8 Regional Disparities and Efficient Transport Policies

Stef Proost and Jacques-François Thisse

Abstract

This chapter addresses the economics of regional disparities and transport policies in the European Union, offering an explanation for the uneven development of regions. We show that recent developments in spatial economics highlight the fact that trade is costly and location still matters. Since the drop in transport costs and the emergence of a knowledge-based economy, the proximity to natural resources has been replaced by new drivers of regional growth that rely on human capital and cognitive skills. Regions with a high market potential – those where demand is high and transport costs low – are likely to attract more firms and pay higher wages, which leads to sizable and lasting regional disparities. As a consequence, investments in interregional transport policies may not deliver their expected effects. In addition, new information and communication devices foster the fragmentation of the supply chain and the decentralization of activities.

8.1 Introduction

This chapter addresses the economics of regional disparities and regional policies in the European Union (EU). The fundamental challenge is to explain the uneven development of regions in both the EU and within EU member states. The purpose is not to delve into concrete regional policies and judge their results but rather to understand the main drivers of contemporary regional development. Earlier explanations evolved around natural resources and transport systems. But since the emergence of a knowledge-based economy, traditional location factors have been replaced with new drivers of regional growth that rely on human capital and cognitive skills. This chapter is organized in seven sections. In the second one, we focus on the concepts and tools of spatial economics that are necessary as a backdrop to regional economics. In the third section, we analyse the main forces driving the allocation of economic activity across regions: firms' market access and labour mobility. The fourth section examines these two forces to see whether they generate over or under-agglomeration. The

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

fifth section is devoted to the effects of investments in interregional transport policies, while the sixth section briefly analyses current interregional transport policies. The seventh section concludes.

8.2 What Is Spatial Economics About?

The Industrial Revolution exacerbated regional disparities by an order of magnitude that was unknown before. The recent development of new information and communication technologies is triggering a new regional divide of which governments and the public should be aware. What economic tools can we use to understand those evolutions? As spatial economics deals with bringing location, distance, and land into economics, its aim is to explain where economic activities are located. This makes spatial economics one of the main economic fields that can be employed to understand how the new map of economic activities is being drawn. Yet, at first glance, the steady (actually spectacular) drop in transport costs since the mid-nineteenth century – compounded by the decline of protectionism post-World War II and, more recently, by the near-disappearance of communication costs – is said to have freed firms and households from the need to be located near one another. Therefore, it is tempting to foresee the '*death of distance*' and the emergence of a '*flat world*' in which competition is thought of as a race to the bottom, with the lowest-wage countries as the winners.

But – and it is a big but – while it is true that the importance of proximity to natural resources has declined considerably, this does not mean that distance and location have disappeared from economic life. On the contrary, recent work in regional and urban economics indicates that new forces, hitherto outweighed by natural factors, are shaping an economic landscape that, with its many barriers and large inequalities, is anything but flat. Empirical evidence shows that sizable and lasting differences in income per capita and unemployment rates exist. In brief, *the fundamental task of spatial economics is to explain the existence of peaks and valleys in the spatial distribution of wealth and people*. This is what we aim to accomplish in this chapter. Most graduate or undergraduate students in economics have barely come across the words 'cities', 'regions', and 'transport' during their studies. We therefore will define the basic concepts of spatial economics that are not part of the tool box of most economists. In particular, we show how the tools of modern economic theory can illuminate the main issues of spatial economics, and how modern empirical methods have helped measure them. Conversely, introducing space into economic modelling allows one to revisit existing theories and suggest new solutions to old problems. In particular, we highlight some of the findings that reveal the increased importance of space in the modern economy.

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

8.2.1 Location Does Matter

Why do economic activities cluster in a few places? There is no satisfactory answer to this question in the dominant paradigm of economic theory, which combines perfect competition and constant returns to scale. In the absence of scale economies, fragmenting production into smaller units at different locations does not reduce the total output available from the same given inputs, but transport costs decline. In the limit, if the distribution of natural resources is uniform, the economy is such that each individual produces for his or her own consumption. This strange world without cities has been called '*backyard capitalism*'. To put it differently, each location would become an autarky, except it is possible that trade between locations might occur if the geographic distribution of natural resources is uneven. Admittedly, different locations do not a priori provide the same exogenous amenities. However, using the unevenness of natural resources as the only explanation for the existence of large cities and for regional imbalance seems weak. Rather, as noted by Koopmans (1957) almost 60 years ago, *increasing returns are critical to understanding how the space-economy is shaped*.

A simple example will illustrate this fundamental idea. Suppose a planner has to decide where to locate one or two facilities to provide a certain good to a population of immobile users who are evenly distributed between two regions. Individual demands are perfectly inelastic and normalized to one; the marginal production cost is constant and normalized to zero. Consumers in the domestic region may be supplied at zero cost, whereas supplying those living in the foreign region entails a transport cost of T euros. If two facilities are built, the cost of building a facility is equal to F euros in each region. If only one facility is made available, the planner must incur cost F; if two facilities are built, the cost is $2F$. A planner who aims to minimize total costs will choose to build a facility in each region if, and only if, $F + T$ is more than $2F$, that is, $T > F$. This will hold when F is small, T is high, or both. Otherwise, it will be less expensive to build a single facility that supplies all people in both regions. In other words, weak increasing returns – F takes on low values – promote the scattering of activities, whereas strong increasing returns foster their spatial concentration. As a consequence, the intensity of increasing returns has a major implication for the spatial organization of the economy.

The first law of spatial economics: If many activities can be located almost anywhere, few activities are located everywhere.

It is in this sense that location matters: although a large number of activities become '*footloose*', a relatively small number of places in many countries account for a large share of the national value added, whereas many large areas account for no or little economic activity. The difficulty economists encounter

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

when they take into account scale economies in general equilibrium theory probably explains why spatial economics has been at the periphery of economics for so long.

Nevertheless, it must be kept in mind that accounting for increasing returns often yields a message that differs from the standard neoclassical paradigm of perfect competition and constant returns to scale. Even though transport costs must be positive for space to matter, one should not infer from this observation that location matters less when transport costs decrease—quite the opposite. Spatial economics shows that lower transport costs make firms more sensitive to minor differences between locations. To put it another way, *a tiny difference may have a big impact on the spatial distribution of economic activity*.

8.2.2 Moving Goods and People is Still Costly

Transportation refers to the movement of people, goods, information, or anything else across space. Ever since the beginning of the Industrial Revolution, there has been spectacular progress in terms of the speed and cost for interregional *and* international transport. According to Bairoch (1997), 'overall, it can be estimated that, from 1800 to 1910, the decrease in (weighted) real average transport costs was on the order of 10 to 1'. For the US, Glaeser and Kohlhase (2004) observe that the average cost of moving a ton a mile in 1890 was 18.5 cents, as opposed to 2.3 cents today (in 2001 dollars). Yet, as will be seen, estimating the gravity equation reveals that distance remains a strong impediment to trade and exchange. What is more, the current concentration of people and activities in large cities and urban regions fosters steadily increasing congestion both in private and public transport as capacity is not easy to expand. In the regional context, transportation consists of interregional and international freight trips of inputs and outputs, as well as passenger trips. Unlike an urban environment, larger interregional passenger and freight flows tend to reduce rather than increase the average transport costs because of the presence of economies of density in scheduled transport and because capacity expansion (physical and/or frequency) is easier to implement. Therefore, transportation faces different challenges at the urban and interregional levels.

In the wonderful dimensionless world of some analysts and journalists, transport costs are zero, and thus any agent is equally connected to, or globally competes with, any other agent. If the monetary cost of shipping goods has dramatically decreased, other costs related to the transport of goods remain significant. For example, the opportunity cost of time rises in a growing economy, so that the time cost wasted in moving certain types of goods steadily rises. Similarly, doing business at a distance generates additional costs, even within the EU, due to differences in business practices, political and legal climates, or culture. One of the most robust empirical facts in economics is the Gravity Law: 'Holding

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

constant the product of two countries' sizes, their bilateral trade will, on average, be inversely proportional to the distance between them' (Head and Mayer, 2013). To put it differently, distance between locations still matters because it affects the economic life under different disguises.

Doing a back-of-the-envelope calculation, Cheshire and Magrini (2006) find that despite smaller regional disparities and larger average distances in the US than in the EU, the net migration rate between areas having a comparable population size is almost 15 times higher in the US than in the EU. (The areas are the 50 US states plus Washington, DC versus the EU-12 large countries – France, Germany, Spain, and the UK – divided into their level 1 regions [in Germany, the Länder] and the EU's smaller countries, treated as single units.) These authors conclude that 'in Europe, urban population growth seems likely to be a rather imperfect signal of changes in welfare in cities'. This is to be contrasted with a recent macroeconomic study by Beyer and Smets (2015), who show that, once they control for country factors, labour mobility across 41 EU regions would account for almost 50 per cent of the long-run adjustment process to negative regional shocks, which is more or less the same as in the US where mobility has been decreasing since 1980. However, it takes much longer in Europe than in the US for this adjustment to unfold.

On the other hand, more disaggregate spatial studies strongly suggest that, even within European countries, migration is sluggish and governed by a wide range of intangible and time-persistent factors. For example, controlling for the geographical distance and several other plausible effects, Falck et al. (2012) show that actual migration flows among 439 German districts (the NUTS 3 regions) are positively affected by the similarity of dialects that were prevalent in the source and destination areas more than 120 years ago. In the absence of such dialects, which are seldom used today, internal migration in Germany would be almost 20 per cent higher than what it is. In the same vein, Dahl and Sorenson (2010) find that Danish scientists and engineers, who exhibit a more substantial sensitivity to wage differences than other Danish workers, have even stronger preferences for living close to their family and friends. Further evidence of the low mobility of workers is provided by Bosquet and Overman (2015). Using the British Household Panel Survey that involved 32, 380 individuals from 1991 to 2009, these authors observe that 43.7 per cent of workers worked only in the area where they were born. Among the unskilled workers, this share grows to 51.7 per cent but drops to 30.5 per cent for workers having a college degree. Such low lifetime mobility provides empirical evidence that migration costs are an important determinant of the space-economy. Furthermore, 44.3 per cent of the panel retirees live where they were born, revealing a high individual degree of attachment to their birthplace. Such studies suggest that labour markets operate at a local level, implying that even sizable wage differences between regions can persist for long periods of time.

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

To sum up, the transport of (some) goods remains costly, while many services used by firms and households are nontradable. Moreover, we will see that proximity remains critical for the diffusion of some information. European people are sticky; this means that the model widely used in the US to study urban and regional growth, which relies on the perfect mobility of people and the search for amenities, has very limited application within the EU, not to say within all European countries. These facts have a major implication for the organization of the (European) economic space:

The second law of spatial economics: The world is not flat because what happens near to us matters more than what happens far from us.

Combining the first and second laws of spatial economics leads us to formulate what we see as the fundamental trade-off of spatial economics:

The spatial distribution of activities is the outcome of a trade-off between different types of scale economies and the costs generated by the transfer of people, goods, and information.

We may thus already conclude that high transport costs promote the dispersion of economic activities, while strong increasing returns act as an agglomeration force, and the other way round. This trade-off is valid on all spatial scales (city, region, country, and continent), which makes it a valuable analytical tool. We will return to this in the next two sections.

At the interregional level, locations are aggregated into subnational units that are distant from each other. Regardless of what is meant by a region, the concept is useful if, and only if, a region is part of a broader network through which various types of interactions occur. In other words, any meaningful discussion of regional issues requires at least two regions in which economic decisions are made. Hence, space is the substratum of activities, but land is not a fundamental ingredient of regional economics. Furthermore, as repeatedly stressed by Ohlin (1967), if we do not want the analysis to be confined to trade theory, we must also account explicitly for the mobility of agents (firms and/or consumers) and for the existence of transport costs in trading commodities. However, how well a region does also depends on the functioning of its local markets and institutions. The surge of *new economic geography* (NEG) has allowed us to rethink regional economics by combining the trade of goods and the mobility of production factors. In NEG, a region is assumed to be dimensionless and is described by a node in a transport network. The objective of *regional economics* is then to study the distribution of activities across a regional system. Figure 8.1 shows the geographical distribution of the GDP per capita per NUTS 3 region in the EU. We note striking differences across countries but also within countries. Understanding these differences and what policies make sense is one of the principal motivations for this survey.

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

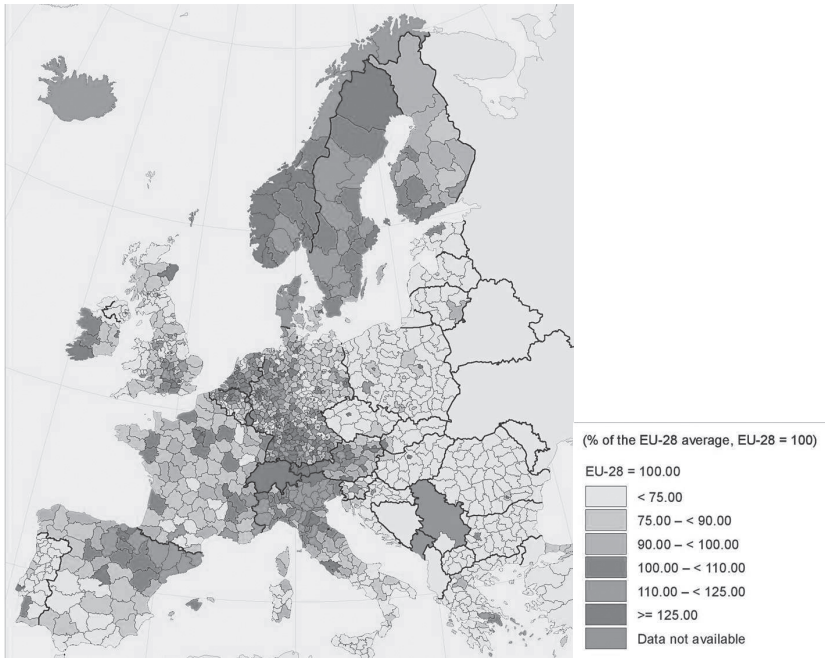

Figure 8.1 Geographical distribution of the GDP per capita per NUTS 3 region in the EU (Eurostat, 2015b).

Before proceeding, observe that the persistence of sizeable regional differences does not provide evidence of a lack of economic integration. Even in the world's largest and most integrated economy, 'local labour markets in the US are characterized by enormous differences in worker earnings, factor productivity and firm innovation' and these differences do not seem to go away (Moretti, 2011).

8.3 The Drivers of Regional Agglomeration

The EU has a wide diversity of cultures and a wide range of incomes at the interregional level. Cultural diversity is an asset that has its costs and benefits, but sizable income differences are a source of concern. Article 158 of the Treaty on European Union states that 'the Community shall aim at reducing disparities between the levels of development of the various regions and the backwardness of the least favoured regions or islands, including rural areas'. European integration is supposed to lead to the convergence of income levels across countries through more intense trade links. However, this process is slow and may be

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

accompanied by widening interregional income gaps despite EU regional policy efforts.[1] The lack of regional convergence may lead to cohesion problems that, when combined with cultural differences, can contribute to secessionist tendencies and threaten the future both of countries and of their membership in the EU. Whether or not there is convergence across the European regional system remains a controversial issue that also raises various unsuspected methodological difficulties (Magrini, 2004).

The idea of spatial interaction is central to regional economics. Broadly defined, spatial interaction refers to a wide array of flows subject to various types of spatial frictions. Examples of these flows include traded goods, migration, capital, interregional grants, remittances, as well as the interregional transmission of knowledge and business-cycle effects. The bulk of NEG has been restricted to the movement of goods and production factors. NEG remains in the tradition of trade theory as it focuses on exchanges between regions to explain why some regions fare better than others. Furthermore, NEG models regions as dimensionless economies without land. In contrast, an approach that would build on urban economics would rather choose to focus on the internal functioning of a region. Both approaches are legitimate, but a full-fledged model of the regional system taking both into account is still missing.

The economic performance of regions is affected not only by their industrial mix and their relative position in the web of relations, but also by the interregional and international mobility of commodities and production factors (e.g., capital and labour). In particular, lowering transport and trade costs changes the incentives for both firms and workers to stay put or move to another location. Therefore, to assess the full impact of market integration and the monetary union, it is crucial to have a good understanding of how firms and workers react to lower trade and transport costs. In this respect, it should be stressed that European policy-makers often overlook the fact that market integration affects the locational choices of firms and households. In particular, as will be seen, NEG highlights the fact that a rising mobility of goods and people does not necessarily reduce spatial inequality. Even though regional development agencies typically think of spatial inequality as '*temporary disequilibrium*' within the economy, stable spatial equilibria often display sizable and lasting differences in income and employment, a fact that agrees with anecdotal evidence. Furthermore, we will see that regional disparities need not be bad because they can be the geographical counterpart of greater efficiency and stronger growth.

On interregional and international scales, accessibility to spatially dispersed markets drives the location of firms; this has long been recognized in both spatial economics and regional science (Fujita and Thisse, 2013). Accessibility is itself measured by all the costs generated by the various types of spatial frictions that economic agents face in the exchange process. In the case of goods

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

and services, these frictions are called *trade costs*.[2] Spulber (2007) refers to them as 'the four Ts': (i) *transaction costs* that result from doing business at a distance due to differences in customs, business practices, as well as political and legal climates; (ii) *tariff and nontariff costs* such as different pollution standards, anti-dumping practices, and the massive number of regulations that still restrict trade; (iii) *transport costs* per se because goods have to reach their destination, while many services remain nontradable; and (iv) *time costs* because, despite the Internet and video-conferencing, there are still communication barriers across dispersed distribution and manufacturing facilities that slow down reactions to changes in market conditions. Because they stand for the cost of coordinating and connecting transactions between the supplier's and customer's locations, trade costs are crucial to the global firm and therefore are likely to stay at centre stage. The relative importance of the 'four Ts' obviously varies enormously from one sector to another, from one activity to another, from one commodity to another.

Anderson and van Wincoop (2004) provide a detailed estimate of trade costs, concluding that these costs would climb to approximately 170 per cent of the average mill price of manufactured goods, but the variance across goods is high. This estimate can be broken down as follows: 55 per cent internal costs, which include all logistics costs; and 74 per cent international costs ($1.7 = 1.55 \times 1.74 - 1$). International costs in turn are broken down as 21 per cent for transport costs and 44 per cent for costs connected with border effects ($1.74 = 1.21 \times 1.44$). Tariff and nontariff barriers account for 8 per cent of the border effects (exceptionally, this is 10 or 20 per cent in the case of developing countries); language difference, 7 per cent; currency difference, 14 per cent; and other costs, including information 9 per cent (all in all, $1.44 = 1.08 \times 1.07 \times 1.14 \times 1.09$). Therefore, it is not an exaggeration to say that the share of trade costs in the consumer price of several manufactured goods remains high. Note that there are also big differences from one trading area to another. For example, Head and Mayer (2004) convincingly argue that North American integration is significantly deeper than European integration.

8.3.1 The Home-Market Effect

The neoclassical theory of the mobility of production factors and goods predicts a market outcome in which production factors receive the same reward regardless of the place of operation. Indeed, when each region is endowed with the same production function that exhibits constant returns to scale as well as a decreasing marginal productivity, capital responds to market disequilibrium by moving from regions where it is abundant relative to labour and receives a lower return towards regions where it is scarce and receives a higher return. If the price of consumption goods were the same everywhere (perhaps because

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

obstacles to trade have been abolished), the marginal productivity of both capital and labour in equilibrium would also be the same everywhere due to the equalization of capital – labour ratios. Therefore, the free mobility of goods and capital would guarantee the equalization of wages and capital rents across regions and countries. In this case, the size of markets would be immaterial to people's welfare.

However, we are far from seeing such a featureless world. To solve this contradiction, NEG takes a radical departure from the standard setting. NEG assumes that the main reason why there is no convergence is that firms do not operate under constant returns but under *internal* increasing returns. This point was made by Krugman (1980) in a paper now famous because it highlights how *market size* and *market accessibility* interact to determine the location of an industry. The idea that size matters for the development of a region or country was emphasized by the economic historian Pollard (1981) for whom 'it is obviously harder to build an industrial complex without the solid foundation of a home market'. In contrast, economic integration and regional trade agreements lower the importance of domestic markets and allow small regions and countries to supply larger markets.

Both economists and geographers agree that a large market tends to increase the profitability of firms established there. The idea is that locations with good access to several markets offer firms a greater profit because these locations allow firms to save on transport costs and lower their average production cost by selling more. In sum, firms would seek locations with the highest market potential where demand is high and transport costs are low. Most empirical works use the concept of *market potential*, introduced by the American geographer Harris (1954) and defined as the sum of regional GDPs weighted by the inverse of the distance to the region in question where the sum includes the region itself and its internal distance as a reduced-form expression derived from general equilibrium trade theory. Econometric studies suggest that market potential is a powerful driver of increases in income per capita (Mayer, 2008). In other words, larger and/or more centrally located regions or countries are, on average, richer than regions or countries with small local markets and few neighbours or neighbours that are also small.

Nevertheless, as firms set up in the large regions, competition is also heightened, thereby holding back the tendency to agglomerate. Indeed, revisiting Hotelling's (1929) pioneering work, d'Aspremont et al. (1979) show that *spatial separation allows firms to soften price competition*. However, by relaxing competition, product differentiation permits firms to seek the most accessible location. Consequently, the interregional distribution of firms producing a tradable good is governed by two forces that pull in opposite directions: the agglomeration force generated by firms' desire for market access, and the dispersion force generated by firms' desire to avoid market crowding. Thus, the

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

equilibrium distribution of firms across regions can be viewed as the balance between these two forces.

The intensity of the agglomeration force decreases with transport costs, whereas the dispersion force gets stronger through tougher competition between regions. Although it is the balance of these forces that determines the shape of the spatial economy, there is no clear indication regarding their relative intensity as transport costs decrease. This is why the main questions that NEG addresses keep their relevance: When do we observe an agglomerated or a dispersed pattern of production at the interregional level? What is the impact of decreasing transport and trade costs on the intensity of the agglomeration and dispersion forces operating at that spatial scale?

Location and Market Size

The standard model involves two regions (North and South) and two production factors (capital and labour). The global economy is endowed with K units of capital and L units of labour. Each individual is endowed with one unit of labour and K/L units of capital. Capital is mobile between regions and capital owners seek the higher rate of return; the share $\lambda \geq 1/2$ of capital located in the North is endogenous. Labour is immobile between regions but perfectly mobile between sectors; the share of workers located in the North is exogenous and equal to $\theta \geq 1/2$. Both regional labour markets are perfect. Capital and labour are used by firms that produce a CES-differentiated product under increasing returns and monopolistic competition (Dixit and Stiglitz, 1977). Let $f > 0$ be the fixed capital requirement and $c > 0$ the marginal labour requirement needed for a firm to enter the market and produce one variety of the differentiated good. Capital market clearing implies that the number of firms is exogenous and given by K/f. Finally, shipping the differentiated good between the two regions is costly.

The above system of push and pull reaches an equilibrium when the capital return is the same in both regions. In this event, *the North hosts a more-than-proportionate share of firms*, a result that has been labeled the '*home-market effect*' (HME).[3] Since the North is larger in terms of population and purchasing power, it seems natural that North should attract more firms than the South. What is less expected is that the initial size advantage is magnified, that is, the equilibrium value of λ exceeds θ. What the HME shows is that the market-access effect dominates the market-crowding effect. Since $(\lambda - \theta)K > 0$ units of capital move from the South to the North, capital does not flow from the region where it is abundant to the region where it is scarce.

How does a lowering of interregional transport costs affect this result? At first glance, one could expect the market-access effect to be weaker when transport costs are lower. In fact, the opposite holds true: more firms choose to set up in the North when it gets cheaper to trade goods between the two regions.

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

This somewhat paradoxical result can be understood as follows. On the one hand, lower transport costs makes exports to the smaller market easier, which allows firms to exploit their scale economies more intensively by locating in the North; on the other hand, lower transport costs also reduce the advantages associated with geographical isolation in the South where there is less competition. These two effects push towards more agglomeration, implying that, as transport costs go down, the smaller region becomes deindustrialized to the benefit of the larger one. The HME is thus prone to having unexpected implications for transport policy: *by making the transport of goods cheaper in both directions, the construction of new infrastructure may induce firms to pull out of the smaller region*. In other words, connecting lagging regions to dynamic urban centres may weaken their industrial base. This result may come as a surprise to those who forget that highways run both ways. What is more, the intensity of competition in domestic markets matters for trade. Since large markets tend to be more competitive, penetrating such markets is more difficult than exporting to small regions, making the former regions even more attractive than the latter. But how robust is the HME?

Wages and Market Size

Although it is convenient to assume equal wages across regions because this allows the impact of falling transport costs to be isolated, the assumption clashes with anecdotal evidence. How wages vary with firms' location is best studied in a full-fledged general equilibrium model where wages are endogenous. As firms congregate in the larger region, competition in the local labour market intensifies, which should lead to a wage hike in North. Since consumers in the North enjoy higher incomes, local demand for the good rises and this makes the North more attractive to firms located in the South. However, the wage hike associated with more firms establishing in the North generates a new dispersion force, which lies at the heart of many debates regarding the deindustrialization of developed countries, that is, their high labour costs. In such a context, firms are induced to relocate their activities to the South when the lower wages in the South more than offset the lower demand. Takahashi et al. (2013) have shown that the equilibrium wage in the North is greater than the equilibrium wage in the South. Furthermore, the HME still holds. In other words, although the wages paid in the North exceed those paid in South, market access remains critical when determining the location of firms.

Furthermore, if the size of the larger region grows through the migration of workers from the South to the North, the interregional wage gap widens. Therefore, *fostering the mobility of workers could well exacerbate regional disparities*. Nevertheless, Takahashi et al. (2013) showed that the magnification of the HME discussed above no longer holds: as transport costs steadily decrease, both the equilibrium wage and manufacturing share first rise and

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

then fall because competition in the larger labour market becomes very strong. Despite this caveat, market integration and factor mobility favour the agglomeration of activities within a small number of large regions.

It is commonplace in macroeconomics and economic policy to think of unemployment as a national problem, the reason being that labour market institutions and demographic evolutions are often country-specific. Yet empirical evidence reveals the existence of a strong correlation between high unemployment rates and a low GDP per capita, and the other way round, across regions belonging to the same EU country. This should invite policy-makers to pay more attention to the regional aspects of unemployment. In particular, is higher interregional labour mobility the right solution for large regional employment disparities? Not necessarily. As migrants get absorbed by the labour market of the core region, the agglomeration economies discussed in the companion chapter come into play, which reduces the number of job seekers. Such a scenario is more likely to arise when migrants are skilled. In contrast, the opposite evolution characterizes the lagging region, which loses its best workers. Epifani and Gancia (2005) illustrate this contrasting pattern by introducing job search frictions à la Pissarides in a standard NEG set-up and conclude that 'migration from the periphery to the core may reduce unemployment disparities at first, but amplify them in the long run'. This result clashes with the widespread idea that geographical mobility is the solution to regional unemployment disparities. Even though it would be daring to draw policy recommendations from a single paper, it is clear that more research is needed to fully understand the impact of labour mobility on the functioning of local labour markets when market size and agglomeration economies are taken into account.

Heterogeneous Firms

The evidence is mounting that firms differ vastly in productivity. This is reflected in their ability to compete in the international marketplace. For example, Mayer and Ottaviano (2007) observe that the top 1 per cent of European exporters account for more than 45 per cent of aggregate exports, while the top 10 per cent of exporting firms account for more than 80 per cent of aggregate exports. In short, a few firms are responsible for the bulk of exports. Having such numbers in mind, it is thus legitimate to ask what the HME is when firms are heterogeneous and also when they are, or are not, sorted out across regions according to their productivity. So, it is legitimate to ask what the HME becomes when firms are heterogeneous.

Heterogeneous workers are sorted between cities along educational lines (see Chapter 9). A comparable process is at work in the case of heterogeneous firms: *the more productive firms locate in the larger region*, whereas the less productive firms seek protection against competition by setting up in the smaller region (Nocke, 2006). Furthermore, despite the greater competition in the North, the

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

HME still holds. Nevertheless, the mechanism that selects firms differs from the sorting of workers. Indeed, the gathering of the more productive firms renders competition very tough in the North, which leads inefficient firms to locate far apart to avoid the devastating effects of competition with efficient firms. This sparks a productivity gap between regions, which is exacerbated when the size difference between regions increases. Using US data on the concrete industry, Syverson (2004) observes that inefficient firms barely survive in large competitive markets and tend to leave them. This result is confirmed by the literature that follows Syverson.

To sum up, *large markets tend to offer more and better opportunities to firms and workers*.

Care is Needed

Can the HME help explain strong regional disparities? First of all, the above results were obtained using specific models so their robustness remains an open question. Second, the share of the manufacturing sector has shrunk dramatically in developed economies. So one may wonder what the HME becomes when we consider the location of nontradable services. In this case, the HME still holds if the North is sufficiently large to overcome the competition effect. Otherwise, the larger region no longer provides a sufficiently big outlet to host a more-than-proportionate share of firms. In this case, the smaller region accommodates a larger share of firms (Behrens, 2005).

Third, and last, the HME is studied in a two-region setting. Unfortunately, it cannot readily be extended to multi-regional set-ups because there is no obvious benchmark against which to measure the '*more-than-proportionate*' share of firms. A multi-regional setting brings about a new fundamental ingredient – the variability in regions' accessibility to spatially dispersed markets. In other words, the relative position of a region within the network of exchanges (which also involves cultural, linguistic, and political proximity) matters. Any global (local) change in this network, such as market integration or the construction of major transport links, is likely to trigger complex effects that vary in non-trivial ways with the properties of the graph representing the transport network (Behrens and Thisse, 2007). For example, in a multi-regional setting, the greater specialization of a few regions in one sector does not necessarily mean that this sector becomes more agglomerated, and vice versa. Therefore, it is hardly shocking that empirical evidence regarding the HME is mixed (Davis and Weinstein, 2003, Head and Mayer, 2004).

However, intuitively, it is reasonable to expect the forces highlighted by the HME to be at work in many real-world situations.[4] But how can we check this? There are two possible ways. First, since there is no hope of deriving general results for multi-regional economies, it is reasonable to try to solve numerically spatial general equilibrium models where transport networks are selected

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

randomly. For this, one needs a mathematical framework that is tractable but yet rich enough to analyze meaningful effects. Working with a NEG model that encompasses asymmetric regions, costly trade, and transport tree-networks that are generated randomly, Barbero et al. (2015) confirm that local market size (measured by population) and accessibility (measured by centrality in the trading network) are crucial in explaining a region's wage; the authors also confirm that local market size (measured by industry expenditure share) explains well the location of firms. Using Spanish data and computed transport costs, Barbero et al. (2015) find that the model is good at predicting the location of industries but less accurate concerning the spatial pattern of wages. The authors also observe that, after three decades of major road investments, the distribution of industries had not changed much in Spain. This might suggest that, once a few key connections exist, the supply of transport links obeys the law of decreasing returns.

The second method is to study empirically the causality between market access and the spatial distribution of firms. There is plenty of evidence suggesting that market access is associated with firms' location, higher wages, and employment. Starting with Redding and Venables (2004), various empirical studies have confirmed the positive correlation between the economic performance of territories and their market potential. Redding and Sturm (2008) exploit the political division of Germany after World War II as a natural experiment to show how the loss of market access for cities in West Germany located close to the border made these cities grow much less. After a careful review of the state of the art, Redding (2011) concludes that 'there is not only an association but also a causal relationship between market access and the spatial distribution of economic activity'. For example, one of the more remarkable geographical concentrations of activities is what is known as the '*manufacturing belt*' in the US. This 'belt' accommodated around four-fifths of the US manufacturing output for a century or so within an area that was one-sixth of the country's area. Klein and Crafts (2012) conclude that 'market potential had a substantial impact on the location of manufacturing in the USA throughout the period 1880–1920 and …was more important than factor endowments'. In the same vein, Head and Mayer (2011) summarize their analysis of the relationship between market proximity and economic development over 1965–2003 by saying that 'market potential is a powerful driver of increases in income per capita'.

All of this only seems a paradox: inexpensive shipping of goods makes competition tougher, thus firms care more about small advantages than they did in a world in which they were protected by the barriers of high transport costs. In other words, *even at the interregional level, proximity matters*, but the reasons for this are not the same as those discussed in Chapter 9. However, both sets of results hinge on the same principle: *small initial advantages may be*

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

translated into large ex post advantages once firms operate under increasing returns.

The HME explains why large markets attract firms. However, this effect does not explain why some markets are bigger than others. The problem may be tackled from two different perspectives. First, the two regions are supposed to be the same size and the internal fabric of each region (e.g., the magnitude of agglomeration economies) determines the circumstances in which a region accommodates the larger number of firms. Second, workers are allowed to migrate from one region to the other, thus leading to some regions being larger than others. The former case – when the two regions are a priori identical – is studied below, while the latter case is investigated in Section 8.3.3 because the mobility of labour generates effects that differ from those observed under the mobility of capital.

8.3.2 *Agglomeration Economies and the Emergence of Asymmetric Clusters*

According to Porter (1998), the formation of industrial clusters depends on the relative strength of three distinct forces: the size of intrasectoral agglomeration economies, the intensity of competition, and the level of transport costs. Despite the existence of a huge empirical – and inconclusive – literature devoted to industrial clusters, how the three forces interact to shape the regional economy has been neglected in NEG. This is probably because working with a model that accounts for the main ingredients of urban economics and NEG seems out of reach. Yet the formation of clusters can be studied by adopting a 'reduced-form' approach in which a firm's marginal production cost in a region decreases with the number of firms locating in the region. In doing this, one captures the effect of agglomeration economies and can study how agglomeration economies operating at the local level interact with the dispersion force generated by market competition in the global economy through lower trade costs (Belleflamme et al., 2000). In a spatial equilibrium, firms earn the same profits. However, if firms observe that one region offers higher potential profits than the other, they want to move to that region. In other words, the driving force that sustains the relocation of firms is the profit differential between the North and the South.

To show why and how a hierarchy of clusters emerges, we look at the interplay among the above three forces as a *symmetry-breaking device*. Therefore, we start with a perfectly symmetric set-up in which firms and consumers are evenly dispersed between the North and the South. When trade costs start decreasing, trade flows grow but, in the absence of agglomeration economies, firms stay put because spatial separation relaxes competition between firms. Things are very different when agglomeration economies are at work. In this

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

case, when trade costs fall enough, some firms choose to produce in the North, say rather than in the South in order to benefit from a lower marginal cost while maintaining a high volume of export. As trade costs keep decreasing, a growing number of firms choose to set up in the North where the marginal cost decreases further. Note that firms tend to gather in one region despite the fact that the two markets where they sell their output are the same size. What now drives firms' agglomeration is no longer the size of the product market but the endogenous level of agglomeration economies.

But where does agglomeration occur? Will it be in the North or in the South? Consider an asymmetric shock that gives a region a small initial advantage. If this shock remains fixed over a long period, firms will attune their behaviour accordingly. The region benefiting from the shock, however small, will accommodate the larger cluster. Hence, *regions that were once very similar may end up having very different production structures* as market integration gets deeper. Once more, lowering trade costs drives the economy toward more agglomeration in one region at the expense of another.

Are growing regional disparities necessarily bad in this context? The answer is no. A planner whose aim is to maximize global efficiency sets up more asymmetric clusters than the market delivers. To explain, at the first-best optimum prices are set at the marginal cost level while locations are chosen to maximize the difference between agglomeration economies and transport costs. In contrast, at market equilibrium, firms take advantage of their spatial separation to relax price competition and do not consider the positive externalities associated with their location decision. So the optimal configuration tends to involve a more unbalanced distribution of firms than the market outcome. If agglomeration economies become increasingly important in some sectors, their uneven geographical distribution need not signify a wasteful allocation of resources. On the contrary, *the size of the clusters could well be too small*. However, the region with the larger cluster benefits from lower prices through larger agglomeration economies, more jobs, and a bigger fiscal basis.

8.3.3 The Core–Periphery Structure

The mobility of capital and the mobility of labour do not obey the same rules. First, while the movement of capital to a region brings with it production capability, the returns to capital do not have to be spent in the same region. In contrast, when workers move to a new region, they take with them *both their production and consumption capabilities* (putting aside remittances). As a result, migration affects the size of the labour and the product markets in both the origin and the destination regions. Second, while the mobility of capital is driven by differences in nominal returns, workers care about their *real wages*. In other words, differences in costs of living matter to workers but not to capital owners.

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

The difference in the consequences of capital and labour mobility is the starting point of Krugman's celebrated 1991 paper that dwells on the idea that the interregional economy is replete with *pecuniary externalities* generated by the mobility of workers. Indeed, when some workers choose to migrate, their move affects the welfare of those who stay behind because migration affects the size of the regional product and labour markets. These effects have the nature of pecuniary externalities because they are mediated by the market, but migrants do not take them into account when making their decisions. Such effects are of particular importance in imperfectly competitive markets as prices fail to reflect the true social value of individual decisions. Hence, studying the full impact of migration requires a full-fledged general equilibrium framework, which captures not only the interactions between product and labour markets, but also the double role played by individuals as workers and consumers.

To achieve his goal, Krugman (1991) considers the classical $2 \times 2 \times 2$ setting of trade theory. There are two goods, two types of labour, and two regions. The first type of labour (workers) is mobile and the only input in the first (manufacturing) sector, which operates under increasing returns and monopolistic competition; shipping the manufactured good is costly. The second type of labour (farmers) is immobile and the only input in the second (farming) sector, which produces a homogeneous good under constant returns and perfect competition; shipping the agricultural good incurs no cost. What drives the agglomeration of the manufacturing sector is the mobility of workers. For this, Krugman considers a setting in which both farmers and workers are symmetrically distributed between the North and the South and asks when this pattern ceases to be a stable spatial equilibrium.

Two main effects are at work: one involves firms, and the other workers. Assume that the North grows slightly bigger than the South. At first, this increase in market size leads to a higher demand for the manufactured good, thus attracting more firms. The HME implies that the hike in the number of firms is more than proportional to the increase in market size, thus pushing nominal wages upward. In addition, the presence of more firms means that a greater number of varieties are produced locally, so prices are lower in the North because competition there is tougher. As a consequence, real wages rise so that the North should attract a new flow of workers. Therefore, there is *circular cumulative causation* à la Myrdal in which these two effects reinforce each other. This snowball effect seems to lead inevitably to the agglomeration of the manufacturing sector in the North, which then becomes the core of the global economy.

But the snowball may not form. Indeed, the foregoing argument ignores several other effects triggered by the migration of workers. On the one hand, the increased supply of labour in the North tends to push wages down. On the other hand, since new workers are also consumers, there will be a hike in local

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

demand for the manufactured good, which leads to a higher demand for labour. But this is not yet the end of the story. As more firms enter the local market, there is increased competition to attract workers, so the final impact of migration on nominal wages is hard to predict. Likewise, there is increased competition in the product market, as well as greater demand. Combining these various effects might well lead to a '*snowball meltdown*', which results in the spatial dispersion of firms and workers.

Krugman's (1991) great accomplishment has been to integrate all these effects within a single framework and to determine precisely the conditions under which the above prediction holds or not. Starting from an arbitrarily small difference between regions, Krugman singles out the cases in which there is agglomeration or dispersion of the manufacturing sector. He shows that the value of transport costs is again the key determining factor. If transport costs are sufficiently high, the interregional shipment of goods is low. In this event, firms focus on regional markets. Thus the global economy displays a symmetric regional pattern of production. In contrast, when transport costs are sufficiently low, then all manufacturers will concentrate in the North; the South will supply only the agricultural good and will become the *periphery*. In this way, firms are able to exploit increasing returns by selling more in the larger market without losing much business in the smaller market. Again, lowering trade costs fosters the gathering of activities. The core–periphery model therefore allows for the possibility of *convergence* or *divergence* between regions, whereas the neoclassical model based on constant returns and perfect competition in the two sectors predicts only convergence. Consequently, Krugman presents a synthesis of the polarization and neoclassical theories. His work appeals because the regional disparities associated with the core–periphery structure emerge as a *stable equilibrium* that is the involuntary consequence of decisions made by a large number of economic agents pursuing their own interests.[5]

Despite its great originality, the core–periphery model has several shortcomings. The following list, while not exhaustive, covers a fair number of issues. (i) The model overlooks the various congestion costs and agglomeration economies generated by the concentration of activities, discussed in Chapter 9. (ii) It only accounts for two sectors and two regions. (iii) The agricultural sector is given a very restricted role, its job being to guarantee the equilibrium of the trade balance. Along the same line, it is hard to see why trading the agricultural good costs nothing in a model seeking to determine the overall impact of trade costs. All these features have attracted a lot of attention, but the '*dimensionality problem*' is the most challenging one.

Having said that, we must stress the work by Helpman (1998) who argues that decreasing freight costs may trigger the dispersion, rather than the agglomeration, of economic activities when the dispersion force lies in the supply of nontradable services (housing) rather than immobile farmers. In this case,

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

various congestion and market-crowding effects put a brake on the agglomeration process, and thus *Krugman's prediction is reversed*. The difference in results is easy to understand. Commuting and housing costs rise when consumers join the larger region/city, which strengthens the dispersion force. Simultaneously, lowering transport costs facilitates interregional trade. By combining these two forces, we see why dispersion arises. In other words, *land use appears to be a major dispersion force in the making of the space-economy*.[6] By neglecting the fact that the agglomeration of activities typically materializes in the form of cities where competition for land acts as a strong dispersion force, the core–periphery model remains in the tradition of trade theory. Therefore, conclusions drawn from this model are, at best, applicable only to very large areas.[7]

The econometric analysis undertaken by Crozet (2004), together with the observations made in Section 8.2, suggests that the low mobility of European workers makes the emergence of a Krugman-like core–periphery structure within the EU very unlikely. Therefore, moving beyond the Krugman model in search of alternative explanations appears to be warranted in order to understand the emergence of large industrial regions in economies characterized by a low spatial mobility of labour – such as the EU. A second shortcoming of the core-periphery model is that it ignores the importance of intermediate goods. Yet the demand for consumer goods does not account for a very large fraction of firms' sales, often being overshadowed by the demand for intermediate goods.[8]

8.3.4 Input–Output Linkages and the Bell-Shaped Curve of Spatial Development

The agglomeration of economic activities also arises in contexts in which labour mobility is very low, as in most European countries. This underscores the need for alternative explanations of industrial agglomeration. One strong contender is the presence of *input–output linkages between firms*: the output of one firm can be an input for another, and vice versa. In this case, the entry of a new firm in a region not only increases the intensity of competition between similar firms; it also increases the market of upstream firm-suppliers and decreases the costs of downstream firm-customers. This is the starting point of Krugman and Venables (1995).

Their idea is beautifully simple and suggestive: *the agglomeration of the final sector in a particular region occurs because of the concentration of the intermediate industry in the same region, and conversely*. Indeed, when firms belonging to the final sector are concentrated in a single region, the local demand for intermediate inputs is very high, making this region very attractive to firms producing these intermediate goods. Conversely, because intermediate goods are made available at lower prices in the core region, firms producing final goods

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

also find that region very attractive. Thus, a cumulative process may develop that leads to industrial agglomeration within the core region.

In this alternative setting, new forces arise. Indeed, if firms agglomerate in a region where the supply of labour is inelastic, then wages must surely rise. This in turn has two opposite effects. On the one hand, *consumers' demand for the final product increases because they have a higher income*. This is again a market expansion force, now triggered by higher incomes rather than larger populations. On the other hand, such wage increases also generate a dispersion force. When the wage gap between the core and the periphery becomes sufficiently large, some firms will find it profitable to relocate to the periphery, even though the local demand for their output is lower than in the core. This is especially true when transport costs are low, because asymmetries in demand will then have a weaker impact on profits.

The set of equilibrium patterns obtained in the present setting is much richer than in the core–periphery model. In particular, if a deepening of economic integration triggers the concentration of industrial activities in one region, then beyond a certain threshold, an even deeper integration may lead to a reversal of this tendency. Some firms now relocate from the core to the periphery. In other words, *the periphery experiences a process of re-industrialization and, simultaneously, the core might start losing firms, thus becoming deindustrialized*. As Fujita et al. (1999) put it, 'declining trade costs first produce, then dissolve, the global inequality of nations'.

Therefore, economic integration would yield a *bell-shaped curve of spatial development*, which describes a rise in regional disparities in the early stages of the development process, and a fall in later stages (Williamson, 1965, Puga, 1999). Such a curve may be obtained in several extensions of the core-periphery model – surveyed in Fujita and Thisse (2013) – and seems to be confirmed by several empirical and historical studies.[9] However, owing to differences in data, time periods, and measurement techniques, it is fair to say that the empirical evidence is still mixed (Combes and Overman, 2004). Furthermore, this self-correcting effect can take too long in the face of some regions' urgent economic and social problems and the time horizon of policy-makers, which leads them to look for policies whose effects are felt more rapidly.

Note that the following coordination failure may prevent the redistribution of activities: many prices are not known in advance in the South. Lack of adequate information may then prevent the development of a network of service and intermediate goods suppliers, which leads to a vicious circle and persistent underdevelopment. In the presence of external effects, this problem is particularly acute. One solution is to have an agent who 'internalizes' the various costs and benefits arising during the first stages of the take-off process and who plays an entrepreneurial role facilitating individual decisions, so that a cluster in the South can form en masse.

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

8.3.5 Communication Costs and the Relocation of Plants

A major facet in the process of globalization is the *spatial fragmentation* of a firm associated with vertical investments. Vertical investments arise when firms choose to break down their production process into various stages spread across different countries or regions. Specifically, the modern firm organizes and performs discrete activities in distinct locations, which together form a *supply chain* starting at the conception of the product and ending at its delivery. This spatial fragmentation of the firm aims to take advantage of differences in technologies, factor endowments, or factor prices across places. We now turn our attention to this problem.

Besides transport costs, spatial separation generates another type of spatial friction, namely '*communication costs*'. Indeed, coordinating activities within the firm is more costly when the headquarters and its production plants are physically separated because the transmission of information remains incomplete and imperfect. Furthermore, more uncertainty about production plants' local environment is associated with conducting a business at a distance. Again, this implies higher coordination costs, hence higher communication costs between the headquarters and its plants. In the same vein, monitoring the effort of a plant manager is easier when the plant is located near the headquarters than across borders. Lower communication costs make the coordination between headquarters and plants simpler and therefore facilitate the process of spatial fragmentation.

For the international/interregional fragmentation of firms to arise, the intrafirm coordination costs must be sufficiently low so the operation of a plant at a distance is not too expensive; at the same time, transport costs must decrease substantially to permit the supply of large markets at low delivery costs from distant locations. To make low-wage areas more accessible and attractive for the establishment of their production, firms need the development of new information and communication technologies, as well as a substantial fall in trade costs. In this case, a certain number of firms choose to go *multinational*, which means that their headquarters are located in prosperous areas where they find the skilled workers they need and their plants are set up in low-wage areas, whereas the other firms remain spatially integrated (Fujita and Thisse, 2013).

Manufacturing firms started to relocate their production plants to regions where labour and land are cheaper than in large cities long ago (Henderson, 1997, Glaeser and Kohlhase, 2004). However, transport and communication costs for a long time imposed a limit to the distance at which plants could operate. The ongoing revolution in information and communication technologies freed some firms from this constraint, thus allowing them to move their plants much further away to countries where wages are a lot lower than in the peripheral regions where they used to establish their plants. Hence, the

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

following question: *Which* 'South' *can accommodate firms' activities that are being decentralized?*

8.4 Does the Market Yield Over or Under-agglomeration?

Whether there is too much or too little agglomeration is unclear. Yet speculation on this issue has never been in short supply and it is fair to say that this is one of the main questions that policy-makers would like to address. Contrary to general beliefs, the market need not lead to the over-agglomeration of activities as competition is a strong dispersion force. We have discussed above two basic mechanisms that may outweigh this force and lead to the spatial clustering of activities. The former is the home-market effect (HME), which points to the relative agglomeration of firms in the large regions. The latter is related to the joint concentration of firms and workers in a few regions to form big markets. Since the mobility of capital and labour is driven by different forces, there is no reason to expect the answer to the question 'Does the market yield over or under-agglomeration?' to be the same.

8.4.1 Does the Home-Market Effect Generate Excessive Agglomeration?

Because spatial separation relaxes price competition, everything else being equal, firms earn higher profits by locating in different geographical markets. What the HME tells us is that the size of markets may outweigh this effect, leading to the concentration of firms in a few regions. When firms move from one region to another, they impose negative pecuniary externalities on the whole economy. More precisely, firms ignore the impact of their move on product and input markets in both destination and origin regions. The social surplus is lowered because location decisions are based on relative prices that do not reflect the true social costs. However, the inefficiency of the market outcome does not tell us anything about the excessive or insufficient concentration of firms in the big regions. In fact, *the HME involves too many firms located in the larger region*. The intuition is easy to grasp. A profit-maximizing firm chooses the location that minimizes its transport costs to serve foreign markets. Therefore, since firms absorb more freight when exporting from the smaller to the larger region than vice versa, they are incentivized to locate in the larger region. Tougher competition there holds back the agglomeration process, but this dispersion force is not strong enough for a sufficiently large number of firms to set up in the smaller region. However, it is worth noting that the first-best distribution of firms still involves a share of firms exceeding the relative size of the larger region (Ottaviano and van Ypersele, 2005).

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

8.4.2 Is the Core-Periphery Structure Inefficient?

Thus far, NEG has been unable to provide a clear-cut answer to this fundamental question. However, a few results seem to show some robustness. In the core–periphery model, the market outcome is socially desirable when transport costs are high or low. In the former case, activities are dispersed; in the latter, they are agglomerated. In contrast, for intermediate values of these costs, the market leads to the over-agglomeration of the manufacturing sector (Ottaviano and Thisse, 2002). Furthermore, when transport costs are sufficiently low, agglomeration is preferred to dispersion in the following sense: people in the core regions can compensate those staying in the periphery through interregional transfers, whereas those staying in the periphery are unable to compensate those workers who choose to move to what becomes the core regions (Charlot et al., 2006). This suggests that interregional transfers could be the solution for correcting regional income disparities. It is worth stressing that such transfers do not rest here on equity considerations, but only on efficiency grounds. However, implementing such transfers, paid for by those who reside in the core regions, may be politically difficult to maintain in the long run. In addition, they may give rise to opportunistic behaviour in the periphery.

Tackling this issue from a dynamic perspective sheds additional light on the problem. It has long been argued that growth is localized, the reason being that technological and social innovations tend to be clustered while their diffusion across places would be slow. For example, Hirschman (1958) claimed that 'we may take it for granted that economic progress does not appear everywhere at the same time and that once it has appeared powerful forces make for a spatial concentration of economic growth around the initial starting points'. And Hohenberg and Lees (1985) argued similarly that, 'despite the rapid growth of urban industries in England, Belgium, France, Germany and northern Italy after 1840 or so, economic development was a spatially selective process. Some regions deindustrialized while others were transformed by new technologies'.

Fujita and Thisse (2013) revisit the core–periphery model in a set-up combining NEG and endogenous growth theory; the high-skilled, who work in the R&D sector, are mobile whereas the low-skilled, who work in the manufacturing and agricultural sectors, are immobile. These authors show that the growth rate of the global economy depends positively on the spatial concentration of the R&D sector. Furthermore, the core–periphery structure in which both the R&D and manufacturing sectors are agglomerated is stable when transport costs are sufficiently low. This result gives credence to the idea that global growth and agglomeration go hand in hand. But what are the welfare and equity implications of this geography of innovative activities? The analysis undertaken

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

by Fujita and Thisse supports the idea that the additional growth spurred by agglomeration may lead to a Pareto-dominant move: when the growth effect triggered by the agglomeration of the R&D sector is strong enough, even those who live in the periphery are better off than under dispersion.

It is worth stressing that this Pareto-optimal move does not require any interregional transfer; it is a pure outcome of market interaction. However, the gap between the unskilled who live in the core and those who live in the periphery enlarges. Put differently, the rich get richer and so may the poor, but without ever catching up. The welfare gap between the core and the periphery expands because of the additional gains generated by a faster growth spurred by the agglomeration of skilled workers. This in turn makes the unskilled residing in the core region better off, even though their productivity is the same as the productivity of those living in the periphery.

8.5 Do EU Interregional Transport Investment Policies Fulfil their Role?

This question may seem odd because the absence of good transport infrastructure is known to be one of the main impediments to trade. This is why international organizations such as the European Commission and the World Bank have financed a large number of transport projects. As the key objective of the EU is deeper market integration among member countries, the construction of big and efficient transport infrastructures was seen as a necessary step towards this goal. However, this does not mean that one should keep increasing the supply of transport infrastructure: its economic performance can be improved by selecting investments more carefully and by using the existing infrastructure better. Whether interregional transport infrastructure is beneficial in terms of welfare and whether it generates economic growth at the macroeconomic level are two different issues.

Another important question often forgotten in the debates over the interregional effects of a new transport infrastructure is that the development of new transport technologies has vastly changed the way in which distance affects transport costs. This history is briefly as follows. The long period during which all movement was very costly and risky was followed by another one during which, thanks to technological and organizational advances, ships could cross longer distances in one go, thus reducing their number of stops. On land, it was necessary to wait for the advent of the railroad for appreciable progress to occur, but the results were the same. In both cases, long-distance journeys became less expensive and no longer demanded the presence of relays or rest areas. This evolution has favoured places of origin and destination at the expense of intermediate places. In other words, *increasing returns in transport explain why*

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

places situated between large markets and transport nodes have lost many of their activities (Thomas, 2002). Having this in mind, it is hardly shocking that not much happened in those transit regions, despite the high expectations of the local populations.

The policy intervention also involved the design of pricing and regulation policies for interregional transport. All this has led to an appreciable increase in the volume of both freight and passenger transport. Nevertheless, transport policies are still formed by individual member countries. Using a NEG set-up in which transport costs between regions of the same country differ from trade costs between countries, Behrens et al. (2007) show that the welfare of a country increases when its internal transport costs are lowered because domestic firms increase their market share at the expense of foreign firms, while the foreign trading partner is affected adversely for the same reason. As a consequence, we have something like a '*fortress effect*' in that *accessing the increasingly integrated national market becomes more difficult*, which may generate conflicts of interest between member countries.

In the EU, transport policy has two main objectives. The first is to decrease trade costs as the aim of transport policy is to build the EU internal market. The second objective is to promote the economic development and structural adjustment of lagging regions. Arbitrage possibilities arising from competition and factor mobility are expected to generate greater-than-average growth in lagging regions. Having the economic engine in a higher gear would eventually make these regions reach the standard of living realized elsewhere. Where convergence does not arrive quickly, an insufficient stock of public infrastructure is often blamed. The EU and national governments have responded by pouring huge quantities of concrete in lagging regions.

The EU has sent rather mixed signals in terms of transport policy. In the first phase, the integration of markets for goods was the priority; later, the emphasis shifted to environmental and resource efficiency. As a result, the development of rail and waterways was favoured over road and air transport. Yet road freight transport in the EU remains by far the dominant mode; the EU has a very different modal split from that in the US. International freight in the EU relies on road transport for 45 per cent of traffic, on sea transport for 37 per cent, on rail transport for 11 per cent, and on inland waterways and pipeline transport for the remainder. In the US, rail transport at 41 per cent is more important than road transport (32 per cent), followed by pipeline (15 per cent), and inland waterways. International passenger transport inside the EU also has a different modal split from that in the US. The US relies on car and air transport, while the EU also relies on high-speed rail (HSR). Thus, in the US, rail has an important share of the freight market while, in Europe, rail is more important for the passenger market.

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

Assessing the benefits of transport investments is difficult both ex ante and ex post, for two reasons. First, transport investments have a multitude of effects. They reduce trade barriers and so affect the pattern of trade for both freight and for services (via lower costs for business and tourism trips). As seen above, the outcome of a transport investment is difficult to predict ex ante in a world where economic activities are increasingly footloose. Second, the effect of an investment is also difficult to evaluate ex post because there is no obvious counterfactual. A transport investment is often located where decision makers expect it to produce the largest benefits. But then it becomes unclear whether it is the transport investment itself or the favourable pre-conditions that cause the observed effects.

As performance of transport infrastructure is an empirical question, we have chosen to discuss both ex ante and ex post methods. In particular, we consider three approaches: the econometric approach, the model-simulation approach, and the case-study approach.

8.5.1 Assessing Transport Investments Using Econometric Models

In the post-Reagan period, public investments were expected to stimulate economic growth. In an influential paper, Aschauer (1989) used a reduced-form estimation and found high rates of return for public investments. This was the start of a series of macroeconomic studies that produced fairly mixed evidence about the impact of transport investments on national growth (Gramlich, 1994). Melo et al. (2013) conducted a meta-analysis of the existing empirical literature on the output elasticity of transport infrastructure. They show that the productivity effects of transport infrastructure vary substantially across industries, tend to be higher in the US than in the EU, and are higher for roads compared with other transport modes of transport. The variation in estimates of the output elasticity of transport is also explained by differences in the methods and data used in the various studies. Failing to control for unobserved heterogeneity and spurious correlations tends to result in higher values, while failing to control for urbanization and congestion levels leads to omitted variable bias. In addition, Puga (2002) highlights several pitfalls of an aggregate approach. First, it could well be that transport investments happen just because economic growth allows the government to spend more money on infrastructure, not the other way around. Second, the first links of a transport network could well be very productive, whereas the productivity of adding new links decreases strongly.

Redding and Turner (2015) develop a general equilibrium framework in the spirit of Helpman to assess the effects of transport investments on the location of production and population, as well as on variables such as wages and prices. This framework allows the authors to construct the necessary counterfactuals

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

to assess the effects of new transport investments. They find only limited evidence on the effect of interregional investments in the EU. Ahlfeldt and Feddersen (2015) study the impact of HSR on a corridor in Germany by comparing the effects on smaller towns with a HSR stop and those without such a stop. They find that, as HSR decreases the cost of human interaction but trade costs remain unchanged, this type of project has another effect on the core-periphery balance. Peripheral regions tend to experience negative effects through projects that reduce freight costs via a trade channel, as in NEG, but could benefit from HSR projects via Marshallian externalities.

Comparing the impact of transport investments in different and non-EU parts of the world, Redding and Turner find that, across a range of countries and levels of development, new transport infrastructures seem to generate similar effects. First, population density falls between 6 and 15 per cent with a doubling of the distance to a highway or railroad, while highways decentralize urban populations and, to a lesser extent, manufacturing activity. Second, different sectors respond differently to different transport modes. Another forceful piece of evidence is Faber (2014) who shows that the construction of new highways in China decreased trade costs but, as suggested by NEG, reinforced the core cities at the expense of the periphery.

One limitation of the econometric assessment approach is that transport investments are chosen in a political process, which may lead to the selection of poor investments. For example, Knight (2004) has found that, for the US Federal Highway Fund, about half of the investment money was wasted. Therefore, any econometric ex post assessment has the tough task of distinguishing between poor political selection mechanisms and the potential effects of a well-selected transport investment.

8.5.2 Assessing Transport Investments Using Model Simulations

When a reliable multi-regional simulation model is available, one can simulate the effects of transport investments and discriminate between the effects of the selection process and the productivity of a transport infrastructure. Only a handful of such models exist in the world. To this end, the European Commission has developed a spatial computable general equilibrium model (SCGE), RHOMOLO, where different policy shocks can be simulated at the regional level to obtain an ex ante impact assessment. The spatial implications of the general equilibrium approach followed in RHOMOLO have been investigated by Di Comite and Kancs (2015) who describe how the main agglomeration and dispersion forces of NEG enter the model: agglomeration is driven by increasing returns to scale, the use of intermediate inputs, and localized externalities; dispersion is driven by costly trade and locally produced varieties entering consumer utility asymmetrically (calibrated on observed trade flows). Capital

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

and labour are mobile, and vertical linkages are accounted for using regionalized international input-output matrices. The model is implemented for the 267 NUTS 2 regions of the EU and used to assess the effect of investments that reduce trade costs. The properties of this model are tested by simulating the impact of planned Cohesion Policy investments in infrastructure, whose main targets are the poorer, peripheral regions. The aim of the exercise is to isolate the effect of the different economic mechanisms identified in Section 8.3, for which three scenarios are simulated.

Scenario 1: Isolating the Effect of Capital Mobility

By switching capital mobility on and off, allowing savings in one region to be invested in other regions, the authors find that the tendency toward the equalization of the rates of return on investments spreads the growth effects of the transport investments more equally. This is the home-market effect at work: although the poorer (peripheral) regions received a larger share of the transport investment, the relocation of capital leads to more growth in other EU regions.

Scenario 2: Isolating the Effect of Labour Mobility

By switching labour mobility on and off, allowing workers to relocate where their real wages are higher according to estimated elasticities, the authors find that the region receiving the initial investment will benefit from a lower cost of living. This attracts more workers and increases the size of the region, its production, and its consumption, which should foster agglomeration. However, since consumer tastes are calibrated in each region based on the observed trade flows in the base year, the growing regions also demand more from the peripheral regions, which bids up prices and prevents a strong agglomeration effect. The cost-of-living effect is found to be stronger than the labour market-crowding effect, thus magnifying the beneficial effect of local investments and making the lagging region better off, but the effect is very localized.

Scenario 3: Isolating the Effect of Vertical Linkages

By switching interregional consumption of intermediates on and off, it can be noted that higher demand for intermediate goods in regions with improved accessibility attracts producers of intermediate goods, which lowers the production costs for the producers of the final goods. In the absence of vertical linkages, the benefits of Cohesion Policy investments are more localized. However, when vertical linkages are allowed, the productivity improvements in one region spread to all the regions using its output as an input in their productive processes. Therefore, the benefits of allocating resources to a region are felt beyond its borders.

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

These models are powerful tools to check ex ante the potential effects of different transport policies. However, they suffer from several shortcomings. First, the model is calibrated but not econometrically tested. Second, the mechanisms are so complex and the model so big that it is impossible to isolate and identify the drivers of agglomeration and dispersion when all the features are included together. Last, the way workers' mobility is modelled is critical as European workers are very sticky, while mobility habits may change over time and respond to specific policies (which are impossible to capture accurately in the model). It should also be noted that the administrative capacity of local authorities and the quality of planned investments are key determinants of the success of a policy, but these aspects cannot be captured in a general equilibrium model. For this reason, the following approach should complement the ones based on econometric analysis and model simulations.

8.5.3 Assessing Transport Investments Using Case Studies

In the late 1990s, the EU selected a priority list of transport investments – the '*Trans European Network*' investments – whose total value accounted for some €600 billion. These investment projects are the first that should receive European subsidies. In an attempt to assess the benefits of the 22 priority freight projects, Bröcker et al. (2010) developed a model in the tradition of the new trade theories with 260 European regions. In this model, firms produce a differentiated good and operate under increasing returns and monopolistic competition; interregional trade is costly while capital and labour are immobile. Since production factors are immobile, one major ingredient of NEG is missing, that is, the endogenous formation of clusters. A particular transport investment decreases transport costs between specific regions, which translates into changes in production activities, trade patterns, and ultimately the welfare level of consumers residing in different regions.

There are three main findings for this first round of EU transport priority projects (Proost et al., 2014). First, *only* 12 *of the* 22 *projects pass the cost-benefit analysis test*. Second, most projects benefit only the region where the investment takes place, so that the '*EU value added*' – or the positive spillover argument – does not seem to warrant the investment. Finally, the projects do not systematically favour the poorer regions. These findings illustrate the role of political economy factors in the selection of projects. Knight's (2004) study suggests that substantial amounts money are spent inefficiently on interregional transport infrastucture. To avoid such a waste of resources, the EU should rely on independent project assessment. There has been great progress in this area over the last decade. The group of countries with a strong tradition of independent project assessment (Netherlands, Sweden, and the UK)

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

has been widened and the methods are being refined to allow for relocation effects.

A second round of EU transport priority projects was approved in 2015. The selection of the projects is based on expert judgments, which refer to a wide range of objectives, but it is not clear how many projects would pass the cost-benefit-analysis test. In total, 276 proposals were recommended for funding.

When it comes to passenger transport, *the EU has put a strong emphasis on HSR investments*. This contrasts with the choice made in the US where air transport for medium to long-distance travel is used much more, but where HSR projects have never taken off. On average, Americans travel almost 3000 km a year by air inside the US, while the average EU citizen travels slightly more than 1000 km per a year by air inside Europe and some 200 km by HSR (Eurostat, 2015a). Both Americans and Europeans also make long-distance trips by car, but Europeans clearly have a lower demand for long-distance trips than Americans.

The EU probably opted for HSR because of the presence of strong (public) national railway companies wanting to preserve their market share. Air transport has grown robustly, and the liberalization of passenger air transport has led to lower prices, higher frequencies, and the loss of market share for rail. HSR networks require a large upfront investment in infrastructure (tracks, locomotives). Compared with air transport, HSR has high fixed costs, while infrastructure construction is almost fully subsidized. Maintenance and operation are supposed to be paid for by passenger fares. More investment subsidies are spent on rail than on roads, so it is crucial to have a good ex ante appraisal of the different transport modes.

De Rus and Nombela (2007) use standardized cost-benefit estimates to determine the level of demand that is needed to make a HSR link socially beneficial. They find that *a link needs some* 10 *million passengers a year and many new HSR links do not meet this target*. Adler et al. (2010) use a full-network model where EU passengers have the choice between HSR, air, and car for medium to long-distance trips. The reactions of the air transport sector are taken into account in order to avoid the mistake made when the Channel Tunnel was assessed without anticipating the reaction of competing ferries. When HSR has to cover all its costs, these authors have found that there will be an insufficient number of passengers for the project to be economically viable. When trips are priced at marginal cost, the HSR has a better chance of passing the cost-benefit test. But charging the marginal cost requires high government subsidies. In addition, the government must be able to pick the right project and cannot serve all regions equally. France and Spain have the largest HSR networks, and part of their network would probably not pass the cost-benefit test. The UK and the Netherlands have almost no HSR network. Finally, the EU defends HSR projects on environmental grounds, but sensitivity analysis shows that one

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

needs extremely high carbon values to make HSR better than air transport on these grounds.

8.6 Is the EU Moving to a Better Utilization of Its Existing Transport Policy?

8.6.1 Competition on Diesel Fuel Taxes Leads EU Countries to Revise Their Pricing of Road Freight

Trucks are responsible for climate damage, conventional air pollution, accidents, congestion and road damage. The main determinant of road damage is the axle weight of a truck. In Europe, trucks pay for the use of roads via excise taxes on diesel fuel but this is changing fast as a result of intense fuel tax competition. Because trucks can cover 1000 to 2000 km with a single tank of fuel, countries or regions engage in fuel tax competition. The difference in distances covered implies that tax competition is much more important for trucks than for cars. Within the EU, some small countries (Luxemburg being the most obvious example) choose a strategy of low diesel excise tax rates to make international haulers fuel up in their country, generating large excise tax revenues for these countries. This strategic behaviour has prompted the EU to negotiate a minimum level of excise taxes.

New pricing technologies have allowed countries with a lot of transit traffic, such as Switzerland (2001), Austria (2004), Germany (2005), the Czech Republic (2007), Slovakia (2010), Poland (2011), and Belgium (2016), to introduce distance-based charging. The vignette system (a low fixed annual fee) is then replaced by a kilometre tax that charges trucks much more than before.

Replacing diesel taxes by distance charges is not necessarily welfare-improving (Mandell and Proost (2016)). When a country uses distance charges, it can replace part of the diesel fuel tax by a distance charge. In this way, it undercuts the diesel tax of its neighbours and increases its revenues. As a consequence, the neighbouring countries also have to implement a distance charge if they want to preserve their tax revenues. The end result will be low diesel taxes and high distance charges. Furthermore, when passenger cars also use diesel fuel, taxes are too low for diesel cars while diesel taxes and distance charges are too high for freight transport. Accounting for the inefficient levels of taxes and charges and for the high implementation costs of distance charges, tax competition could lead to a less efficient equilibrium than the fuel tax equilibrium. So the revolution in truck taxes, which is a priori an instrument for more efficient pricing, may end up with massive tax exporting.

To some extent, the EU has anticipated that the introduction of distance charges in countries with transit freight traffic may lead to charges that are too high. The EU constitution does not allow discriminatory charges, but this

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

is no guarantee against too high truck charges in transit countries. It therefore requires distance charges for trucks to be based on external costs. This may be viewed as a principal–agent problem in which the EU is the principal, and the country is the agent with better information about external costs. For this reason, distance charges are capped by the EU on the basis of average infrastructure costs. Interestingly, this turns out to be a smart policy: when road congestion is an important share of external costs, and road building is governed by constant returns, this cap can guarantee efficient pricing, and there is no need for the regulator to know the external congestion costs (Van der Loo and Proost, 2013). Distance charges for trucks have, up to now, been used chiefly as a simple distance toll with some environmental differentiation. However, the charges can become much more efficient when they are more closely geared to the external costs such as congestion, local air pollution, and accidents. The current revolution in the pricing of trucks may pave the way for a very different charging system for cars.

Finally, we observe that this evolution in the pricing of trucks is largely a European phenomenon. In the US, the 'stealing' of fuel tax revenues from neighbouring states is avoided by a complex system of regularization payments among states, which allows the US to function as an efficient trade zone.

8.6.2 Europe Does Not Make the Best Use of Its Rail and Air Transport System

The EU is still confronted with an archaic rail and air transport system. For rail, there are powerful national regulators and powerful national companies. Rail freight activity has been more or less stable but rail passenger activity has been decreasing substantially over the last 20 years. Rail freight could play a bigger role in freight transport; its market share is 11 per cent compared with 41 per cent in the US. There are probably two reasons for this difference: the lack of consolidation among national companies, and the lack of harmonization in operation. Ivaldi and McCullough (2008) study the integration of freight activities in the private US rail market and found that this leads to an important gain in consumer surplus. In the EU, together with the lack of consolidation, there is a lack of harmonization in the rail business. Harmonization of operating standards is an extremely slow process as the national producers all want to protect their own rail and equipment market.

In the air space, similar mechanisms are at work. In the US, there is a single regulator for the management of air space while in Europe, there are 37 national, and partly regional, monopolies managing air traffic. All regional monopolies function under a cost-plus rule, but an effort is being made to shift to a price-cap system. As a result, costs are almost twice as high as they are in the US.

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

Consolidation of different air traffic control zones is possible, which should also lead to important cost reductions. However, it is blocked by the national monopolies.

8.7 What Have We Learnt?

1. Owing to the strength of market forces shaping the spatial economy, regional development is inevitably unequal. Given the first law of spatial economics, not all regions may have a large market populated by skilled workers employed in high-tech industries. To a large extent, the unevenness of regional development may be viewed as the geographical counterpart of economic growth, which is driven mainly by large and innovative cities. The cumulative nature of the agglomeration process makes the resulting pattern of activities particularly robust to various types of shocks, thus showing why it is hard to foster a more balanced pattern of activities. Regions may be similar at the outset, but they can diverge considerably later on. What makes the agglomeration forces so powerful is the combination of a drastic fall in transport and communication costs, together with the cumulative nature of the agglomeration process to give rise to a new type of economic geography in which space is '*slippery*', whereas locations are '*sticky*'. Affluent regions enjoy the existence of agglomeration rents that single-minded policies cannot easily dissipate. Consequently, if the aim of the European Commission is to foster a more balanced distribution of economic activities across European regions, it should add more instruments to its policy portfolio.

2. We show in Chapter 9 that people comprise a significant part of the wealth of regions. As a consequence, training people and investing in human capital are often better strategies than building new transport infrastructure, for this heightens the probability of individuals finding a job, maybe in places other than their region of origin. As observed by Cheshire et al. (2014), *regional disparities are driven more by differences between individuals than by differences between places*, although worker and place characteristics interact in subtle ways that require more investigation. After all, Toulouse initially did not seem a great place for the creation of a top school in economics. So there is hope for many places to develop new and creative activities.

If some regions are richer, it follows that others are less rich or poorer. Thus, at first sight, it seems logical to make spatial equity a criterion of economic policy. However, the underlying principles of spatial equity are ambiguous vis-à-vis the principles of social justice. Interpersonal inequality is often larger than interregional inequality. *Helping poor regions does not necessarily mean helping poor people*. The poor or unemployed in major urban areas today probably have more right to assistance than the inhabitants of poorer regions with a substantially lower cost of living. The job of the

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

welfare state is to reduce interpersonal inequalities that run counter to the principles of social justice, and these principles do not refer to particular spatial entities.

3. A key difficulty highlighted by NEG is that small differences may be sufficient to trigger regional divergence. This leads to the following question: *When do small differences matter?* As pointed out by Duranton et al. (2010), great places are great because 'they have managed to periodically reinvent themselves after losing an important part of the economic fabric'. Since the reasons for the success of these cities are often region or country-specific, it would be futile to seek a universal recipe. Yet a few general principles may serve as a guide. The historical and social background of a region, its economic strengths and weaknesses, its education system, and its portfolio of amenities are the fundamental ingredients to be accounted for when designing local development policies.

Very much like firms that differentiate their products to relax competition, regions must avoid head-to-head (fiscal) competition with well-established areas. Instead, regional development strategies should identify areas of specialization that exploit local sources of uniqueness. The aim of these strategies is to strengthen regions' comparative advantages and to give priority to finding sustainable solutions to regions' weakest links. For example, by differentiating the infrastructure services they provide, regions can create niches that make them attractive to a certain type of firms, which need not be high-tech firms. The scope for such a strategy is increasing as the revolution in information and communication technology has shifted firms' needs towards more specialized inputs. Implementing such a policy requires precise assessments of the strengths and weaknesses of the regional socio-economic and political fabric. For this to be possible, better data must be made available at various levels (regional, local, household).

4. One should also bear in mind that *a spray-gun distribution of increasing-returns activities results in high investment expenditure and/or underutilization of infrastructure and facilities*. Spatial dispersion of public investments is often inefficient because it prevents activities from reaching the critical mass needed to be efficient enough to compete on the national or international marketplace. What is more, for infrastructures to have an impact on the space-economy, they must be available in only a few places. Once they become widespread across locations, their impact is negligible because they no longer matter when firms and workers compare different locations. This is one more reason for giving up spray-gun policies. Regional policies fail to recognize that regional income differences are often the result of scale economies. To a certain extent, this explains the disillusion regarding the effectiveness of policies that aim for a more balanced distribution of activities across the EU, which in turn affects the credibility of the EU.

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

A related and unsolved question is the lasting decay that characterizes several regions that used to be the engines of the Industrial Revolution. All industries must one day decline, and examples abound in Europe of old industrialized regions that have succeeded in attracting sizable subsidies to finance inefficient firms. These regions have thus delayed any possibility of the region finding a new niche in which to specialize. Polése (2009) uses the expression 'negative cluster' to describe situations where the (regional) government is essentially captured by a declining cluster dominated by a few big employers and trade unions. In addition, it is well documented that the performance of regions in a country also depends on institutions that may be deeply rooted in the past. This leads Polése (2009) to write 'It is not by accident that the traditional European centres of coal and steel became strongholds of socialist and sometimes communist parties. The era of violent social conflict and divisive labour disputes is today – hopefully – over. But that era has left a legacy from which some regions have found it more difficult to escape than others. …I can find no other explanation of why seemingly well-located regions in northern France and in southern Belgium – in the European heartland – should continue to perform so poorly.' This is a strong claim but part of the story. However, as convincingly argued by Breinlich et al. (2014), we still have a poor understanding of *regional decline* as it *is not the mirror image of regional growth*.

5. One would expect the market-access effect to be weaker when transport costs are lower. But the opposite often holds true: more firms choose to set up in the large markets when it gets cheaper to trade goods between regions. Lower transport costs render competition tougher, leading firms to pay more attention to small differences between locations. They also make exports to small markets easier, which allows firms to exploit their scale economies more intensively by locating in large markets. Finally, lower transport costs reduce the advantages associated with geographical isolation in small markets where there is less competition. These various effects push toward more agglomeration. Hence, *connecting lagging regions to dynamic urban centres may weaken the lagging regions' industrial base*. This result may come as a surprise to those who forget that highways run both ways.

6. Regarding transport investment, there are at least three main research questions that are unsolved. First, given a major transport project, what share of the benefits is triggered by the resulting interregional shift in economic activity – and when does this shift unfold? If it is 10 per cent or less, this is within the margin of error of a conventional cost-benefit analysis of a transport project. In contrast, if the share is about 50 per cent, a conventional cost-benefit analysis is insufficient and must be supplemented by new econometric tools borrowed from regional and urban economics. Second, if small differences in accessibility can have a large impact on the location of economic activity, where is this more likely to happen? And third, how can we make sure that the right

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

transport investments are selected? For example, the EU has been promoting HSR for medium-distance travel, but the selected investments were far from optimal. Another related issue is to make sure that the capacity we currently have is used efficiently.

7. At present, *most interregional road, rail, inland waterways, and air networks are not priced efficiently*. European rail and air networks are still run largely by national monopolies that fail to comply with the principles of European integration. Furthermore, as member countries and regions do not take into account the full benefits of international and transit traffic, they are incentivized to charge too much for networks used intensively by foreign companies.

Acknowledgements

We thank F. di Comite, G. Duranton, M. Lafourcade, F. Robert-Nicoud, J. Südekum, A. Turrini, K. Van Dender, E. Viladecans-Marsal, as well as R. Lindsey, J-Cl. Prager, E. Quinet, K. Small, R. Evers, and C-M. Fung for their comments on previous drafts. We are grateful to M. Ivaldi, W. Leininger, and M. Paasi for their encouragement and the organization of the COEURE and DG-REGIO workshops, and acknowledge the funding by the FP7/320300 grant under the FP7 SSH research programme.

Notes

1. See Boldrin and Canova (2001), Midelfart-Knarvik and Overman (2002), and Puga (2002) for early critical assessments of the EU regional policies.
2. We follow the literature and view market integration as a gradual reduction in the costs of shipping goods and services.
3. See Baldwin et al. (2003), Fujita and Thisse (2013), and Zeng (2014) for a discussion of the HME in different set-ups.
4. Using a simple NEG model, a dataset including 250,000 randomly selected potential city locations, as well as all actual cities during the period 800–1800, Bosker and Buringh (2017) observe the two factors critical in explaining the location of cities: firstly, the proximity of waterways and land transport, and secondly, the relative position within the existing urban system. As suggested by NEG, being too close or far away from a large city reduces a place's chances to attract new activities.
5. See Fujita et al. (1999) and Baldwin et al. (2003) for more detailed analyses of NEG models.
6. See Fujita and Thisse (2013) for more details.
7. Rossi-Hansberg (2005) considers a set-up with a continuum of regions distributed along a one-dimensional space, several sectors, and positive transport costs. As transport costs decrease, firms become less sensitive to distance, which implies that peripheral locations will have better access to the core region and so will produce more than before. Thus, as in Helpman (1998), lowering transport costs fosters the geographical dispersion of activities. Desmet and Rossi-Hansberg (2014) propose

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

a dynamic version of this model in which technology diffuses across a continuous space to develop a spatial endogenous growth theory.

8. Intermediate goods represent 56 per cent of total trade in goods, while final consumption goods represent only 21 per cent of total trade in goods (Miroudot et al., 2009).
9. See Barrios and Strobl (2009) and Combes et al. (2011) and references therein.

References

Adler, N., Pels, E., and Nash, C. 2010. High-speed rail and air transport competition: Game engineering as tool for cost-benefit analysis. *Transportation Research Part B: Methodological*, **44**, 812–833.

Ahlfeldt, G. M., and Feddersen, A. 2015. From periphery to core: Measuring agglomeration effects using high-speed rail. *SERCDP0172. London, UK: Spatial Economics Research Centre, LSE.*

Anderson, J., and van Wincoop, E. 2004. Trade costs. *Journal of Economic Literature*, **42**, 691–751.

Aschauer, D. A. 1989. Is public expenditure productive? *Journal of Monetary Economics*, **23**, 177–200.

Bairoch, P. 1997. *Victoires et dboires. Histoire é conomique et sociale du monde du XVIe siĕcle ŕ nos jours*. Paris: Editions Gallimard.

Baldwin, R. E., Forslid, R., Martin, P., Ottaviano, G. I. P., and Robert-Nicoud, F. 2003. *Economic Geography and Public Policy*. Princeton, NJ: Princeton University Press.

Barbero, J., Behrens, K., and Zofio, J. L. 2015. *Industry location and wages: The role of market size and accessibility in trading networks*. CEPR Discussion Paper No. 10411.

Barrios, S., and Strobl, E. 2009. The dynamics of regional inequalities. *Regional Science and Urban Economics*, **39**, 575–591.

Behrens, K. 2005. Market size and industry location: Traded vs non-traded goods. *Journal of Urban Economics*, **58**, 24–44.

Behrens, K., and Thisse, J.-F. 2007. Regional economics: A new economic geography perspective. *Regional Science and Urban Economics*, **37**, 457–465.

Behrens, K., Gaigné, C., Ottaviano, G. I. P., and Thisse, J.-F. 2007. Countries, regions and trade: On the welfare impacts of economic integration. *European Economic Review*, **51**, 1277–1301.

Belleflamme, P., Picard, P., and Thisse, J.-F. 2000. An economic theory of regional clusters. *Journal of Urban Economics*, **48**, 158–184.

Beyer, R. C. M., and Smets, F. 2015. Labour market adjustments and migration in Europe and the United States: How different? *Economic Policy*, **30**, 643–682.

Boldrin, M., and Canova, F. 2001. Inequality and convergence in Europe's regions: Reconsidering European regional policies. *Economic Policy*, **16**, 206–253.

Bosker, M., and Buringh, E. 2017. City seeds: Geography and the origins of the European city system. *forthcoming in Journal of Urban Economics*.

Bosquet, C., and Overman, H. 2015. Home versus hometown: What do we mean by spatial sorting? *Paper presented at the RIETI Workshop on Spatial Economics, Tokyo, April 14, 2015*.

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

Breinlich, H., Ottaviano, G., and Temple, J. 2014. Regional growth and regional decline. In: Aghion, P., and Durlauf, S. N. (eds), *Handbook of Economic Growth*, vol. 2. Amsterdam: Elsevier, pp. 683–779.

Bröcker, J., Korzhenevych, A., and Schürmann, C. 2010. Assessing spatial equity and efficiency impacts of transport infrastructure projects. *Transportation Research Part B: Methodological*, **44**, 795–811.

Charlot, S., Gaigné, C., Robert-Nicoud, F., and Thisse, J.-F. 2006. Agglomeration and welfare: The core-periphery model in the light of Bentham, Kaldor, and Rawls. *Journal of Public Economics*, **90**, 325–347.

Cheshire, P., and Magrini, S. 2006. Population growth in European cities: Weather matters – but only nationally. *Regional Studies*, **40**, 23–37.

Cheshire, P., Nathan, M., and Overman, H. 2014. *Urban Economics and Urban Policy*. Cheltenham, U.K.: Edward Elgar Publishing.

Combes, P.-P., and Overman., H. 2004. The spatial distribution of economic activities in the European Union. In: Henderson, J. V., and Thisse, J.-F. (eds), *Handbook of Regional and Urban Economics*, vol. 4. Amsterdam: Elsevier, pp. 2845–2909.

Combes, P.-P., Lafourcade, M., Thisse, J.-F., and Toutain, J.-C. 2011. The rise and fall of spatial inequalities in France. A long-run perspective. *Explorations in Economic History*, **48**, 243–271.

Crozet, M. 2004. Do migrants follow market potentials? An estimation of a new economic geography model. *Journal of Economic Geography*, **4**, 439–458.

Dahl, M. S., and Sorenson, O. 2010. The migration of technical workers. *Journal of Urban Economics*, **67**, 33–45.

d'Aspremont, C., Gabszewicz, J. J., and Thisse, J.-F. 1979. Hotelling's 'Stability in Competition'. *Econometrica*, **47**, 1145–1150.

Davis, D. R., and Weinstein, D. E. 2003. Market access, economic geography and comparative advantage: An empirical test. *Journal of International Economics*, **59**, 1–23.

De Rus, G., and Nombela, G. 2007. Is investment in high speed rail socially profitable? *Journal of Transport Economics and Policy*, **41**, 3–23.

Desmet, K., and Rossi-Hansberg, E. 2014. Spatial development. *American Economic Review*, **104**, 1211–1243.

Di Comite, F., and Kancs, A. 2015. *Modelling of agglomeration and dispersion in RHOMOLO*. Mimeo.

Dixit, A. K., and Stiglitz, J. E. 1977. Monopolistic competition and optimum product diversity. *American Economic Review*, **67**, 297–308.

Duranton, G., Martin, P., Mayer, T., and Mayneris, F. 2010. *The Economics of Clusters: Lessons from the French Experience*. Oxford: Oxford University Press.

Epifani, P., and Gancia, G. 2005. Trade, migration and regional unemployment. *Regional Science and Urban Economics*, **35**, 625–644.

Eurostat. 2015a. EU Transport in Figures – Statistical Pocketbook 2015. *Luxembourg: Publications Office of the European Union*.

Eurostat. 2015b. Regional Statistics. http://ec.europa.eu/eurostat/cache/RSI/#?vis=nuts2.economy.

Faber, B. 2014. Trade integration, market size, and industrialization: Evidence from China's national trunk highway system. *Review of Economic Studies*, **81**, 1046–1070.

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

Falck, O., Heblich, S., Lameli, A., and Südekum, J. 2012. Dialects, cultural identity, and economic exchange. *Journal of Urban Economics*, **72**, 225–239.

Fujita, M., and Thisse, J.-F. 2013. *Economics of Agglomeration: Cities, Industrial Location, and Globalization*. 2 edn. New York: Cambridge University Press.

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

Glaeser, E. L., and Kohlhase, J. E. 2004. Cities, regions and the decline of transport costs. *Papers in Regional Science*, **83**, 197–228.

Gramlich, E. M. 1994. Infrastructure investment: A review essay. *Journal of Economic Literature*, **32**, 1176–1196.

Harris, C. 1954. The market as a factor in the localization of industry in the United States. *Annals of the Association of American Geographers*, **44**, 315–348.

Head, K., and Mayer, T. 2004. The empirics of agglomeration and trade. In: Henderson, J. V., and Thisse, J.-F. (eds), *Handbook of Regional and Urban Economics*, vol. 4. Amsterdam: Elsevier, pp. 2609–2669.

Head, K., and Mayer, T. 2011. Gravity, market potential and economic development. *Journal of Economic Geography*, **11**, 281–294.

Head, K., and Mayer, T. 2013. What separates us? Sources of resistance to globalization. *Canadian Journal of Economics*, **46**, 1196–1231.

Helpman, E. 1998. The size of regions. In: Pines, D., Sadka, E., and Zilcha, I. (eds), *Topics in Public Economics: Theoretical and Applied Analysis*. Cambridge, UK: Cambridge University Press, pp. 33–54.

Henderson, J. V. 1997. Medium size cities. *Regional Science and Urban Economics*, **27**, 583–612.

Hirschman, A. O. 1958. *The Strategy of Development*. New Haven, CN: Yale University Press.

Hohenberg, P., and Lees, L. H. 1985. *The Making of Urban Europe (1000–1950)*. Cambridge, MA: Harvard University Press.

Hotelling, H. 1929. Stability in competition. *Economic Journal*, **39**, 41–57.

Ivaldi, M., and McCullough, G. 2008. Subadditivity tests for network separation with an application to U.S. railroads. *Review of Network Economics*, **7**, 1–13.

Klein, A., and Crafts, N. 2012. Making sense of the manufacturing belt: Determinants of U.S. industrial location, 1880–1920. *Journal of Economic Geography*, **12**, 775–807.

Knight, B. 2004. Parochial interests and the centralized provision of local public goods: Evidence from congressional voting on transportation projects. *Journal of Public Economics*, **88**, 845–866.

Koopmans, T. C. 1957. *Three Essays on the State of Economic Science*. New York: McGraw-Hill.

Krugman, P. R. 1980. Scale economies, product differentiation, and the pattern of trade. *American Economic Review*, **70**, 950–959.

Krugman, P. R. 1991. Increasing returns and economic geography. *Journal of Political Economy*, **99**, 483–499.

Krugman, P. R., and Venables, A. J. 1995. Globalization and the inequality of nations. *Quarterly Journal of Economics*, **110**, 857–880.

Magrini, S. 2004. Regional (di)convergence. In: Henderson, J. V., and Thisse, J.-F. (eds), *Handbook of Regional and Urban Economics*, vol. 4. Amsterdam: Elsevier, pp. 2741–2796.

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

Mandell, S., and Proost, S. 2016. Why truck distance taxes are contagious and drive fuel taxes to the bottom. *Journal of Urban Economics*, **93**, 1–7.

Mayer, T. 2008. Market potential and development. *CEPR Discussion Paper No. 6798*.

Mayer, T., and Ottaviano, G. I. P. 2007. *The Happy Few: The Internationalisation of European Firms. New Facts Based on Firm-Level Evidence*. Brussels: Bruegel Blueprint Series.

Melo, P., Graham, D. J., and Brage-Ardao, R. 2013. The productivity of transport infrastructure investment: A meta-analysis of empirical evidence. *Regional Science and Urban Economics*, **43**, 695–706.

Midelfart-Knarvik, K. H., and Overman, H. G. 2002. Delocation and European integration: Is structural spending justified? *Economic Policy*, **17**, 321–359.

Miroudot, S., Lanz, R., and Ragoussis, A. 2009. *Trade in intermediate goods and services*. OECD Trade Policy Working Paper No. 93, TAD/TC/WP(2009)1/FINAL, OECD.

Moretti, E. 2011. Local labor markets. In: Card, D., and Ashenfelter, O. (eds), *Handbook of Labor Economics*, vol. 4b. Amsterdam: Elsevier, pp. 1237–1313.

Nocke, V. 2006. A gap for me: Entrepreneurs and entry. *Journal of the European Economic Association*, **4**, 929–956.

Ohlin, B. 1967. *Interregional and International Trade*. Rev. ed. Cambridge, MA: Harvard University Press. (Originally published in 1933.)

Ottaviano, G. I. P., and Thisse, J.-F. 2002. Integration, agglomeration and the political economics of factor mobility. *Journal of Public Economics*, **83**, 429–456.

Ottaviano, G. I. P., and van Ypersele, T. 2005. Market size and tax competition. *Journal of International Economics*, **67**, 25–46.

Polése, M. 2009. *The Wealth and Poverty of Regions: Why Cities Matter*. Chicago: University of Chicago Press.

Pollard, S. 1981. *Peaceful Conquest: The Industrialization of Europe 1760–1970*. Oxford: Oxford University Press.

Porter, M. E. 1998. Clusters and the new economics of competition. *Harvard Business Review* (November-December), 77–90.

Proost, S., Dunkerley, F., van der Loo, S., Adler, N., Bröcker, J., and Korzhenevych, A. 2014. Do the selected Trans European transport investments pass the cost benefit test? *Transportation*, **41**, 107–132.

Puga, D. 1999. The rise and fall of regional inequalities. *European Economic Review*, **43**, 303–334.

Puga, D. 2002. European regional policies in the light of recent location theories. *Journal of Economic Geography*, **2**, 373–406.

Redding, S. J. 2011. Economic geography: A review of the theoretical and empirical literature. In: Bernhofen, D., Falvey, R., Greenaway, D., and Kreickemeier, U. (eds), *The Palgrave Handbook of International Trade*. London: Palgrave Macmillan, pp. 497–531.

Redding, S. J., and Sturm, D. 2008. The cost of remoteness: Evidence from German division and reunification. *American Economic Review*, **98**, 1766–1797.

Redding, S. J., and Turner, M.A. 2015. Transportation costs and the spatial organization of economic activity. In: Duranton, G., Henderson, J. V., and Strange, W. (eds), *Handbook of Regional and Urban Economics*, vol. 5. Amsterdam: Elsevier, pp. 1339–1398.

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press

Redding, S. J., and Venables, A. J. 2004. Economic geography and international inequality. *Journal of International Economics*, **62**, 53–82.

Rossi-Hansberg, E. 2005. A spatial theory of trade. *American Economic Review*, **95**, 1464–1491.

Spulber, D. F. 2007. *Global Competitive Strategy*. Cambridge: Cambridge University Press.

Syverson, C. 2004. Market structure and productivity: A concrete example. *Journal of Political Economy*, **112**, 1181–1222.

Takahashi, T., Takatsuka, H., and Zeng, D.-Z. 2013. Spatial inequality, globalization, and footloose capital. *Economic Theory*, **53**, 213–238.

Thomas, I. 2002. *Transportation Networks and the Optimal Location of Human Activities: A Numerical Geography Approach*. Cheltenham, UK: Edward Elgar.

Van der Loo, S., and Proost, S. 2013. The European road pricing game: How to enforce optimal pricing in high-transit countries under asymmetric information. *Journal of Transport Economics and Policy*, **47**, 399–418.

Williamson, J. G. 1965. Regional inequality and the process of national development: A description of the patterns. *Economic Development and Cultural Change*, **13**, 1–84.

Zeng, D.-Z. 2014. The role of country size in spatial economics: A survey of the home market effects. *Journal of Economics and Business*, **32**, 379–403.

https://doi.org/10.1017/9781316636404.010 Published online by Cambridge University Press