This implies the availability of even more varieties supplied at even lower prices in the region in question, so that the process keeps going on.

In order to figure out how this works, we add a new sector to the model developed before. This is shown in Box 2. This analysis shows that *agglomeration occurs when transport costs pass fall below some threshold level* (denoted by t^* in the model). Agglomeration rises because agents (firms and workers)

belonging to the modern sector are able to benefit from the exploitation of increasing returns within a single region without losing much sales in the peripheral region. This is shown in Figure 2. The horizontal axis shows the value of t, while the vertical axis gives the fraction of firms in the modern sector that are located in one of the two regions. The solid lines in the figure depict the stable equilibria.

Figure 2. The transition toward a core-periphery structure



When increasing returns are stronger, the threshold transportation costs at which this occurs (t^*) rises because starting the production of a variety requires more workers. Hence, agglomeration is more likely, the stronger are the increasing returns at the firm's level. In addition, this threshold also increases with product differentiation. In other words, more product differentiation fosters agglomeration because firms fear less the negative impact of competition.

Interestingly, the market provides excessive agglomeration only for intermediate values of the transport costs.

To which extent is such an agglomeration efficient? Considering again a benevolent planner who maximises total welfare, it appears to be efficient to have agglomeration once (i) transport costs are low, (ii) increasing returns are strong and/or (iii) the output is sufficiently differentiated. Hence, the optimum displays a pattern similar to that of the market equilibrium. This implies that there is a threshold value t° below (above) which optimality entails agglomeration (dispersion).

Interestingly, it can be shown that t° is such that to $t^\circ < t^*$. Such a discrepancy may be explained as follows. In the first place, the individual demand elasticity is much lower at the optimum (marginal cost pricing) than at the equilibrium (market equilibrium pricing), so that regional price indices are less sensitive to a decrease in t. The fall in transport costs must therefore be sufficiently large to make the agglomeration of workers socially desirable. In the second, workers do not internalise the negative external effects they impose on the workers who stay behind. However, the magnitude of these effects tends to decline as transport costs decrease by a sufficiently large amount. Hence, for agglomeration to be socially efficient, transport costs must be sufficiently low for the home market effect to be strong enough.

When transport costs are low ($t < t^\circ$) or high ($t > t^*$), no regional policy is required from the efficiency point of view, although equity considerations might justify such a policy when agglomeration arises. On the contrary, *for intermediate values of the transport costs* ($t^\circ < t < t^*$), *the market provides excessive agglomeration*, thus justifying the need for an active regional policy in order to foster dispersion on both the efficiency and equity grounds.

Box 2. The core-periphery in equilibrium

The model of Box 1 can be extended by having two factors (e.g., unskilled and skilled workers), denoted L and H. Factor L is evenly distributed between regions A and B and is spatially immobile. Factor H is mobile and λ denotes its share in region A.

There are two goods in the economy. The first good is homogeneous and is produced in the traditional sector, using factor L as the only input under constant returns to scale and perfect competition. In this sector, technology requires one unit of factor L in order to produce one unit of output. The good can be traded at zero cost between the two regions. In the traditional sector, equilibrium wages are therefore equal to one in both regions. The second good is a differentiated product supplied by a large number of firms belonging to the modern sector. These firms use H as the only input under increasing returns to scale and imperfect competition. In this sector, technology requires f units of factor H in order to produce any amount of a variety.

All consumers are endowed with preferences that are given by a utility identical to the one used in Box 1. Each variety can be traded at a positive cost of t units of the homogeneous good for each unit transported from one region to the other.

Labour market clearing implies that the number of workers in each region is just sufficient to allow the local firms in the modern sector to operate:

$$(1) \quad N_A = \lambda H/f \quad \text{and} \quad N_B = (1-\lambda)H/f$$

In this sector, the equilibrium wages corresponding to the above equation are determined by a bidding process between firms for labour, which ends when no firm can earn a strictly positive profit at the equilibrium market prices. In other words, all operating profits are absorbed by the wage bills. Firms and workers move together so that it is sufficient to focus on workers' migration. The distribution $(\lambda,1-\lambda)$ is a spatial equilibrium when no worker may get a higher utility level by changing location. It arises at an interior solution $0 < \lambda < 1$ when the utility differential $\Delta V(\lambda)$ is zero, or at a corner $\lambda = 0$ $(\lambda = 1)$ when the utility differential at the corresponding point is nonpositive (nonnegative).

Evaluating consumers' utility at the equilibrium prices and wages, we obtain the following regional utility differential:

$$(2) \quad \Delta V(\lambda) = C_2 (t^* - t).(\lambda - {}^1/_2)$$

where C_2 is a positive constant and $t^* > 0$ when there are increasing returns $(f > 0)$. It follows immediately from this expression that full dispersion $(\lambda = {}^1/_2)$ is always an equilibrium (as in Section 3.1). For $\lambda \neq {}^1/_2$, C_2 being positive and the above equation linear in λ, the utility differential has always the same sign as $\lambda - {}^1/_2$ if and only if $t < t^*$; otherwise it has the opposite sign. Hence, for large transport costs $(t > t^*)$, it is readily verified that the symmetric configuration is the only stable equilibrium. In contrast, when $t < t^*$, this equilibrium becomes unstable and workers agglomerate in region A (B) if the initial fraction of workers residing in this region exceeds ${}^1/_2$.

3.3 Some tentative conclusions

The agglomeration processes exhibit a "chaotic" behaviour that explains why it is so hard to make predictions about regional development.

There are striking analogies between the general trends pushing toward agglomeration in the two settings discussed in 3.1 and 3.2. In both cases, the nature of the process of agglomeration is similar, even though the former model involves partial equilibrium and the latter general equilibrium. In addition, we have seen that agglomeration processes exhibit a "chaotic" behaviour that could well explain why it is so hard to make relevant predictions about regional development.

Among the critical parameters, it appears that the current fall in the various components defining transport costs is likely to play a significant role in shaping modern space-economies by fostering more agglomeration. Similarly, a higher degree of product differentiation favours more geographical concentration (5).

By contrast, the conclusions in terms of welfare are different. On the one hand, when the process of interaction generating technological externalities goes both ways (as in Section 3.1), the equilibrium distribution of agents turns out to be less concentrated than the optimal distribution. On the other, when pecuniary externalities are at work (as in Section 3.2), it is reasonable to believe that there exists a domain of transport cost values for which there is too much agglomeration. Hence, if clusters might well be too small, there might be too much agglomeration at the macro-spatial level. Among other things, this implies that *different agglomeration mechanisms acting at different spatial scales may lead to different conclusions in terms of efficiency*. This invites us to say that there is no general presumption regarding the direction in which governments should move in their regional policies; instead, any policy recommendation should rest on a detailed analysis of the main agglomeration forces at work.

Yet, we have seen that co-ordination failure leaves scope for an active regional policy. However, governments often lack the relevant information and are influenced by local lobbies with vested interests. What is needed is a combination involving private agents and the public sector at the local level. In this perspective, the rapid growth of a land development industry in the United States or in South East Asia is worth noting. It reveals that private/public developers may succeed in correcting inefficiencies through the co-ordination of residential location choices within small areas. Together with housing or offices, developers also provide different public goods (Henderson and Mitra, 1996). It is reasonable to believe that similar operations could be undertaken in the case of industrial parks supplying specific technological infrastructure.

Technological infrastructure generally takes the form of a local network of specialised providers of technological services, characterised by economies of scope (Justman and Teubal, 1995). These firms provide technological services to small and medium firms in low and mid-tech industries. These might include: product design services; identification, screening and testing of new production technologies; quality control and application of international standards; solution of ecological problems facing firms in a particular sector or region; and so on. Stated differently, the purpose is to create on a lower scale the diversified advantages that a large metropolis naturally offers to entering firms.

5) It is worth pointing out that the agglomeration forces discussed here are also encountered in many other models in which they generate similar trends (Fujita and Thisse, 1996).

Even though a case-by-case approach can barely be avoided, I am convinced that the design of sensible regional policies requires a deep understanding of the making of the various types of economic agglomerations that fashion the European economic space. In all cases, we come to the conclusion that relatively micro analysis is needed.

4. How to divide space into regions?

4.1 The indetermination of regions

When comparing the economic performance of regions, one should strive to compare the like. This is far from being the case.

In addition to the difficulties encountered in the foregoing discussion, there is a fundamental problem with regions. While nation-states seem to be admissible reference entities in that they clearly distinguish themselves from each other by the scale of their domestic trade and the instruments of national sovereignty (at least until recently), this is not the case for the regions. The systematic use of the region as a solution to the different questions raised in the discussion on decentralisation had the effect of concealing the vagueness of the concept. Even if the definitions often vary with the authors, they nonetheless share two important characteristics. On the one hand, the places that form a region are, in one way or another, considered as sufficiently similar to be grouped together within the same spatial entity. On the other hand, whatever definition is opted for, a look at the literature leaves one with a sense of frustration, a feeling shared, and often expressed, by the authors of the studies themselves.

The similarity of the places is judged in light of the principle of homogeneity or of functionality. The former emphasises the similarities between the places, while the latter concerns itself with certain relationships between the various places forming a region. The principle of homogeneity underlies the emergence of the concept of natural region, developed in the 19th century by geographers. The principle of functionality led to the concept of an economic region developed later by both geographers and economists. Nevertheless, much of this continues to be arbitrary.

It is, indeed, often forgotten that *the region is a relative concept*. It should be clear to the reader that the regions in a given territory are the outcome of a process of division of that territory into a finite number of sub-sets of places based on a particular binary relation defined over the corresponding territory. Some places will be taken as "equivalent" when considered from the point of view of this relation, even if they differ in other respects, and will be grouped together within the same spatial entity called a region. In other words, the places forming a region depend totally on the relation chosen in order to compare them. This relation is called an equivalence in set theory, that is, a binary relation that is reflexive, symmetric and transitive. As territorial divisions can be made on the basis of different relations of equivalence, *the regions change if the relation changes*. This difficulty has led several economists and geographers to doubt the existence of regions *per se* (Isard, 1956).

4.2 Do we compare the like or the unlike?

When comparing the economic performance of regions, one should strive to compare the like. This is far from being the case because not enough attention has been paid by economists to the drawing of regions. Indeed, one may wonder what can be the meaning of having Ile-de-France or Lombardia together with Ireland, on the one hand, and Hainaut (Belgium), on the other, in the NUTS2 classification of European regions. Let me just illustrate the nature of the difficulty we

encounter with such regional data. For example, the Ile-de-France is formed by several "départements" whose size is comparable to that of the Belgian "provinces" such as Hainaut. Seine-Saint-Denis and Hauts-de-Seine are two such districts that are included in the Ile-de-France. While the former district is likely to have an income per capita below the European average that makes it somewhat comparable to Hainaut, the latter could well become the leading area of the EU, especially if the Quartier de la Défense is included in it. So, at the very least, one should be very careful when making comparisons based on such a regional division. Economists should become aware that geographers are much more careful than they are when building their data banks, and should revise their work accordingly.

In the second place, one should not be fooled by the numbers. First, the per capita Gross Domestic Product of a region is a relevant indicator only if the resident population is more or less the population actually working in the region (this is true for Ile-de-France but not for the Hamburg and Brussels regions). A region, therefore, has a minimum possible size. Also, the differences between nominal per capita incomes are not of any great significance. Real incomes would need to be compared and, to do so, regional price indices would have to be available. This is especially important for strongly urbanised regions since we know that land rents are significantly higher in larger cities than in small cities, and that housing expenditures stand for a large fraction of individual consumption. Moreover, even the real per capita income does not take into account the many social and environmental variables that also influence the well being of people (after all, individuals are welfare-maximisers and not income-maximisers). In particular, living conditions and amenities are too often neglected in interregional economic comparisons, although they have a considerable influence on the well being of the inhabitants of a given area.

The work of geographers and regional scientists allows us to shed light on some interesting facts. First, given the existence of an urban hierarchy and the fact that big cities tend to be more productive than small cities, one should expect regional products per capita to be different and not to converge. For example, regions including major cities are likely to have a higher product per capita than the others. Not surprisingly, the leading European areas are precisely the large urban regions (Ile-de-France, Lombardia, Great London,...). Second, we also know that focussing on smaller spatial units is often associated with a widening of spatial inequalities. Therefore, the decentralisation of nations into regions makes regional disparities more transparent. This is likely to be accompanied by a higher demand for spatial equity than before. Third, one should keep in mind that the magnitude of spatial disparities is very sensitive to the design of the regional borders. This is true especially for small regions, such as those confined to the limits of a city. We encounter here what geographers call a *modifiable areal unit problem* (Goodchild, 1979; Openshaw and Taylor, 1979). By paying insufficient care to the design of regions, one runs a serious risk: arbitrary or partial indicators of regional discrepancies are likely to exacerbate the demand for spatial equity, thus threatening the political stability of the European construction.

Finally, I am not sure that all these interregional comparisons make any good economic sense. For example, by enlarging the European Union to some large and relatively poor countries (e.g., the Eastern European countries), most regions of the EU-15 would have a per capita regional product that would exceed the average product of the enlarged Union (6). Instead, it seems to me that the relevant

6) This poses again the question of the space to be used as a reference.

question is: *did a particular region succeed in improving upon its previous economic and social performances?* Although we all know the economic process is not a zero-sum game, by focussing too systematically on regional ranking we give ground to the idea that the growth of some regions takes place at the expense of others and that economic development resembles to a horse race.

5. The metropolitan region as a foundation for public policy

A region should be organised around a major city or a network of smaller cities.

The arbitrary nature of administrative divisions is likely to cause a reduction in the advantages associated with the agglomeration of private activities and the decentralisation of public decision-making if they cannot be deployed within coherent political and economic entities. In defining regions, there should be (at least) two guiding principles. First, *a region should be organised around a major city* (or a network of smaller cities) because this offers a greater potential for a better management of the various agglomeration forces at work in a metropolitan area. Second, *a region should internalise most local public policy effects*, thus achieving a better balance between the public and private spheres. We call *metropolitan regions* spatial areas organised on the basis of these two principles (Thisse and van Ypersele, 1999). Let me explain the reasons that have motivated these choices.

5.1 Agglomeration forces in cities

There is a great deal of evidence suggesting that cities or clusters play a growing role in modern economies. Besides the reasons mentioned in Sections 2.2 and 3.2, there are several others that explain this fact. In the first place, cities offer a very wide range of services and intermediate products that permit an increase in the productivity of private inputs (Hansen, 1990). This effect is particularly significant for the labour factor (Peri, 2000; Rauch, 1993). Moreover, firms find a wider range of skilled workers in cities; in the same way, workers face a large number of differentiated job opportunities, which enables them to enhance the value of their skills. The result is *a better matching between jobs and workers in urban labour markets* which, simultaneously, tend to fragment and diversify (Hamilton *et al.*, 2000). In short, the division of labour becomes finer in major cities as a result of the diversification and specialisation of tasks.

Contemporary forms of corporate organisation also contribute to the increasing role of cities in the process of economic development. Since the 1980s, one witnesses a drastic reduction in the internal share of production (7). The growth of outsourcing is the result of firms' policy to refocus their activities on their core competencies. At the same time, a policy of product customisation has developed, aimed at exploiting economies of scope by appealing to flexible production techniques. Such a sales policy, operated in conjunction with just-in-time management, results in the customer being supplied much faster than previously. All of this has resulted in a substantial growth in logistic services whose costs can be reduced by the formation of *clusters* (Porter, 1998). Clusters are very naturally nested in large cities (8), although they may also emerge outside of major cities as shown by many of the Italian industrial districts (see Pyke *et al.*, 1990).

7) For example, in Germany, the internal share has dropped from 67% in 1981 to less than 50% in 1990 (Conference Européenne des Ministres des Transports, 1997).

8) For the United States, Porter (1998) cites the examples of New York, Los Angeles, Boston, Detroit and Seattle. This list is not exhaustive.

5.2 Urban public goods

I now turn to local public policies. It is well known that a major difficulty in the supply of public goods lies in the incentive that each individual has not to reveal her true preferences with regard to the quantity of public goods she wishes to consume. Tiebout (1956) has pointed out that most public goods are local: their effect is only felt within service areas of a limited size. Therefore, the existence of communities supplying different quantities of public goods on the basis of different local taxes would lead households to reveal their preferences by their residential choice. The supply of public services would then also be carried out by a whole set of communities competing with each other on the political market. Although attractive and ingenious, Tiebout's suggestion is in many ways incomplete and its limits have been known for a long time.

It is possible to reconsider this approach from the territorial viewpoint. In this perspective, Hochman *et al.*, (1995) have proposed grouping the service areas of various local public goods within the same territory, while financing them by means of a tax on the land rent prevailing within this territory. These authors show that the constitution of such entities make it possible to achieve efficiency through competition between local governments. Geographical considerations would then seem to impose an institutional system having a territorial rather than a functional basis. In such a context, the relevant decision making entities should be consolidated and incorporated into areas sufficiently large to allow them to internalise as much as possible the effects of local public policies. This involves extending both the geographical base of local governments and their powers. Allowing for too fragmented territorial units fosters spatial/social segregation and runs, therefore, against the basic principles of redistribution.

Because it tries to incorporate both agglomeration economies and public policies within the same reference area, the concept of metropolitan region, while remaining arbitrary, seems to be less open to criticism. It must be emphasised, however, that the formation of regions must be endogenous. When arbitrary administrative boundaries are imposed without economic justifications, the expected benefits of agglomeration and decentralisation might be called into question. In addition, the regional boundaries should be revised regularly because of the continuous decrease in communication and transport costs. Such a proposal certainly goes against the customary habits that consider administrative boundaries permanent. Let me say that I am fully aware that this rigidity responds, at least partially, to the individual's need to belong to a lasting community whose geographic contours have to remain stable. Nevertheless, such a feeling should not deter an effective co-operation between neighbouring regions, especially when borders are considered (at least for the time being) as permanent.

6. Concluding remarks

There is a regional question whose origin lies in the inability of competitive markets to deal with the process of agglomeration of economic activities, due to the direct interdependence between locational decisions. But, apart from that, what have we learned? In what follows, I do not intend to be comprehensive but will restrict myself to summarise the main points.

1. Modern economies encapsulate a strong system of forces pushing toward more agglomeration in economic activities (see also Fujita *et al.*, 1999). What makes these forces so powerful is the

combination of a drastic fall in transport and trade costs, which combines with the cumulative nature of the agglomeration process. This gives rise to a new type of economic geography in which space is "slippery" while locations are "sticky". Furthermore, technological progress brings about new types of innovative activities that benefit from being agglomerated and, therefore, tend to arise in developed areas (Audretsch and Feldman, 1996). Consequently, the wealth or poverty of people seems to be more and more related to the existence of prosperous and competitive clusters of specific industries, as well as to the presence of diversified and large metropolitan areas.

There is a risk of excessive agglomeration if the mobility of the skilled labour force keeps rising. Still, firms are likely to be attracted by cheaper areas offering a niche.

There is a risk of excessive agglomeration at the level of the EU if the mobility of the skilled labour force keeps rising. Yet, one would go too far in predicting that the European space will be much more polarised than what it is today. Urban systems are characterised by a strong inertia that favours dispersion. In addition, the growing concentration of activities in a few large regions is likely to be accompanied with higher urban costs (such as land rent, commuting costs, pollution) that will make these regions eventually less attractive (Ottaviano *et al*., 1999). Finally, even though innovative activities often benefit from being agglomerated, firms are likely to be attracted by cheaper areas when technologies are well monitored, thus offering a *niche* to less diversified areas that can specialise in the production of specific goods (Commissariat Général du Plan, 1999). At the very least, this is a scenario suggested by the American example (Henderson, 1997). In this perspective, European cities would do well by improving their provision of public goods and services used directly by firms and by co-operating more with their hinterland.

2. Local clusters may emerge in very different places, thus opening the door to possible local development within depressed regions. However, one should resist to the temptation of planning and organising such clusters from above. Indeed, they often rest on informal processes such as discussions among workers within firms, inter-firm mobility of skilled workers, exchange of ideas within families or clubs, and bandwagon effects. The proliferation of externalities within cities leads Anas *et al*. (1998, p.1458) to conclude as follows: "only very comprehensive and detailed planning can overcome the resulting inefficiencies. Because the externalities are so poorly understood, however, attempted cures may well do more harm than the disease". It is my contention that the situation is very similar when we come to the case of regional clusters, although the nature of externalities to take into account may be different (Soubeyran and Thisse, 1999).

   Still, there is a lot to be learned from the many successful experiences undertaken. In particular, they concur in saying that the *efficiency and quality of local institutions* that facilitate communication and social co-ordination are critical in successful local development stories. This is a far too much neglected factor in development plans designed for lagging regions. The European Commission should be more active in detecting such inefficiencies and in making its regional aids conditional upon significant improvements in local (nonmarket) institutions.

3. Globalization of investment and decentralisation of public policy intensify fiscal competition between regions seeking to attract private investment. Offering subsidies is a common policy, but it raises the risk of a "race to the bottom". Yet, the rapid rate of technological change, following on revolutionary developments in information technologies, life sciences, and new materials, has

introduced new dimensions of differentiation, which offer regions new opportunities to avoid the excesses of fiscal competition by developing *technological infrastructure*.

The European Commission should be more active in detecting institutional inefficiencies and in making its aid conditional upon improvements in local institutions.

Technological infrastructure is aimed at developing advanced scientific and engineering capabilities, often through generic research, that answer the needs of firms operating at the leading edge of technological innovation. Technological infrastructure introduces a dimension of quality that is far more pronounced than in conventional infrastructure, as firms differ widely in their ability to exploit it. *This offers regions an opportunity to avoid head-to-head competition by differentiating the characteristics of their infrastructure*, much as firms differentiate their products to relax price competition. The European Commission could launch an autonomous agency whose role would be to foster a better co-ordination among development plans established by regional governments.

References

Abdel-Rahman, H. and Fujita, M. (1990). "Product Variety, Marshallian Externalities, and City Sizes". *Journal of Regional Science* 30, pp. 165-83.

Anas, A., Arnott, R. and Small, K.A. (1998). "Urban Spatial Structure". *Journal of Economic Literature* 36, pp. 1426-1464.

Audretsch, D.B. and Feldman, M.P. (1996). "R&D Spillovers and the Geography of Innovation and Production". *American Economic Review* 86, pp. 630-40.

Belleflamme, P., Picard, P. and Thisse, J.-F. (2000). "An Economic Theory of Regional Clusters". *Journal of Urban Economics*, forthcoming.

Commissariat Général du Plan (1999). *Scénario pour une nouvelle géographie économique de l'Europe*. Economica, Paris.

Conférence Européenne des Ministres des Transports (1997). *Les nouvelles tendances de la logistique en Europe*. Editions de l'OCDE, Paris.

d'Aspremont, C., Gabszewicz, J.J. and Thisse, J.-F. (1979). "On Hotelling's Stability in Competition". *Econometrica* 47, pp. 1045-50.

European Commission (1996). *First Report on Economic and Social Cohesion*. Office for Official Publications of the European Communities, Luxembourg.

Feldman, M.P. (1994). *The Geography of Innovation*. Kluwer Academic Publishers, Dordrecht.

Fujita, M., Krugman, P. and Venables, A.J. (1999). *The Spatial Economy. Cities, Regions and International Trade*. MIT Press, Cambridge (Mass.).

Fujita, M. and Thisse, J.-F. (1996). "Economics of Agglomeration". *Journal of the Japanese and International Economies* 10, pp. 339-378.

Glaeser, E.L. (1999). "Learning in Cities". *Journal of Urban Economics* 46, pp. 254-277.

Goodchild, M. (1979). "The Aggregation Problem in Location-Allocation". *Geographical Analysis* 11, pp. 240-256.

Hamilton, J., Thisse, J.-F. and Zenou, Y. (2000). "Wage Competition with Heterogeneous Workers and Firms". *Journal of Labor Economics*, forthcoming.

Hansen, N. (1990). "Do Producer Services Induce Regional Development?". *Journal of Regional Science* 30, pp. 465-478.

Henderson, J.V. (1988). *Urban Development. Theory, Fact and Illusion*. Oxford University Press, Oxford.

Henderson, J.V. (1997). "Medium Size Cities". *Regional Science and Urban Economics* 27, pp. 583-612.

Henderson, J.V. and Mitra, A. (1996). "The New Urban Landscape: Developers and edge cities". *Regional Science and Urban Economics* 26, pp. 613-43.

Hochman, O., Pines, D. and Thisse, J.-F. (1995). "On the Optimal Structure of Local Governments". *American Economic Review* 85, pp. 1224-1240.

Isard, W. (1956). "Regional Science, the Concept of Region, and Regional Structure". *Papers of the Regional Science Association* 2, pp. 13-26.

Justman, M. and Teubal, M. (1995). "Technological Infrastructure Policy (TIP): Creating capabilities and building markets". *Research Policy* 24, pp. 259-281.

Krugman, P.R. (1991). "Increasing Returns and Economic Geography". *Journal of Political Economy* 99, pp. 483-499.

Matsuyama, K. and Takahashi, T. (1998). "Self-defeating Regional Concentration". *Review of Economic Studies* 65, pp. 211-243.

Openshaw, S. and Taylor, P. (1979). "A Million or So Correlation Coefficients: Three experiments on the modifiable areal problem". In J. Wrigley (ed.), *Statistical Applications in the Spatial Sciences*. Pion, pp. 127-144, London.

Ottaviano, G., Tabuchi, T. and Thisse, J.-F. (1999). "Agglomeration and Trade Revisited". University of Tokyo, CIRJE-F-65.

Peri, G. (2000). "Human Capital Externalities in US Cities". *Journal of Urban Economics*, forthcoming.

Porter, M.E. (1998). *On Competition*. Boston, A Harvard Business Review Book.

Pyke F., Becattini, B. and Sengenberger, W. (1990). *Industrial Districts and Inter-firm Cooperation in Italy*. International Institute for Labour Studies, Geneva.

Rauch, J. E. (1993). "Productivity Gains from Geographical Concentration of Human Capital: Evidence from the Cities". *Journal of Urban Economics* 34, pp. 380-400.

Samuelson, P.A. (1952). "Spatial Price Equilibrium and Linear Programming". *American Economic Review* 42, pp. 283-303.

Saxenian, A. (1994). *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Harvard University Press, Cambridge (Mass.).

Scitovsky, T. (1954). "Two Concepts of External Economies". *Journal of Political Economy* 62, pp. 143-151.

Soubeyran, A. and Thisse, J.-F. (1999). "Learning-by-doing and the Development of Industrial Districts". *Journal of Urban Economics* 45, pp. 156-176.

Starrett, D. (1978). "Market Allocations of Location Choice in a Model with Free Mobility". *Journal of Economic Theory* 17, pp. 21-37.

Thisse, J.-F. and van Ypersele, T. (1999). "The Challenge Raised by Metropolisation and Fiscal Competition in Economic Development". *World Economy* 22, pp. 1201-1220.

Tiebout, M. (1956). "A Pure Theory of Local Expenditures". *Journal of Political Economy* 64, pp. 416-424.

The role of public policy in the process of regional convergence

Philippe Martin

1. Introduction

Spatial concentration of economic activities is one of the most salient features of economic development. The almost parallel urge by policymakers to counteract such a trend through public polices is also striking. This is not only reserved to those countries, especially in Europe, which have a long tradition of public intervention. To a lesser extent, the United States has (for example during the New Deal period) put into place policies aimed at correcting uneven patterns of regional development. Public intervention is usually defended on either efficiency or equity grounds. In the case of economic geography, a justification in terms of efficiency implies identifying the various market failures, specific to the issue of space, that make the optimal economic geography differ from the one induced solely by market forces. Although one also needs to show that public intervention will make a better job than market forces, the identification of market failures is a necessary first step to justify public intervention on efficiency grounds. Whereas this type of analysis has been standard for public intervention in the fields of education, technology, pollution, etc., the counterpart for regional policies is much less developed. There are two ways forward: the first is to analyse how some "standard" market failures are modified by the introduction of space and distance and how in turn, this should affect the definition of public policies; the second is to understand how space and distance themselves can be at the origin of market failures.

Another way to justify public intervention is to do it on equity grounds. Some economic agents, workers and consumers, are not mobile and are stuck in poor or declining regions, regions from which mobile factors, some labour and capital, have left. Because of the lower demand for labour in those regions, real wages will either adjust downwards or if real wages cannot adjust due to rigidities on the labour market, unemployment will increase. As consumers, these agents will also see their welfare decrease because some of the goods and services formally produced locally will be produced in the core, richer region. In this case, they will either have to pay a higher price for those goods and services because of the transaction cost involved in importing them from the rich region. In some cases, in particular for services, the transaction cost will become so high that they will become non-tradable so that the diversity of available services will decrease. Also, if the mobile agents are those with the highest human capital and if positive spillovers exist between workers due to localised social interactions, then as mobile agents move away from the poor region, immobile workers will also lose the benefits of these positive spillovers which may imply a decrease in their productivity and therefore in their equilibrium wage. One can say that the root of this problem is then the lack of concentration and the lack of mobility of agents rather than concentration itself. This is partially right and we want to analyse some policy implications of this interpretation. However, one could not go too far along this road because some economic agents will always remain immobile so that the equity motive behind regional policies remains. This raises the question whether regional policies are best equipped to deal with this issue and how to co-ordinate them

Philippe Martin is with the Centre d'enseignement et de recherche en analyse socio-économique (CERAS) of the Ecole National des Ponts et Chaussées, Paris. He was awarded First Prize in the 1999 EIB Prize *competition for his essay,* Are European regional policies delivering? *(see* EIB Papers, *4(2), 1999).*

with other redistributive policies. Finally, we want to analyse how regional policies affect economic geography and regional inequalities. One important difficulty is that these policies, due to the very nature of the self-sustaining agglomeration forces at work in economic geography, have extremely complex long-term effects.

2. Searching for market failures in economic geography

Physical space has a strong impact on both technological and pecuniary externalities.

Externalities are the best friend of an economist who wants to defend public intervention, and regional policies are no exception. Both *technological* and *pecuniary* externalities can be put forward in the case of economic geography because physical space has a strong impact on both. The first category occurs when there are technological spillovers that are spatially localised. Several reasons can be advanced. One possibility is the existence of localised technological spillovers such as those studied by Jacobs (1969) and by Henderson and *et al.*, (1995). For instance, the proximity of numerous firms might enable the innovative sector greater scope for observing and analysing the production process and thereby facilitate the creation of new production processes. Silicon Valley is the most successful example of the effect of such interactions between producers and innovators in a particular domain, that of information technology (1). Northern Italian regions are other examples of the force of such localised spillovers. Also, if the innovative sector uses manufacturing sector inputs, its concentration will enable transaction costs and hence the cost of innovation to be reduced. In this case, the positive externality arising from spatial concentration is pecuniary, operating through an effect on prices (see Martin and Ottaviano 1996, for such a model).

A further type of externality comes from the fact that firms (and in general owners of mobile factors) do not take into account the welfare of other agents when they choose where to locate. In particular, they do not take into account the welfare of those agents who are immobile. The reason is that they do not get the whole benefits linked to their location decisions. Here the market failure is due ultimately to the fact that certain agents do not move. If no congestion effects appear, then full concentration would not create any problem. Hence, if this were the only market failure, public policies that promote mobility of workers should be enough to respond with problems caused by agglomeration. Indeed, the fact that mobility (both between regions of a given country and between countries) is much lower in Europe than in the US explains why the location of economic activities has become a more important policy issue on this side of the Atlantic. From the policy point of view, housing and tax policies that facilitate the mobility of workers should therefore be regarded as part of the regional toolkit. The fact that regions can be specialised in specific industries also suggests that low inter-sectoral mobility of workers adds to the welfare cost of spatial concentration. This means that policies that facilitate inter-sectoral mobility such as education and training policies in poor regions should be reinforced.

In a recent paper, however, Matsuyama and Takahashi (1998), show that the freedom to move can in fact be self-defeating in certain circumstances. They show, in fact, that agents would be better off if their freedom to move were taken away. The reason is that as agents move to the agglomeration in pursuit of a better life because of the diversity of services and goods provided there, the production of the goods from the poor region (now in "the middle of nowhere") declines and the standard of living of all agents drops. Here, the market failure is the absence of co-ordination between the

1) The work of Jaffe, Trajtenberg and Henderson, (1993) shows that the citation and use of patents is very localized. This is very strong evidence that knowledge spillovers are themselves very localized.

different agents rather than immobility *per se*. Another market failure that is not solved by the mobility of agents, and can even be aggravated by it, is congestion externalities.

Space itself can be at the origin of market failures because it leads to imperfect competition.

Finally, space itself can be at the origin of market failures because it leads to imperfect competition. The reason, first analysed in the context of the Hotelling model, is that distance between producers gives firms a relative market power over consumers who are located nearby. In this case, as transaction costs go down, competition between firms is reinforced and firms will react by differentiating their products along non-geographical characteristics. This important insight has been analysed by Gabszewicz and Thisse (1986) and Scotchmer and Thisse (1992). Some of its implications for regional policies may not have been entirely exploited yet. If regional policies reduce transaction costs, then they take away the monopolistic power of firms that is based on distance, and increase the incentive to regain monopolistic power through product differentiation. This latter type of differentiation may have some positive impact on welfare if consumers value diversity.

3. Equity considerations

What is the impact of economic geography on equity, and can regional policies be justified on this ground? This depends very much upon: the relative distortion effects of regional policies and of redistributive fiscal policy on individuals; the mobility of factors (capital and labour); and, the extent of inequality among individuals in the population of both the poor and the rich regions.

To make equity considerations a possible justification for regional policy, we must assume first that non-distortionary lump sum transfers are not possible. Otherwise, if a region experiences a delocation of its economic activities and could be compensated in this way by another region, then the question of regional inequalities would be easy to solve. The standard view is however that such redistribution is indeed not possible due, in particular, to information problems. In that case it can be argued that regional policies are less distortionary than the income taxes needed to compensate individual losers of changes in economic geography. However, regional policies add a supplementary distortion in the sense that they alter economic geography through the location decisions of firms. In recent papers (Martin 1999a and 1999b), I have argued that a trade-off exists between equity and efficiency at the spatial level so that public policies which, through taxation and subsidies, induce firms to relocate in poor regions may reduce the overall efficiency of the economy. An indication of these arguments is given below.

The spatial equity problem also depends very much on income inequality in the population. The more inequality among the individuals, especially between workers and capital owners, the more acute the problem of spatial inequality will be. This can be seen in a simple model with two regions and two factors, mobile capital and immobile agents such as in Martin and Rogers (1995). Workers of the monopolistic manufacturing sector and of the perfect competition sector earn the same nominal wage in the two regions. This is because the goods produced in the perfect competition sector are traded with no costs so that nominal wage rates are equalised. What determines their welfare is their real wage which also depends on the number of firms in each region. Workers in the region with the highest number of firms gain because they pay lower transaction costs as many of the goods are produced locally. This decreases the price index and therefore increases the real wage in that region.

Facilitating capital mobility between the two regions, for example by eliminating legal barriers to plant closures, have a different impact on the welfare of the different agents. If firms relocate from the poor to the rich region, the price index will increase in the poor region and decrease in the rich region. However the return (profits in the monopolistic sector) to capital will increase in the poor region and decrease in the rich region. The reason is that as firms move out of the poor region, local competition will decrease and the opposite will occur in the rich region. Another way to say this is that firms from the poor region will move out if profits are higher in the rich region up to the point where returns are equalised in the two regions. Unambiguously welfare of workers in the poor region decreases: their nominal wage is tied by factor equalisation due to free trade in the perfectly competitive sector, but the price index increases so that their real wage decreases. The inverse happens for workers in the rich region so that inequality between workers of the two regions increases when firms are free to choose location.

The equity motive behind regional policies is not as straightforward as it seems because it requires a choice between types of inequality.

The situation for capital owners is more ambivalent. The nominal income of capital owners in the poor region rises. The relocation of some firms to the rich region lowers competition and increases the profits of the firms that they own. However, as consumers, they may loose because the price index increases in the poor region. Following the methodology of Martin and Rogers (1995), it is possible to show that capital owners in the poor region will gain with relocation if transaction costs are low enough and if the extent of competition (measured by the inverse of the degree of elasticity of substitution between varieties in the monopolistic sector) is not too high. The exact reverse result holds for capital owners in the rich region. However, because the nominal income of capital owners in the poor region rises with free relocation, the inequality between workers and capital owners in the poor region (measured in terms of real income or welfare) will always increase when firms choose freely their location. This may be an important argument in favour of regional policies. However, note that the concentration process in the rich region will, by the same reasoning, decrease inequality between workers and capital owners in that region because as competition increases, profits will decrease (to equalise those in the poor region by an arbitrage process) as well as incomes of capital owners. This implies that regional policies that would tend to impede this relocation process will benefit immobile workers of the poor region but will harm immobile workers of the rich region. It would decrease inequality in the poor region and increase it in the rich region. The equity motive behind regional policies is thus not as straightforward as it seems because it requires a choice on reducing one type of inequality at the expense of another type of inequality.

Another important result is that the extent to which inequalities will be increased by letting the concentration process free will depend crucially on the distribution of factors of production. The more unequal the distribution of mobile capital in the population the more the concentration process will exacerbate inequalities in the population. The reason is that if immobile workers can relocate some capital, then the welfare loss due to higher a price index when firms relocate outside the region will in part be compensated by an increase in their income from the higher return to capital outside the poor region.

Equity considerations are important for analysing regional policies. However, the question: "Do regional policies decrease inequality between poor and rich regions?", is not the same as: "Do regional policies improve welfare of agents in the poor regions?". To see this we will use two simple examples.

In a similar framework as the one described above, suppose that we look at the welfare impact of a decrease of transaction costs between a poor and a rich region. This could be the result, for

example, of building a new highway. The impact on the two regions of decreasing transaction costs in this way is modelled in more detail in Box 1. In this model an improvement of infrastructure facilitating trade leads to relocation of firms from the poor to the rich region. Firms can now better exploit economies of scale in the larger market and still export to the poor region as trade is facilitated between the two regions (2). Hence, if we were to look at regional GDP we would see a fall in the poor region and an increase in the rich one. From that point of view, one could interpret this policy as increasing inequalities between the two regions.

But what is the impact on welfare of a worker in the poor and the rich region? Lower transaction costs affect welfare in two different ways. The direct effect, lower costs for imported goods, is always positive for the poor region (3). At the same time, industrial location from the poor to rich region has a negative indirect impact on welfare in the poor region as more goods must be imported at a cost. In this particular model, the direct benefit is always greater than the indirect loss for the poor region. Hence, the example shows that, even though on *equity* grounds a policy of lowering transaction costs may not be called for, it can be defended on the grounds that it *increases welfare* of the poor region. *A contrario*, even though new economic geography insists on the concentration effects of lower transaction costs, its normative implications are certainly not to promote higher transaction costs.

Nevertheless, it is true in such models that if a planner could change economic geography, that is, could choose the number of firms in each region, equity considerations would entail to increase the number of firms in the poor region at the expense of the rich region (this again assumes that no lump sum transfers are possible as these may dominate such a distortionary policy). However, this result itself is not general. Martin and Ottaviano (1999) show that the existence of localised technology spillovers introduces an ambiguity. In this case, higher concentration in the rich region increases the extent of technology spillovers (firms being close learn more from each other) which increases the growth rate and therefore benefits the poor region. Martin and Ottaviano (1999) report that the net effect on welfare in the poor region depends in particular on the level of transaction costs, the importance of localised spillovers and on the inequality in capital endowments between the two regions. When transaction costs between the two regions are low, the positive effect of concentration will dominate because in this case, the fact that more goods have to be imported from the rich region is not very important. The net effect of concentration is also positive when spillovers are strong enough. Finally, if the poor region has initially little capital (or the inequality in capital endowments is high), then the positive effect will again dominate. This is because higher growth decreases profits of existing firms due to stronger competition: as the poor region has little capital the negative effect of lower profits is weak and the positive effect of stronger competition is important. Hence, the existence of localised spillovers, which induces a trade-off between regional equity and efficiency, may be an important factor in choosing the type of regional policies to implement.

The existence of localised spillovers, which induces a trade-off between equity and efficiency, may be an important factor in choosing the policies to implement.

To summarise, we have seen that a policy that reduces transaction costs between regions may improve welfare in the poor region even though it induces more spatial concentration and inequality. Moreover, regional policy that induces firms to move to the poorer location (for example through subsidies) may not be always welfare improving for the poor region, especially if spillovers are

2) See Combes and Lafourcade (1999) for a study that shows that the reduction of transaction costs in France has indeed led to more industrial concentration.

3) This is an overstatement because the infrastructure must be paid for. Implicitly, we assume that infrastructure projects are paid for by the rich region.

strong, inter-regional transaction costs are low and inequality in capital endowments between regions is high. If one believes that this characterises the European situation, then regional policies that focus on reducing transactions can be legitimate, but not for the reasons that are usually advanced by policy makers. Reducing transaction costs between regions will induce more concentration, but will weaken the detrimental effects of spatial concentration. It will increase efficiency and growth and therefore improve welfare in the poorest regions. However, if the ultimate goal of regional policies is not only to improve welfare, but also to decrease inequalities between European regions, then policies that focus on human capital (education and training) would be more appropriate.

Reducing transaction costs between regions will induce more concentration, but will weaken the detrimental effects of spatial concentration.

Box 1. A model of lowering transactions costs

The important assumption of this model (see Martin and Rogers, 1995, for further details) is that the manufacturing sector experiences increasing returns due to the fact that each firm requires a fixed amount of capital. Because capital is perfectly mobile, firms can choose to locate production in either a rich (r) or poor (p) region. K_r and K_p are the respective stocks of capital owned by the rich and the poor region and L_r and L_p are the number of immobile workers in those regions. We assume that $K_r > K_p$ and $L_r > L_p$. There are iceberg transaction costs τ on trade on manufacturing goods between the two regions and $\rho = \tau^{1-\sigma} < 1$, is a usual transformation (σ is the elasticity of substitution between goods in the monopolistic sector) so that an increase in ρ implies an improvement in infrastructure facilitating trade between the two regions. In equilibrium, the number of firms locating production in each region is:

$$(1) \quad n_r = \frac{K_r+K_p}{L_r+L_p}\left(\frac{L_r-L_p\cdot\rho}{1-\rho}\right);\ n_p = \frac{K_r+K_p}{L_r+L_p}\left(\frac{L_p-L_r\cdot\rho}{1-\rho}\right)$$

This equilibrium location is found by equating supplies and demands on goods markets and by an arbitrage condition that requires that the profit of a unit of capital be equal in both regions so that no relocation can be profitable. Equation (1) shows that more firms locate in the rich region than in the poor one. It is easy to check that an increase in ρ leads to relocation of firms from the poor to the rich region.

Welfare is given by the equations:

$$(2) \quad V_r = C\,(n_r + n_p\cdot\rho)^{\frac{\alpha}{\sigma-1}};\ V_p = C\,(n_p + n_r\cdot\rho)^{\frac{\alpha}{\sigma-1}}$$

where C is a constant and α is the share of manufacturing goods in the utility function. These equations just say that welfare depends on industrial location (n_r and n_p) and on transaction costs. Because ρ is less than 1 (some of the goods are lost in the process of transporting them between the two regions), welfare increases with the number of firms located in one's own region (as $n_r + n_p$ is constant and equal to the total capital stock, K_r+K_p).

Using equations (1) and (2), welfare in the poor region is:

$$(3) \quad V_p = C\left[\frac{L_p\,(K_r+K_p)}{L_r+L_p}\right]^{\frac{\alpha}{\sigma-1}}(1+\rho)^{\frac{\alpha}{\sigma-1}}$$

Hence, even though lower transaction costs (higher ρ) induces industrial relocation from the poor to the rich region, the net effect is always positive for welfare in the poor region.

4. Demand and supply effects of regional policies

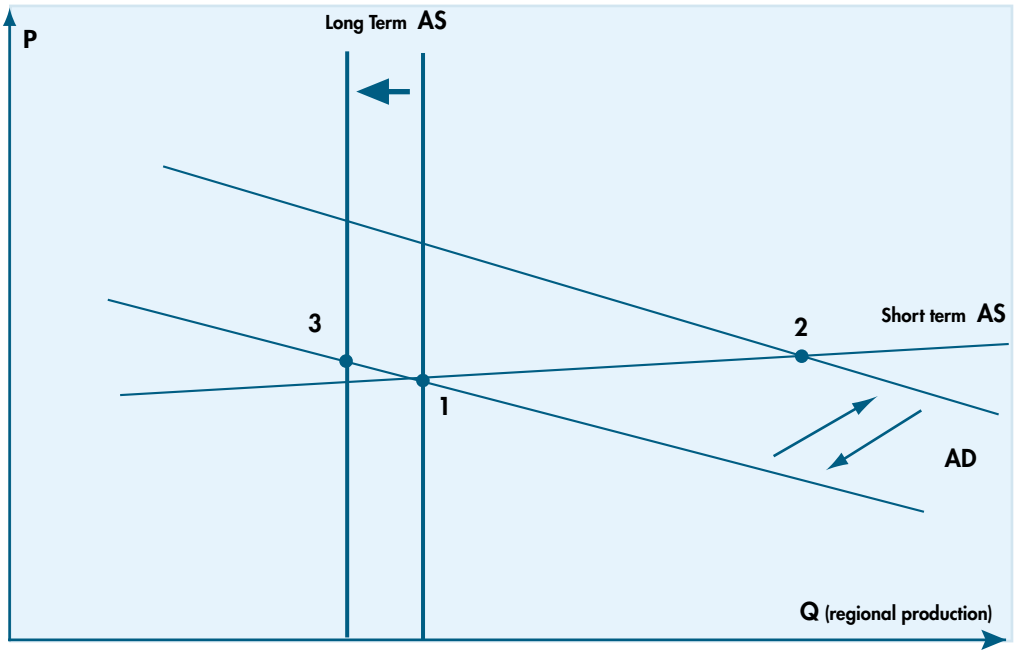
Regional policies that finance infrastructure projects have both demand and supply effects. The demand effects are mostly short-term effects whereas supply effects are more medium- to long-term. The demand effects of infrastructure projects such as roads, highways and other heavy infrastructures that are often financed through regional funds are quite clear. With a simple Keynesian framework in mind, it is easy to understand that this localised spending increases aggregate demand in the region. The effect is both direct and indirect through the Keynesian multiplier. The effect will be stronger the higher the unemployment rate and the lower the utilisation rate of factors of production such as capital in the region. Of course, the demand effects are not permanent, and once the infrastructure projects are over, the demand effects are reversed. However, they are certainly the most visible and the easiest to analyse and quantify. Indeed, the European Commission (1999) insists on these effects and use a Keynesian econometric model at the level of the country to quantify them. They find that for the period from 1989 to 1999 the contribution of the EU transfers has been to increase the average growth rate by a maximum of 1 percentage point (Greece and Portugal during the period from 1994 to 1999) and a minimum of 0.3 percentage points (Spain during the period from 1989 to 1993).

Regional policies that finance infrastructure projects have both demand and supply effects. The demand effects are mostly short-term.

These results are very difficult to interpret for two reasons. First, they measure at best the upper limit of the effect of regional policies. The reason is that they attribute any gap to the past trend of growth to the effect of regional policies. But we know that during this period where the integration process was very strong the convergence process was also very strong due to large private capital inflows to these countries (except for Greece). These inflows can well be explained in a simple neo-classical model with capital movements and convergence. Second, these studies look only at the impact on countries and not regions. This also is problematic because several studies (such as Neven and Gouyette, 1994) insist on the fact that convergence in Europe occurs between countries, but not between regions. De la Fuente and Vives (1995), for instance, building on the work of Esteban (1994), suggest that around *half the income inequality between the regions of the EU is accounted for by domestic inequality between regions within individual countries*. Thus, during the 1980s and 1990s per capita income differentials have been narrowing between countries, but widening between regions within individual countries (Martin, 1998). The EU studies provide very little information on the impact of regional policies on *regional* inequalities in Europe.

Furthermore, in the context of regional policies, it is more important to study the supply effects. As the earlier discussion has revealed, the long-term supply effects may be exactly opposite to the short-term demand effects. The dynamics of this evolution can be seen with a traditional aggregate demand/aggregate supply graph. In Figure 1 we illustrate this for a poor region that receives funds to finance infrastructure connections with a richer region. In the short-term, the aggregate supply curve in the poor region may be almost horizontal because of slack capacity and because some capital will move to the poor region when aggregate demand increases. Hence, the new infrastructure spending has a high positive short-run impact on output (output goes from point 1 to 2). However, this is temporary. The long-term effect is uncertain: The economic geography message is that the reduction of transaction costs may induce firms to concentrate in the rich region so that aggregate supply in the poor region is reduced (in which case output goes to point 3).

Figure 1.



The long-term supply effects may be exactly opposite to the short-term demand effects.

The distinction of demand and supply effects is also important for political economy reasons. Because the demand effects are short-term effects and they are most important for heavy infrastructure, and because the political horizon is also a short-term one, the strong bias in favour of heavy transport infrastructure in regional policies can be explained easily.

5. Conclusion

We have seen that public policies aimed at altering economic geography and regional development have multiple and sometimes contradictory impacts. The reason is that economic geography is key for many economic issues. It is important as a determinant of welfare, inequalities, productivity, growth and innovation. Moreover, economic geography is itself endogenous and public policies that influence transaction costs, innovation, or mobility of factors will change the location decisions of economic agents. Because of these potential self-reinforcing mechanisms at work, analysed earlier by Faini (1983) and Krugman (1991), it is also likely that regional policies have compound effects. If the dynamics of economic geography can be interpreted as one equilibrium loosing suddenly its stability at the benefit of an another equilibrium, this implies that regional policies will be most of the time useless, though extremely powerful in some rare circumstances. If agglomeration is due to a self-sustaining mechanism, through vertical linkages for example, then giving a small advantage to the poor region (for example through subsidies) will in no case alter the stability of the equilibrium. However, in the case where a new economic geography is in the process of being made, because of some drastic exogenous change in the economic environment or because new activities are created, then public policies may be the exogenous force that gives a key advantage to one region or to one stable equilibrium out of many stable and possible equilibria.

Regional policies may have very little impact most of the time, but a strong one in some very specific circumstances. This means the policy mistakes are going to be numerous.

It may be that the process of European integration is exactly such a moment where previously stable equilibria are redefined and where new equilibria emerge. The experience of call centres in France is also revealing. This is a rather new activity which by itself does not require to be close to a specific region. The city of Troyes in Champagne has been relatively successful to attract call centres by a specific training policy and a real estate policy aimed at favouring this activity. To a certain extent the example of Brittany with some information technologies linked to telecommunications is similar; training policy was again a key element. If, indeed, regional policies have very little impact most of the time and a strong one in some very specific circumstances, then policy mistakes are going to be numerous because the information requirement is too severe. This does not imply that regional policies have no use, but that these compound effects should be carefully integrated in the choices made.

References

Combes, P. and Lafourcade, M. (1999). "Transportation Costs Decline and Regional Inequalities: Evidence from France, 1978-1993, mimeo, CERAS.

De la Fuente, A. and Vives, X. (1995). "Infrastructure and Education as Instruments of Economic Policy: Evidence from Spain". *Economic Policy*, 20, pp. 11-40.

Esteban, J. (1994). "La Desigualdad interregional en Europe y en Espana: descripcion y analisis", in: *Crecimento y Convergencia Regional en Espana y en Europa*. *Vol. II*, Instituto de Analisi Economico, Barcelona.

European Commission. (1999). The European regions: Sixth periodic report on the socio-economic situation in the regions of the European Union. Official Publication Office, Luxembourg.

Faini, R. (1983). "Cumulative Process of Deindustrialization in an Open Region: The Case of Southern Italy, 1951-1973. *Journal of Development Economics*, 12, pp. 277-301.

Gabszewicz, J. and Thisse, J. (1986). "Spatial Competition and the Location of Firms", in: Gabszewicz, J., Thisse, J., Fujita M., and Schweizer, U. (eds), *Location Theory*, Harwood Academic Publishers.

Henderson, V., Kuncoro, A. and Turner, M. (1995) "Industrial Development in Cities". *Journal of Political Economy*, 103(5), pp. 1067-90.

Jaffe, A., Trajtenberg, M. and Henderson, R. (1993). "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citation". *Quarterly Journal of Economics*, 108, pp. 557-98.

Jacobs, J. (1969). *The Economy of Cities*, New York: Vintage.

Krugman, P. (1991). "Increasing Returns and Economic Geography". *Journal of Political Economy* , 99, pp. 483-499.

Martin, P. (1998). "Can Regional Policies Affect Growth and Geography in Europe?" *The World Economy*, 21, pp. 757-74.

Martin, P. (1999a). "Public Policies, Regional Inequalities and Growth". *Journal of Public Economics*, 73, pp. 85-105.

Martin, P. (1999b). "Are European Regional Policies Delivering ?", *EIB Papers*, Vol.4 No 2, pp. 10-23.

Martin, P. and Rogers, C. (1995). "Industrial Location and Public Infrastructure". *Journal of International Economics*, 39, pp. 335-51.

Martin P. and Ottaviano, G . (1996). Growth and Agglomeration. CEPR discussion paper 1529.

Martin P. and Ottaviano, G. (1999). "Growing Locations: Industry Location in a Model of Endogenous Growth". *European Economic Review*, 43, pp. 281-302.

Matsuyama, K. and Takahashi, T. (1998). "Self-Defeating Regional Concentration". *Review of Economic Studies*, 65, pp. 211-234.

Neven, D. and Gouyette, C. (1994). "Regional Convergence in the European Community". Université de Lausanne, Cahiers de Recherches Economiques N° 9311.

Scotchmer, S. and Thisse, J. (1992). "Space and Competition: A Puzzle". *Annals of Regional Science*, 26, pp. 269-286.

How productive are capital investments in Europe?

When the only tool you own is a hammer,
every problem begins to resemble a nail.
Abraham Maslow

Patrick Vanhoudt

Thomas Mathä

Bert Smid

1. Introduction

Economists agree at least on this: it is difficult to find evidence for, or merely to imagine any growth mechanism that does not work through the increase of a stock of capital in one way or another. From a more policy point of view - in particular in terms of the debate of economic development and convergence in standards of living - an important question then is when spending on investment is best done by the government itself, and when public funds should be used to support investment by the private sector.

A key concept in this issue is the degree of external benefits, or spillovers, of investment. These notions refer to the fact that sometimes a certain action by an economic agent results indirectly in productivity gains for others that cannot be completely captured by the principal investor in his price setting behaviour.

Under the standard textbook assumptions, property rights are complete and possible externalities are supposed not to be important enough to offset the decreasing marginal productivity of capital and labour. The market outcome will then be such that production factors are optimally allocated - the role for active policy intervention is in that case henceforth rather limited. However, if some investments yield external benefits, their full rate of return for society will exceed that of comparable yet perfectly excludable and rival assets. There may consequently be scope for public intervention from an efficiency (as opposed to redistributive) point of view (see the contributions of Martin and Thisse, this volume). An important policy instrument here is public investment.

Substantial positive external benefits associated with physical capital provide indeed a basic rationale for public investment. The intuition behind this is as follows. Private agents will be hesitant to invest in those goods for which they cannot completely exclude others from benefiting from them - by doing so the investor would give a free input to competitors. The inability to charge a market price for this indirect service implies that the agent will not be able to reap the full profit of his investment. Thus, the larger the external benefits, the less likely private agents are to invest in the considered goods. Public finance theory stresses therefore that the government should supply such capital.

Public investment is not only motivated because of possible external effects. One can also make a plausible theoretical case in favour of public investment along the lines of market failures. For instance, long-term unemployed that want to invest in education may find it difficult to borrow against their future productivity to finance the investment. Reasons for this are uncertainty and asymmetric information, so that providing public schools solve, at least in part, some of these market failures.

Yet, even though publicly provided goods may be characterised by positive externalities, they have to be financed with tax money. Levying taxes presumably reduces the private savings and investment behaviour, so that public investment may crowd out private investment. However, some public goods may at the same time make private investments more attractive and productive. From

Patrick Vanhoudt is an economist at the Chief Economist's Department. Thomas Mathä and Bert Smid are with the University of East Anglia Norwich (UK) and the University of Groningen (the Netherlands), respectively. Both visited the EIB during this research.

An important question is when spending on investment is best done by the government, and when public funds should be used to support private investment.

a theoretical point of view, the net effect is not evident and empirical investigations will have to provide an answer.

This leads to a number of questions. Is there any evidence of external benefits? Do public and private capital investments equally contribute to increases in the standard of living? Are there differences between productiveness at the national and regional level? Although considerable effort has been undertaken to investigate these questions for the United States and different specifications have been applied to a cross-section of both developed and less developed countries, they have not extensively been studied for economies within the European Union. The current paper tries to bridge that gap.

We have organised the remainder of the paper as follows. In the next section we will review some methodological issues and different approaches that have been suggested in the economic literature to investigate the issue. Section 3 thereafter reports the empirical regularities for Europe, both on the national and the regional level. Section 4 summarises and concludes.

2. Methodological issues and approaches

2.1 Comparing measures of standard of living

When talking about benefits for society, it is good to clarify what one means by that. A natural interpretation would be to argue that economic growth translates into an increase in the quantity and quality of available goods through an improvement of efficiency by which the production factors are employed. In that way people can buy more and improved goods in the next period with the income they earn from participating in the production process. Macroeconomists therefore take the per capita income as a fairly suitable proxy for the standard of living (1).

Comparing levels of standards of living among countries or regions by using these indicators can nonetheless be difficult a task, and one has to be rather careful in doing so. The obvious method would be to value each country's production of final goods and services at domestic prices, to apply the GDP deflator in order to express these numbers in real terms, and thereafter to convert these figures into a common monetary currency using the relevant exchange rates.

In theory, exchange rates should adjust through the action of the market so that the local currency prices of a group of identical goods and services indeed represent equivalent value in every nation. In practice such adjustments can, however, lag far behind rapidly changing economic circumstances. There are consequently often large and systematic departures of exchange rates from the "purchasing power parities" (PPP) - a given amount of euro will buy different bundles of goods in different countries. International comparisons based on market exchange rates can hence greatly over- or understate the purchasing power value of a nation's real economic activity.

An alternative approach is therefore based on estimates of the purchasing power of different currencies, rather than their market exchange rates. The construction of national accounts in purchasing power parities that are comparable across space and time is not easy though. It relies on obtaining price data for a wide range of goods, and building suitable aggregation procedures

1) While income per employed person is used as a measure for average labour productivity. However, since age-structures may differ substantially between economies, a more appropriate indicator for cross-economy comparisons is the income per person of working age, where the working age is defined as 15 up to 64.

to obtain a regional or national PPP adjusted GDP. A blend of extrapolated and regression-based numbers for this concept is available from different sources such as Eurostat, the OECD, the World Bank and the so-called Penn World Tables. However, these data are primarily designed to compare countries within the same year, and should strictly spoken not be used in time-series comparisons (2). Doing so might very well result in an intuitively attractive rank ordering of countries, it yields, however, implausible results for annual average growth rates. This is often overlooked when analysing determinants of economic growth in empirical exercises (3).

The current paper focuses on the productivity of public and private investment in European countries and regions.

Thus, we must return to the first option to employ GDPs expressed in prices and exchange rates of a particular year. Although in this approach the growth rates reflect more consistently the changes in the volume of real output, one has to keep in mind that real exchange rate appreciations over time are not captured here. As an illustration Table 1 compares the OECD PPP data with this major alternative. The table suggests that the major drawback of the methodology we propose is that it significantly overvalues the standard of living in Sweden and Finland, and ranks Belgium and the Netherlands too low.

Table 1. Comparison of real and PPP GDP per person of working age. EU-15 countries, 1990-97.

Countries are ranked according to their situation in 1997							
	Exchange Rates and Prices of 1990, USD				PPPs USD		
	Income per worker in		Average Annual Growth		Income per worker in		Average Annual Growth
Country	1990	1997	1990-97	Country	1990	1997	1990-97
Luxembourg	39.145	51.015	3.86%	Luxembourg	32.917	49.558	6.02%
Denmark	38.510	45.112	2.29%	Denmark	25.383	37.844	5.87%
Sweden	41.653	42.946	0.44%	Belgium	24.888	35.313	5.13%
Finland	40.168	42.089	0.67%	Austria	24.818	34.127	4.66%
France	31.980	34.104	0.92%	France	26.329	32.551	3.08%
Austria	30.638	33.573	1.32%	Netherlands	23.154	32.481	4.95%
Germany[a]	30.660	32.692	1.08%	Germany[a]	24.786	32.271	4.50%
Belgium	29.388	32.575	1.48%	Sweden	26.414	32.034	2.79%
Netherlands	27.528	31.622	2.00%	Ireland	18.550	31.814	8.01%
Ireland	21.165	30.969	5.59%	UK	24.256	31.512	3.81%
Italy	28.135	30.042	0.94%	Italy	23.720	31.089	3.94%
UK	25.942	28.691	1.45%	Finland	24.046	30.780	3.59%
Spain	19.031	20.772	1.26%	Spain	17.714	23.426	4.07%
Greece	12.264	13.060	0.90%	Portugal	14.537	21.398	5.68%
Portugal	10.579	11.923	1.72%	Greece	13.800	20.451	5.78%

Source: GDP: OECD National Accounts. Number of people of working age: AMECO.
[a]: 1991-97

2) To quote the World Bank's World Resources (1996, p. 171): "Although considerable effort has been made to standardize economic data according to the UN system of National Accounts, care should be taken in interpreting them. Intercountry and intertemporal comparisons using economic data involve complicated technical problems that are not easily resolved; therefore, readers are urged to read these data as characterizing major differences between economies rather than as precise, quantitative measurements".

3) e.g. Quah 1993, 1996, 1997a,b. The regional PPP data by Eurostat (Nuts 2 and 3 level) show for instance typical average annual growth rates of per capita GDP of 8 percent and much more. Some researchers do not seem to take this problem into account, nor that the definition of the PPS has changed with every enlargement of the EU.

2.2 Assessing the impact of input factors on the standard of living: What does economic theory suggest?

To produce their output, economies have access to similar types of inputs: a certain amount of the population - the active labour force which represents a mixture of skills or human capital - and the stock of physical capital. All of them are combined with a particular efficiency. Economic growth consequently results from continuous increases in these variables.

A natural question henceforth is whether it is efficient to reduce private saving through taxation and to inject those means as public investment to improve regional development. Although many policy maker in the 1950s believed so, the enthusiasm for active interventions among economists waned, as the neo-classical paradigm became more prevalent during the sixties and beyond. It has somewhat revived however, with the recent arrival of theories on endogenous growth and "new" economic geography.

How can one think about the issue of making backward economies catch-up? Growth theories provide useful frameworks here. A key concept and starting point in these theories is always the production function. In its most simple form, it is assumed that there exists a link between the total amount of goods and services that an economy can produce per unit of time (Y), the available input factors such as domestic physical capital (K) and labour (L), and their total factor productivity or efficiency (A).

$$(1) \quad Y_t = f(L_t, K_t, A_t)$$

Economic growth results from continuous increases in the stocks of capital and efficiency with which these stocks are used.

In this respect, a frequently used specific form for this function is the Cobb-Douglas type:

$$(2) \quad Y_t = A_t \cdot K_t^{\alpha_1} \cdot L_t^{\alpha_2}$$

which in terms of per capita income reads:

$$(3) \quad y_t \equiv \frac{Y_t}{L_t} = A_t \cdot \left[\frac{K_t}{L_t}\right]^{\alpha_1} \cdot L_t^{\alpha_1+\alpha_2-1} = A_t \cdot k_t^{\alpha_1} \cdot L_t^{\alpha_1+\alpha_2-1}$$

Of particular interest here are the coefficients α_1 and α_2, which, in fact, indicate how responsive the standard of living is to changes in the input factors. When they sum up to one, production takes place under constant returns to scale. That is, doubling every factor of production will result in twice the amount of total output, and income per worker remains unchanged. If they are also each strictly positive and smaller than one, capital and labour are characterised by diminishing marginal productivity. Investing a constant fraction of output every period will then consequently result in decreasing additions to output over time. In that way, a boost in the investment share causes a jump in income, but the growth effect will fade out. These are the well-known standard neo-classical predictions. Moreover, when markets are competitive so that each factor is valued according to its productivity, α_1 and α_2 represent the shares of GDP which are made as a payment to the capital and wage bill, respectively. According to the national accounts, the total wage bill in economies is typically about two-thirds of GDP while approximately one-third of it goes to the remuneration for capital. Thus, $\alpha_1 \approx 0.3$ and $\alpha_2 \approx 0.7$.

Substantial positive external benefits associated with physical capital would provide a basic rationale for government intervention.

Now let us turn back to the issue of external benefits. If externalities exist for physical capital, investment by one agent will benefit productivity and output for other agents as well. Thus, if one were able to double both capital and labour, total output would in that case more than double. The sum of α_1 and α_2 will then exceed one. Increasing returns to scale provide a growth bonus for the economy and are a necessary condition for sustained growth in the absence of (exogenous) efficiency increases (see the contribution of de la Fuente, this volume).

It is the merit of "new" growth and geography theory of providing economic underpinned rationale for possible external benefits. Learning-by-doing is one example (see Arrow, 1962; Romer, 1986). Overcoming problems and improving equipment accordingly results in technological change, new investment, and productivity gains. These investments will in turn lead to new experiences and new solutions, and will end up in the realisation of further increases in efficiency. If the breakthroughs are important, other firms will follow soon so that the initial investment benefits the whole sector. It is clear that the learning-effect will be more substantial if the human capital carried by each worker is higher.

Thus, external benefits do not solely appear as a side effect from accumulating *physical* capital. *Human* capital accumulation is another factor which has been put forward to explain sustained growth (Uzawa, 1964; Lucas, 1988; Stokey, 1988) (4). On the one hand, human capital must have some effect that is internalised otherwise no one would spend valuable resources on schooling, training, business seminars, etc. On the other, people learn from one another so that the total gains of investments in human capital cannot be completely captured by the agent investing in it.

2.3 Quantifying the contribution of input factors to the standard of living: Empirical evidence

Research on economic growth and convergence has proceeded through several stages, each of them characterised by a specific empirical methodology. A very basic one is the *accounting approach* to economic growth (5), which simply checks whether changes in total output are identical to the sum of the changes in the stocks of capital and labour weighted with their factor shares as reported in the national accounts. The major contribution from these kind of exercises is that a substantial part (often over 50%) of economic growth remains unexplained. This residual is referred to as the Solow residual, total factor productivity (TFP) growth, or technological change. Growth accountants have, however, made little progress in answering what economic variables correlate with total factor productivity. As such they have not brought us particularly further in understanding why TFP growth rates may differ across time and space (6).

Also, by imposing the elasticities of output to inputs, rather than estimating them, one typically presumes that there are no external benefits, that production takes place under constant returns to scale and that there is perfect competition. In fact, from a theoretical point of view it is very difficult to motivate the use of this technique to assess the importance of public capital: as there is no (market driven) remuneration for this sort of capital, its factor share - the crucial weighting factor for public

4) Externalities may also arise from an activity such as research and development (R&D). See Romer, 1990.

5) See Solow, 1957; Denisson, 1967; Maddison, 1982, 1989, 1991.

6) An interesting application, however, is a contribution by Coe and Helpman, 1995. They compute TFP in the traditional way, and then regress it on measures for domestic and international R&D stocks as suggested by new growth theory. The authors find that variation in total factor productivity growth in OECD countries can be attributed to variations in these stocks, and they find strong evidence of international knowledge spillovers.

capital in the accounting exercise - would be zero. Consequently, it cannot make a contribution to growth of the standard of living, by definition. In general it is therefore difficult to extract any policy recommendations from growth accounting results.

Aschauer introduced in 1989 the obvious, but until then neglected, notion that the stocks of public infrastructures and private capital may be a key to explaining the level of national output in the private sector. We can illustrate his methodology by distinguishing between public (K_g) and private (K_p) infrastructure in the earlier shown type of production function:

$$(4) \quad Y_{P_t} = A_{P_t} \cdot K_{P_t}^{\alpha_1} \cdot L_{P_t}^{1-\alpha_1} \cdot K_{g_t}^{\alpha_3}$$

where Y now refers to a measure of real output in the private sector, rather than in the whole economy, and we impose that all markets in the private sector are perfectly competitive and free of external benefits ($\alpha_2=1-\alpha_1$). Dividing both sides by the total stock of private capital yields an expression for the amount of output produced by each unit of private capital - the private capital productivity. After taking logarithms one obtains:

$$(5) \quad \ln\left[\frac{Y_{P_t}}{K_{P_t}}\right] = \ln\left[A_{P_t}\right] + (1-\alpha_1) \ln\left[\frac{L_{P_t}}{K_{P_t}}\right] + \alpha_3 \ln\left[K_{g_t}\right].$$

Thus, output per unit of capital relates positive to the labour-to-capital ratio. At the same time and other things being equal will economies with an extensive stock of public infrastructure experience a high capital productivity in the private sector, at least in theory. If the coefficient α_3 is strictly positive, one can conclude that public infrastructure result in external benefits (7).

Variations to this basic equation have subsequently been taken to the data in a number of studies, mainly focusing on the US. Interestingly, in virtually all cases the level of public capital is found to be significantly productive. Results from *Aschauer type of regressions* have, however, been widely criticised as being implausible because of their sheer magnitude: the reported production elasticities imply a stratospheric marginal product of government capital of over 100% per annum or more.

Table 2. Some results from Aschauer type of regressions

Authors	Output Elasticity of *Private* Capital	Output Elasticity of *Public* Capital	Level of aggregation
Aschauer, 1989	0.56	0.38	US, national level, time series
Munnell, 1990a	0.62 - 0.64	0.31 - 0.37	US, national level, time series
Munnell, 1990b	0.31	0.15	US, regions (States), panel data
Berndt and Hansson, 1991	0.37 - 0.66	0.68 - 1.60	Sweden, time series
Garcia-Mila and McGuire, 1992	0.37 - 0.45	0.10 - 0.03	US, regions (States), panel data
Holtz-Eakin, 1994	0.11 - 0.50	-0.12 - 0.20	US, regions (States), panel data
Karras, 1997	0.12 - 0.23	0.15 - 0.18	20 OECD countries, national level, panel data

7) Aschauer also provides a specification in which he controls for congestion effects, and an equation in which total factor productivity is related to the stock of public infrastructure.

As a result controversy arose about the method of estimation and about the interpretation of the elasticities. Three major critiques have appeared in the literature. First, the time series estimates that show a positive and significant effect of the public capital stock on private sector productivity do so because of a statistical fallacy: they result from a "spurious regression". Trends in time series - in this case output and the stock of public infrastructure - may exhibit an apparent statistical relationship, even though no economic relationship between them exists. A solution to this problem - namely removing the trends by taking first differences - mostly yield results showing that public capital's effect is rather small and in general not statistically significant. Second, it is argued that the wide range of estimates reported in various studies renders the coefficients suspect. Finally, the direction of causality is doubtful: causation may not run from the stock of public capital to output, but rather in the opposite way. The empirical linkage between output and public capital based on this approach has therefore been discredited and said to be fragile at best.

Other than that, it is hard to employ the Aschauer approach if the focus is to assess the speed of convergence to the long-run equilibrium. Neo-classical growth models offer more appropriate testable equations in this respect. In fact, they take the production function technique one step further in the sense that the stock of capital is determined endogenously. In its simplest form, it is assumed that in every period of time, a fraction s_k of total output is forgone, and re-injected in the economy as physical capital. Production factors are, however, characterised by diminishing marginal productivity. Investing a constant fraction of output over time henceforth leads to smaller additions as time evolves. Moreover, production is assumed to take place under constant returns to scale ($\alpha_1+\alpha_2=1$). In other words, the (testable) hypothesis is that possible externalities are not sufficiently large to offset the decreasing marginal products. Economies will therefore converge towards a per capita income that grows solely because of (exogenous) factors that influence the productivity of the inputs, such as technological change or improvements in the management system, etc.

The neo-classical convergence property has been the subject of many debates. In the late 1980s it was studied under the implicit assumption that all countries would converge towards the same long-run *levels* of per capita income (Baumol, 1986, is a good example). This became later known as *absolute* convergence. Barro and Sala-i-Martin, 1992, and Mankiw, Romer and Weil, 1992, introduced, and rigorously defined the concept of *conditional* convergence. It was emphasised that growth theory did not imply *identical* long-run per capita incomes for all countries, and that one has to control for factors that influence this long-run level of income other than the initial condition, such as the investment share and population growth. Yet, when empirically tested, the basic model that only included physical capital did not yield theory-consistent results for the output elasticities, α_1 and α_2. It was the merit of Mankiw, Romer and Weil, 1992, to point out that the data behave consistent in a neo-classical manner only if one includes human capital in the production function. In that case, α_1 is indeed estimated to be about one-third in a comprehensive sample of developed and less developed countries - an appealing value for it is comparable to capital's share in GDP (8).

8) The model nonetheless still performs rather poor in explaining variations in standards of living in OECD countries. Nonneman and Vanhoudt, 1996, argue that this might be due to the fact that technological knowledge capital is important in such highly developed countries, yet it is not included as a factor of production in the Mankiw et al., *paper. The authors consequently augment the model with investment in research and development and show that this considerably improves the explanatory power of the regressions (measured by the adjusted* R^2*).*

Box 1. Research on EU convergence and the productivity of capital

In its "Periodic Reports" the Commission provides an indepth description of the economic performance in the EU regions, as well as of the role of the EU structural actions in assisted areas. In addition, the "Single Market Review - Regional Growth and Convergence: Aggregate and Regional Impact" attempted to analyse the sources of differences in standard of living within European countries. The latter research has assessed the impact of total investment in physical and human capital at the national level by means of traditional cross-section growth regressions. A major problem in doing cross-country growth regressions for the EU is, unfortunately, the low number of observations.

However, no serious attempt so far has been made to address the productivity of different types of physical capital for European economies. Following Neven and Gouyette, 1995, and Sala-i-Martin, 1996, research at the European regional level has remained rather limited to testing for conditional convergence based on panel data regressions that control for differences in countries' economic situation by including dummy variables. When it comes to public infrastructure, sometimes an ad hoc and debated index (e.g. the Biehl index) for public infrastructure (see for instance Capron, 1997), or physical indicators for infrastructure - such as motorways per capita, air freight per capita, or air passengers per capita - are included. A recent unpublished study for the Commission indicates, however, that infrastructure indicators explain only a very small part (about 8%) of the variation in standards of living (Pinelli, 1998). Based on the latter study the 6th Periodic Report mentions that four other factors are possibly linked with regional differences in GDP measures: 1) the structure of the economic activity, 2) the extent of innovative activity, 3) regional accessibility and 4) the skills of the work force.

The main advantage of employing specifications from a model is that such implicit hypotheses can be tested. However, researchers have also tried to explain growth performances in often highly *ad hoc* ways. In these *Barro-type* of regressions - named after the methodology proposed by Barro, 1991 - growth rates are regressed on whatever data or indicators are available (9). Although not rigorous from a theoretical point of view, these kind of exercises can be useful in discovering robust correlations and stylised facts when combined with a sensitivity analysis.

The latter approach has been employed to investigate issues such as the social rate of returns to various types of investment (equipment, infrastructure, etc.) (10). It has also been used to examine the productivity of public capital in various samples of countries (11). In these studies the main consensus so far seems to be that the effect of private investment on nation-wide growth is robustly positive while the one for public investment highly depends on the sample used. For instance, Barro, 1991, finds in a cross-section of 76 countries that public and private investment have similar effects on growth. Easterly and Rebelo, 1993, report that there is an important role for infrastructure capital, especially transportation and communication, in a broader sample of one hundred countries. Based on the US experience, Holtz-Eakin, 1994, argues however, that such a finding

9) "I just ran two million regressions" by Sala-i-Martin, 1997, illustrates this pure empiricism very well.
10) See e.g. De Long and Summers, 1991, 1992, and Auerbach, Hassett and Oliner, 1994.
11) See for instance Easterly and Rebelo, 1993, Crihfield and Panggabean, 1995. Aschauer, 1998, and Vanhoudt, 1999, derive their specifications from the neo-classical theory. A good review of growth evidence can be found in Temple, 1999. Gramlich, 1994, provides a review of the literature on infrastructure investment.

crucially depends on the estimation technique used and that the positive effect from public capital largely disappears when one allows for unobserved, state specific characteristics. In addition, Hulten, 1996, finds little or no support in the US data for an important effect of public capital on productivity after controlling for the efficiency use of public capital.

EU related research so far has failed to take into account the effect of different forms of capital on real convergence.

The cross-section methodology on which these analyses are based can clearly not be subject to the critique of spurious regression results due to non-stationary variables. However, there are other caveats. Although simultaneity issues also plagued the time-series literature, an additional important shortcoming in the cross-section literature is the assumption of strict homogeneity of the technology shift parameter in economies' production functions (i.e. the parameter A in equation (3)). Panel data techniques that control for fixed or random effects provide a useful tool and solution in this respect (12).

In the empirical part of this study we will continue to take the neo-classical paradigm as the reference, and derive empirically testable equations from an extended Solow model (see Box 1). Our research will be focused on the EU-15 countries and regions. Although the European Commission has put considerable effort in investigating the issue of convergence in Europe (see Box 2), EU related research so far has failed to take into account the effect of different forms of capital on real convergence at the *national* level, nor has an analysis of the impact of investment at the *regional* level been presented. With the current paper we would like to pick up these issues and contribute to the debate on this subject.

12) See Islam, 1995. The coefficient for human capital in this study is however estimated with a rather implausible yet statistically not significant negative sign for all the samples (98 non-oil countries, 75 countries with "high quality data", and 22 OECD countries). See Lee et al., *1998, for a comment and Islam, 1998, for a reply.*

Box 2. Deriving the extended neo-classical model

For simplicity we assume that countries are specialised in producing a domestically manufactured good (or composite of goods), symbolised by Y. We presume that output of this product (or composite of products) at time t is generated according to the following Cobb-Douglas production function:

$$(6)\quad Y_t = \left[A_t \cdot L_t\right]^{\left(1-\sum_i \alpha_i\right)} \cdot \prod_i K_{i_t}^{\alpha_i} \qquad \sum \alpha_i < 1$$

with L being employment and K_i capital of type i. In the absence of market imperfections, the α_i's reflect the respective shares of the production factor in total output (Euler's theorem). Finally, A is a shift variable representing other, currently unspecified, economic 'environment' conditions that may be important to the production process. Traditionally this variable has been interpreted as an indicator of the state of the art of the technology, management or government efficiency etc.

Equation (6) implies that income per (effective) worker reads:

$$(7)\quad y_t = \prod_i K_{i_t}^{\alpha_i} .$$

Lower case letters in expressions throughout the paper denote variables per unit of effective worker: $y = Y/\left[AL\right]$ and $k_i = K_i/\left[AL\right]$. Further, the rates of growth of labour and technology are assumed to be out of the direct influence of the policy maker - they evolve at an exogenous rate of n and x respectively.

Every period in time society forgoes an amount of consumption and chooses to re-inject it in the economy as capital formation. Let us denote the fraction of output that is invested in capital component i by s_i. By definition, the difference (per unit of time) of the stock of capital of type i equals new investment minus total deprecation: $dK_i/dt = s_i.Y - \delta_i.K_i$ with δ_i being the capital specific rate of depreciation. The laws of motion that govern the evolutions of the different stocks of capital per effective worker accordingly are:

$$(8)\quad \frac{dk_i}{dt} = s_i \cdot y - (n+x+\delta_i) \qquad \forall i = 1 \ldots m$$

which yields a system of m differential equations. Long-run equilibria or "steady state values" for the relative capital stocks can be found by substituting the production function (7) into the set of differential equations. After taking logs, the resulting log-linear system can be easily solved. To be more precise (see e.g. Vanhoudt, 1999, for formal proof), the closed form solution for these steady state values are:

$$(9)\quad k_{i*} = \left[\left(\frac{s_i}{n+x+\delta_i} \right)^{1-\sum_{\substack{j=1 \\ j \neq i}}^{m} \alpha_j} \cdot \prod_{\substack{r=1 \\ r \neq i}}^{m} \left(\frac{s_r}{n+x+\delta_r} \right)^{\alpha_r} \right]^{\frac{1}{1-\sum_{j=1}^{m} \alpha_j}}$$

It is noteworthy that the equilibrium value for capital component i is in this framework not only determined by its own investment share, but also by the shares invested in all the other components. Injecting capital in one particular sector of the economy may henceforth lead to capital formation by agents in others sectors as a side effect. Equation (8) shows why: an increase in a particular capital component will boost income per worker (y), so that all other steady state capital-to-worker ratios will

start to change as well. This dynamic process will continue until the economy has reached its new equilibrium values.

Replacing the k_is in equation (7) by their long-run values reported in (9), results in an expression for the long run output per worker. After taking logs, this expression reads:

$$(10)\quad \ln\left[\frac{Y}{L}\right]_* = \ln\left[A_0\right] + x\cdot t + \sum_{i=1}^{m} \frac{\alpha_i}{1-\sum_{j=1}^{m}\alpha_j}\cdot\left(\ln\left[s_i\right] - \ln\left[n+x+\delta_i\right]\right).$$

This equation says that the long-term standard of living in a particular economy will be higher, the higher the investment shares in the different types of capital, other things being equal.

The estimated values for the α_is can thereafter be applied to compute steady-state marginal products for each capital component, which are indicators for the social rates of return. In this kind of model the latter are equal to:

$$(11)\quad MP_i = \alpha_i \cdot \frac{n+x+\delta_i}{s_i}.$$

The assumption that countries are in their steady state may be too stringent and not appropriate for some countries. However, it is possible to derive a "dynamic" specification that holds regardless of the deviation from the long-run equilibrium and is henceforth applicable to all economies. By using a log-linearization of the growth rate, evaluated at the steady state, it is possible to derive a growth equation (see Barro and Sala-i-Martin, 1992, 1995, or Mankiw, Romer and Weil, 1992):

The growth equation

$$(12)\quad \ln\frac{(Y/L)_t}{(Y/L)_0} = (1-e^{-\lambda t})\left(\ln\left[\frac{Y}{L}\right]_* - \ln\left[\frac{Y}{L}\right]_0\right)$$

in which $\ln\left[Y/L\right]_*$ can be replaced by equation (10). In this equation, the left-hand side variable is the cumulative growth rate over a period of t years, and the parameter λ indicates the speed at which economies converge (conditionally) towards *their* individual long-run equilibrium.

Note that for one particular production factor - human capital - investment shares over a long period of time are not readily available. We can overcome this lack of data by including the accumulated end of period stock of human capital per worker in the analysis (see also Islam, 1995) and presume that this stock is the equilibrium one, or that deviations from the equilibrium value are at least random. Equation (10) will in that case read:

The level equation

$$(13)\quad \ln\left[\frac{Y}{L}\right]_* = \ln\left[A_0\right] + x\cdot t + \sum_{\substack{i=1\\ i\neq n}}^{m} \frac{\alpha_i}{1-\sum_{\substack{j=1\\ j\neq n}}\alpha_j}\cdot\left(\ln\left[s_i\right] - \ln\left[n+x+\delta_i\right]\right) + \frac{\alpha_n}{1-\sum_{\substack{j=1\\ j\neq n}}\alpha_j}\cdot \ln\left[h_*\right]$$

This approach differs in three ways from Aschauer's. Firstly, variation in economic performances are here no longer explained in terms of differences in capital stocks, but rather in terms of differences in

fundamentals that drive the evolution of the stocks. These fundamentals are of direct relevance for policy. Secondly, the methodology allows us to distillate a speed of convergence (λ). Knowledge on the magnitude of this speed is quite important since the scope for policy may be more substantial if convergence is a slow process. Thirdly, growth rates and investment shares are usually integrated of order zero, so that it is unlikely that the obtained estimations are subject to the spurious regression critique in a time series or panel data set-up.

Nonetheless some caveats remain. For instance, the neo-classical and Aschauer models typically depart from a closed economy - it is implicitly assumed that there are no systematic current account surpluses or deficits. This may not be a very appropriate theoretical simplification, especially when considering regions as the unit of analysis. Secondly, the underlying production function presumes that public and private capital are close substitutes. Although economic theory provides no guidelines on this issue, one can equally well argue that publicly provided goods are part of the 'environmental' variable (A). Finally, in the neo-classical models there is no room for unemployment.

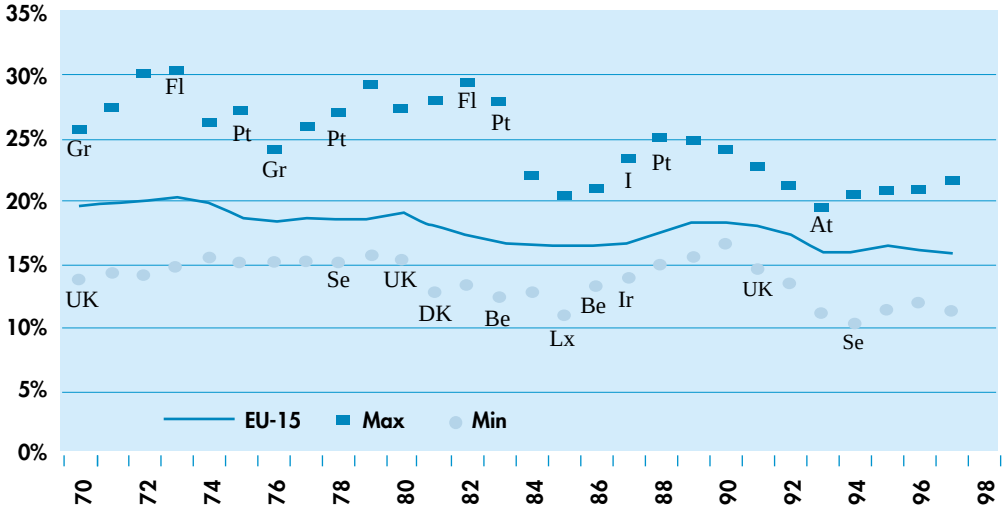
3. Empirical implementation

3.1 A closer look at the European data

Before we assess the impact of the different fundamentals on economic performance, it is useful to define the key data precisely, and to provide some descriptive statistics on them.

Figure 1 presents the private investment share, i.e. gross fixed capital formation done by households and the business sector relative to GDP. This variable refers to investments in market sector activities, and captures as such also capital outlays by quasi-corporate enterprises. On average there has been a slight decline in the private investment share. While the EU-15 average mounted to 20% of EU-15 GDP in 1970 it has reduced to about 16% of EU-15 GDP today. The variation around the average is however, substantial. Currently, the business sector with the highest investment relative to its country's GDP is located in Austria while the country with the lowest private investment share is Sweden. The difference is roughly two to one.

Figure 1. Private investment relative to GDP



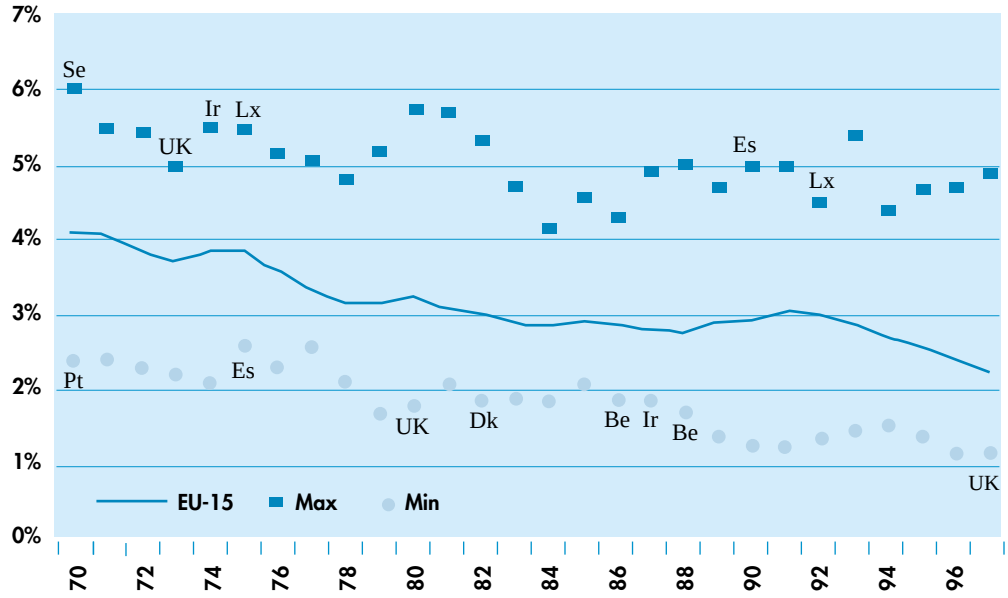
Note: the country labels as maximum or minimum remain valid over time until another country takes over.

By public investment, on the other hand, we mean capital formation by the general government. The majority of these outlays go mainly to four categories: the most important one is the provision and maintenance of transport infrastructure (in particular roads and bridges) which accounts for roughly a third, about 10 to 15% goes to education and health (school buildings, hospitals, etc.), approximately 10% is devoted to the provision of housing and community amenities, and another 10% is taken up by general public services. The remaining part is split amongst defence, public order and safety, recreational, cultural and religious affairs, etc.

Public investment is typically much lower than capital formation by the business sector. Figure 2 shows that the Union's average at the moment is approximately 2% of the EU-15 GDP, and has shown a continuous downward trend over the last decades. This negative trend has accelerated since the early 1990s due to the fiscal constraints of the Maastricht criteria.

The spread around this average is, again, substantial. While the UK's public investment share these days is only half the EU average, the Luxembourgish one is about two and a half times as high.

Figure 2. Public investment relative to GDP



Note: the country labels as maximum or minimum remain valid over time until another country takes over

13) These administratively defined regions vary substantially in size. In terms of surface, the largest region in our sample is Castilla y León (Es) with 94 010 km², the smallest one is Bremen (De) with only 404 km². The average region accounts for 21 930 km². In terms of population, the largest region is Nordrhein-Westfalen (De) with roughly 17.8 m inhabitants while Valle d'Aosta (It) has a population of only slightly over 0.1 m. The average region in the sample has approximately 2.9 m people. As for population density, Castilla-la Mancha (Es) has about 25 people per km² and is as such the least densely populated region in the sample. Hamburg (De), on the other hand, has some 2200 people living on one squared kilometre. The average population density is roughly 250 persons per km².

At the regional level data are not so readily available.

At the regional level data are not so readily available and comparable. After carefully examining and selecting the figures we were able to compose a sample of 78 administrative regions among eight countries: Denmark (Nuts 1), Luxembourg (Nuts 1), France (Nuts 2), Germany (Nuts 1), Ireland (Nuts 1), Italy (Nuts 2), The Netherlands (Nuts 2) and Spain (Nuts 2)(13). Public investment at the regional level refers to gross fixed capital formation in the non-market sector as reported by Eurostat. As becomes clear from Figure 3, poorer regions have shown a tendency to receive more public investment than richer ones. The two outliers are Valle d'Aosta, the smallest and less dense populated area situated in the north-west of Italy, and Zeeland, a south-western region in The Netherlands. Geographical obstacles characterise both. While Valle d'Aosta is a mountainous area, Zeeland was exposed to the threat of flooding from the North Sea. The governments have consequently carried out and maintained major infrastructure works to overcome these inconveniences. Over the considered time span, a stretch of about 50 kilometres of highway towards the Mont Blanc passage was built in Valle d'Aosta, consisting purely of bridges and tunnels. Zeeland hosts what has been the most expensive environmental investment in the world: an ingenious system to prevent the region from flooding, known as the Delta-water works.

At the same time, there is a clear positive correlation between the regional standard of living and the level of human capital per worker - see Figure 4.

Figure 3. Regional public investment vs. initial GDP per worker

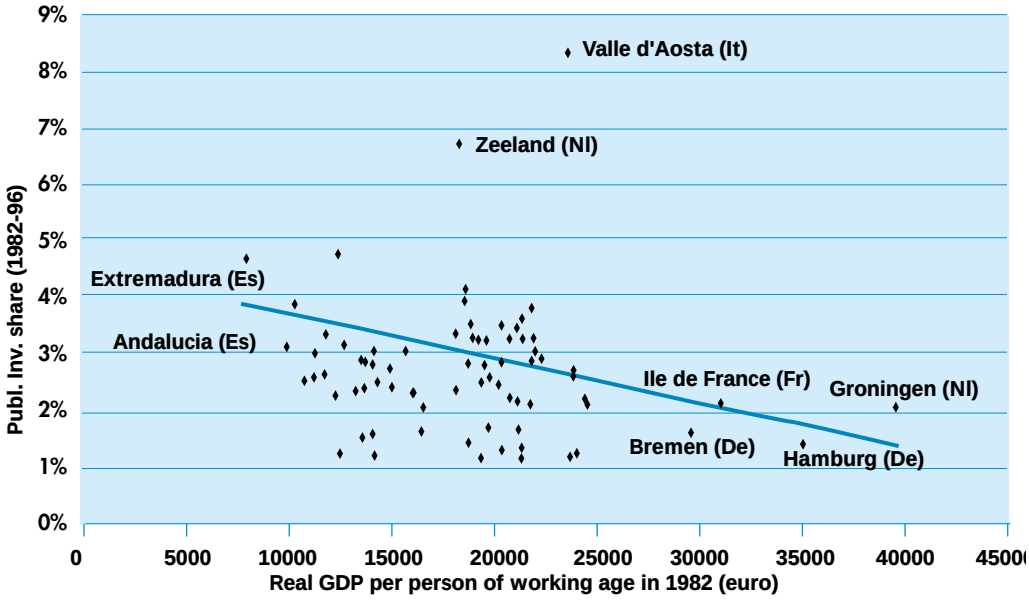
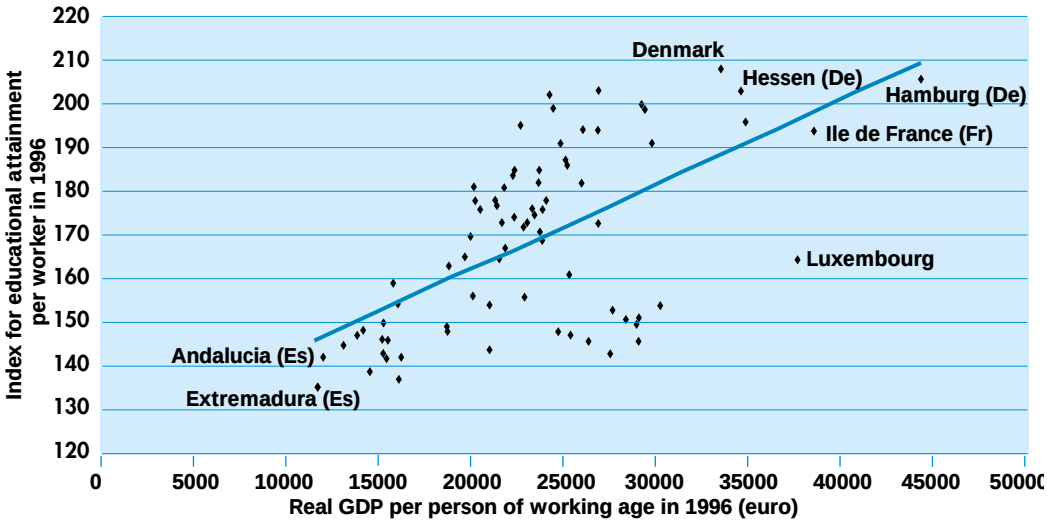


Figure 4. Human capital per worker vs. regional GDP per worker

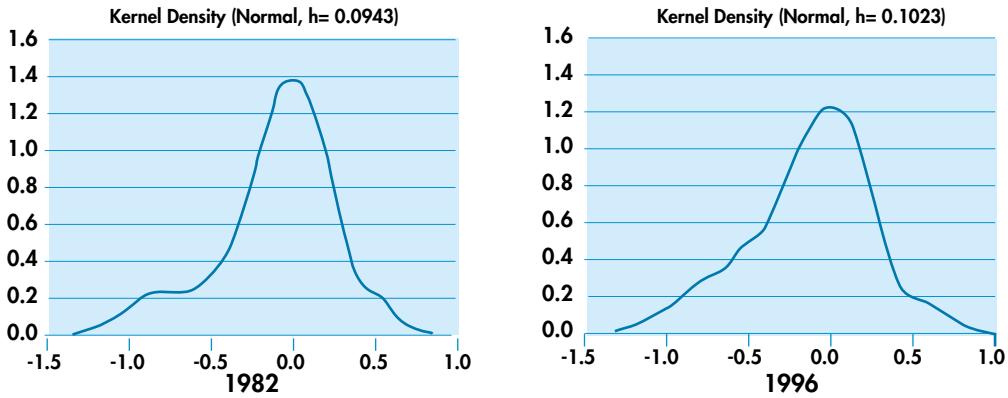


Note: the human capital indicator is a weighted index based on Eurostat data for 1996, computed as: 1x the fraction of the work force with basic education + 2x the fraction with secondary education + 3x the fraction with higher education.

While the standard deviation has remained virtually constant, there have been important changes within the distribution.

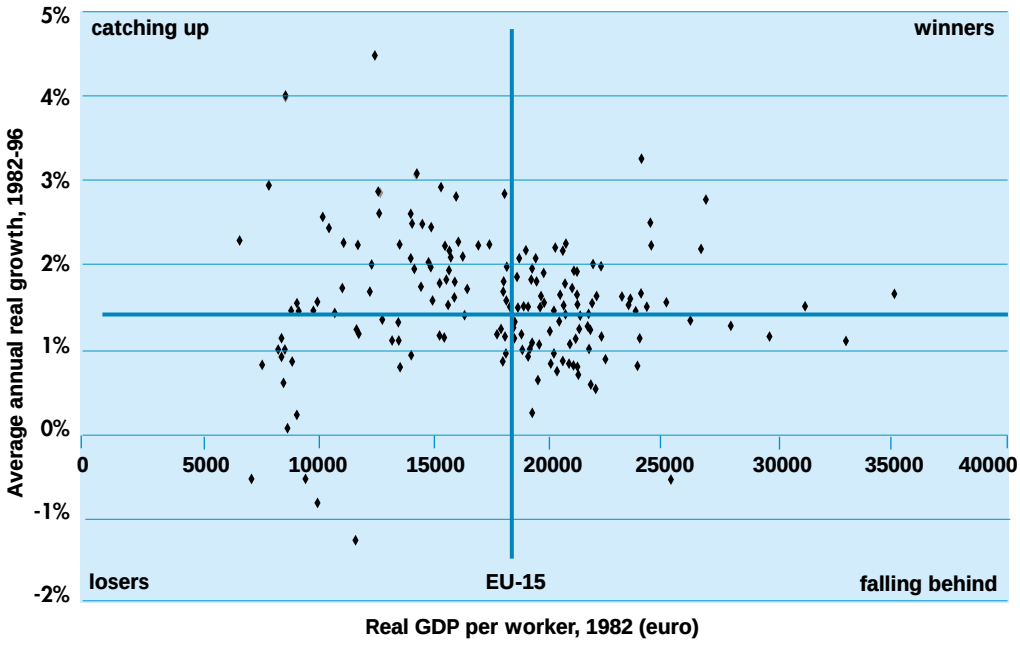
Finally, Figures 5 and 6 give an impression about the economic dynamics that have been at play in NUTS 2 regions. From the snapshots in Figure 5 one would be inclined to conclude that not much convergence has actually taken place. Indeed, the standard deviation of those distributions - which is taken as a measure of dispersion - has remained virtually constant over time. The small increase is attributable to the entry of the New German Länder in the Union. However, plotting the average annual growth performance of regions against their starting position nonetheless indicates that there have been important changes *within* the distribution. To be more precise, one can observe "*winners*" (i.e. regions that had an initial per capita income higher than the EU average and have grown faster than the average), "*losers*" (with an initial standard of living below the EU average and growing at a slower pace than the average), "*catch-ups*" (also regions with an initial below average income per head, yet growing faster than the average), and "*fall-behinds*" (initially situated above the average, but growing towards a per capita income below average).

Figure 5. Distributions compared



Note: The charts plot the standardised distribution of the deviation from the European average income per person of working age.

Figure 6. Within-distribution changes



3.2 Econometric evidence on the role of public and private investment

How are those economic dynamics related to the discussed fundamentals? To investigate this, we will test the level and growth equations as derived in Box 2. The data allowed us to carry out these exercises both at 14 national levels (that is, the EU-15 minus Luxembourg, which lacks a reliable time series on human capital), and for the above-mentioned 78 administrative regions. While we will employ a panel data set-up at the national level, the data availability forced us to apply a more traditional cross-section investigation at the regional level.

Following conventional practice in the panel data growth empirics, every observation for the explanatory variables in our national database represents a five-year average (14) while for the subset of regions the independent variables are averages over the considered time span. Table 3 presents the source and definition of the variables in more detail.

A few remarks are in place before we present the results. A first one concerns the level regressions. When we ran the steady-state regression for the whole EU sample, the results were neither satisfactory nor plausible. Leaving out the Cohesion countries improved the performance of the model drastically, indicating that the assumption of equilibrium is less suitable for these countries than for the more advanced ones. As for the regional sample, we left out the Dutch province of Groningen since variation in income in that region is purely related to the gas production.

In addition, we experimented with different rates of depreciation for the capital components (taken from Beutel, 1996, and Verughese *et al.*, 1997). The results, were, however, not sensitive to reasonable changes in these parameters. For comparisons with earlier studies, we followed the

14) Except for the three years from 1995 to 97, which is added in the level regressions as a separate observation for each country.

standard setting in these kinds of regressions and assumed that the rate of improvements in technological efficiency (x) plus depreciation (δ) equals five percent for every capital component. Concerning fixed effects in the panel data regressions we did not find evidence that country specific effects differed significantly between EU countries during the econometric procedures. The reported panel data results consequently refer to estimations with only time specific fixed effects. Finally, the index for the stock of human capital per worker did not, unfortunately, allow us to compute a plausible rate of return to human capital.

Table 3. Description and source of the data

Variable	Description	Source
human	Stock of human capital per worker	National: Average schooling years in total population over age 25, taken from Barro and Lee, 1992. Figures for 1990, 1995: staff estimates. 5 year intervals.
		Regional: Weighted index based on Eurostat data for 1996: 1x fraction of work force with basic education + 2x fraction of work force with secondary education + 3x fraction of work force with higher education.
n	Growth rate of the work force (people of working age)	National: AMECO, Commission. 5 year averages 1960-94, three year average 1995-97.
		Regional: Regio, Eurostat. Average 1982-96, where available.
s	Total investment share	National: National Accounts, Eurostat, gross fixed capital formation divided by GDP, except for Luxembourg: OECD (in current prices and national currencies). 5 year averages 1960-94, three year average 1995-97.
		Regional: Regio, Eurostat. Average 1982-96, where available.
s_{pub}	Public investment share	National: Figures from the Commission's DG Ecfin, updated with Eurostat values. 5 year averages 1960-94, three year average 1995-97.
		Regional: Gross fixed capital formation in the non-market sector as a percentage of GDP, Regio, Eurostat. Average 1982-96, where available.
y	Income per person of working age	National: National Accounts, OECD, Table 1 (Main Aggregates), USD in prices and exchange rates of 1990. 5 year intervals.
		Regional: Regio, Eurostat. Euro in prices and exchange rates of 1990.
y_{EU}	EU average income per person of working age	

Notes: The private investment share (s_{priv}) is computed as the difference between the total and the public one. East and West Germany are treated as two separate entries in the panel of data. The interested reader can find a detailed description of the regional data used here in a technical EIB working paper (Mathä and Smid, 2000)

Public investment may result in short-term Keynesian effects but the more structural and long-term supply-side impact is debatable.

Tables 4 and 5 present the results of our estimations, which in general provide the same broad messages. Let us first consider the level regression at the national level. From the last lines in the second column of Table 4 we observe that the output elasticity for capital formation by the private sector (α_{priv}) has been larger than the one for capital formation by the government (α_{pub}). The total capital elasticity

($\alpha_{priv}+\alpha_{pub}$) - which we expected to be roughly one third based on the theory and the evidence reported in the national accounts - amounts to an acceptable 31%. We cannot, moreover, reject the hypothesis of a share of human capital of about one-third at the national level (15).

According to the formulas obtained in Box 2, the estimated coefficients imply a social rate of return to private capital of about three percent, which is about the average real interest rate observed over the considered time span. This is consistent with the neo-classical assumption that this sort of capital receives its marginal product, and indicates that there is little evidence to presume large externalities at the aggregate level for private capital. The obtained elasticity for public capital is somewhat lower than values put forward in studies for other samples of countries, implying a rate of return for public capital of approximately ten percent - more than three times the value obtained for private investment. Consequently, this would be a strong indication that public investment takes place in capital goods that induce large externalities.

However, one has to be careful in drawing conclusions solely based on the level regression. When we turn to the growth regression (last column of Table 4), it becomes clear that public investment is negatively related to growth performances. In other words: countries that had high public investment shares have in general experienced low growth, other things being equal. The combined regressions teach us that an important issue of reverse causality is likely to be at play here: richer countries have been able to provide more public capital, but it came at an opportunity cost of lower growth. In that respect, public investment is simply a form of demand-driven consumption: at high levels of income, one is prepared to pay for having utility from better infrastructure. As such, public investment may result in short-term Keynesian effects, the long-term supply-side impact is, however, debatable (see also Martin's contribution, this volume). Capital formation by the government does presumably not directly contribute to productivity growth, but it lowers private saving and growth through the distorting effects from taxation with which those investments are financed.

Causality does not run from public investment to growth, but rather the opposite.

Reverse causality leaves the rate of return to public capital, as obtained from the level regression, meaningless. It is a mistake to argue the existence of large externalities based on the estimate for public investment by looking only at a level regression. Unfortunately this is the case in most of the production function based literature.

In addition, the growth regression implies a speed of convergence of only about one percent. This is half the value reported for the United States, and indicates that factor mobility in Europe has been extremely low. In fact, earlier studies have indeed reported that labour mobility in the US is currently about twice as high as in Europe, which would explain the difference in the estimated speed of convergence rather well (16).

15) Gauging a reasonable value for the share of human capital in GDP is difficult, however. Mankiw, Romer and Weil, 1992, report that the minimum wage - roughly the return to labour without human *capital - has averaged about 30 to 50% of the average wage in manufacturing. This suggests that 50 to 70% of total labour income represents the return to human capital, implying that α_{human} should be estimated in the interval one-third to one-half. Since the minimum wage in Europe is higher than the one in the US, α_{human} should presumably be estimated closer to one-third than to one-half based on European data.*

16) These results are reinforced when an efficient (i.e. GMM) estimation technique is used. The Gauss procedure to apply this method on a panel of data was provided by the courtesy of Rien Wagenvoort, EIB.

Table 4. Regressions for the EU countries (panel data, 1960-97, GLS)

Variable	Dependent variable: $\ln[y]_t$ Level regression	Growth regression
Constant	+9.121 (0.112)**	+0.374 (0.124)**
$\ln[s_{priv}]-\ln[n+x+\delta]$	+0.327 (0.065)**	+0.088 (0.026)**
$\ln[s_{pub}]-\ln[n+x+\delta]$	+0.128 (0.039)**	-0.028 (0.013)**
$\ln[human]$	+0.487 (0.023)**	-0.004 (0.011)
$\ln(y_{t-5})$	-	+0.965 (0.014)**
Time Effects:		
60-64	-0.669 (0.029)**	-
65-69	-0.517 (0.034)**	+0.007 (0.009)
70-74	-0.365 (0.028)**	-0.007 (0.010)
75-79	-0.220 (0.026)**	-0.032 (0.012)**
80-84	-0.165 (0.024)**	-0.119 (0.019)**
85-89	-0.084 (0.023)**	-0.021 (0.016)
90-94	-0.056 (0.025)**	-0.099 (0.017)**
95-97	-	-
R^2	76.95%	99.41%
s.e.r.	0.131	0.042
# obs	66 (non-Cohesion)	81 (EU-14)

Note: White heteroskedasticity-consistent standard errors in parentheses.
* denotes significance at the 10% level;
** denotes significance at the 5% level or better.

	Implied speed of convergence	
Implied λ	-	0.9 %

	Implied output elasticities	
Implied α_{priv}	0.225	-
Implied α_{pub}	0.088	-
Implied α_{human}	0.335	-

Note: The fixed time effects should be interpreted relative to the constant term. In the level regression they indicate as such that per capita income has continuously increased over time. In the growth regression the time effects suggest that increases in the standard of living were significantly lower after 1975 (with the exception of 1985-89) than before. This observation is often referred to as the productivity slow down.

Similar messages can be concluded from the regional regressions, although one has to keep in mind that a major number of regions in the sample belong to just three countries (France, Italy and Spain) (17). Nonetheless, in this case the estimated coefficients in the level regression (second column in Table 5) imply a total capital share that is implausibly low, mainly because of a negative public capital elasticity. Proponents of increased spending argue that this finding may be due to the fact that disaggregating the data reduces the potential to capture geographical spillovers very precisely (18). A negative capital elasticity is, however, only conceivable if the social rate to return is negative. This could for instance be the case if resources are withdrawn from high productivity regions and injected in low productivity ones, without affecting productivity in the latter. Recall from Figure 3 that public investment has indeed been higher in poorer regions. Another explanation may very well be mis-judgement of policy makers in the sense that they did not optimise an economic objective function: independent of the question whether or not public capital formation contributes to productivity, new infrastructures are likely to attract votes, especially when local agents are involved in the construction process.

From the regional growth regression (last column in Table 5) we learn moreover that public investment has not been a source of economic growth, on the contrary. Again, regions that have had a high investment share have known low growth, other things being equal. In addition, this regression reveals a speed of convergence that is exactly the same as the one found at the national level: one percent. Regional policy may consequently have the capacity to be more effective in Europe than suggested in the literature (e.g. by Sala-i-Martin, 1996).

In other words, our results indicate that public capital investments have mainly been used as an instrument for redistribution in Europe; they have not, however, been an engine of regional growth and convergence.

What can we conclude on human capital? The tables point out that the level of human capital is an important factor in explaining the variation in the *levels* of standards of living between regions and countries. In fact, its elasticity is at least as important as the one for private physical capital. However, the stock of human capital has only had a significant impact on *regional growth rates*, not on *nation-wide* economic growth in the EU. This finding has been reported in greater detail in previous growth studies for a sample of highly developed countries such as the OECD (19), and can easily be understood. The average level of human capital has increased substantially for all countries, yet in spite of this there has been a nation-wide productivity slowdown since the early 1970s - a trend that does not unequivocally hold at the regional level. It should therefore not surprise us that the stock variable we employed has had no significant explanatory power in the national growth regression. From a policy point of view our results suggest a regional approach: focussing on improving the human capital of people living in backward regions by providing the appropriate incentives may be an effective tool to promote regional convergence.

Capital formation by the business sector has been effective in stimulating growth. Targeting education in lagging regions seems to be another key towards convergence.

17) We did not find evidence of strong differences in country specific intercepts in the regional regressions.

18) See e.g. Munnell, 1992: "As the geographic focus narrows, the estimated impact of public capital becomes smaller. The most obvious explanation is that, because of leakages, one cannot capture all of the payoff to an infrastructure investment by looking at a small geographic area".

Table 5. Regressions for the EU regions (cross-section, 1982-96, OLS)

Variable	Dependent variable: $\ln[y]_{1996}$ Level regression	Growth regression
Constant	+9.178 (0.477)**	+0.761 (0.362)**
$\ln[s_{priv}]-\ln[n+x+\delta]$	+0.187 (0.239)**	+0.216 (0.070)**
$\ln[s_{pub}]-\ln[n+x+\delta]$	+0.161 (0.079)**	-0.093 (0.033)**
$\ln[human]$	+0.182 (0.090)**	-0.118 (0.033)**
$\ln(y_{1982})$	-	+0.880 (0.040)**
R^2	23.35%	90.36%
s.e.r.	0.244	0.086
# obs	78	78

Note: White heteroskedasticity-consistent standard errors in parentheses.

* denotes significance at the 10% level;

** denotes significance at the 5% level or better.

Implied speed of convergence		
Implied λ	-	0.9 %
Implied output elasticities		
Implied α_{priv}	0.183	-
Implied α_{pub}	-0.156	-
Implied α_{human}	-0.178	-

4. Summary and conclusion

Are public and private capital investments equally productive, and - if so - is there a difference between productivity at the national and regional level? These were the main motivational questions behind our investigations.

At the regional level, public capital investments have mainly been used as an instrument for redistribution in Europe. They have failed to close the productivity gap.

Our research contributes in two ways to the existing literature. First, we explicitly focused on the European countries and regions, rather than at a mixture of developed and less developed nations. Second, we carried out regressions that are consistent with the theory: we have properly controlled for different investment shares while we have also allowed human capital to be a factor of production.

19) E.g. Levine and Renelt, 1992, Mankiw, Romer and Weil, 1992; Islam, 1995; Nonneman and Vanhoudt, 1996; Vanhoudt, 1999. The empirical evidence from cross-country growth regressions on the effect of human capital is indeed ambiguous, see Pritchett, 1997, for a survey.

The qualitative results from our exercises are clear and consistent. Three main findings are noteworthy.

- The first one is a message of reverse causality. Our results indicate that causality does not run from public investment to growth, but rather in the opposite way. What seems to be the case is that richer countries have been able to invest more in public capital, thereby willing to accept a lower pace of growth. Put bluntly, public investment strongly resembles demand driven consumption. As such it undoubtedly has had short-run Keneysian effects, but it can hardly be considered as an engine for long-run - structural - growth. At the regional level, public capital investments have mainly been used as an instrument for redistribution in Europe; they have not, however, closed the productivity gaps. We are able to conclude that private capital, on the other hand, has been effective in stimulating growth and reducing disparities.
- A second one reinforces human capital theory: the formation and quality of human capital is an important economic factor to make regions catch up. Targeting the level of schooling in lagging regions seems to be a key towards convergence.
- Thirdly, the speed of (conditional) convergence in Europe is only half the size of the one reported for the US. This may reflect a low degree of factor mobility - especially labour mobility in Europe.

Our conclusions should nonetheless be somewhat put into perspective. It would be wrong and a departure from common sense to deduce from the analyses that the large stock of public capital and infrastructures provides no benefits (broadly defined) or utility to society. The main message from the regressions in this essay rather is that the use of aggregated national and regional data does not reveal sufficiently large linkages between public sector investment and positive developments in average income per person of working age. Put differently: even though public capital may yield benefits to citizens, it has not been a significant driving force behind regional development and convergence from an aggregate point of view.

However, there is presumably a wide array of public capital investments that would survive a rigorous economic cost-benefit analysis, even ex-post, and that would contribute to local growth. Project selection and performance need to be studied in more detail. This clearly calls for a complementary bottom-up approach that links actual realised economic rates of return at the micro-level to project selection criteria, while one controls at the same time for the quality of economic policies in the receiving region (see Rossert's contribution in the previous volume).

Though public capital has not been a significant force behind convergence from an aggregate view, a wide range of investments are likely to achieve this at the local level.

Finally, our regional database covered a time period of only fourteen years, which may be too short for the purpose we had in mind. Indeed, time horizons for public projects are typically much longer than those of private capital investments, which respond faster to market signals and needs. We were moreover unable to include all European administrative regions because of either lack of (investment) data or data inconsistencies.

Therefore, even though the conclusions we obtained for this sample point in the same direction as those reached from national sources, we are aware that our regional results are indicative at best. The exercise should consequently be carried out again when the regional accounts for all member countries will have been standardised and made available by Eurostat.

References

Arrow, K. (1962). The Economic Implications of Learning By Doing. *Review of Economic Studies*, 29, pp. 155-173.

Aschauer, D. (1989). Is Public Capital Productive. *Journal of Monetary Economics*, 23(2), pp. 177-200.

Aschauer, D. (1998). Public Capital and Economic Growth: Issues of Quantity, Finance, and Efficiency. Working Paper Nr. 233, The Jerome Levy Institute.

Auerbach, A., Hassett, K. and Oliner, S. (1994). Reassessing the Social Returns to Equipment Investment. *Quarterly Journal of Economics*, 109(3), pp. 789-802.

Barro, R. (1991). Economic Growth in a Cross Section of Countries. *Quarterly Journal of Economics*, 106(2), pp. 407-443.

Barro, R. and Lee, J. (1994). Data Set for a Panel of 138 Countries, NBER.

Barro, R. and Sala-i-Martin, X. (1992). Convergence. *Journal of Political Economy*, 100(2), pp. 223-51.

Barro, R. and Sala-i-Martin, X. (1995). Economic Growth. McGraw Hill, New York.

Baumol, W. (1986). Productivity Growth, Convergence and Welfare: What the long-Run Data Show. *American Economic Review*, 76(5), pp. 1072-85.

Berndt, E. and Hansson, B. (1991). Measuring the Contribution of Public Infrastructure Capital in Sweden. NBER Working Paper, 3842.

Beutel, J. (1996). Capital Stock data for the European Union - Report to the Statistical Office of the European Union , *mimeo*, University of Constance.

Capron, H. (1997). Evaluation de L'Impact Reel Des Politiques Structurelles Sur la Cohesion Economique et Sociale De L'Union: Etude de Convergence. Study for the European Commission by the Université Libre de Bruxelles, Unité Economie Spatiale et de la Techologie.

Coe, D. and Helpman, E. (1995). International R&D Spillovers. *European Economic Review*, 39, pp. 859-87.

Crihfield, J. and Panggabean, M. (1995). Is Public Infrastructure Productive? A Metropolitan Perspective Using New Capital Stock Estimates. *Regional Science and Urban Economics*, 25, pp. 607-30.

De Long, B. and Summers, L. (1991). Equipment Investment and Economic Growth. *Quarterly Journal of Economics*, 106(2), pp. 445-502.

De Long, B. and Summers, L. (1992). Equipment Investment and Economic Growth: How Strong is the Nexus?. *Brooking Papers on Economic Activity*, 0(2), pp. 157-99.

Dennison, E. (1967). Sources of Postwar Growth in Nine Western Countries. *American Economic Review*, 57, pp. 325-32.

Easterly, W. and Rebelo, S. (1993). Fiscal Policy and Economic Growth. An Empirical Investigation. *Journal of Monetary Economics*, 32(3), pp. 417-58.

European Commission, (1997). Regional Growth and Convergence: Aggregate and Regional Impact. *The Single Market Review*, Subseries VI: Volume 1. Office for Official Publications of the European Communities.

European Commission, (1998). 6th Periodic Report. Office for Official Publications of the European Communities.

Eurostat (1999). Regio Dataset, New Cronos CD-Rom.

Eurostat. National Accounts, various issues.

Eurostat (1996). Statistical Yearbook.

Garcia-Mila, T. and McGuire, T. (1992). The Contribution of Publicly Provided Inputs to States' Economies. *Regional Science and Urban Economics*, 22(2), pp. 229-41.

Gramlich, E. (1994). Infrastructure Investment: A Review Essay. *Journal of Economic Literature*, 32(3), pp. 1176-96.

Holtz-Eakin, D. (1994). Public-Sector Capital, and the Productivity Puzzle. *Review of Economics and Statistics*, 76(1), pp. 12-21.

Hulten, C. (1996). Infrastructure Capital and Economic Growth: How Well You Use It May be More Important than How Much You Have. NBER Working Paper #5847.

Islam, N. (1995). Growth Empirics: A Panel Data Approach. *Quarterly Journal of Economics*, 110(4), pp. 1127-70.

Islam, N. (1998). Growth Empirics: A Panel Data Approach - A Reply. *Quarterly Journal of Economics*, 113(1), pp. 325-29.

Karras, G. (1997). On the Optimal Government Size in Europe: Theory and Eméirical Evidence. *The Manchester School of Economic and Social Studies*, 65(3), pp. 280-294.

Lee, K., Pesaran, H. and Smith, R. (1998). Growth Empirics: A Panel Data Approach - a Comment. *Quarterly Journal of Economics*, 113(1), pp. 319-24.

Levine, R. and Renelt, D. (1992). A Sensitivity Analysis of Cross-Country Growth Regressions. *American Economic Review*, 82(4), pp. 942-63.

Lucas, R. (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics*, 22(1), pp. 3-42.

Maddison, A. (1982). Phases of Capitalist Development. Oxford University Press, Oxford.

Maddison, A. (1989). The World Economy in the Twentieth Century. OECD, Paris.

Maddison, A. (1991). Dynamic Forces in Capitalist Development. Oxford University Press, Oxford.

Mankiw, G., Romer, D. and Weil, D. (1992). A Contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics*, 107(2), pp. 407-37.

Mathä, T. and Smid, B. (2000). Data Manual and Descriptive Statistics for the Study on Regional Convergence, EIB.

Munnell, A. (1990a). Why has Productivity Growth Declined? Productivity and Public Investment. *New England Economic Review*, Federal Reserve Bank of Boston, January-February, pp. 3-22.

Munnell, A. (1990b). How Does Public Infrastructure Affect Regional Economic Performance?. *New England Economic Review*, Federal Reserve Bank of Boston, September-October, pp. 22-32.

Munnell, A. (1992). Policy Watch: Infrastructure Investment and Economic Growth. *Journal of Economic Perspectives*, 6(4), pp. 189-98.

Neven, D. and Gouyette, C. (1995). Regional Convergence in the European Community. *Journal of Common Market Studies*, 33(1), pp. 47-65.

Nonneman, W. and Vanhoudt, P. (1996). A Further Augmentation of the Solow Model and the Empirics of Economic Growth for OECD Countries. *Quarterly Journal of Economics*, 3, pp. 943-53.

OECD. National accounts, various years.

Pinelli, D. (1998). Regional Competitiveness Indicators. European Commission, DG16. Unpublished study by Cambridge Econometrics.

Pritchett, L. (1997). Where Has All the Education Gone?. *Policy Research Working Paper*, World Bank.

Quah, D. (1993). Galton's Fallacy and Tests of the Convergence Hypothesis. *Scandinavian Journal of Economics*, 95(4), pp. 427-43.

Quah, D. (1996). Empirics for Economic Growth and Convergence. *European Economic Review*, 40(6), pp. 1353-76.

Romer, P. (1986). Increasing Returns and Long-Run Growth. Journal of Political Economy, 94(5), pp.1002-37.

Romer, P. (1990). Endogenous Technological Change. *Journal of Political Economy*, 98(5), pp. S71-S102.

Sala-i-Martin, X. (1996). Regional Cohesion: Evidence and Theories of Regional Growth and Convergence. *European Economic Review*, 40(6), pp. 1325-52.

Sala-i-Martin, X. (1997). I Just Ran Two Million Regressions. *American Economic Review*, 87(2), pp. 178-83.

Solow, R. (1957). Technical Change and the Aggregate Production Function. *Review of Economics and Statistics*, 39, pp. 312-20.

Stokey, N. (1988). Learning by Doing and the Introduction of New Goods. *Journal of Political Economy*, 96(4), pp. 701-17.

Temple, J. (1999). The New Growth Evidence. *Journal of Economic Literature*, 37(1), pp. 112-56.

Uzawa, H. (1964). Optimal Growth in a Two-Sector Model of Capital Accumulation. *Review of Economic Studies*, 31, pp. 1-24.

Vanhoudt, P. (1999). Are Public and Private Outlays for Physical and Knowledge Capital Equally Productive? *Applied Economics*, 31, pp. 1401-10.

Verughese, J., Subhash, S. and Grabowski, R. (1997). Capital Stock Estimates for Major Sectors and Disaggregated Manufacturing in Selected OECD Countries. *Applied Economics*, 29(5), pp. 563-79.

World Bank. (1996-97). World Resources. Washington D.C.

Economic and Financial Reports

Report 2000/01 Dermot O'Brien & Rien Wagenvoort, *"Some empirical findings on the characteristics of cost-efficient credit institutions"*.

Report 99/P3 Carsten Sprenger, *"Direct investment and the eastern enlargement of the EU"*.

Report 99/P2 Daniel Piazolo, *"The credibility and growth effects of EU institutions for Eastern Europe"*.

Report 99/P1 Hendrik Hassheider, *"Private sector investment as a means of financing Trans-European Networks"*.

Report 99/02 Rien Wagenvoort & Paul Schure, *"The recursive thick frontier approach to estimating efficiency"*.

Report 99/01 Paul Schure & Rien Wagenvoort, *"Economies of scale and efficiency in European banking: New evidence"*.

Report 98/05 Daniel Ottolenghi, *"Walking a tightrope: Financial liberalisation in Mediterranean countries"*.

Report 98/04 Christian Keuschnigg, *"Venture capital: A case for investment promotion"*.

Report 98/03 Roman Arjona, Pascale Viala & Rien Wagenvoort, *"Small business finance and the European financial environment: Some empirical evidence"*.

Report 98/02 Alfred Steinherr & Roman Arjona, *"The impact of European Monetary Union on the development of the Mediterranean zone"*.

Report 98/01 Marco Buti, Lucio R. Pench & Paolo Sestito, *"European unemployment: Contending theories and institutional complexities"*.

Report 97/05 Mireille Fischbach & Maria Scattaglia, *"Financial flows to Eastern Europe: Determinants, volatility and sustainability issues"*.

Report 97/04 Mirela Keuschnigg, *"Banking in Central and Eastern Europe"*.

Report 97/03 Gregorio Impavido, *"Pension funds and stock market development in Eastern Europe: Is there a link?"*.

Report 97/02 Agnès Belaisch & Kristian Kjeldsen, *"Value-at-risk models for determining capital adequacy: Does allowance for correlation within broad risk categories matter?"*.

Report 97/01 Agnès Belaisch, *"Why are long-term interest rates so high in Europe?"*.

Economic and Financial Reports are preliminary material circulated to stimulate discussion and critical comment.

The above reports are available free of charge from:

European Investment Bank
Chief Economist's Department
100, boulevard Konrad Adenauer
L-2950 LUXEMBOURG

FAX: (352) 4379-3492
E-mail: h.halahan@eib.org

EIB Papers 1999 - 2000

Volume 5, Number 1, 2000: Regional development in Europe: An assessment of policy strategies

Yannis Ioannides & George Petrakos, "*Regional disparities in Greece: The performance of Crete, Peloponnese & Thessaly*".
Rodolfo Helg, Giovanni Peri & Gianfranco Viesti, "*Abruzzo & Sicily: Catching up and lagging behind*".
Andrés Rodríguez-Pose, "*Economic convergence and regional development strategies in Spain: The case of Galicia & Navarre*".
Margarethe Quehenberger, "*Lessons from Eastern Germany*".
Bertrand Rossert, "*Contributing to regional development through project selection*".

Volume 4 Number 2, 1999: The 1999 EIB Prize

First Prize; Philippe Martin, "*Are European regional policies delivering?*".
Second Prize; Rudi Dornbusch & Francesco Giavazzi, "*Hard currency and sound credit: A financial agenda for Central Europe*".
Third Prize; Jordi Gual, "*Deregulation, integration and market structure in European banking*".
Special Topic Prize; Jens Verner Andersen & Jørn Henrik Rasmussen, "*Solving the unemployment puzzle in Europe*".

Volume 4, Number 1, 1999: European banking after EMU

Christopher Hurst & Rien Wagenvoort, "*European banking after EMU*".
Martin Brookes, "*The impact of EMU on portfolio management*".
Graham Bishop, "*New capital market opportunities in Euroland*".
Colin Mayer, "*European capital markets: Competition between systems*".
Luigi Spaventa, "*Comments on the future of the EMU capital markets.*"
Daniel Gros & Karel Lannoo, "*The structure of financial systems and macroeconomic instability*".
Clas Wihlborg, "*Supervision of banks after EMU*".
Christopher Hurst, Eric Perée & Mireille Fischbach, "*On the road to wonderland? Bank restructuring after EMU*".
Rien Wagenvoort & Paul Schure, "*Who are Europe's efficient bankers?*".
Philip Molyneux, "*Increasing concentration and competition in European banking: The end of anti-trust?*.
Jean Dermine, "*The case for a European-wide strategy*".
Ingo Walter, "*Financial services strategies in the euro-zone*".

The above publications are available free of charge from:

European Investment Bank
Chief Economist's Department
100, boulevard Konrad Adenauer
L-2950 LUXEMBOURG

FAX: (352) 4379-3492
E-mail: h.halahan@eib.org

EIB Papers 1996 - 1997 - 1998

Volume 3, Number 2, 1998: International financial institutions in the 21st century

Christopher Hurst & Eric Perée, "*Only a mid-life crisis? The future for IFIs in an integrated world*".
Michael Klein, "*One hundred years after Bretton Woods: A future history of the World Bank Group*".
Jannick Lindbaek, Guy Pfeffermann & Neil Gregory, "*The evolving role of multilateral development banks: History and prospects?*".
Jean-François Rischard, "*Multilateral development banks and global public networks: Speculations for the next century*".
Pasquale Scandizzo, "*The purpose, management and governance of IFIs: A case study of ambiguity*".
Nicholas Stern & Hans Peter Lankes, "*Making the most of markets: The role of IFIs*".
Joseph Stiglitz, "*IFIs and the provision of international public goods*".
A discussion between Alfred Steinherr & Jacques Attali.

Volume 3, Number 1, 1998: Employment in Europe

Christopher Hurst, "*Adiscussion of the papers and an overview of the role of investment*".
Jørgen Elmeskov, "*The unemployment problem in Europe: Lessons for implementing the OECD Jobs Strategy*".
Gilles Saint-Paul, "*Does technical progress create unemployment?*".
Michael Burda, "*The consequences of EU enlargement for Central and Eastern European labour markets*".
José Viñals & Juan Jimeno, "*The impact of EMU on European unemployment*".
Jacques Drèze, Alain Durré & Henri Sneessens, "*Investment stimulation, with the example of housing*".
Ole Rummel & Thomas Schröder, "*The employment effects of investment projects*".
Pascale Viala, "*Financing young and innovative enterprises in Europe: Supporting the venture capital industry*".

Volume 2, Number 2, 1997: The 1997 EIB Prize

First Prize; Bernadette Frédérick & Peter Vanden Houte, "*What is the optimum strategy for freezing exchange rate parities at the start of EMU*".
Second Prize; Jean Dermine, "*Eurobanking, a new world*".
Joint Third Prize; Daniel Gros, "*The dynamics of European integration*".
Joint Third Prize; Martin Kolmar, "*On the efficiency of national social security systems in the European Union*".
Awards for Authors under 30 for essays by Pablo de Andrés Alonso &
Félix J. Lopez Iturriaga; Volker Köllmann & Andri Kopperschmidt; and Jérôme Vacher.

Volume 2, Number 1, 1997: Banking in an uncertain world

Edward Gardener & Philip Molyneux, "*The too-big-to-fail doctrine revisited*".
Agnès Belaisch, "*High real interest rates: A long-term risk?*".
Pier-Luigi Gilibert, "*Expected loss as a credit risk measure*".
Jacques Girard & Christopher Hurst, "*The capital structure of private infrastructure projects and the risk of default*".
Kristian Kjeldsen, "*Value-at-risk and capital adequacy: The challenge for financial regulation*".

Volume 1, Number 1, 1996: Special Issue on EMU

Daniel Gros, "*The steeple chase towards EMU*".
Ole Rummel, "*On the feasibility of EMU with diverse European economic transmission mechanisms*".
Bernhard Winkler, "*Is Maastricht a good contract?*".
Paul de Grauwe, "*How to fix conversion rates at the start of EMU*".
Luis Gonzalez-Pacheco & Alfred Steinherr, "*A proposal to stabilise the value of the ECU*".

The above publications are available free of charge from:

European Investment Bank
Chief Economist's Department
100, boulevard Konrad Adenauer
L-2950 LUXEMBOURG

FAX: (352) 4379-3492
E-mail: h.halahan@eib.org

European Investment Bank

Economics and Information Directorate

100, boulevard Konrad Adenauer L-2950 LUXEMBOURG
Fax: (352) 43 79 - 34 92