European policy makers may have asked too much from regional policies: to decrease inequalities between regions, to increase efficiency at the national and European levels and to decrease inequalities between countries. This paper argues that these policies face a trade-off between equity and efficiency at the spatial level. If the existence of positive localised spillovers and of returns to scale explain the phenomenon of self-sustaining agglomeration, then agglomeration must have some positive efficiency effects. We also argue that because infrastructure financed by regional policies have an impact on transaction costs and therefore on the location decision of firms, the long-term effect of certain regional policies may be unexpected and unwelcome. Policies that finance infrastructure to reduce transaction costs on goods between regions lead to more agglomeration but higher growth at the national level. We show that policies that reduce agglomeration (transfers, financing of transport infrastructure inside the poor regions) may then also reduce efficiency and growth. On the contrary, a policy that reduces the cost of innovation or increases the diffusion of innovation reduces regional income inequality, agglomeration and increases growth.

Philippe Martin is a researcher at the Centre d’enseignement et de recherche en analyse socio-économique (CERAS) at the Ecole Nationale des Ponts et Chaussées, Paris and an assistant professor in international economics at the Graduate Institute of International Studies. He is also a Research Fellow at the Centre for Economic Policy Research in London.

A graduate from Sciences-Po in Paris, he has a Ph.D. in economics from Georgetown University in Washington DC. He specialises on subjects such as trade and monetary integration in Europe, growth theory, and economic geography. He has worked as consultant for both the World Bank and the European Commission.
1. Introduction

Does trade and monetary integration in Europe entail the risk of widening inequalities between the regions? To judge by the sums devoted to regional policies in Europe, which now account for one third of the Community budget and form the second largest item after the Common Agricultural Policy, the reply given by governments and the European Commission is clearly yes. The sharp expansion in regional policy spending has taken place since the accession of Spain and Portugal. This, following the admission of Greece, led to a widening of income disparities between the poor and rich regions of what was then called the European Community. The negotiations on the accession of the two Iberian countries resulted in an increase in the resources for regional policies from ECU 3.7 billion in 1985 to ECU 18.3 billion in 1992. The figure will reach ECU 33 billion in 1999, equal to 0.45% of Europe’s GDP. These transfers represented up to 3% of GDP for some countries of the Cohesion Group (Greece, Ireland, Portugal and Spain). The actual amounts spent on regional policies in these countries were much larger as the EU requires that its transfers be matched by national spending. National regional policies have also been very important in certain countries such as France, Italy (the transfers to Mezzogiorno) and Germany (the transfers to the New Länder). The enlargement of the European Union to the Central and Eastern European countries, where per capita GDP levels are much lower than in the four Cohesion countries will imply a major overhaul of European regional policies.

Compared with the scale of expenditure on these regional policies, the observer is struck by the weakness of the conceptual framework used to justify them. It is scarcely a caricature of the Commission’s position to portray this as a belief that transfers to the poorest European regions are beneficial to them, bringing about a reduction in regional inequalities which in turn is bound to benefit Europe as a whole. The inspiration behind regional policies is to be found in Article 130a of the Treaty of the European Union, which speaks of “harmonious development” with the aim of “reducing disparities between the levels of development of the various regions”. The justification is not meant to be solely political. It is also economic, since the report says that “the disequilibria indicate under-utilisation of human potential and an incapacity to take advantage of the economic opportunities that could be beneficial to the Union as a whole”.

The economic reasoning underlying this thesis is not, however, clear. The neo-classical theory of international trade tells us that a low level of productivity (a disadvantage in absolute terms) is no impediment to benefiting from trade gains based on comparative advantages. Furthermore, the neo-classical theory of growth with decreasing returns of scale predicts that trade integration and liberalisation of capital movements will accelerate convergence: because of decreasing returns, regions with low incomes and low availability of capital should, other things being equal, have a high return on capital and this should therefore attract capital movements in an integrated area such as the EU. Active policies to help the most disadvantaged regions cannot be justified in a neo-classical framework of perfect competition and without economies of scale, since within such a framework the process of integration should accelerate convergence between regions.
The new theories of economic geography and the new theories of endogenous growth – a common feature of which is their emphasis on the importance of economies of scale, imperfect competition and phenomena of localised spillovers – seem more appropriate. Contrary to the neo-classical paradigm, the theories of endogenous growth do not predict convergence between rich and poor regions even when movements of goods and capital are free: Indeed, by abandoning the hypothesis of decreasing returns on capital, these models exclude the economic mechanism that generates the process of convergence. Moreover, the recent models of geographic economics show that regional integration, by reducing transaction costs between the regions, may lead to self-sustaining inequality.

The new theories of economic geography and endogenous growth can therefore serve as a conceptual framework for regional policies, since they offer explanations for self-sustaining phenomena of regional inequalities. However, with its emphasis on the positive effects of local spillovers and on economies of scale, this framework also implies that there are positive effects from agglomeration and hence from regional inequalities (see Fujita and Thisse, 1996; Jayet, Puig and Thisse, 1996). If economies of scale and localised spillovers explain phenomena of increased regional inequalities, this necessarily implies that efficiency gains (in terms precisely of economies of scale or spillovers) accrue from the existence of economic agglomeration. The existence of these beneficial effects of agglomeration suggest rather that, in certain respects, Europe’s economic geography is insufficiently agglomerated and specialised (for example in comparison with American geography). It is therefore illogical to claim that the diminution of regional inequalities supposedly facilitated by regional policies will generate efficiency gains at pan-European level. To oppose concentration and geographical specialisation is also to renounce their beneficial effects.

2. What links are there between efficiency and territorial equity?

2.1 A simple theoretical framework

To illustrate this tension between the countervailing effects of agglomeration – positive in terms of efficiency, but potentially negative in terms of equity – I shall use a two-region theoretical scheme. Firms can locate either in the capital-rich North region or in the South region. The geographical concentration of firms in the rich region increases when transaction costs between the regions fall. The logic (which is common to the new theories of international trade and to the models of economic geography) is that it is always more profitable to produce in the richer area, the larger market, in order to maximise the benefits of economies of scale. When transaction costs between the regions fall, businesses can then exploit these economies of scale while also selling on the “small market” which is less “protected” by high transaction costs. In addition, when regional inequality in terms of income increases, regional inequality in terms of spatial distribution of firms (industrial agglomeration) likewise increases, since economies of scale give firms an incentive to locate where demand is strongest and income consequently highest. Equilibrium geography is such that the profits of businesses are identical in both regions, which eliminates any incentive to relocate. This equilibrium relationship (profits are equal in both regions) can be encapsulated in the following relationship:

Equation No. 1: \( A = A(R) \).
where \( A(R) \) is a growing function of \( R \) and where \( A \) is an agglomeration index (for example, the ratio of the number of firms in the rich region to the total number of firms). \( R \) is an index of inequality of regional incomes (for example, the ratio of income in the rich region to income in the poor region).

Spatial concentration in turn has an impact on the rate of innovation and hence on the long-term growth of the overall economy, because the cost of innovation in the richer region falls as the agglomeration of economic activities increases.

Several reasons can be advanced. First, 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. Another possibility is the existence of localised technological spillovers such as those studied by Jacobs (1969) and by Henderson and others (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.

In both cases, geographical concentration of production activities increases opportunities to reduce the cost of innovation and consequently to increase its rate of growth, with beneficial effects for the territory as a whole. In endogenous growth models this is an equilibrium relationship, because when the cost of innovation falls this induces new entrepreneurs/researchers to enter the innovation market which is regarded as being competitive. This equilibrium relationship between the long-term growth rate and the agglomeration index will be summed up by the following relationship:

**Equation No. 2:** \( g = g(A) \),

where \( g(A) \) is an increasing function of \( A \), the index of industrial agglomeration.

The rate of innovation itself has an impact on regional income inequalities since a high rate of innovation accelerates market entry by new businesses, which then compete with existing businesses and hence reduce their profits. One effect therefore is to reduce existing incomes. From this point of view, an increase in the rate of innovation reduces income disparities between regions by reducing the profits of monopolistic firms, which are more numerous in a rich than in a poor region. This last equilibrium relationship is summed up by the following relationship:

**Equation No. 3:** \( R = R(g) \),

where \( R(g) \) is a negative function of the growth rate \( g \).

In Figure 1, which sums up these different equilibrium relationships, the upper part shows the spatial equilibrium where income inequalities and industrial agglomeration are determined. The curve AA shows that the phenomenon of agglomeration tends to increase when income inequalities increase, because firms locate in markets with high purchasing power (Equation 1).
**Figure 1.** Relationship between innovation, regional income inequalities and agglomeration

The curve RR shows that when industrial agglomeration increases competition intensifies, thereby tending to reduce the profits of monopolistic businesses and income inequality between regions (Equations 2 and 3). The equilibrium level of agglomeration and the equilibrium level of income inequality is indicated by the intersection of the two curves AA and RR. The lower part of the graph shows how spatial equilibrium in its turn influences the rate of innovation. The equilibrium level of agglomeration A is given by the spatial equilibrium. The curve SS shows the positive relationship between innovation and agglomeration, due to the existence of localised spillovers (Equation 2). The equilibrium rate of innovation and the equilibrium level of income inequalities are indicated by the intersection of the line A and the curve SS.

### 2.2 What is the empirical link between efficiency and geographical equity in Europe?

Quah’s (1996) results suggest that there is indeed a trade-off between regional equity and a country’s aggregate growth. He finds that, among the Cohesion group of countries (Greece, Ireland, Portugal and Spain, though there are no Irish regional data), the two countries that have achieved a high rate of growth and converged in per capita income terms towards the rest of Europe (Spain and Portugal) have also experienced the most marked regional divergence, Portugal being the country to have exhibited the sharpest increase in regional inequalities. By contrast Greece, which has a low growth rate and has not benefited from a tendency to converge with the rest of Europe, has not experienced a rise in regional inequalities. A recent study by INSEE (1998) shows also that the countries with a per capita GDP level above the European Union average also experience above-average regional disparities.
Another way of asking the question is to study the nature of the convergence process in Europe. Taking Europe as a whole, a slow, long-run process of convergence does indeed exist between the European regions (at NUTS2 level). Thus, Sala-i-Martin (1996) finds that, over a long period (1950-1990), the average growth rates of the regions are negatively correlated with initial income. The speed of convergence is 2%, meaning that on average 2% of the per capita income difference between regions is eliminated and that it takes more than 30 years to eliminate half of the initial income difference.

Neven and Gouyette (1994), however, find that, over the more recent period starting in the 1980s that has witnessed major advances in European integration, a process of divergence has been appearing between regions of the North and the South. In addition, even a slow long-run convergence at the European level may mask a process of regional divergence within countries. De la Fuente and Vives (1995), for instance, building on the work of Esteban (1994), suggest that around half the income inequality existing 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). This would suggest that Europe is experiencing a process of convergence between countries simultaneously with one of divergence between regions within individual countries. This possibility can be illustrated by the following Figure, where each dot represents a region.

**Figure 2.** Pan-European convergence, local divergence

In this example, there is indeed convergence between regions at European level, since the initially poor regions tend to grow faster than the rich regions: A negative relationship does indeed exist on average between growth rate and initial per capita GDP. However, no process of convergence exists within each country. In fact, in country 1, the poorest, which is growing the fastest and converging towards the others, there is a process of domestic regional divergence.
These results seem therefore to suggest that the neo-classical growth model holds at country level whereas a model of endogenous growth with elements of geographic economics holds for the regions of individual countries; and that the economic mechanisms which generate increasing returns, and hence the possibility of divergence, are therefore more powerful at local than at national level. Several hypotheses can be advanced to explain this difference.

- Spillovers deriving from increasing returns are geographically limited phenomena, since they depend on social interactions between individuals (1).
- Migration may be the origin of agglomeration phenomena (see Krugman, 1991a,b) and, as is known, labour migration is low between European countries on account of cultural and linguistic barriers.
- It is possible that transaction costs between regions within each country are much lower than transaction costs between regions of different countries, notably because of the existence of an exchange risk between countries which has only just disappeared with Monetary Union. The empirical studies by Engel and Rogers (1996) on the “cost” of the frontier between the United States and Canada seem to support this. They find that the frontier has the same effect, in price change terms, as a domestic distance of nearly three thousand kilometres.

3. Public policies, regional inequalities and growth

3.1 What can theory tell us?

Why is intervention necessary, that is to say why is market-driven geography not optimal? In the first place, when firms decide where to locate they do not take into account the impact of this choice on the well-being of immobile economic agents. From this point of view, equilibrium geography will be too concentrated because people remaining in the disadvantaged region will be penalised both as workers and as consumers. Secondly, in deciding where to locate, businesses will also not take into account the positive effects of agglomeration on the rest of the economy, particularly the innovation sector. From this point of view, market-driven geography will be insufficiently concentrated in the sense that it will generate too low a rate of innovation and growth. There is therefore a difficult choice between these two considerations which regional policies should take into account.

Let us first assume that a simple monetary transfer is made from the rich to the poor region. In Figure 3, this transfer therefore produces a leftward shift of curve RR (for a given industrial geography, the income inequality will be lower). The induced effect on the geography will be to weaken the agglomeration phenomenon, since the increase in incomes in the poor region (and the diminution of income in the rich region) will stimulate relocation of firms to the region that has relatively increased its purchasing power. In the lower part of the graph, this decline in agglomeration is reflected in a more dispersed economic geography less conducive to spillovers, and hence in a lower growth rate.

1) The work of Jaffe, Trajdenberg and Henderson (1993) shows that the citation and use of patents is very localised. This is very strong evidence that knowledge spillovers are themselves very localised.
Figure 3. Effect of a transfer to the South

However, European regional policies aspire to be more than simple transfers. In fact, their objective is to transform supply conditions. Thus some 30% of the Structural Funds is allocated to the financing of infrastructure, largely transport infrastructure. This concentration on infrastructure expenditure is justified by the Commission on the basis that inter-regional disparities in infrastructural terms are more marked than inequalities in terms of income.

Within the analytical framework presented here, the main consequence of financing such public infrastructure (particularly of the transport kind) is to reduce transaction costs. This is moreover the paramount objective sought by the Commission, which wishes in this way to enable the poor regions to benefit from the advantages of the Single Market. However, lowering transaction costs has a widely differing impact on economic geography depending on whether the reduction is mainly in costs within the region or between the regions (see Martin and Rogers, 1995 and Martin, 1999).

1. In the first case, a reduction in transaction costs within the poor region, by increasing the effective local demand for locally produced goods, will have the consequence of attracting new firms into this region. In Figure 4, this leads to a leftward shift of curve A.A (given the same level of income inequality, agglomeration diminishes). Industrial agglomeration has diminished to the benefit of the poor region but this leads to a lower rate of innovation and greater income inequality as businesses in the North, now facing less competition, increase their profits. Therefore it is not certain that such a policy is to be recommended, whether from the standpoint of efficiency or equity. This example may seem paradoxical, but it highlights the fact that industrial location inequality does not always exactly mirror income inequality, since economic geography has an impact on the rate of innovation which can itself influence income inequalities.
2. The exactly opposite effect occurs if regional policy tends to reduce transaction costs between the two regions (2). The reason is that such a reduction offers firms an incentive to relocate to the richer region where they can now benefit from economies of scale, while selling in the poor region thanks to lower inter-regional transaction costs. This result recurs in numerous models of economic geography (see in particular Krugman, 1991a,b). Thus this type of regional policy accentuates the phenomenon of agglomeration; it thereby raises the long-term growth rate and brings down income inequality since it reduces monopolistic business profits. The result may seem paradoxical at first sight: facilitating access to a poor region may increase agglomeration. The example of motorway building between northern and southern Italy, which was supposed to unlock the south but has spurred agglomeration in the north, shows that this paradox is not a purely theoretical one (see Faini, 1983). However, induced agglomeration is not necessarily unfavourable at the national level insofar as the rate of innovation of the economy as a whole is boosted. There again, the Italian example of a high rate of innovation in the North illustrates the positive impact of agglomeration.

As can be seen from these examples, the effects of regional policies are rather complex and may at times seem paradoxical. In all events the situation is fairly distant from the very simple logic of regional policies based on the idea that transfers or infrastructure financing always favour the poor regions and that this in turn is bound to benefit the country or Europe as a whole.

In all the examples looked at, regional policy has an unfortunate consequence: a reduction in the rate of growth (direct transfer), or the same effect coupled with an increase in income inequalities (infrastructure financing within the poor region), or relocation of firms to the rich region (financing

2) The reduction in transaction costs within the rich region has the same effects, in our context, as a reduction in inter-regional transaction costs.
of infrastructure between the poor and the rich region). Hence, regional policies face a trade-off between equity and efficiency. In the case of the Cohesion countries (and even more so in the future for Eastern European countries), this suggests that it will be difficult to attain through these policies the objective of higher national growth (and therefore fast convergence towards the rest of Europe) and at the same time the objective of a decrease in regional inequalities.

However, a policy aimed at reducing regulatory barriers to innovation or the costs of innovation makes it possible simultaneously to achieve objectives of reducing regional inequalities and increasing the rate of growth. The policies involved could be R&D subsidies, education infrastructure, lowering barriers to entry on goods markets, making capital markets more conducive to new start-ups.

In this case (Figure 5), it is the dynamic equilibrium (lower part of the graph) which is first affected. A reduction in the cost of innovation tends to increase the rate of growth: The curve SS shifts downwards (the rate of growth increases for a given level of agglomeration). By boosting competition, this increases in the rate of innovation, reduces business profits and hence income inequalities between the two regions. This induced effect means that spatial equilibrium is also affected: The curve RR shifts leftward and industrial agglomeration in the rich region diminishes.

In the final equilibrium state, agglomeration and income inequality have diminished while the growth rate has risen. The apparent paradox is therefore that the public policy which is least "regional" in its application enables the regional policy objectives to be achieved.

**Figure 5.** Effect of a reduction in the cost of innovation or of an increase in diffusion of innovation

Another policy, closer to the traditional vision of regional policies, can also have the same effects. As has been seen, infrastructural policies that reduce the inter-regional or intra-regional cost of goods led either to a more unequal geography or to a decline in the rate of innovation. However, when infrastructure-improvement policy focuses on lowering the cost of conveying information rather than the cost of transporting goods, the effect is quite different: By fostering the effects of inter-regional
spillovers, such a policy enables the rate of innovation for a given geography to be stepped up, since the innovation sector benefits more from spillovers generated by geographically remote firms. These policies would have the objective of increasing the capacity of poor regions to absorb new technologies and to increase spatial diffusion of innovation. This could be done by financing infrastructure in telecoms and in education. The impact is then similar to that illustrated in Figure 5.

Finally, we have seen that the main equity consideration justifying the objective of regional policies to counter agglomeration is the existence of immobile economic agents who are penalised by the concentration of economic activities. 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 policy issue only on this side of the Atlantic. From the regulatory point of view, housing and tax policies that facilitate their mobility should therefore be regarded wholly as regional policies. 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 suggests that policies that facilitate inter-sectoral mobility such as education and training policies should be reinforced.

3.2 The effect of regional policies: Empirical results

De la Fuente and Vives (1995) have obtained results that are somewhat disappointing for regional policies as applied to Spanish circumstances. They find that the contribution made by public investment to regional convergence of incomes was small, accounting for around 1% of the inequality reduction during the 1980s. The Commission (1996) used macroeconomic input-output models and found that, in the absence of the structural and cohesion funds, GDP growth in the four Cohesion group countries (Spain, Portugal, Ireland and Greece) would have been $\frac{1}{2}$% lower than the actual outcome. It must be stressed that these numerical estimates focus on the positive short term Keynesian effect on local demand and not on the long-term supply effects. Moreover, they tell us nothing about the effects on convergence between regions within a given country. Finally, the bias that local policy makers have in favour of large infrastructure projects (especially transport infrastructure) can certainly be explained by these positive short term effects on local demand and output. It is however important to stress that in the case of infrastructure that lowers the cost of transaction between regions (such as highways) the long term location and supply effect is negative for the poor region and therefore exactly inverse to the short-term positive impact. Hence, relying on results that emphasise short-term demand effects can be misleading.

Combes and Lafourcade (1999) indeed find that the decrease of transport costs in the last 20 years has indeed led to more agglomeration. Martin (1998), using data on regional stocks of infrastructure (transport, telecommunications, energy, and education) finds that, if the regional growth regressions are linked to these stocks, the speed of convergence across regions in Europe increases for telecommunications infrastructure. Thus, if the telecommunications infrastructure had been similar in all the regions of Europe, the speed of convergence (that is to say, the average annual percentage reduction in the inter-regional income gap) would have been 4.1% as against an actual 1.3% over the period 1978-1992. Calculated on transport infrastructure, the speed of convergence would have been 2%. These figures cannot, however, be interpreted as being very promising for the effects of regional policies. All in all the gain is fairly small in terms of
convergence and must be set against the huge cost of an infrastructural programme that would equalise infrastructure stocks between the European regions. Above all, when the regressions are carried out for the regions of an individual country, it is found that, with the exception of communications infrastructure, stocks of public infrastructure have no significant impact on the speed of convergence between the regions within a given country. It will be noted that the positive effect on the convergence of communications infrastructure is consistent with the theoretical idea presented here that a reduction in the cost of conveying information is theoretically more favourable to regional equity than a reduction in the cost of transporting goods.

4. Conclusion

A standard principle in economics is that with one policy instrument it is difficult - to say the least - to attain different objectives. In some sense, European policy makers have asked too much from regional policies: To decrease inequalities between regions, to increase efficiency at the national and European levels and to decrease inequalities between countries. A key point of this paper, both from a theoretical and empirical point of view, is that this may contradictory. Moreover, policy makers often expect that transfers in the form of the financing of infrastructure will have both a positive short-term demand effect and a positive long-term supply effect. The first one is quite obvious and certainly contaminates the data and the debate on regional policies but because infrastructure has an impact on transaction costs and therefore on the location decision of firms, the long term supply effect on the region may be opposite to the short-term effect.

This suggests that the objectives of regional policies need to be carefully redefined. First, if the ambition of regional policies is to affect the long-term economic geography of Europe, then only the supply effect should be considered in the allocation of funds. Second, policy makers should decide whether their main objective is to decrease inequalities between the different countries and therefore give priority to national growth and efficiency or to decrease inequalities between the different regions inside countries and therefore to give priority to growth in poor regions and to spatial equity. This is a crucial political question in view of the enlargement to the East as these countries have average incomes per capita that are much below those of the four Cohesion countries and also have growing regional inequalities.

A final point is that to justify public intervention a necessary (but not sufficient) condition is that a market failure is clearly identified. In the case of economic geography, we have seen that the market failure may come from the externality due to immobility of agents (firms do not take into account the impact of their location choice on immobile private agents) and to the positive externalities linked to technological spillovers (in this case the market driven geography may be insufficiently agglomerated and specialised). Usually, the best policy is the one that intervenes at the source of the externalities. In the first case, facilitating the interregional (and inter-industrial) mobility of workers seems to be the most direct way to diminish the negative social impact of that spatial externality. If workers were more mobile (both between regions and between sectors), they would suffer less from the effects of the location decision of firms. Facilitating mobility is not very much considered to be in the realm of regional policies but this should be reconsidered especially in view of the very small interregional mobility in Europe (compared to the US for example) and of its social costs. The second externality comes from the fact that technological spillovers are
localised. Here, the aim of the public policy should be to make these spillovers less localised so that they benefit the whole of Europe. This is however a very different policy than the one that consists in the spatial dispersion of innovation activities. We have seen that theoretically, and to a certain extent also empirically, this suggests that more emphasis should be put on the financing of telecommunications and education infrastructure.
References


