Agglomeration and the Effects of Regional Transfer Schemes

Inaugural-Dissertation

zur Erlangung des Grades

Doctor oeconomiae publicae (Dr. oec. publ.)

an der Ludwig-Maximilians-Universität München

2009

vorgelegt von

Maximilian von Ehrlich-Treuenstätt

Referent:	Prof. Dr. Dr. h.c. Hans-Werner Sinn
Korreferent:	Prof. Dr. Robert Fenge
Datum der mündlichen Prüfung:	22. Januar 2010
Promotionsabschlussberatung:	10. Februar 2010

Acknowledgements

First and foremost I would like to thank my supervisors Hans-Werner Sinn and Robert Fenge for giving me the opportunity to pursue this dissertation. Hans-Werner Sinn not only provided me with guidance and an inspiring environment at the Center for Economic Studies (CES) but initiated my interest in economics and shaped my way of economic thinking. Robert Fenge encouraged me to study issues in regional economics and provided me with continuous support and advice. Our joint projects gave me the opportunity to learn a lot from his experience and knowledge.

I am also very grateful to Sascha Becker and Peter Egger who substantially improved my econometric knowledge and introduced me to the latest stata techniques. Of course I want to thank my colleagues and friends at CES, at the department of economics, and at the Ifo Institute with whom I shared a great time. Their comments and advices at numerous seminars contributed strongly to this thesis. Moreover, I benefited from many visitors at the CES of whom I want to mention Harald Badinger, representatively.

During my doctoral studies I received support from the German Research Foundation (DFG) within its priority program "Institutionelle Gestaltung Föderaler Systeme" which I appreciated a lot. Various conference and workshops of the DFG program stimulated my work. Among the members of the DFG program I want to thank particularly Matthias Wrede, Wolfgang Eggert and Günther König.

Last but not least I want to thank Tanja and my family for their patience and continuous support as well as for helpful proofreading. I owe my largest debt to my parents, to whom I dedicate this thesis.

Maximilian von Ehrlich-Treuenstätt

Munich, September 2009

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Regardless of which country or federation one considers, the spatial distribution of economic activity is far from uniform. Due to political decisions like the fall of the iron curtain or structural changes like the shift to a service-based economy, the spatial distribution of economic activity adjusts from time to time. These adjustments represent a constant challenge for politicians as they might lead to social tensions. In Europe for example, the ongoing integration of the Single Market yields a substantial reorganization of economic activity in space. Yet mainstream economic models pay little attention to the pure spatial dimension. In contrast, politics is very concerned with the spatial allocation of economic activity. The active reduction of regional disparities in Europe is regarded as a requirement of solidarity and it is argued that the utilization of the total territory contributes to strengthening the competitiveness of the European economy as a whole (see European Commission, 2008, p. 4). Hence, regional cohesion – interpreted as the reduction of regional disparities – is one of the aims of the Lisbon Treaty.[1] Enforcing regional cohesion through a multitude of instruments plays a prominent role in the agenda of European politics. In 2009, for instance, about 36 percent of the total EU budget is allocated to regional cohesion policies (see European Commission, 2009). These funds represent transfers from rich to poor regions which finance public infrastructure investments or subsidize private investments.

In Germany, the aim for equivalent regional living-conditions is part of the federal constitution (see articles 72 and 106). However, the exact interpretation of this principle is up for debate and the persistent disparities cast doubt on the effectiveness of

[1] Article 174 of the *Treaty on the Functioning of the European Union* states: "[...] the Union shall aim at reducing disparities between the levels of development of the various regions and the backwardness of the least favoured regions" (see Official Journal C 115/127 09/05/2008).

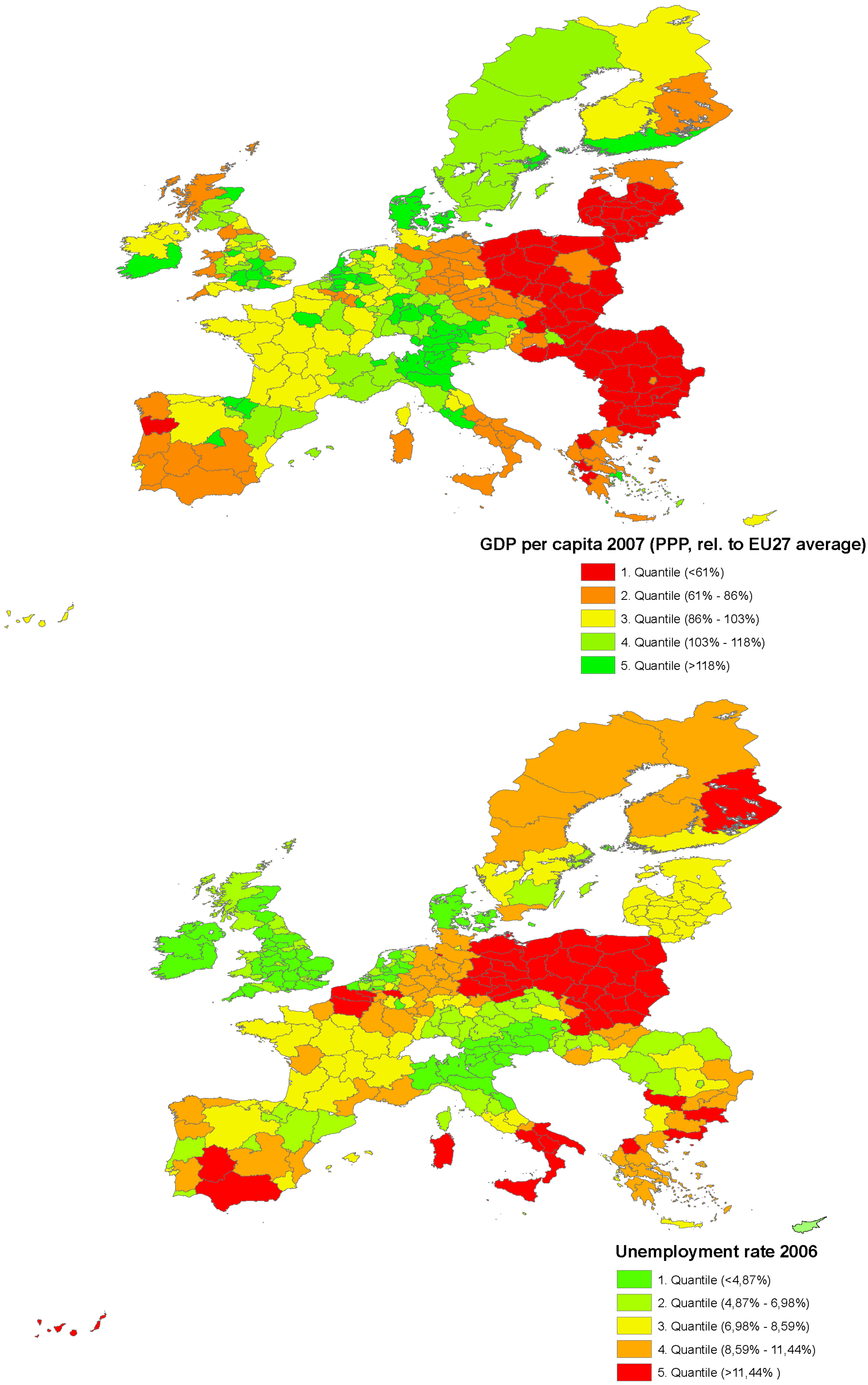
regional policies.[2] Moreover, demographic studies predict a severe reduction of active population in Germany's lagging regions such that economic activity will be concentrated to an increasing degree in the agglomerated core-regions. Given this scenario, traditional regional policy is regarded as being ineffective and the aim for equivalent regional living-conditions is regarded as illusionary (see Berlin Institut, 2009).
This thesis contributes to the debate on regional disparities and the role of regional policy by analyzing the driving forces of agglomeration and discussing its welfare implications. Moreover, the effectiveness of regional policy aiming at a more even distribution of economic activity is evaluated theoretically as well as empirically.

In fact, regional disparities are considerable in Europe. Today, 43 percent of the economic output is concentrated on just 14 percent of the European territory (see European Commission, 2008, p. 4). Figure 1.1 illustrates the degree of regional disparities in Europe. The maps depict regional GDP per capita in purchasing power parity terms (PPP) relative to the EU27 average as well as regional unemployment rates. The level of regional aggregation chosen in these maps is NUTS2 which corresponds to groups of counties and unitary authorities with a population of 0.8-3 million inhabitants. Note that Eurostat, the statistical office of the European Commission, distinguishes between three sub-national regional aggregates: NUTS1 (large regions with a population of 3-7 million inhabitants), NUTS2 and NUTS3 regions (counties of 150-800 thousand inhabitants).[3] To begin with, it is useful to consider NUTS2 aggregates since the European regional policy targets NUTS2 regions.
The poorest regions with a GDP per capita in the first quantile of the distribution, which corresponds to less than 61 percent of the EU27 average, lie mainly in the new member states of East-Central Europe. However, even after the EU's enlargement to the east, the poorest regions of the former EU15 – which have received substantial transfers for many years – remain in the second quantile of per-capita GDP which corresponds to less than 86 percent of the EU27 average. These are regions in Spain, Portugal, Southern Italy, Greece and East Germany. The most prosperous regions are

[2] The German Federal Ministry of Transport, Building and Urban Affairs states that the aim for equivalent regional living-conditions does not imply identical living-conditions but has to be interpreted as equivalent opportunities and requires the provision of certain minimum standards of public goods at any place (see BMVBS, 2006, p. 57).

[3] NUTS is the acronym for *N*omenclature des *U*nités *T*erritoriales *S*tatistiques coined by Eurostat. The highest level of regional aggregation (NUTS1) corresponds to Germany's *Bundesländer*, France's *Zones d'Études et d'Aménagement du Territoire*, the United Kingdom's *Regions of England/Scotland/Wales* or Spain's *Grupos de Comunidades Autónomas.* At the other end of the NUTS classification scheme, NUTS3 regions correspond to *Landkreise* in Germany, to *Départements* in France, to *Unitary Authorities* in the UK or to *Comunidades Autónomas* in Spain.

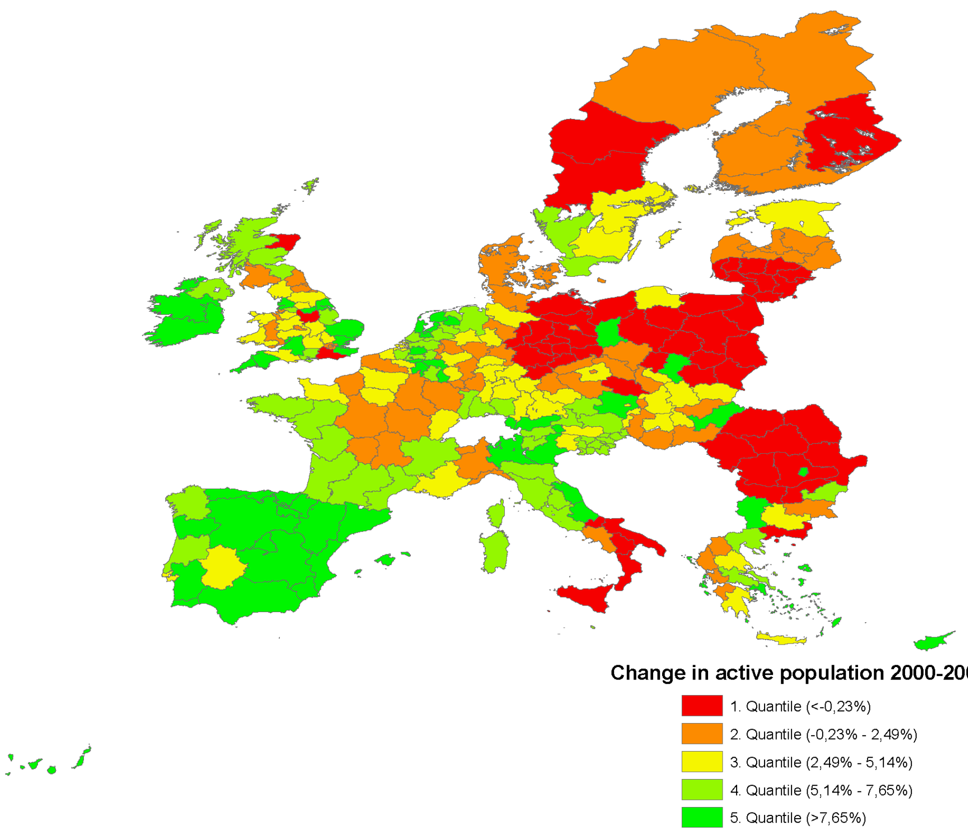
Figure 1.1: Regional Disparities in the EU (NUTS2)



Data Source: Cambridge Econometrics Regional Database.

located in central Britain, the Benelux countries, southern Germany, parts of Scandinavia and northern Italy.[4] When comparing the distribution of GDP per capita to the distribution of unemployment a correlation becomes apparent. With few exceptions the regions in the lower quantiles of per-capita GDP tend to be the regions in the higher quantiles of unemployment. Yet, unemployment rates seem to be driven to a larger degree by national characteristics. For instance, all British and Irish regions feature unemployment rates that correspond to the lowest two quantiles whereas their per-capita GDP ranges from the second to the fifth quantile. In other countries like Germany or Italy, regional disparities are more apparent in unemployment rates than in per-capita GDP. This is partly because productivity differences reflect in unemployment differences if collective wage-agreements prevent wage disparities.

Figure 1.2: POPULATION DEVELOPMENT IN THE EU (NUTS2)

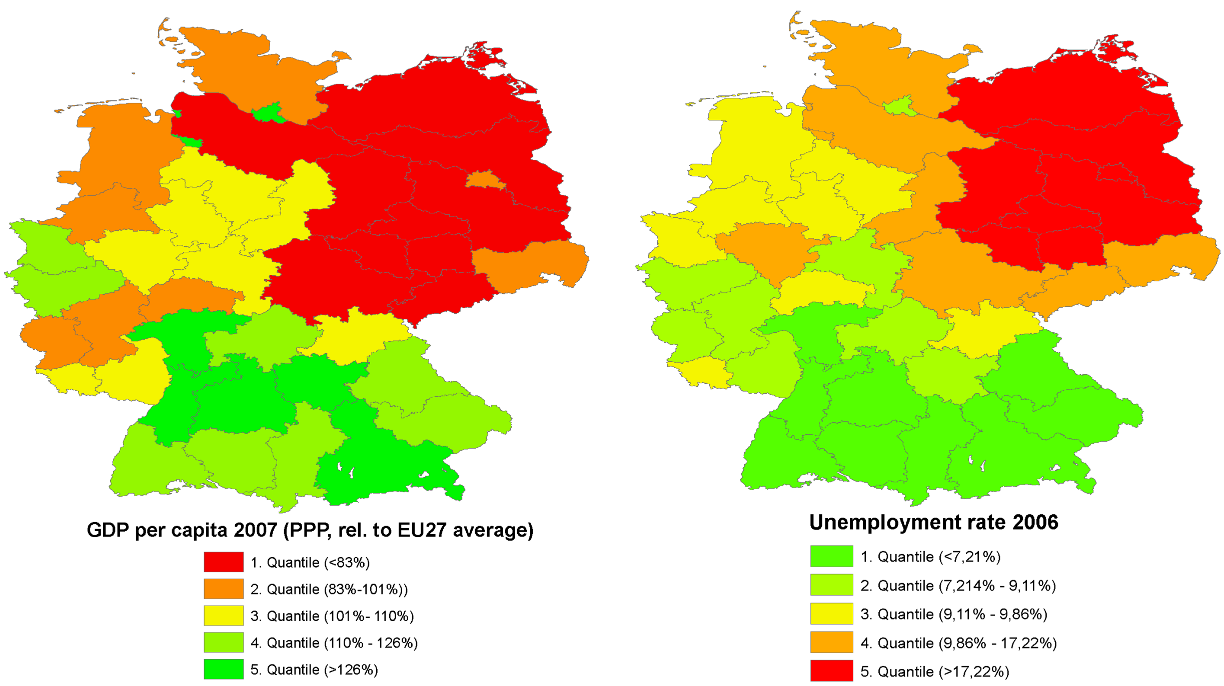


Note: The active population comprises individuals who either are employed or are actively seeking employment. *Data source*: Cambridge Econometrics Regional Database.

[4] More detailed descriptive statistics on the distribution of per-capita GDP are provided in Chapter 5, table 5.2.

Regardless of whether disparities in unemployment or per-capita GDP dominate, such differences in regional prosperity certainly induce migration. When considering figure 1.2 one observes a tendency of the active population to shift to the prosperous regions. This process may induce convergence in per-capita GDP but also constitutes a trend of further regional agglomeration of economic activity in Europe.
Focusing on Germany, which has strived for an equalization of living conditions between the eastern and the western regions for almost twenty years, shows that regional disparities are still dominant. However, against the general perception, Germany is not characterized by a dichotomy of prosperity but rather exhibits three spatially separated levels of prosperity.

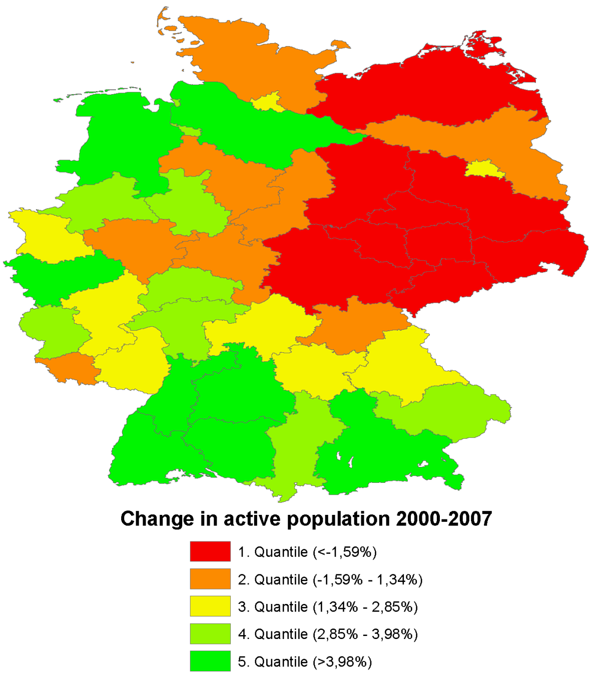
Figure 1.3: Regional Disparities in Germany (NUTS2)



Data Source: Cambridge Econometrics Regional Database.

Figure 1.3 shows that GDP per capita is the highest in the south of Germany where unemployment is the lowest. Overall, the poorest regions with the highest unemployment rate are still located in East-Germany but some regions in Lower Saxony or Rhineland-Palatinate have similar unemployment rates and almost the same per-capita GDP. Just as in Europe, a correlation between the levels of prosperity and the change in active population can be observed. The lagging East-German regions are characterized by a significant decline in active population as is shown in figure 1.4.

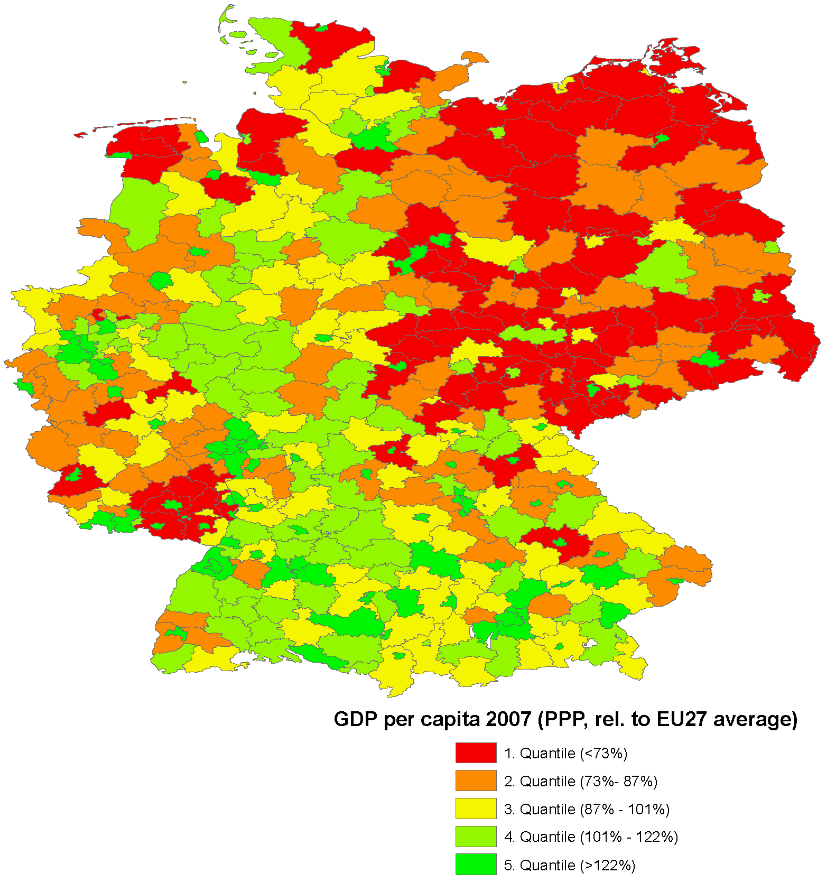
Figure 1.4: POPULATION DEVELOPMENT IN GERMANY (NUTS2)



Note: The active population comprises individuals who either are employed or are actively seeking employment. *Data source*: Cambridge Econometrics Regional Database.

However, the pictures illustrated above depend on the level of regional aggregation. As mentioned before, European regional policy targets NUTS2 regions. Yet, the Lisbon Treaty does not state explicitly the level of regional aggregation that should provide the basis of cohesion policy. Accordingly, the choice of NUTS2 remains somewhat arbitrary. The variation in per-capita GDP increases significantly when considering the more disaggregated level of regional aggregation (NUTS3) as illustrated in figure 1.5. In this case, even the comparatively rich areas in southern Germany comprise comparatively poor NUTS3 regions with a per-capita GDP in the lowest two quantiles of the German distribution of GDP per capita. This suggests that the reduction of disparities obviously has its limits and that certain spatial disparities are inevitable.

Figure 1.5: Regional Disparities in Germany (NUTS3)



Data Source: Cambridge Econometrics Regional Database.

The following chapters analyze the reasons for regional disparities to arise and evaluate the spatial outcome of market processes from a welfare perspective. Moreover, the effectiveness of regional policy in reducing disparities is assessed. Before going into detail, a brief outline of the central ideas and major contributions of each single chapter – except for the summary in Chapter 6 – will be provided. Note that the following chapters may all be read separately but follow one line of reasoning.

Chapter 2: Theories of Economic Geography

In order to evaluate regional policy and to discuss the optimal spatial allocation it is essential to understand how regional concentration of economic activity arises in the first place. What are the driving forces of agglomeration and which natural mechanisms countervail agglomeration tendencies? Only if the market mechanisms that determine the spatial equilibrium are known, can policy measures intervening in the spatial allocation be assessed.

Over recent years, the theory of economic geography experienced a revival associated with the development of what is called the New Economic Geography. However, there is a long tradition of economic geography which is closely related to scholars like Johann Heinrich von Thünen, Alfred Weber, Harold Hotelling, Walter Christaller and August Lösch. Their theories of location and central places are neglected in a large part by the recent approaches. This chapter brings together the recent micro-foundations of agglomeration economies and the traditional approaches of economic geography. The theories are connected in a way that they complement one another in deriving a comprehensive spatial pattern of economic activity. While the traditional theories lack a convincing foundation of what drives firms and consumers to cluster, the more recent approaches simplify in the sense that they consider a world of only two regions – an agglomeration and a peripheral region – while neglecting the hierarchical system of multiple agglomerations that we observe in reality. As stated in the central place theory, each territory features a few higher-order agglomerations (higher-order central places) which supply a broad range of goods and services and many lower-order agglomerations (lower-order central places) which supply only a narrow range of goods and services. The synthesis put forward in this chapter derives the economic determinants of the distance between agglomerations (central places) of the same order.

It is shown that sectors which are characterized by a low level of trade costs, a low elasticity of substitution, a high share in aggregate expenditure, strong knowledge spillovers, and a high volatility exhibit strong agglomeration economies and will locate at only few higher-order agglomerations. The agglomeration economies described in this chapter allow for gains from variety, realizing knowledge spillovers, or sharing risk and thus generally benefit aggregate welfare. Yet an explicit welfare analysis is postponed to Chapter 3.

Chapter 3: Public Input Competition and Agglomeration

This chapter analyzes the impact of public input competition in a New Economic Geography framework. It is shown that regional competition represents a prominent agglomeration mechanism in addition to the mechanisms introduced in Chapter 2. As long as trade costs are sufficiently high, public input competition ensures a dispersion of industry but as soon as trade costs fall below a certain threshold level, public input competition induces agglomeration even for ex-ante identical regions.

The welfare analysis provided in this chapter shows that regional competition yields an overprovision of public inputs if trade costs are sizable while it leads to underprovision if regions are highly integrated. This result is related to the neoclassical literature on fiscal competition between individual regions. Moreover, it is shown that agglomeration becomes efficient for sufficiently low trade costs such that a trade-off between the political goals of regional cohesion and aggregate efficiency arises. A central planner's efficient distribution of regional infrastructure requires full agglomeration for sufficiently low trade costs. The critical trade costs that render agglomeration efficient can be translated into the optimal distance between individual agglomerations (central places) as introduced in the previous chapter. Finally and most importantly, it is shown that market processes will cause agglomeration to arise not before it becomes efficient. In fact, the uncoordinated competition-equilibrium is characterized by agglomeration only at levels of trade costs that fall below the efficient level.

Hence, the chapter suggests that if regional policy interferes with the regional distribution of economic activity, it should essentially promote regional concentration instead of regional dispersion.

Chapter 4: Infrastructure and the Location of Business

The fourth chapter represents on the one hand an analysis of the effects of infrastructure provision on the regional distribution of economic activity and on the other hand acts as an empirical test of the New Economic Geography.

First of all, the effects of intra- and inter-regional infrastructure provision are separated from each other. As opposed to inter-regional infrastructure which facilitates the movements of goods across the administrative boundaries of regions, intra-regional infrastructure affects only the productivity or the trade costs within the borders of a region. It is shown analytically that intra-regional infrastructure unambiguously benefits the local share of economic activity while inter-regional infrastructure is more effective in raising the local share of economic activity if the trade costs to a region with a smaller home market are reduced. The analysis demonstrates that regions with little local expenditure can be cast under the hub-shadow of a neighbor region with a major home market. This is because firms prefer to locate in the hub-region with the larger home market and serve the small neighbor region through exports. Accordingly, regions with little local expenditure and a major home market in the neighborhood exhibit a disproportionately low level of economic activity. Moreover, the theoretical analysis predicts that regions with a high quality of intra-regional infrastructure, and a high accessibility accommodate a high share of economic activity. The hub-shadow effect as well as the importance of market accessibility originate precisely from the New Economic Geography's agglomeration mechanism such that the empirical test of the theoretical predictions allows us to infer the appropriateness of the New Economic Geography in general. Since infrastructure investments represent a major instrument of regional policy, this chapter also gives important insights for policy makers. The analysis suggests that the political goal of regional cohesion will be reached more effectively by providing intra-regional infrastructure in peripheral regions than by improving the connections between core and peripheral regions.

The second part of the chapter is devoted to an empirical test of the theoretical predictions. In order to distinguish the effects of inter-regional infrastructure from intra-regional infrastructure, a novel measure of trade-freeness is introduced. In addition, this measure accounts for the quality of infrastructure and potential spillover effects. The empirical results provide evidence for the New Economic Geography's predictions regarding the spatial distribution of the service sector. However, with respect to the manufacturing sector, the empirical results are much less in line with the theoretical hypotheses. Consequently, it is argued that the New Economic Geography's agglomeration mechanism is underlying the spatial distribution of the service sector while other mechanisms seem to be more adequate in explaining the spatial distribution of the manufacturing sector.

Chapter 5: The Effect of EU Structural Funds on Regional Performance

Finally, the actual effects of regional policy are evaluated empirically. By doing so, the most important instrument of regional policy in Europe – the Structural Funds Program – is analyzed.

The European Union (EU) provides grants to disadvantaged regions of member states to allow them to catch up with the EU average. Under the Objective 1 scheme, NUTS2 regions with a GDP per capita level below 75% of the EU average qualify for structural funds transfers from the central EU budget. This rule gives rise to a regression-discontinuity design that exploits the discrete jump in the probability of EU transfer receipt at the 75% threshold. Additional variability arises for smaller regional aggregates – so-called NUTS3 regions – which are nested in a NUTS2 mother-region. Whereas some relatively rich NUTS3 regions may receive EU funds because their NUTS2 mother region qualifies, other relatively poor NUTS3 regions may not receive EU funds because their NUTS2 mother region does not qualify.

The analysis provides evidence for positive growth effects of Objective 1 funds, but finds no employment effects. A simple cost-benefit calculation suggests that Objective 1 transfers are effectively contributing to the political goal of regional cohesion and are – in net terms – not wasteful.

Chapter 2

Theories of Economic Geography

2.1 Introduction

At the center of economics is the analysis of factor and resource allocation. Yet, most economic models account only for the allocation of factors and resources to different activities, individuals, or points in time but neglect the spacial dimension of the allocation problem. In reality we observe that factors and resources are far from being equally distributed in space. Therefore, economic models should take mechanisms which are able to explain this phenomenon into account. Moreover, analyzing the determinants underlying the spatial allocation permits to infer the efficiency of market processes with respect to the spatial dimension. Only if the causes for spatial disparities are known, can the political instruments aiming at a reduction of regional disparities be evaluated.

The little attention which is devoted to the spatial dimension in most economic models may be attributed to the fact that generally accepted methods for incorporating space into economic models are missing. Regarding the time dimension there are standard procedures of intertemporal consumption and production optimization and almost every popular static model is translated into a dynamic context in order to verify its implications. In contrast, spatial analyses are often constrained to the mere assumption of a force that drives factors to agglomerate in space. Yet, in recent years the analysis of economic geography experienced a revival due to novel ways of formalizing agglomeration forces explicitly summarized under the heading *New Economic Geography*.

This chapter provides a synthesis of the most important theories of economic geography. It brings together the recent micro-foundations of agglomeration with the established theories of location and central places. In doing so we focus exclusively on the pure economic reasons driving the spatial dimension of economic activities. This means that we ignore natural comparative advantages in the sense of natural irregularities and endowments such as a beneficial location next to the seaside or exceptionally fertile soil.

Since chapters 2 and 3 build on New Economic Geography models, we will relate each theory introduced in the following to the predictions and assumptions made by the New Economic Geography.

2.2 Location Theories

The first theories which analyze the role of purely economic mechanisms in driving the geography of economic activity were particularly concerned with the location choice of individual firms. They already acknowledged the importance of agglomeration economies and transportation costs for the location decision. The intention of firms to minimize the overall transportation costs is confronted with their advantages from being close to other firms.

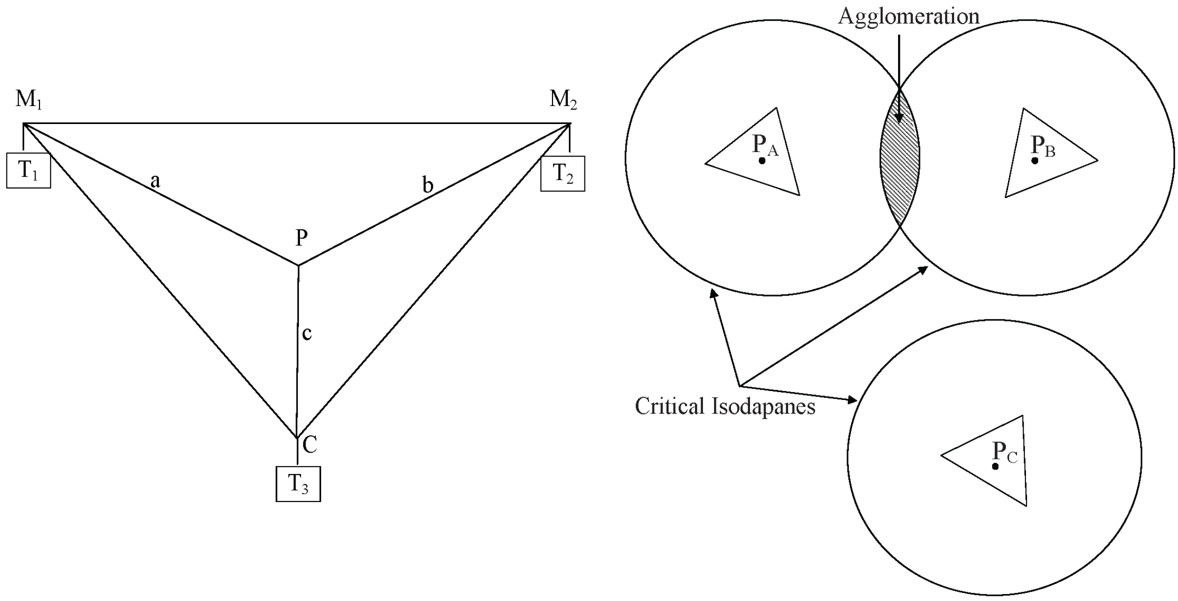
The seminal work by Weber (1909) thinks of agglomeration economies as the availability of labor at lower costs and/or better quality in an agglomerated region than in the periphery. The underlying assumptions of Weber's model are a punctiform market for the final good which is produced under perfect competition and a punctiform supply of the raw material which is necessary for the production of the final good.

In the model's first stage, each firm minimizes the transportation costs between the locations of the raw materials and the market place of the final good. Weber introduced the location-production triangle which is illustrated in figure 2.1. The points M_1 and M_2 represent the locations of the raw materials, while the point P represents the production plant where the raw materials are assembled. The market place where the final good is to be sold is denoted by C. The model presumes that the final good is made up of a constant amount of each raw material and rules out any substitution between the raw materials.[1] The weights of the raw materials necessary to produce one unit of output are denoted by T_1 and T_2, respectively, while the final good weighs

[1]Moses (1958) extends the Weber model by allowing for substitution between the inputs. He shows that the choice over the production technologies is closely related to the location choice.

T_3. The distance between the production place and the locations of the raw materials are a and b, respectively, and the distance between the production plant and the final market is denoted by c. With this information, the total transportation costs, which are thought to be linearly related to the weight of the transported goods and the distance of the transports, become $T = aT_1 + bT_2 + cT_3$. A profit-maximizing firm will choose its location such that the total transportation costs are minimized. This requires that in equilibrium $aT_1 + bT_2 = cT_3$ has to be satisfied.[2]

Figure 2.1: Weber's Model



Location-Production Triangle **Isodapanes**

Note: The figures are based on Weber (1909, p. 54 and p.134).

In the second stage, circles are constructed which are commonly referred to as Isodapanes. The Isodapanes link all points exhibiting the same increase in total transportation costs per output unit compared with the minimum transportation-cost location

[2]If $aT_1 + bT_2 < cT_3$ holds true for production place P, the point of minimum transportation costs is shifted towards C. In contrast, if $aT_1 + bT_2 > cT_3$ holds true the point of minimum transportation costs is shifted towards the location of raw materials. Weber distinguishes between a market oriented and a material oriented location of the firm, that is whether the minimum cost location is closer to the final market (in case the transportation the final good is relatively expensive) or to the raw material locations (in case the transportation the raw material is relatively expensive).

(see figure 2.1). Put differently, the Isodapanes illustrate the transportation-cost increase due to the distance from the minimum transportation-cost location. However, being close to other firms constitutes an agglomeration advantage which in monetary terms is equal to v. Therefore, each firm will be willing to relocate its production plant from the minimum transportation-cost optimum until the critical Isodapane of level v is reached, if the relocation allows for spatial concentration with another firm.

An infinite number of Isodapanes can be drawn around the minimum transportation-cost location, but beyond the critical Isodapane of level v, production is no longer profitable. Hence, firms A and B will relocate their plants to the area where their critical Isodapanes intersect. The territory where the two firms concentrate their production is represented by the shaded area in figure 2.1. The relocation raises the firms' profits because it allows them on the one hand to realize the agglomeration benefit v, while on the the other hand the increase in transportation costs remains below v (i.e., the shaded area is within both critical Isodapanes). In contrast, firm C will stay at its transportation-cost minimum since the critical Isodapane of C does not intersect with the critical Isodapanes of firms A and B.

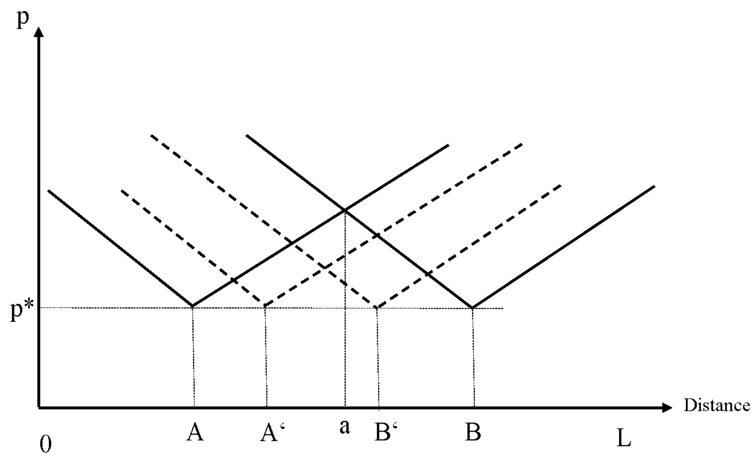
In this simple setting of punctual demand and punctual raw material supply, Weber provides a first argument why firms tend to concentrate in space despite the existence of transportation costs. An interesting extension is to introduce some dynamics into the model by varying the average transportation costs. A lower level of average transportation costs raises the radius of the critical Isodapanes, such that even areas which are further distant from the minimum cost optimum become potential agglomeration areas. Hence, reducing the average transportation costs yields a higher level of agglomeration. In figure 2.1, for example a sufficiently low level of transportation costs will induce firm C to cluster too. This implication is in line with the predictions of the New Economic Geography which state that the amount of regional concentration is inversely related to the level of trade costs.

The assumption of a constant punctual demand is of course very unrealistic and drives the results of Weber's model. Hotelling in his famous work from 1929 provides an argument why firms may still cluster if demand is not punctual but uniformly distributed. In such a scenario, the presence of transportation costs gives rise to a spatial division of markets. In figure 2.2, the abscissa defines a linear market with consumers being uniformly distributed between 0 and L.[3] The demand for the homogenous good which is supplied by two separate firms, located at points A and B, is assumed to be entirely

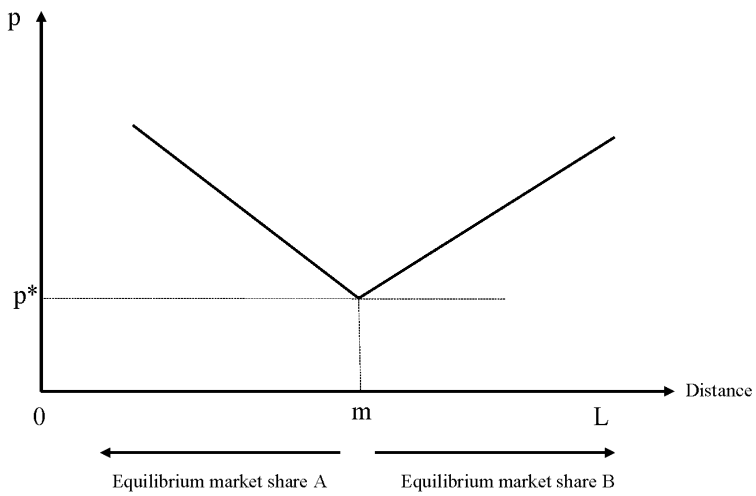
[3] A variation of the model locates firms around a circle and allows for a second good (see Salop, 1979).

inelastic to price changes. That is, consumers will always demand the same quantity, irrespectively of the price.[4]

Figure 2.2: The Hotelling Model



Initial Locations and Adjustment Process



Equilibrium

Note: The figures are based on Capello (2007, p. 35).

Due to the transportation costs τ, the consumer price of the homogenous good $p = p^* + \tau d$ increases linearly with the distance d from the respective firm. The producer

[4]The model is frequently illustrated by relying on an example with two ice-cream sellers at a beach. Tourists' demand for ice-cream is almost completely inelastically and they are located quite uniformly along the coastal line. Hotelling uses the Main Street of a town for illustration purposes.

price p^* is the same for both firms and is kept constant such that firms do not compete in terms of prices but only for market shares. Consumers closely located to the firms have the advantage of paying a lower price for the good, while consumers distant from the firms face a high price. Rational behavior implies that all consumers located between 0 and a buy the homogenous good from firm A, whereas all consumers located between a and L buy the good from firm B. Hence, the spatial structure separates the market between the two firms and gives the firms some monopoly power which, in principle, would allow them to charge higher prices.
However, for given producer prices p^*, firm A will decide to move to A' in order to gain market shares from firm B. Likewise, firm B will choose to move to B' in order to increase its market share. This process will continue until both firms are located at the middle of the market m, which represents the Nash equilibrium. In equilibrium, each of the two firms serves half of the market and the economic activity is concentrated at one point in space. Since the new equilibrium may also lead to an increase in the average transportation costs, Hotelling points out the inefficiency of the equilibrium, which gave an early justification for regional policy relocating industries.[5]
Yet, the model builds upon very strict assumptions. It can easily be shown that the result of spatial concentration breaks down if a third competitor enters the market. In order to maximize market shares, a new firm would locate in the periphery. Moreover, the assumption of price inelastic demand is crucial for the result. If demand declines with the price, firms will minimize transportation costs in order to maximize their revenues.

Another notable analysis of economic geography is von Thünen's model of the monocentric city (see von Thünen, 1826). Compared to Hotelling's model it virtually reverses the assumptions as it relies on a punctiform demand and a production site which is spread over space. The original model considers an isolated town which is supplied with agricultural goods by farmers who settle around the center of the town. By assumption, all goods have to be traded in the town's center. In contrast to the models described above, von Thünen is concerned with the allocation of land to different activities. In Weber's model, the agglomeration economies and the countervailing dispersion force – represented by the transportation costs – determined the location decisions of firms, whereas in von Thünen's model the allocation of activities is exclusively determined by the scarcity of the input factor land. The price farmers are willing to pay for land is the remainder left after subtracting the transportation and production costs from the

[5] It can be shown that the efficient distribution of firms, which minimizes total transportation costs requires that one firm is located at 1/4 of the range $0L$ and the other firm at 3/4 of $0L$.

price of the final good. Therefore, farmers are willing to pay a higher rent for locations close to the center of the town. Presuming that the quality of land is homogenous, a gradient of land rents is derived which declines from its maximum at the town's center to zero at a distant place where production becomes unprofitable due to the immense costs required for delivering the goods to the town's market place.

Agricultural goods vary in their perishability which implies different transportation costs per unit of distance for different goods. Von Thünen shows that, in equilibrium, the activities which achieve a higher price on the market and face higher transportation costs are located closer to the center than the activities which produce goods that are less perishable and achieve only a lower price. The equilibrium use of land can be determined by drawing a pattern of concentric rings around the center of the town, commonly known as *Thünenschen Ringe*. Within the first ring the activities with the highest transportation costs and the highest prices will be located, while in the most distant ring the activities with the lowest transportation costs and the lowest prices will be located.

The uncoordinated equilibrium of the von Thünen model can be interpreted as being efficient in so far as the market allocation of agricultural activities is such that it minimizes the town's total transportation costs. Maybe the most remarkable result of the model is that land rents are not simply determined by the quality i.e. fertility of the land but by the accessibility to the market. This is an important insight which is adopted by the New Economic Geography models where the factor rewards increase with the proximity to large markets.[6]

The von Thünen model was applied to the scenario of a modern city with commuters and a central business district by Alonso (1964) and is regarded as the starting point for a strand of literature mostly called Urban Economics. These models explain how production will be allocated around the center of a city or a metropolitan area but they cannot explain why agglomeration arises in the first place. Given the existence of a city center, the efficient allocation of activities is addressed and compared to the market outcome. However, in this thesis we are not concerned with the allocation of economic activity within an a priori specified city or metropolitan area, but we want to analyze the driving forces for the emergence of regional concentration and especially regional disparities. Moreover, we want to discuss the implications for regional policy instead of the consequences for urban policies. The former is concerned with a much broader regional scale where the policy makers try to create incentives for industries

[6] Chapter 4 proves empirically that accessibility plays an important role for location decisions.

to relocate from one region to another. The latter emphasizes the urban economic environment which for example includes preventing the degeneration of city districts or the emergence of city districts with a problematic population composition. The real estate market is almost always in the focus of urban economic policies, while the scarcity of land plays an inferior role for regional policy. The most important dimensions for regional policy are regional institutions, public infrastructure, and taxes. Moreover, regional policy implies that it has be governed on a federal level whereas urban policy operates in a decentralized manner.

Having laid out the differences, we will not go into the details of Urban Economics but introduce theories of Regional Science which analyze the hierarchical system of agglomerated regions; that is, the frequency we observe agglomerations in space or the optimal distance between individual agglomerations.

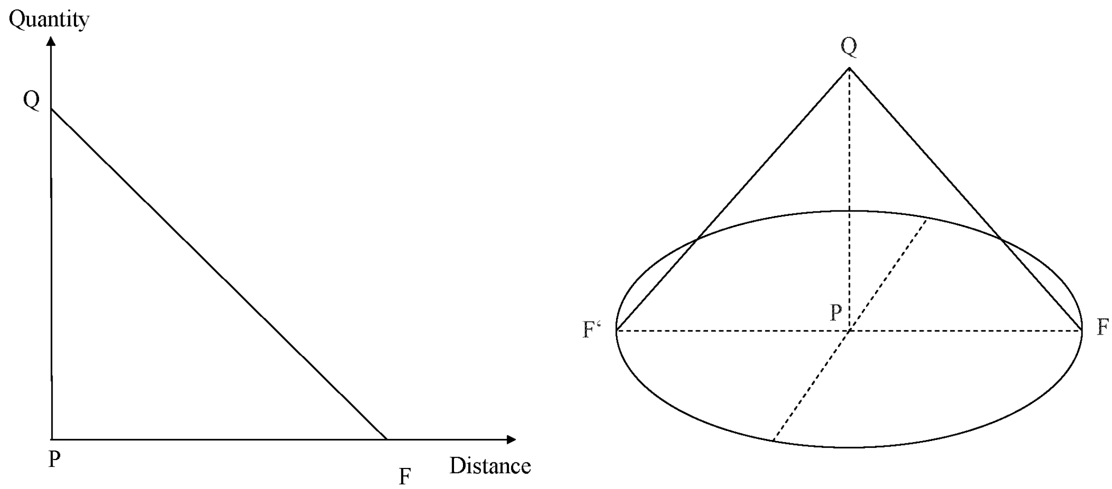
2.3 Central Place Theory

The location theories analyze the decision of firms over where to locate their production unit and how to allocate their various activities in space. It is shown that there is a clear tendency for firms to cluster. Yet, these theories ignore the existence of several agglomerations which represent alternative location options for firms. In reality, we observe numerous agglomerations and the questions of what determines the firms' decision in choosing a specific agglomeration and of why we observe different levels of agglomeration with different functions arise. Given that there are strong motives for firms to concentrate in space, the central place theory which was introduced by Christaller (1933) and Lösch (1940) constructs a hierarchical system which classifies agglomerations according to their size and functions. Based on a rather normative approach the models show that different sized agglomerations perform different roles and that there is an optimal distance between agglomerations of the same level. This optimal distance again increases with the hierarchical level of the agglomerations. Moreover, the central place theory demonstrates that such a hierarchical system should evolve as a result of market processes, despite the assumption of homogeneous space. These models are the first to formulate a complete pattern of economic activity for states, nations and federations.[7] Edwin von Böventer (1969) praises the central

[7]Christaller (1933) proves his theory by comparing the prediction of his model with the pattern of economic activity in Southern Germany. Lösch (1940) applies his theory to the distribution of agglomerations in Iowa.

place theory to rank "[...] among the most original and most important publications that have come out on economic relationships in space during this century."
In the sense that the distribution of activity over a range of different sized agglomerations is determined, the central place theory derives a general equilibrium of economic activity in space. In the following we will characterize subsequently the demand and supply side of these models.
Under the assumptions that the consumer price of a specific good increases linearly with the distance the goods need to be shipped, and that the quantity demanded by an individual consumer decreases linearly with the consumer price, Lösch (1940) constructs an aggregate spatial demand curve which is illustrated in figure 2.3.[8] The consumer price increases with the distance from the supplier location which implies, for uniformly distributed consumers, that the aggregate demand for the specific good decreases with the distance from the supplier's plant.

Figure 2.3: THE SPATIAL DEMAND CURVE



Note: The figures are based on Lösch (1940, p. 66).

At the firm's location P, the total demand will be given by Q, whereas beyond the distance F there will be no demand anymore. Assuming a homogenous space, this

[8]Note also, that the producer price is exogenously fixed.

simple linear market triangle can be rotated such that it translates into a cone with radius PF which reflects the circular market. Hence, the total demand faced by the firm in P is given by the volume of the cone multiplied by the population density. Note also, that the volume of the cone and therefore the total demand is inversely related to the transportation costs per unit of distance.

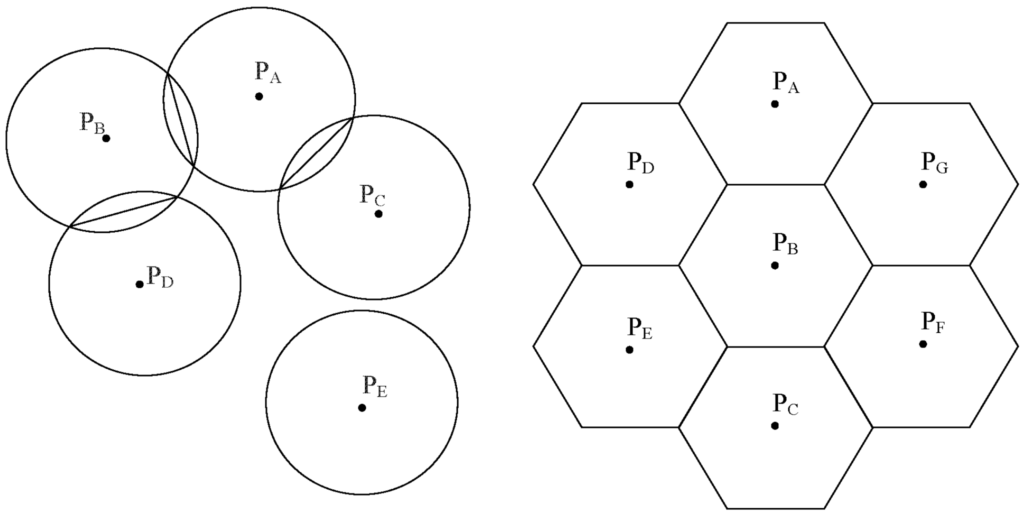
On the supply side, the territory Lösch considers is covered with many producers of the specific good (P_A, P_B, P_C,...) each of them characterized by a circular market area around the respective plant (see left panel of figure 2.4). Consumers are uniformly distributed across the territory. As long as there remains an area which is not covered by firms, new suppliers will enter the market and take the chance of earning a surplus. At the intersection of two circular market areas consumers would in general be willing to consume the good from both firms. However, since the good is homogenous and consumers are rational they will only purchase the good from the closer firm. This way consumers save transportation costs. Therefore, the long-run the equilibrium is characterized by a honeycomb pattern as is illustrated in the right panel of figure 2.4.[9] Each of the regular hexagons reflects the market area of one firm which is located in the middle of the hexagon. Thus, Lösch derives a spatial equilibrium by relying on rational consumers and profit-maximizing firms.

The general spatial pattern Christaller (1933) envisaged is the same as is derived above. However, in contrast to Lösch (1940) Christaller did not trace the market areas back to an individual spatial demand function but simply opted for the hexagon structure, because the regular hexagon represents the highest order polynomial (i.e., the polynomial most similar to a circle) which can completely cover the territory. He considers those features to be important basically for the same reasons as the hexagon emerged from Lösch's analysis. First, complying with consumer's minimization of transportation costs requires a shape which resembles a circle as much as possible, second competition among firms requires that the market areas do not overlap, and third serving every consumer requires that no part of the territory is left out.

Once the market areas of the specific good are set up in the territory under consideration, Christaller defines different categories of goods which he calls the central goods of n^{th} order. The distinctive characteristics between these goods of different orders are the respective sizes of their market areas. The size of a good's market area again is determined by the maximum distance which consumers are willing to travel in order to purchase it. This is called the range of a good.

[9]Note that higher order polynomials as for example heptagons or octagons are not able to cover the territory completely but would leave some areas spare. Therefore, the assumption that the total territory is covered with demand which firms will meet requires regular hexagonal market areas.

Figure 2.4: Market Areas



Initial Situation **Long-Run Equilibrium**

Note: The figures are based on Lösch (1940, p. 69).

Moreover, a good or service will only be provided if a circle with a radius equal to the good's specific range comprises enough population in order to constitute sufficient demand for the firm to become profitable. In this respect the theory reflects increasing returns to scale.

The central good of the highest order has the highest range and accordingly the largest market area which continuously decreases with the good's order. Hence, a higher order good is consumed even if it has to be transported for a long distance whereas consumers are not willing to pay the same amount of transportation costs for a good of lower order. A n^{th} order good constitutes a n^{th} order market area which contains a n^{th} order central place where the plant is located. According to Christaller, central goods and services of high order are, for instance, supplied by media outlets, the banking sector or to a large extent the retail sector. Importantly from a policy point of view he also regards many public services such as theaters, schools, public administration facilities or public health services to be higher order services.

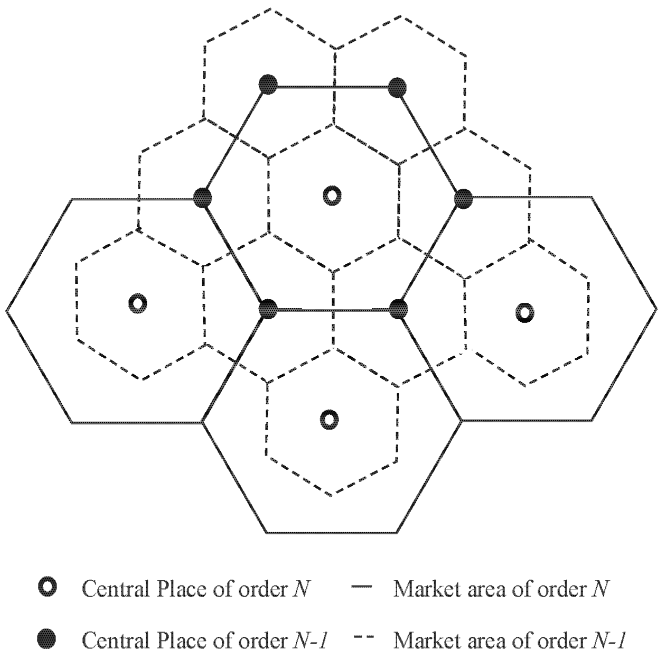
Each of the n^{th} order goods and services is supplied according to the spatial honeycomb pattern shown in figure 2.4 where the sizes of the regular hexagons increase with the respective order. Starting from the highest order good *N*, Christaller argues that the producers of the immediately lower order good *N-1* should choose to supply at the central place where the producers of the higher order good are already in situ. Thereby they can benefit from agglomeration economies. Christaller maintains this hypothesis across all orders of goods, such that the central place which provides the highest order good N will also provide all lower order goods $(N-1, N-2, ..., N-t)$, where $t+1$ represents the total number of orders.

Since the $(N-1)^{th}$ order good has a smaller range than the N^{th} order good, there have to be several $(N-1)^{th}$ order market areas within the N^{th} order market area. The suppliers of the $(N-1)^{th}$ order good will choose the additional locations by minimizing the number of supply spots necessary to cover the whole market area of the higher order good N. This is achieved by locating the supply of the $(N-1)^{th}$ order good at the vertexes of the market areas of the N^{th} order good as is shown in figure 2.5 where the ringlike dots represent the central places of the N^{th} order good and the full dots represent the central places of the $(N-1)^{th}$ order good. Accordingly, six central places of the $(N-1)^{th}$ order develop around each N^{th} order central place. Since those new lower order central places are located exactly at the vertexes of the higher order market area, each of them is assigned by only one third to the higher order market area. In total this means that there are $k = 6/3 + 1$ central places of the immediately lower order within each higher order market area. Note that one central place of the lower order good always coincides with the central place of the immediately higher order good as described above.

Christaller keeps the factor of proportionality k constant throughout the whole hierarchy, that is each n^{th} order market area (n^{th} order central place) corresponds to $k = 3$ market areas (central places) of order *n-1*. [10] Hence, the system of urban hierarchy as it is illustrated in figure 2.5 can easily be extended to an infinite number of orders.

[10] The described hierarchical principle is called the market principle. In addition to this structure, Christaller derives the transportation principle which implies $k = 4$ and the administrative principle which implies $k = 7$. However, the qualitative implications of a hierarchical system are the same as in the market principle.

Figure 2.5: CHRISTALLER'S SYSTEM OF CENTRAL PLACES



Note: The figure is based on Christaller (1933, p. 71).

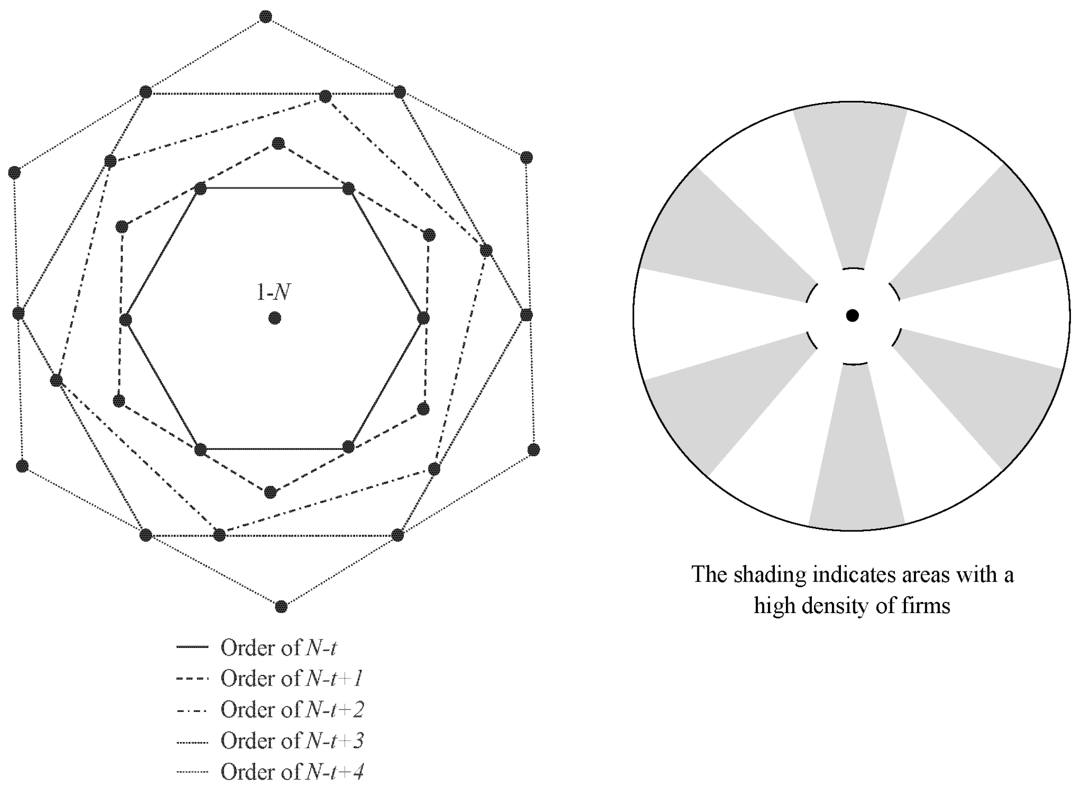
As stated above Lösch proceeds in his analysis closer to the standard economic framework of individual demand and supply than Christaller. However, the main difference between the two models is that Lösch refrains from the assumption of a constant factor of proportionality. In Lösch's approach each category of goods is assigned a specific market-area size such that the ratio between the number of $(n-1)^{th}$ order market areas and the number of n^{th} order market areas does not necessarily equal three anymore. Therefore, Lösch's work can be interpreted as a generalization of Christaller (1933).

Lösch's analysis starts from the lowest level of hierarchy, that is from the production places of the $(N-t)^{th}$ order good. The locations of these firms are again arranged according to a regular hexagon. One of the firms is located in the center of the hexagon and six others on the vertexes of the hexagon. The market area of the immediately higher order good always constitutes a multiple of the lower order market area.[11] Therefore, the location of higher order firms always follows a larger hexagon than the

[11]Lösch uses market areas which are 3-,4-,7-,9-,12,-,13-,16-,19-,21-,25- times the market area of the lowest order good. Note that Christaller's model implies a constant multiple of 3.

location of lower order firms. For the purpose of determining the overall distribution of economic activity, Lösch puts these hexagons of different size on top of each other. As is shown in the left panel of figure 2.6, all hexagons share a common center which produces the complete range of goods. The vertexes of the smallest hexagon reflect the additional locations of the lowest order firms and the vertexes of the largest hexagon reflect the additional locations of the highest order firms.

Figure 2.6: Lösch's System of Central Places



Note: The figures are based on Lösch (1940, p. 75 and 80). The dots on the vertexes represent the firm locations of the respective order.

Even though they do not precisely state the source of agglomeration economies, both Lösch and Christaller assume that firms have a preference for locating close to each other. For this reason, Lösch rotates the superimposed hexagons such that the maximum density of firms in some areas is reached (i.e., the vertexes of the different sized hexagons lie as close as possible to each other or even coincide). Note that the firm

locations remain on the vertexes of the hexagons. This yields a landscape with six sectors characterized by a high density of firms and six sectors with a low density of firms which is illustrated in right panel of figure 2.6. Lösch claims that this spatial structure of economic activity minimizes the overall transportation costs while complying with the firms' requirement for a sufficiently large market area. He describes this system as "[...] the ideal landscape of economic activity" (Lösch, 1940, p.90).
Just as in Christaller's model, the most central agglomeration provides all goods and services. However, the clear interrelation between the agglomeration's order and its specific functions disappears through relaxing the assumption of a constant factor of proportionality. An agglomeration of n^{th} order does not necessarily provide all lower order goods and services $(n-1, n-2, ..., n-t)$ anymore . Hence, compared to Christaller's model the Lösch approach allows more flexibility, but it also comes with much more fuzzyness. Moreover, in an empirical application of his model, Lösch essentially finds that a constant factor of proportionality ($k \simeq 4$) fits very well. Since the qualitative implications of the two models are similar, we will in the following rely on Christaller's workhorse model when we refer to the central place theory.

In recent years, the central place theory has received only little attention even though it represents an appealing way of deriving a general spatial equilibrium by structuring the partial equilibria of individual categories of goods and services. The theory offers a convincing explanation for why we observe certain hierarchical structures of agglomerations with different functions. In addition it allows for the derivation of an optimal distance between individual agglomerations.
The on-going trend of regional integration in Europe provides an interesting application for the central place theory. Rising integration implies a reduction of transportation costs which, according to the central place theory, enlarges the individual market areas and thereby raises the optimal distance between agglomerations. That is, for the territory as a whole we should observe a rising spatial concentration of economic activity. This is completely in line with the more recent New Economic Geography approaches and we will argue that both theories complement each other in explaining the nature of economic geography.
However, the central place theory described above faces notable shortcomings. First of all, these models neglect any changes in the population distribution across space. Consequently, von Böventer (1962) criticizes the central place theory to face a contradiction with respect to migration. On the one hand, the models presume a constant uniformly distributed demand while on the other hand explaining the evolution of agglomerations. In fact one should expect the spatial distribution of demand to

be significantly affected by the evolution of agglomerations. This means that the distribution of demand is by no means exogenous to the formation of agglomerations. Second, the demand side does not allow for any substitution between the goods. Yet, the demand for each good should be influenced by the producer prices of all other goods as well as by the distance from the producers of all other goods. Third, the central place theory misses a profound explanation of why agglomeration economies arise in the first place. It is taken for granted that firms prefer to cluster, but there is neither a micro-foundation of this preference nor an analysis about what determines the magnitude of agglomeration economies. In any case there are also forces that counteract the agglomeration economies, as for example the intensity of competition for scarce input factors which is increased in agglomerations. Hence, a theory of economic geography should discuss the determinants of the relative strength between agglomeration and dispersion forces.
In the following we will introduce models that describe how agglomeration economies emerge from the behavior of individual firms, laborers and consumers. Moreover, we will analyze the crucial parameters which determine the strength of the agglomeration economies. It will be argued that the recent theories on economic geography constitute a complementary analysis to the theory of spatial hierarchy.

2.4 Micro-Foundations of Agglomeration

Agglomeration economies represent any kind of advantage firms, laborers or consumers experience from being close to each other. A common characteristic of all those benefits is that they come, apart from relocation costs, for free. For instance, firms can increase their productivity by being close to each other or their profits by being close to consumers without paying an explicit price for this advantage. Hence, the individual location decisions create external effects. This means that each firm exerts a direct effect on the productivity or profit of other firms which is not internalized by the market (i.e., there is no market price for this productivity or profit enhancing effect). The theories described so far did not analyze the microeconomic determinants of agglomeration economies which are important in order to understand the dynamic evolution of economic geography. Fujita et al. (1999, p.4) call the processes which determine agglomeration economies the "black box" of economic geography.
Marshall (1890) provides in his chapter on industrial organization and the concentra-

tion of specialized industries a conductive classification of the processes that determine agglomeration economies. He distinguishes between three categories of agglomeration mechanisms.[12] First, agglomerated industries benefit from a constant pool of skilled laborers.[13] This allows them to react immediately to changing market conditions and reduces the costs of acquiring qualified personnel. Following Krugman (1991a), we call this agglomeration channel *Labor Market Pooling*. Second, there are knowledge spillovers which increase the productivity of localized industries. This agglomeration force is closely related to Michael Polanyi's definition of tacit knowledge (see Polanyi, 1958), precisely because there are constraints in the transmission of knowledge and skills. The potential of knowledge, skills and information to be shared decreases with the distance. Therefore, firms which are located closer to other firms will benefit much more from the transmission of knowledge than firms which operate in spatial isolation. Clustering of firms allows ideas to be copied by other firms and then even come back to the initiator after being further developed.[14] We will call this agglomeration channel *Knowledge Spillovers*. The last category, which received by far the most attention, is commonly described as the *New Economic Geography*. Here the agglomeration tendency results from economies of scale at the level of individual firms. On the one hand, firms are willing to move to the larger market as this allows them to increase their profits. On the other hand, consumers prefer the market with more firms which allows them to save transportation costs. Hence, a circular causality arises which yields a concentration of economic activity. The same process occurs when replacing the consumers with firms demanding their reciprocal output as intermediate input goods.[15]

2.4.1 New Economic Geography

Over the recent years, a range of models have been added to the New Economic Geography literature, yet the fundamental channels of agglomeration in most instances resemble the seminal work by Krugman (1991b). Therefore, we introduce Krugman's

[12]There are other popular schemes as for example the more technical classification by Duranton and Puga (2004), too.

[13]"[..] a localized industry gains a great advantage from the fact that it offers a constant market for skill" Marshall (1890, p.333).

[14]"[..] if one man starts a new idea, it is taken up by others and combined with suggestions of their own; and thus becomes the source of yet more new ideas" Marshall (1890, p.332).

[15]"[..] the economic use of expensive machinery can sometimes be attained in a very high degree in a district in which there is a large aggregate production [...]. For subsidiary industries devoting themselves each to one small branch of the process of production, and working it for a great many of their neighbors, are able to keep in constant use machinery of the most highly specialized character; and to make it pay its expenses [...]" Marshall (1890, p.332).

core-periphery model and subsequently discuss important modifications and extensions. The detailed formal derivation of the model is provided in the appendix of this chapter.

The core-periphery model features two regions which are initially identical. Each of them produces agricultural and manufactured goods, which can be traded across regions. Trading in manufactured goods is inhibited by trade costs while, for analytical convenience, it is assumed that agricultural goods can be shipped for free. Note that we use the broader term trade costs in the following which is commonly used in the New Economic Geography literature as is accounts not only for all kinds of transportation costs but also incorporates any other trade impediments. The workers are specific to each sector and only the employees of the manufacturing sector are mobile across regions. Since the agricultural sector serves as the anchor of the model, it is constructed in a very simple way. It produces a perfectly homogenous good under perfect competition and marginal cost pricing. The technology is such that workers represent the variable input and fixed costs do not occur. Moreover, exporting free of trade costs equalizes the prices of agricultural goods as well as the agricultural sector's wages between both regions.[16]

In contrast to the agricultural sector, firms in the manufacturing sector face increasing returns to scale and monopolistic competition as in Dixit and Stiglitz (1977). On the demand side, consumers are characterized by love-of-variety preferences, which means that they prefer to consume as many differentiated varieties of the manufactured good as possible. These varieties can be substituted according to a constant elasticity such that individuals consume higher quantities of the local varieties which come free of trade costs. This type of preferences allow firms to charge a mark-up on the marginal costs of production. Moreover, fixed costs in combination with love-of-variety preferences imply that each variety of the manufactured good will be produced by only one firm preferably at the location with the highest demand.

The agglomeration process takes place in the manufacturing sector while the agricultural production as well as the agricultural workers are at all times symmetrically distributed across the two regions. Consequently, agglomeration does not imply total depopulation of peripheral regions, but just the closure of certain industries and migration of a fraction of the total population. This is a realistic setting since we observe concentration of economic activity and mobile citizens in certain areas of Europe and other federations but we do not observe a complete depopulation.[17]

[16] Note that the technologies are identical in both regions, too.

[17] However, the anchor sector which can be interpreted as a the primary sector is often disabled due to labor market rigidities.

Given an outset with identical population in both regions, the migration of one worker due to any exogenous reason can have drastic effects. If a worker of the manufacturing sector in region *1* moves to region *2*, the demand for manufactured goods in region *2* will exceed region *1*'s demand since every local worker represents also a local consumer. Hence, firms operating under increasing returns to scale have an incentive to move to region *2* which features now a slightly larger market than region *1*. At the same time the relocation of a manufacturing firm from region *1* to region *2* constitutes an incentive for further migration of workers to region *2*, since the costs of living are lower in the region with more manufacturing firms. This is because consumers save transport costs if they reside in the region with the higher number of manufacturing firms. Taken the two effects together, a self-reinforcing circular causality may evolve which does not end before all manufacturing firms and manufacturing workers are located in core region *2*.

However, Krugman's model features also a dispersion force which may break the circular process. Each additional firm in the core region intensifies competition and thereby reduces the revenue of the other manufacturing firms. Ceteris paribus, the local expenditure is subdivided between more firms which reduces their individual revenue. Note that the model allows for free market entry and workers represent the variable as well as the fixed input factor such that the revenue of each manufacturing firm is passed over to the wages in the manufacturing sector.[18] Hence, intensified competition reduces the workers' incentive to immigrate.[19]

Whether the self-reinforcing agglomeration is set in motion depends on the relative strength of the dispersion force against the agglomeration forces. The crucial parameters determining the countervailing forces are the trade costs (τ), the elasticity of substitution between the manufactured varieties (σ), and the expenditure share of the manufactured goods (μ). In the following we will illustrate these agglomeration and dispersion forces as well as the sensitivity of the long-run equilibrium with respect to changes in the exogenous parameters.

To get back to the central place theory, the following analysis will allow us to infer the determinants of the distance between the central places depicted in figure 2.5 and thereby endogenize the order of goods.

[18]The production of manufactured varieties involves workers as fixed input and as variable input, therefore, the zero pure profit condition implies that the firms' operating profit is passed over to the wages in the manufacturing sector. A variant of the model calls the fix input workers entrepreneurs which facilitates the intuition of the model. These entrepreneurs earn the operating profits of manufacturing firms such that pure profits are zero.

[19]Baldwin et al. (2003) calls this the *market- crowding effect*.

As already mentioned, manufacturing firms prefer the larger market. This is because their revenues increase with local expenditure and – due to increasing returns and monopolistic competition – so does their operating profit. Moreover, holding the wages constant, a one percent increase in local expenditure for manufactured goods requires a change in local manufacturing production, and thereby manufacturing employment, of more than one percent in order to restore zero pure profits. Hence, the inflow of manufacturing firms has to be more than proportional to the rise in local expenditure.[20] The *New Trade Theory* named this effect the home-market effect.

As the firms settle, the number of jobs and the local expenditure increases which encourages further firms to relocate. Krugman (1991b) labels this self-reinforcing mechanism the *backward linkage*.[21] Note that the home-market effect does not hinge on the circular causality but applies also to exogenous expenditure shares as we will see in Chapter 3.

The second agglomeration force develops from the consumer preference for diversified varieties which implies that consumers demand every variety. The more firms are located in the home region, the more varieties can be consumed free of trade costs. If the share of local manufacturing firms increases, the price index of consumers is reduced because less of the varieties have to be imported. This is called the price-index effect. A lower price index again constitutes an incentive for immigration. Due to the assumption that workers represent the fixed input for manufacturing firms, the number of workers is proportional to the number of firms. Hence, immigration of workers is equivalent to an inflow of manufacturing firms which further reduces the regional price index. Krugman (1991b) labels this part of the process *forward linkage*.[22]

As already mentioned, the increased intensity of competition in the regions with the higher number of manufacturing firms countervails the agglomeration force. This way a symmetric equilibrium is stabilized until critical levels of the exogenous parameters are reached. The model's migration equilibrium is reached if either the real wages of manufacturing workers (ω_1, ω_2) are equalized or a corner solution applies where all manufacturing workers are concentrated in one region. Figure 2.7 depicts the real wage difference between the two regions against the share of manufacturing workers in region *1* (s_L) for three levels of trade costs. This kind of illustration allows us to analyze the

[20] Due to increasing returns to scale profits increase with the absolute market size. Accordingly, a regional shift in demand causes not only a shift of manufacturing firms until the ratios between consumers and firms are equalized again but implies additional relocation of firms to the region that is larger in absolute terms.

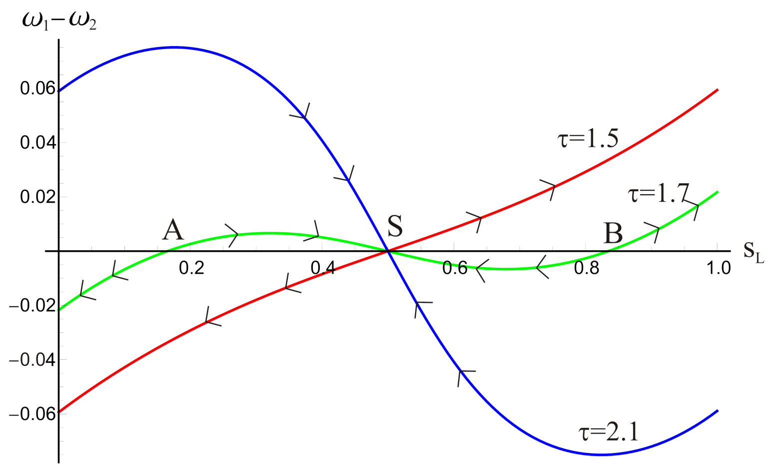
[21] Possibly more to the point is the term *demand-linked circular causality* which is used by Baldwin et al. (2003).

[22] Baldwin et al. (2003) labels it the *cost-linked circular causality*.

effects of any exogenous parameters on the equilibrium distribution of industry and will repeatedly be used in the following.

The blue line in figure 2.7 indicates the case of high trade costs ($\tau = 2.1$). Three potential equilibria occur for high trade costs: one interior at point S (where $s_L = 0.5$) and two corner solutions with full concentration of manufacturing workers in either of the two regions (i.e., $s_L = 0$ or $s_L = 1$). The real wage difference ($\omega_1 - \omega_2$) is positive if less than half of the manufacturing workers are located in region *1*. Therefore, if full agglomeration in region *2* prevails, workers will have an incentive to migrate from region *2* to region *1* such that the corner solution with $s_L = 0$ is unstable. If the share of manufacturing workers in region *1* is greater than one half, the real wage in region *1* is lower than the real wage in region *2* and workers will emigrate from region *1*. Hence, the corner solution with $s_L = 1$ is unstable, too. After all, if the symmetric equilibrium with $s_L = 0.5$ applies, deviations from the equilibrium will always result in a movement back to the equilibrium such that the symmetric equilibrium at point S is stable. Note also that the regional share of manufacturing workers reflects exactly the regional share of manufacturing firms, as the number of workers is proportional to the number of firms.[23]

Figure 2.7: THE CORE-PERIPHERY MODEL - STABILITY OF EQUILIBRIA

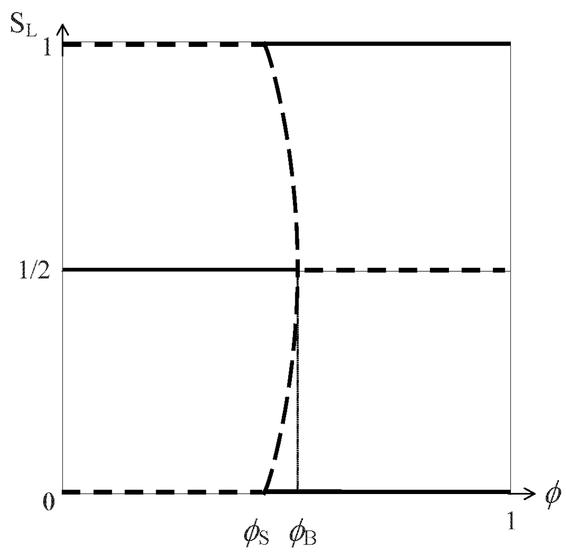


Note: In this figure we set the elasticity of substitution to $\sigma = 5$ and the expenditure share of the manufacturing sector to $\mu = 0.4$. Iceberg trade costs imply $\tau > 1$.

[23]In both regions the same constant number of workers have to be employed in order to open up a manufacturing firm.

In the case of low trade costs ($\tau = 1.5$) which is illustrated by the red line in figure 2.7, the corner solutions represent stable equilibria. For a marginal deviation from $s_L = 0$ the wage rate in region *2* exceeds the wage in region *1*. Consequently, there is no incentive for workers to leave the agglomeration region. The same is true for the equilibrium with full agglomeration in region *1*. In contrast, the symmetric equilibrium is unstable. The positive slope of the red line implies that a region becomes the more attractive, the higher its share of manufacturing. Therefore a marginal deviation from the symmetric equilibrium, that is one worker's decision to migrate, will result in a core-periphery pattern where $s_L = 1$ or $s_L = 0$.
For medium trade costs ($\tau = 1.7$) the green line in figure 2.7 depicts three interior equilibria. However, applying the same reasoning as above shows that only the corner solutions and the symmetric equilibrium are stable. The equilibria at points A and B are unstable.

Figure 2.8: Agglomeration and Trade Freeness



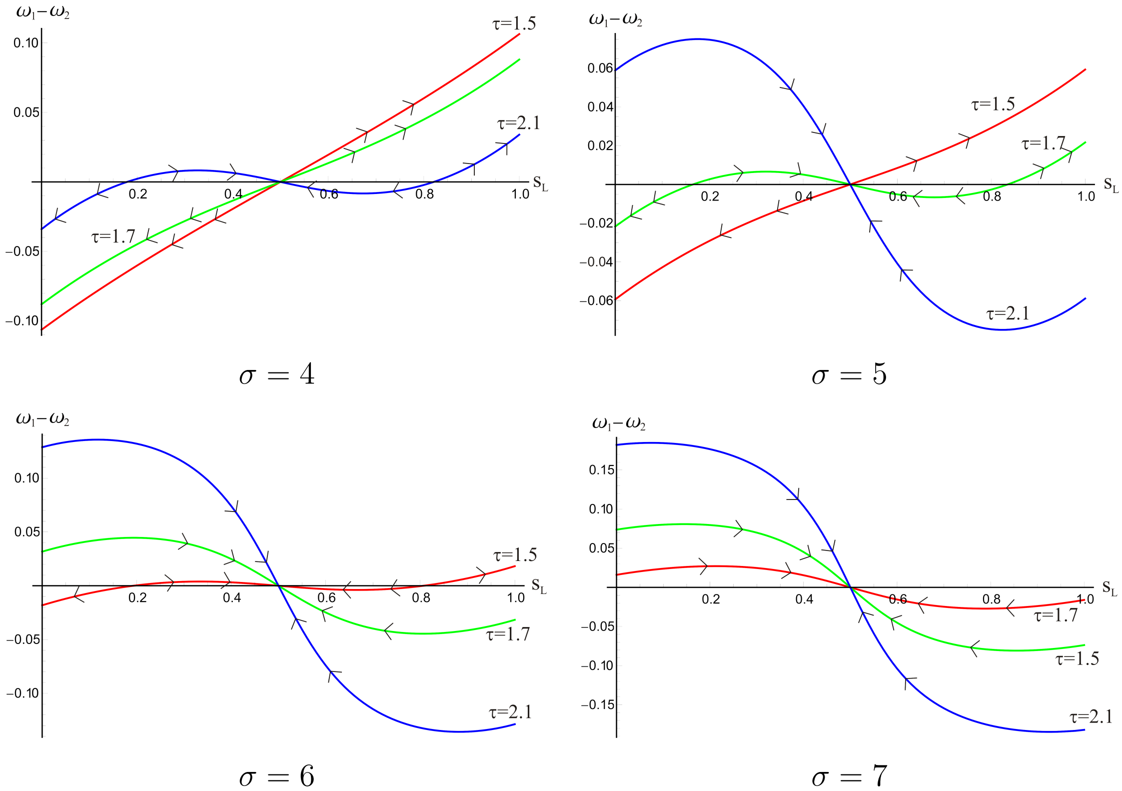
In order to illustrate the relationship between trade costs and the occurrence of agglomeration for the whole range of trade costs, it is necessary to translate the trade costs into a variable which is bounded to a closed interval. For this purpose it is common to define the trade freeness – being the inverse of the trade costs – as $\phi = \tau^{1-\sigma}$ where $\sigma > 1$ denotes the elasticity of substitution. Note that ϕ lies always between zero and one. We distinguish between two important threshold levels of trade freeness. First, the lowest level of trade freeness (i.e., highest level of trade costs) where full agglomeration becomes stable (ϕ_S i.e. $\tau_S = \phi_S^{\frac{1}{1-\sigma}}$) and second the highest level of trade freeness (i.e., lowest level of trade costs) where the symmetric equilibrium is just stable (ϕ_B i.e. $\tau_B = \phi_B^{\frac{1}{1-\sigma}}$).[24] Figure 2.8 illustrates the stability of agglomeration equilibria and symmetric equilibria for the whole range of trade freeness. This kind of diagram is commonly referred to as bifurcation diagram.[25] The dashed lines in figure 2.8 represent unstable equilibria while the solid lines represent stable equilibria. For low trade freeness up to ϕ_S only the symmetric equilibrium is stable and an agglomeration equilibrium will always be dissolved by a marginal exogenous shock. Yet, if the trade freeness crosses the threshold level ϕ_B, the symmetric equilibrium becomes unstable and cannot be maintained. Only for a small range between ϕ_S and ϕ_B do both agglomeration and dispersion represent stable equilibria. Hence, the likeliness of agglomeration increases with a higher trade freeness or equivalently with lower trade costs.

The second exogenous parameter which was held constant above is the elasticity of substitution between the varieties (σ). Figure 2.9 plots the real wage differences for different elasticities of substitution. Presuming a small elasticity of substitution ($\sigma = 4$), we observe strong agglomeration tendencies. The symmetric equilibrium remains stable only for high trade costs ($\tau = 2.1$). The agglomeration forces become weaker when the elasticity of substitution is raised, which manifests in a clockwise rotation of the graphs in figure 2.9. Presuming $\sigma = 7$, the slope of each curve is negative at $s_L = 0.5$ such that the symmetric equilibrium is stable for any level of trade costs, while the agglomeration equilibria are always unstable. Hence, the lower the elasticity of substitution between the varieties the stronger the agglomeration forces. This is because it becomes more important for firms as well as for consumers to reside in the large market if consumers are bound to the specific varieties and cannot substitute between them.

[24]The explicit values can be found by calculating the first derivative of the real wage difference at the respective equilibria. If the derivative is positive, the equilibrium is stable; if it is negative the equilibrium is unstable (see figure 2.7).

[25]The same manner of illustration will be used in Chapter 3, figure 3.1.

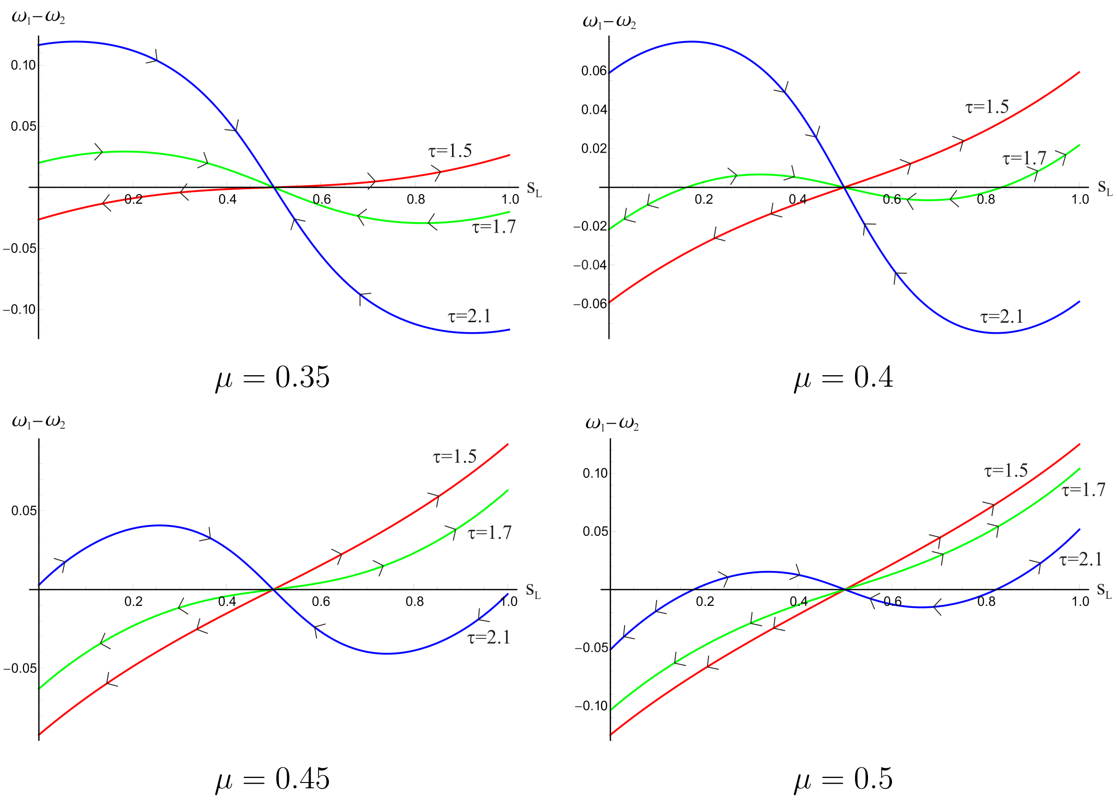
Figure 2.9: AGGLOMERATION AND ELASTICITY OF SUBSTITUTION



Note: In this figures we set expenditure share of the manufacturing sector to $\mu = 0.4$.

Lastly, we vary the income share that is spent on manufacturing goods (μ) while keeping the elasticity of substitution constant. This is shown in figure 2.10. A higher expenditure share of manufactured goods strengthens consumers' interest in residing close to manufactured firms which is reflected by the anti-clockwise rotation of the curves. Remember that the symmetric equilibrium is stable if the curve exhibits a negative slope at $s_L = 0.5$ while it is unstable if the slope is positive. Therefore, the symmetric equilibrium loses stability if the expenditure share is raised. In case $\mu = 0.35$, the symmetric equilibrium is unstable only for very low trade costs while it is unstable for each but the highest level of trade costs if $\mu = 0.5$. Hence, the higher the expenditure share of manufactured goods the stronger the agglomeration forces.

Figure 2.10: Agglomeration and Expenditure Share

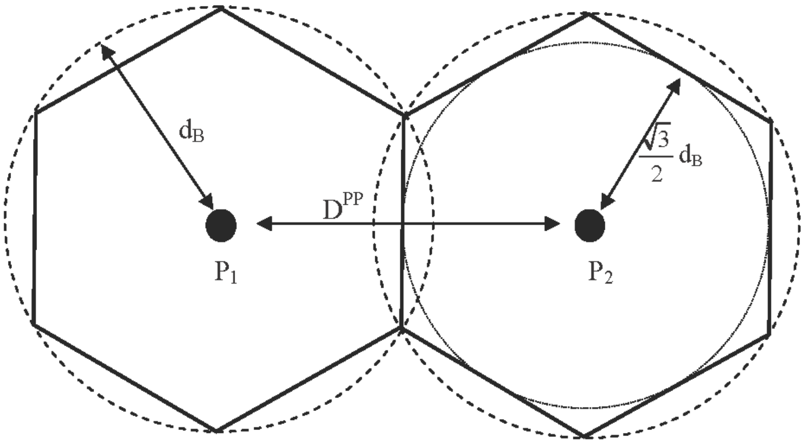


Note: In this figures we set elasticity of substitution to $\sigma = 5$.

Summing up the analysis, sectors with a low elasticity of substitution and a high expenditure share feature strong agglomeration economies. Put differently, these sectors will already concentrate for higher levels of trade costs than sectors with a high elasticity of substitution and a low expenditure share.

In order to relate these findings to the comprehensive spatial pattern provided by the central place theory, we define the trade costs as a function of the spatial distance $\tau = f(d)$ where $f'(d) > 0$. Recalling the honeycomb pattern of economic activity in Section 2.3, we can translate the critical trade costs for which the symmetric equilibrium becomes unstable into the radius of the market area surrounding a central place $d_B = f^{-1}(\tau_B)$. The symmetric equilibrium which is characterized by two central places is unstable for $\tau < \tau_B$ and accordingly for $d < d_B$. Figure 2.11 illustrates the market areas of two central places (P_1, P_2). Within the circles of radius d_B there is no room for a second central place because firms have too strong an incentive to cluster (i.e. the symmetric equilibrium is unstable). Hence, any economic activity located within the

Figure 2.11: The Distance between Central Places



circle of radius d_B surrounding central place P_1 will relocate to P_1. The same is true for central place P_2. Only if the distance from the respective central place exceeds d_B, can an additional central place form. At the intersection of the two circular market areas economic activity will relocate to the closer of the two central places such that regular hexagons emerge. Accordingly, the distance between two central places (D^{PP}) equals two times the radius of the hexagons' incircle, that is $D^{PP} = d_B\sqrt{3}$ where $d_B = f^{-1}(\tau_B)$.[26] If the agglomeration forces become stronger, the symmetric equilibrium becomes unstable already for higher trade costs (i.e., τ_B increases); consequently the radius of the market area (d_B) increases which yields a higher distance between two central places (i.e., D^{PP} increases). From our analysis above we infer that the distance between individual central places (D^{PP}) increases with the sectoral expenditure share and decreases with the elasticity of substitution. Further determinants of D^{PP} will be derived in sections 2.4.2 and 2.4.3. Regarding the hierarchical system, we can can interpret sectors with a lower elasticity of substitution and a higher expenditure share as higher order sectors which are characterized by larger market areas than lower order sectors.

In contrast to the central place theory, the core-periphery model considers demand to be endogenous too. Therefore, a central place as derived above will not only exhibit a

[26]A hexagon can be divided into six equilateral triangles. The radius of the hexagon's incircle is given by the triangles' altitude which is from the Pythagorean theorem $\frac{\sqrt{3}}{2}a$ where a denotes the length of the triangles' sides. These sides are again equal to the radius of the circle circumscribing the regular hexagon which represents the market area.

higher density of firms but also a higher population density which resolves the major inconsistency of the central place theory.

The core-periphery model was modified in various dimensions. An important variation are the vertical-linkage models, which assume that workers are immobile but adopt instead input-output linkages between firms of the manufacturing sector. The product of each firm makes an intermediate good in the production process of other firms. In doing so, a similar circular causality arises which implies agglomeration after crossing a certain threshold level of trade costs (see Krugman, 1995 and Fujita et al., 1999). A shortcoming of the core-periphery model is the prediction of an immediate jump from the symmetric to the full agglomeration equilibrium. Reality is not characterized by only two potential patterns of economic activity – full agglomeration or complete symmetry – but instead we observe a continuous set of allocations. By introducing additional dispersion forces, the model features a continuous process of agglomeration. Puga (1999) introduces diminishing returns in the manufacturing sector which act as a dispersion force, while Pflüger (2004) assumes a quasi-linear utility function which eliminates income effects and thereby weakens agglomeration tendencies. However, regarding public-policy-oriented welfare analysis the main issue of the core-periphery model remains the fact that the model has to be solved numerically and that it becomes analytically intractable when introducing explicit policy instruments. As Baldwin et al. (2003) put it, this category of models always faces a trade-off between tractability and richness of features, therefore we will rely in chapters 3 and 4 on a modification of the core-periphery model which allows us to introduce a fourth process of agglomeration and to identify the effects of federal infrastructure provision on the distribution of economic activity as well as to perform a welfare analysis. Chapter 4 will also provide an empirical test of the New Economic Geography.

2.4.2 Knowledge Spillovers

Knowledge is a nonrival good which can be used by an infinite number of people at the same time. Therefore, acquiring knowledge exerts a positive externality on others. If an individual decides to invest in knowledge, this will not only augment his own productivity but may benefit the productivity of many others too. In the economic literature, the role of knowledge spillovers gained importance with the *endogenous growth theory* which considers knowledge spillovers to be a side product of physical or human capital (see Romer, 1986 and Lucas, 1988). However, the transmission of knowledge is possibly limited by its geographic scale such that only close neighbors can benefit from spillovers. This provides an incentive for firms to settle in the neighborhood of each other. Hence, the possibility to share information and knowledge can act as a strong agglomeration force.

It seems plausible that the transmission of ideas is facilitated in particular within similar industries since specific ideas and techniques are not necessarily useful in every other industry. Therefore, the clustering of certain industries, as for example the information technology industry in the Silicon Valley or the biotech industry in the south-west of Munich, are often attributed to knowledge spillovers. Since these positive externalities augment the overall productivity, governments often actively encourage clustering of specific industries. Yet, knowledge spillovers not only apply to exceptionally innovative high-tech industries but the accumulation of knowledge is important for any industry. A broad definition of productivity-enhancing knowledge also includes organizational expertise which may be copied from nearby firms.[27]

Assuming invisible spillovers is a very popular approach in economic geography but the explicit process of innovation is rarely described.[28] Yet, we argue that the plain assumption of spillovers is sufficient as long as the underlying processes do not qualitatively affect the outcome. In any case it remains virtually impossible to distinguish between invisible mechanisms generating knowledge spillovers by standard econometric methods.

In our analysis we do not emphasize the processes of innovation but show that the assumption of a positive spillover externality leads to a similar agglomeration mechanism as described in Section 2.4.1. Certainly, the determinants of agglomeration will be different from above. In doing so we rely on a simple model which was introduced by Helpman (1998).

The model features two regions $i \in \{1, 2\}$ which are endowed with workers (L_i) and land (G_i). Workers are perfectly mobile whereas the available land is exogenously given and can be consumed only in the region of residence. Hence, the land market acts as a dispersion force. Besides land (g), individuals consume a homogenous manufactured good (c). For the sake of simplicity, shipping the manufactured good from one region to the other does not involve trade costs. The preferences of individuals in both regions

[27]Note that Nelson and Phelps (1966) already recognize the importance of knowledge spillovers and imitation in their model of economic growth.

[28]Duranton and Puga (2001) are a notable exception. They build a model where firms learn from other firms about their ideal production process. At the early stage of the product life cycle firms stay in diversified cities while in later stages of the cycle they move to specialized cities.

are given by:

$$u = g^{\mu} c^{1-\mu}. \tag{1}$$

Land is the numéraire good (i.e., its price is equal to unity) while the price of the manufactured good is denoted by p_c. The total income in region i (Y_i) originates from wage income and land rents. It is presumed that the total amount of land is owned by the workers. Moreover, the ownership of land is distributed equally across workers and regions such that the sum of income in region i is given by the aggregate wages in region i ($w_i L_i$) plus region i's share of total rent income.[29]

$$Y_i = w_i L_i + \frac{\mu L_i}{(1-\mu)L}(w_1 L_1 + w_2 L_2) \text{ where } L = L_1 + L_2. \tag{2}$$

The income per resident in region i is given by Y_i/L_i and the amount of land consumed per resident equals G_i/L_i. According to the above preferences each individual spends a fraction $(1-\mu)$ of his income on manufactured goods of which each costs p_c such that the utility of a representative individual in region i becomes:

$$u_i = \left(\frac{G_i}{L_i}\right)^{\mu} \left(\frac{(1-\mu)Y_i}{L_i p_c}\right)^{1-\mu}. \tag{3}$$

Firms in the manufacturing sector require workers as the only input factor. The production function of a representative firm in region i is given by $c_i = A_i l_i$ where c_i denotes the output of the respective firm, l_i denotes the number of workers employed and A_i represents the marginal product of labor which can vary between the two regions. Profit maximization under perfect competition implies that prices (p_c) equal the marginal costs ($\frac{w_i}{A_i}$), where the latter is given by the ratio of regional wages (w_i) and productivity (A_i).

Due to the existence of knowledge spillovers, the productivity of an individual worker increases with the total number of workers employed in the respective region. Workers learn from the knowledge of others working in the same region. Hence, we presume that the transmission of knowledge is spatially limited such that spillovers occur only within each region's borders. Since the total number of workers in each region is proportional to the total regional output, knowledge spillovers result in a positive correlation

[29] Each individual spends a fraction μ on land such that the aggregate expenses for land are μY where $Y = Y_1 + Y_1$. Aggregate income is $Y = w_1 L_1 + w_2 L_2 + \mu Y \Leftrightarrow Y = \frac{1}{1-\mu}(w_1 L_1 + w_2 L_2)$. Hence, aggregate expenses for land and accordingly rent income are $\mu Y = \frac{\mu}{1-\mu}(w_1 L_1 + w_2 L_2)$.

between regional productivity (A_i) and total output per region (C_i):

$$A_i = C_i^{\frac{\delta}{(1+\delta)}} \quad \text{where} \quad \delta > 0 \tag{4}$$

Note that δ quantifies the intensity of knowledge spillovers. If $\delta = 0$ there are no spillover effects. The higher δ the stronger the external effects. After introducing localized spillover effects, the total regional output of the manufacturing sector (C_i) is given by:

$$C_i = C_i^{\frac{\delta}{(1+\delta)}} L_i \quad \Leftrightarrow \quad C_i = L_i^{1+\delta}. \tag{5}$$

For the individual firm, the marginal product of labor represents a constant but on the regional level knowledge spillovers give rise to increasing returns to scale. Combining the condition that prices have to equal the marginal costs with equations (4) and (5) yields the wage rate in region i:

$$w_i = p_c A_i = p_c C_i^{\frac{\delta}{(1+\delta)}} = p_c L_i^{\delta}. \tag{6}$$

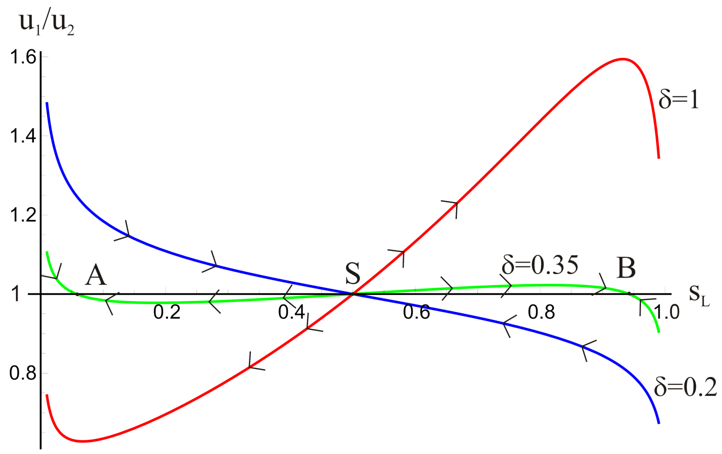
Finally, the migration equilibrium can be derived by comparing the utility levels of representative individuals in regions *1* and *2* (u_1, u_2). An equilibrium is reached either if the utility levels are equalized across both regions (i.e., $\frac{u_1}{u_2} = 1$) or if full agglomeration prevails. The latter implies that either the share of workers in region *1* (s_L) or the share of workers in region *2* ($1 - s_L$) is equal to unity. Inserting (2) and (6) in (3) yields:

$$\frac{u_1}{u_2} = \left(\frac{G_1(1-s_L)}{G_2 s_L}\right)^{\mu} \left(\frac{(1-\mu)s_L^{\delta} + \mu\left[s_L^{(1+\delta)} + (1-s_L)^{(1+\delta)}\right]}{(1-\mu)(1-s_L)^{\delta} + \mu\left[s_L^{(1+\delta)} + (1-s_L)^{(1+\delta)}\right]}\right)^{(1-\mu)}. \tag{7}$$

As long as neither an interior equilibrium (i.e., $u_1 \neq u_2$) nor a corner solution with $s_L = 1$ or $(1 - s_L) = 1$ prevails, there is an incentive for workers to relocate. In order to analyze the stability of potential equilibria, we assume in the following that the two regions are of equal geographic size, that is $G_1 = G_2$ which simplifies the migration condition.

Figure 2.12 illustrates the potential equilibria as well as their stability for three levels of knowledge spillover intensity (δ). The blue line features a low level of spillovers ($\delta = 0.2$), the green line a medium level ($\delta = 0.35$) and the red line a high level of spillovers ($\delta = 0.8$).

Figure 2.12: AGGLOMERATION AND KNOWLEDGE SPILLOVERS



Note: In this figure we set expenditure share of the manufacturing sector to $\mu = 0.2$.

In the first case with weak spillovers the symmetric equilibrium ($s_L = (1 - s_L) = 0.5$) is stable since the utility in region *1* is lower than the utility in region *2* as long as $s_L > 0.5$ while the reverse is true for $s_L < 0.5$. In contrast, the agglomeration equilibria are unstable and will be dissolved by a marginal shift.

In the case of medium spillover intensity there are five potential equilibria: three interior equilibria at points A, S, B and two agglomeration equilibria. The agglomeration equilibria are again unstable since $u_1 > u_2$ for a marginal deviation from $s_L = 0$ and $u_2 > u_1$ for a marginal deviation from $s_L = 1$. That is workers face an incentive to move to region *1*, if full agglomeration in region *2* occurs while they have an incentive to move to region *2*, if full agglomeration in region *1* prevails. Similarly to the core-periphery model derived above, the two additional interior equilibria emerge only for a certain range of δ. In contrast to the core-periphery model, these equilibria in points A and B are stable while the symmetric equilibrium for intermediate spillover intensity ($\delta = 0.35$) is not.

If the spillovers have a strong impact on the regional productivity ($\delta = 0.8$) which is depicted by the red line, only the full agglomeration equilibria are stable. Hence, increasing the intensity of knowledge spillovers yields stronger agglomeration forces and eventually spatial concentration of economic activity.

Since the asymmetric interior equilibria in points A and B are stable, raising the spillover intensity does not cause a jump from the symmetric distribution of activity to full agglomeration as in the core-periphery model but a continuous process of agglomeration.

The second exogenous parameter is the expenditure share of the manufacturing sector $(1-\mu)$. It can easily be shown that an increase in the expenditure share strengthens the agglomeration process as in the New Economic Geography (i.e., rotates the functions in figure 2.12 anticlockwise). This is obvious as a lower μ reduces the importance of land while the role of manufacturing goods in the optimal consumption bundle is strengthened.

To sum up the results derived from a simple model with knowledge spillovers we can state that the intensity of spillovers as well as the expenditure share of sectors which face knowledge spillovers enhance the agglomeration tendencies. Integrating this finding into the hierarchy of central places, we classify products of industries with strong knowledge spillovers as higher order goods which are provided only at few places. Hence, the distance between individual central places (D^{PP}) derived in Section 2.4.1 increases with the intensity of knowledge spillovers.

2.4.3 Labor Market Pooling

Lastly, we discuss the role of labor market pooling for the concentration of economic activity. This mechanism is based on the assumption that firms face idiosyncratic productivity shocks such that their demand for labor is not constant over time but deviates from the average structural demand. Presuming that these firm specific shocks are not perfectly correlated implies that the labor demand of individual firms is not perfectly correlated either. Intuitively, this means that one firm experiences a good year while another firm experiences difficult times. Even though we observe that recessions hit almost all firms just like booms benefit the whole economy this is a realistic assumption, since there remains a considerable variance which avoids a perfect correlation of firm specific shocks. During downswings for instance we observe a decline in productivity but this decline is by far not equally distributed across all sectors and firms. Moreover, the shift in demand is to some extent unpredictable for individual firms as they may anticipate that a recession is going to take place but they neither know the exact timing nor the intensity the recession will affect their branch and specifically themselves. Therefore, this kind of firm specific shock can be regarded as being random and imperfectly correlated across firms.

In such a scenario it is beneficial for firms as well as for laborers to operate and work in an agglomerated region. As an example, consider two firms out of which one experiences extraordinary good times while the other firm is hit by a sharp drop in demand. The latter firm lays off workers while the former firm wants to expand its business. For illustration purposes presume for now that the wage rate is exogenously fixed. If in this case both firms operate in spatially separated labor markets (i.e., workers as well as firms are bound to the labor market of their respective region), the booming firm will be restricted by too little labor supply while some workers of the troubled firm become unemployed. In contrast, if the firms would share a pooled labor market, the booming firm could take over the laid off workers from the struggling firm. Hence, neither will a firm be restricted in its output nor will workers become unemployed. Since workers are immobile in the short-run (i.e., when the deviation from the structural labor demand takes place), labor market pooling is constrained by the geographic distance. Therefore, both workers and firms prefer to concentrate in space.[30]

Despite the fact that we observe significant wage rigidities, completely inflexible wages are unrealistic. However, in the following we will show that the described agglomeration mechanism does not rely on the assumption of a fixed wage rate but holds also true for labor market clearing wages. In doing so, we build on a simple model by Krugman (1991a) which will be somewhat modified.

There are two regions $i \epsilon \{1, 2\}$ endowed with workers (L_i), manufacturing firms (n_i), and an exogenously fixed amount of land (G_i). In contrast, to the model above individuals derive utility only from the manufactured good which can be shipped for free across regions. Land enters on the production side because this allows us to spare an explicit analysis of the consumption side. In general it makes no difference whether the scarce factor land countervails agglomeration tendencies from the consumption or the production side. Hence, the technology of a representative firm in the manufacturing sector is described by the following production function:

$$y(h) = g(h) + [\beta + \varepsilon(h)]l(h) - \frac{1}{2}\gamma[l(h)]^2 \tag{8}$$

where $\varepsilon(h)$ represents a firm specific shock which is independent and identically-distributed with mean zero and variance σ^2. The amount of land utilized by the representative firm is denoted by $g(h)$ and the employed workers are $l(h)$.

The production function faces diminishing returns to scale which allows us to derive a discrete number of firms while relying on a homogenous manufactured good. Moreover,

[30] Workers may be willing to commute a long distance but there are obviously limits which can only be overcome by moving. Yet, moving entails significant costs.

due to diminishing returns, firm owners derive positive operating profits from the intra marginal units of labor.[31] Firm owners will strive to maximize their expected profits by choosing the most promising firm location. The intensity of diminishing returns is denoted by γ.
Since there is an exogenous amount of land, each firm in region i utilizes $\frac{G_i}{n_i}$ units of land such that the production function of a representative firm becomes:[32]

$$y(h) = \frac{G_i}{n_i} + [\beta + \varepsilon(h)]l(h) - \frac{1}{2}\gamma[l(h)]^2 \tag{9}$$

After experiencing the shock firms know their specific productivity and decide upon how many workers to employ. Profit-maximizing behavior requires firms to hire workers until the marginal product of an additional worker equals the regional wage rate (w_i). Yet, the regional wages are taken as given by each single firm such that the realization of the shock manifests in the number of workers the respective firm employs ($l(h)$).

$$w_i = \beta + \varepsilon(h) - \gamma l(h) \quad \Leftrightarrow \quad l(h) = \frac{\beta + \varepsilon(h) - w_i}{\gamma} \tag{10}$$

The regional wage rate (w_i) adjusts until the total demand for labor in region i equals the regional labor supply (L_i):

$$\begin{aligned} \sum_{h=1}^{n_i} l(h) &= n_i \frac{\beta - w_i}{\gamma} + \frac{1}{\gamma}\sum_{h=1}^{n_i}\varepsilon(h) = L_i \\ w_i &= \beta - \frac{\gamma}{n_i}L_i + \frac{1}{n_i}\sum_{h=1}^{n_i}\varepsilon(h) \end{aligned} \tag{11}$$

Workers' location decisions take place before the productivity shocks are realized. Therefore, they form expectations about the wages in the two regions and compare these. Since the mean of the shocks in each region is assumed to be zero, the expected wages are given by:

$$E(w_i) = \beta - \frac{\gamma}{n_i}L_i \tag{12}$$

[31] In principle, the mechanism applies also in a framework with firms producing heterogenous varieties and consumers facing love-of-variety preferences as derived in Chapter 3. The operating profit would stem from firms charging a mark up while the marginal product of labor would be kept constant (apart from firm specific shocks). However, this would unnecessarily complicate the model.

[32] Note that the marginal product of land is constant and the same for all firms. Therefore, each firm in region i demands the same amount of land. We do not model the land market explicitly but assume that rent income is equally distributed across all individuals such that it is irrelevant for location choice.

Hence, the expected wages increase with the demand for labor (i.e., the number of local firms) while they decrease with the aggregate supply of labor (i.e., the size of the local labor force). Workers' migration equilibrium requires that either the expected wages are equalized, that is $E(w_1) = E(w_2)$, or a corner solution with full agglomeration in one of the regions applies. Any other allocation of workers does not represent an equilibrium, since workers expect to benefit from relocation. From equation (12) it becomes apparent that this condition requires the regional share of workers to equal the regional share of manufacturing firms. That is, $s_L = s_n$ has to be satisfied where s_L represents the share of workers in region *1* and s_n region *1*'s share of manufacturing firms.[33]

Firm owners base their location decision upon the firm's operating profits expected in the respective region. Note that they cannot split up their production between the two regions but have to settle for one region. This assumption can be justified by sufficiently high setup costs. The operating profit of a representative firm owner in region i is given by $R(h) = y(h) - w_i l(h)$ where the price of the manufactured good is the numéraire. Using equation (10) and (9) we can rewrite the operating profit as:

$$R(h) = \frac{G_i}{n_i} + \frac{[\beta + \varepsilon(h) - w_i]^2}{2\gamma}. \tag{13}$$

Calculating the expected value of the operating profits in equation (13) yields after some simplifications:[34]

$$\begin{aligned} E(R_1) &= \frac{G_1}{n_1} + \frac{\gamma}{2}\left(\frac{s_L L}{s_n n}\right)^2 + \frac{\sigma^2}{2\gamma}\left(\frac{n_1 - 1}{n_1}\right) \qquad (14) \\ E(R_2) &= \frac{G_2}{n_2} + \frac{\gamma}{2}\left(\frac{(1-s_L)L}{(1-s_n)n}\right)^2 + \frac{\sigma^2}{2\gamma}\left(\frac{n_2 - 1}{n_2}\right) \end{aligned}$$

$$\text{where} \quad L = L_1 + L_2 \ \text{ and } \ n = n_1 + n_2.$$

[33]Note that $E(w_1) = E(w_2)$ implies $\frac{L_1}{n_1} = \frac{L_2}{n_2}$ which can be rewritten as $\frac{s_L L}{s_n n} = \frac{(1-s_L)L}{(1-s_n)n}$ where $n = n_1 + n_2$ and $L = L_1 + L_2$. Hence $s_L = s_n$ holds at the interior equilibrium. Full agglomeration equilibria necessarily require $s_L = s_n = 0$ or $s_L = s_n = 1$ since firm cannot operate without workers.

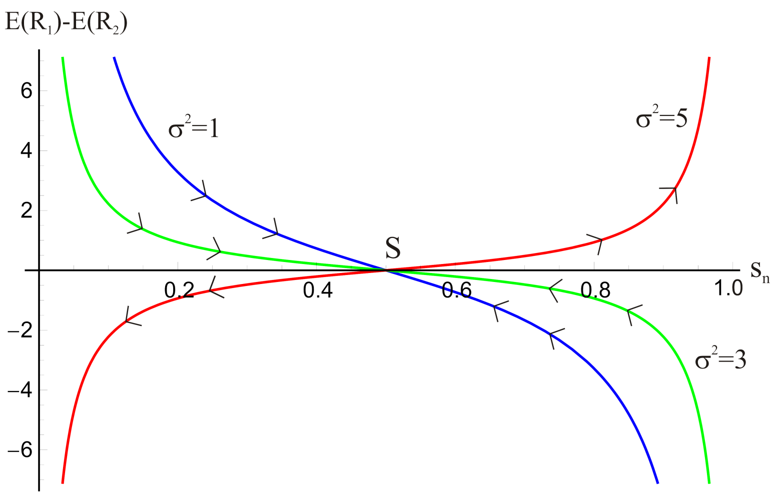
[34]Note that from the computational formula for the variance the expected value of the square of a random variable equals its variance plus the square of its expected value: $Var(X) = E(X^2) - [E(X)]^2 \Leftrightarrow E(X^2) = Var(X) + [E(X)]^2$. The expected value of the shock ($\varepsilon(h)$) is zero such that $[E(\beta + \varepsilon(h) - w_i)]^2 = [\beta - E(w_i)]^2 = \gamma^2 \left(\frac{L_i}{n_i}\right)^2$. The variance is independent of additive constants such that: $Var\,[\beta + \varepsilon(h) - w_i] = Var\,[\varepsilon(h) - w_i]$. Moreover, the variance of a sum (difference) of random variables equals the sum of their variances plus (minus) two times their covariance, such that: $Var\,[\varepsilon(h) - w_i] = Var\,[\varepsilon(h)] + Var(w_i) - 2Cov\,[\varepsilon(h), w_i]$. The variance of the shock is denoted by σ^2 while deriving the variance of the wage rate in region i as well as the covariance yields: $Var(w_i) = Cov\,[\varepsilon(h), w_i] = \frac{\sigma^2}{n_1}$.

The first terms of the expected operating profits in regions *1* and *2* reflect the dispersion force. Due to the scarcity of land, the expected operating profits decrease with the local number of firms which makes agglomeration unfavorable. As mentioned before, the migration equilibrium of workers implies $s_L = s_n$ such that the expressions' second term only reflects the overall level of operating profits which is determined by the ratio between the number of firms and laborers. When setting up the equilibrium condition for firm owners ($E(R_1) = E(R_2)$), these terms drop out. The labor pooling effect which gives rise to agglomeration tendencies is captured in the last term. Note that the labor pooling effect drops out if the variance of the shocks is zero which implies that shocks are nonexistent. In this case, the equilibrium distribution is determined by the relative geographic size $\frac{G_1}{G_2}$, only. However, in the following we will refrain from differences in geographic size such that $\frac{G_1}{G_2} = 1$. In the absence of idiosyncratic shocks, equal regional sizes clearly imply a symmetric equilibrium.
The labor market pooling effect gains importance with the variance of the firm-specific shocks. The higher the uncertainty about the future productivity of labor the more important it is to settle in an agglomerated region. In contrast, the intensity of diminishing returns (γ) weakens the labor market pooling effect. This is because γ reflects the slope of the labor demand function. If the intensity of diminishing returns is high, the demand function is steep and accordingly the elasticity of demand is low. Therefore, if γ is high, idiosyncratic shocks lead to only little adjustment in the labor demand.

Figure 2.13 depicts the difference between expected operating profits in regions *1* and *2* against the share of firms in region *1*. In doing so, we require the migration equilibrium condition of workers to be satisfied at all time. Note also that the total number of firms ($n = n_1 + n_2$) and workers ($L = L_1 + L_2$) does not affect the regional distribution of manufacturing but only the aggregate level of operating profits.

The blue line illustrates a scenario with low variance ($\sigma_2 = 1$) and accordingly a inelastic labor demand. The negative slope of the ($E(R_1) - E(R_2)$) curve at point S (where $s_n = 0.5$) implies that the symmetric equilibrium is stable while the full agglomeration equilibria are not. Raising the variance rotates the curve. For intermediate variance ($\sigma = 3$) the symmetric equilibrium is still stable whereas for high variance ($\sigma = 5$) – illustrated by the red line – only the full agglomeration equilibria are stable.
Note that the labor market pooling argument does not rely on risk aversion but the assumption that workers and firm owners compare expected values reflects risk neutrality. Risk averse workers would take the variance of the wages into account which would essentially strengthen the agglomeration tendencies, since the variance of

Figure 2.13: AGGLOMERATION AND FIRM SPECIFIC SHOCKS



Note: In this figure γ, L and n are set to unity.

regional wages decreases with the number of local firms.

The agglomeration mechanism originating from labor market pooling predicts that firms which face high uncertainty and consequently elastic labor demand have a strong incentive to locate in agglomerated regions. Moreover, in order to realize the beneficial effect, it is necessary that the clustering firms face uncorrelated shocks which means that they are affected to a different extend or point in time by the business cycle. Hence, in contrast to the knowledge spillover mechanism which predicts that especially firms of similar sectors agglomerate, the labor market pooling effect provides a reason for firms of different sectors to locate next to each other. Incorporating this mechanism in the hierarchic system of the central place theory we classify the products of exceptionally volatile sectors as higher order goods which will be provided at only few central places. Hence, the distance between individual central places increases with volatility.

The described agglomeration mechanism is not only a theoretical construct but in a recent paper Overman and Puga (2009) provide empirical evidence for the labor market pooling effect. Using data on the location of UK manufacturing industry they show that sectors whose establishments experience more idiosyncratic volatility than others exhibit a stronger spatial concentration.

2.5 Concluding Remarks

The spatial distribution of individuals and production facilities is to a large extent determined by economic interactions. In order to explain the uneven allocation of economic activities we observe in Germany, Europe and any other federation this chapter forms a synthesis of important theories of economic geography.
First of all it is shown that the location decision of individual firms is driven by the interplay of transportation costs and the intention of firms to realize agglomeration economies. Thereafter, we illustrate how transportation costs give rise to a spatial division of markets between individual firms. Each market area corresponds to one firm which supplies its product at the most accessible place of the market area. The size of a market area again depends on the strength of agglomeration economies and the level of transportation costs. Hence, the territory under consideration can be partitioned into a multitude of individual market areas. Following the central place theory, we merge the market areas of various goods and services in order to derive a hierarchical system of multiple agglomerations. It is shown that agglomerations of different size performing different functions emerge. Sectors facing strong agglomeration economies locate their supply at only a few large agglomerations while lower order goods and services which exhibit negligible agglomeration economies are provided at many locations with little distance between each other. Lastly, we endogenize the agglomeration economies underlying firms tendencies to cluster by introducing three distinct mechanisms.
We show that sectors facing a low level of trade costs, a low elasticity of substitution, a high share in aggregate expenditure, strong knowledge spillovers, or a high volatility will be characterized by strong agglomeration tendencies. This implies a high distance between the central places that provide their respective products.
The agglomeration economies described in this chapter represent positive external effects which generally benefit aggregate welfare as they allow gains from variety, realizing knowledge spillovers, or sharing risk.

However there are also diseconomies of agglomeration. Possibly the most important diseconomy of agglomeration is the increase in aggregate transportation costs compared to a dispersed provision at many places. Yet, as we have shown above, market processes yield agglomeration only for sufficiently low level of transportation costs. Hence, the crucial question which remains topical in the next chapter is whether market processes induce agglomeration already at too high trade costs or only at too low trade costs compared to the efficient scenario.

Other diseconomies of agglomeration are higher factor and consumer prices. Under perfect competition these consequences of agglomeration which work through the price system do not represent a source of market inefficiency but only play a distributional role. However, in a framework with imperfect competition, as will be analyzed in Chapter 3, pecuniary externalities can become relevant for welfare. The third category which acts against agglomeration are congestion costs. Those costs occur for example due to increases in traffic and air pollution or due to overcrowded public facilities. Public facilities, for instance parks or schools, have in many cases no prices such that the effect of an additional firm or consumer on the usage conditions of other residents is not reflected by the market. Hence, the effect is external to the market and the spatial allocation of economic activity induced by market processes is suboptimal.[35] Section 3.5.2 of Chapter 3 analyzes the effects of congestion costs in more detail. Yet, congestion costs become important only after crossing a certain threshold. Moreover, there are efficient ways of internalizing the effects of congestion into the individual location decision. Charging fees and tolls for public facilities and infrastructure, or selling allowances for air pollution make sure that the external effect becomes part of firms' and workers' location decisions.

The following chapters introduce an additional channel of agglomeration, carry out an explicit welfare analysis, and discuss the effects of federal policy instruments aiming at a reduction of regional disparities before empirically evaluating the European Regional Policy.

[35]Note that the scarcity of land which acted as the dispersion force in the preceding models did not generate external effects since the market price of land was part of the individual location decision. The agglomeration economies described above may be interpreted as positive externalities. Likewise, the agglomeration mechanism which is introduced in the next chapter represents a positive externality resulting from the non-rivalry of public goods. Hence, the welfare analysis introduced in the following chapter applies in an isomorphic manner.

Appendix

A1: Formal Derivation of the Core-Periphery Model

In the following a formal derivation of Krugman's core-periphery model is provided.[36] The relationship between the distribution of industry and the trade costs as illustrated in figure 2.7 bases on the equilibrium conditions derived below. Moreover, the comparative static analysis in figures 2.9 and 2.10 simulate these equilibrium conditions for various parameter constellations.
The model features two regions $j \epsilon \{1, 2\}$ which are symmetric in terms of preferences, technology and trade costs. There are two sectors, a manufacturing sector (M) characterized by increasing returns, monopolistic competition and iceberg trade costs, and a perfectly competitive sector labeled agriculture (A) which produces under constant returns a homogenous good that is traded without costs. This good is produced in both regions and is taken as the numéraire, i.e. its price p_A is normalized to one. Individuals consume an agricultural good C_A and a composite manufactured good C_M. The utility of an individual is given by:

$$U = C_M^{\mu} + C_A^{1-\mu} \text{ where } C_M = \left[\int_0^n {m_i}^{\frac{\sigma-1}{\sigma}} di \right]^{\frac{\sigma}{\sigma-1}}. \tag{15}$$

The composite manufactured good again is given by a continuum of n differentiated varieties which can be substituted according to the elasticity of substitution $\sigma > 1$. The expenditure share of the manufacturing sector is denoted by μ and m_i represents the demand for an individual variety. Denoting personal net income by y, the budget constraint is

$$y = C_A + PC_M, \tag{16}$$

where P denotes the consumer price index of the composite manufactured good which can be expressed in terms of the prices p_i of varieties i:

$$P = \left[\int_0^n {p_i}^{1-\sigma} di \right]^{\frac{1}{1-\sigma}}. \tag{17}$$

[36]Note that the framework is very similar to the model introduced in more detail in Chapter 3.

From the individual utility maximization the demand for agricultural goods and individual manufactured varieties is given by:[37]

$$C_A = (1-\mu)Y \quad m_i = \mu Y \frac{p_i^{-\sigma}}{P^{1-\sigma}} \tag{18}$$

The manufacturing sector as well as the agricultural sector use only labor as input factor. However, laborers are sector specific such there is no equalization of wages between the two sectors. Moreover, only manufacturing workers are mobile across regions. The total number of manufacturing workers is L^M and the total number of agricultural workers is denoted by L^A. The latter are assumed to be distributed symmetrically across regions such that the number of agricultural workers in regions *1* and *2* is in each case $\frac{L^A}{2}$. The share of manufacturing workers in region *1* is denoted by s_L while the share of manufacturing workers in region *2* is $(1-s_L)$. Hence, the total number of manufacturing workers in region *1* and *2* is $s_L L^M$ and $(1-s_L)L^M$, respectively.
The agricultural good is produced under a linear technology and perfect competition which leads to marginal cost pricing. In the following we will not model the agricultural sector explicitly since this is not necessary for the formulation of the migration equilibrium. Note that from Walras' law the market for agricultural goods will be in equilibrium if the market for manufactured goods is in equilibrium.
In the manufacturing sector varieties are produced under increasing returns to scale since a fixed input F as well as a variable input c is used such that each firm employs $l_i = F + cM_i$ workers where M_i denotes the output of firm i. Hence, the profit function of a representative firm i in region j is given by:

$$\Pi_i = \hat{p}_i M_i - w_j(F + cM_i) \tag{19}$$

where $\hat{p}_i$ denotes the producer price of firm i, and w_j the wage rate in region j. Facing demand according to equation (18) firms maximize their profits such that the optimal

[37]The maximization problem is solved in two steps. First, for any level of the manufactured composite good C_M the individual varieties m_i are chosen such that the costs for attaining C_M are minimized. Hence, according to an exogenous elasticity of substitution a larger amount of cheaper (i.e., local) varieties is consumed than of more expensive (i.e., imported) varieties. Second, the total expenditure is divided optimally between the agricultural good C_A and the composite manufactured good C_M.

producer prices and profits of a representative firm i in region j become:[38]

$$\hat{p}_i = cw_j \frac{\sigma}{\sigma - 1} \quad \text{and} \quad \Pi_i = w_j \left[\frac{cM_i}{\sigma - 1} - F \right]. \tag{20}$$

Hence, the mark-up on marginal costs firms charge is determined by the elasticity of substitution. From the the zero-profit condition and equation (20) we can solve for the break-even level of output (M_i^*) and the equilibrium labor demand (l_i^*).

$$M_i^* = F\frac{\sigma - 1}{c} \quad \Leftrightarrow \quad l_i^* = F\sigma. \tag{21}$$

Due to increasing returns to scale and love-of-variety preferences each firm will only produce one variety such that the break-even level of a output (M_i^*) represents the total supply for variety i and the number of firms equals the number of varieties. When sipping goods from one region to the other, iceberg trade costs $\tau > 1$ arise. Profit-maximizing behavior in the Dixit-Stiglitz framework of monopolistic competition implies that the consumer prices of imported goods are $p_i = \tau\hat{p}_i$ while consumer and producer prices coincide for local varieties.

From equations (20) and (21) each firm sets the same price, produces the same quantity and demands the same amount of labor. Therefore, we may drop the firm subscript and refer only to regions $j \in \{1, 2\}$ in the following.

Market clearing implies that the total supply of a representative variety in region j equals the total demand which stems from local and foreign consumers.

$$M_j^* = \mu \left[Y_j \frac{\hat{p}^{-\sigma}}{P_j^{1-\sigma}} + Y_k \frac{(\hat{p}\tau)^{-\sigma}}{P_k^{1-\sigma}} \tau \right] \quad \text{where} \quad k \neq j \quad \text{and} \quad j, k \in \{1, 2\}. \tag{22}$$

Inserting the producer prices from equation (20) yields the wage equation:

$$w_j = \frac{\sigma - 1}{c\sigma} \left(\frac{\mu}{M_j^*} \right)^{1/\sigma} \left(Y_j P_j^{\sigma-1} + Y_k P_k^{\sigma-1} \tau^{1-\sigma} \right)^{1/\sigma} \tag{23}$$

$$\text{where} \qquad k \neq j \quad \text{and} \quad j, k \in \{1, 2\}.$$

The wage equation represents the first component of the model's equilibrium condition. The second component is the consumer price index from equation (17) which becomes

[38] Note that in the Dixit-Stiglitz framework it is assumed that individual firms take the consumer price index as given. That is, they do not consider the effect of their individual price decision on the aggregate consumer price index.

after distinguishing between local and foreign varieties:

$$P_j = \left[\hat{p}^{1-\sigma} n_j + (\hat{p}\tau)^{1-\sigma} n_k\right]^{\frac{1}{1-\sigma}}. \qquad (24)$$

$$\text{where} \quad k \neq j \text{ and } j,k \in \{1,2\}.$$

Note that the consumer price for local varieties equals the producer price while the consumer price for imported varieties is τ-times the producer price ($\hat{p}$).
The regional income (Y_j) is given by the sum of aggregate wages in the manufacturing sector ($L_j^M w_j$) and aggregate wages in the agricultural sector (L_j^A). The wage rate in the agricultural sector is unity due to the agricultural good being the numéraire and the input coefficient being normalized to unity. Moreover, the world endowment of agricultural and manufacturing workers is chosen such that $L^A = 1-\mu$ and $L^M = \mu$. Then the aggregate income in the two regions is given by:

$$Y_1 = \mu s_L w_1 + \frac{1-\mu}{2} \qquad Y_2 = \mu(1-s_L)w_2 + \frac{1-\mu}{2}. \qquad (25)$$

Choosing convenient units for the input coefficients in the manufacturing sector, $c = \frac{\sigma-1}{\sigma}$ and $F = \frac{\mu}{\sigma}$ the wage equation (26) can be expressed for both regions as:

$$w_1 = \left(Y_1 P_1^{\sigma-1} + Y_2 P_2^{\sigma-1}\tau^{1-\sigma}\right)^{1/\sigma} \qquad w_2 = \left(Y_2 P_2^{\sigma-1} + Y_1 P_1^{\sigma-1}\tau^{1-\sigma}\right)^{1/\sigma} \qquad (26)$$

Moreover, the price indices from equation (24) can be reformulated as:

$$\begin{aligned} P_1 &= \left[\frac{s_L L^M}{\mu} w_1^{1-\sigma} + \frac{(1-s_L)L^M}{\mu}(w_2\tau)^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \\ P_2 &= \left[\frac{(1-s_L)L^M}{\mu} w_2^{1-\sigma} + \frac{s_L L^M}{\mu}(w_1\tau)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}. \end{aligned} \qquad (27)$$

Lastly, the real wages are given by:

$$\omega_1 = \frac{w_1}{P_1} \qquad \omega_2 = \frac{w_2}{P_2}. \qquad (28)$$

A migration equilibrium is reached either if the real wages are equalized across both regions or if full agglomeration in one of the two regions occurs. Using the simultaneous equations (25), (26), (27), and (28) the equilibria can be determined numerically. In order to analyze the stability of the equilibria, the real wage difference $\omega_1 - \omega_2$ is calculated and plotted (see figures 2.7, 2.9, and 2.10).

Chapter 3

Public Input Competition and Agglomeration*

3.1 Introduction

All over the world we observe the tendency for further economic integration that increases factor mobility and stimulates regional trade. Firms are moving from one region to another rapidly in search of the most favorable conditions. During this process regions actively compete for becoming a center of economic activity.
Besides taxes and subsidies the most important instrument for local governments in attracting industry is certainly the quality of its local public inputs. Hence, jurisdictions may compete by providing public transportation, IT- and communication infrastructure, R&D facilities or other public inputs that increase the factors' productivity. As Borck et al. (2007) have shown, there is clear evidence that regions investing in productivity enhancing infrastructure respond positively to an increased investment in the neighboring regions. Using data on capital expenditures by US majority-owned companies in 18 European countries, Bénassy-Quéré et al. (2007) find a significant positive effect of local public capital on inward FDI. However, their analysis suggests that public infrastructure raises FDI investment only if it is not financed by a tax on the mobile factor. In contrast, Gabe and Bell (2004) come to the conclusion that a strategy of high spendings on local public infrastructure along with high taxes is even more rewarding for attracting businesses than a strategy of low taxes and low local government spending. Bénassy-Quéré et al. (2007) analyze the effects of public infras-

*This chapter is based on Fenge, von Ehrlich, and Wrede (2009)

tructure spending on European national level, whereas Gabe and Bell (2004) look at the location decision of businesses on the municipal level in Maine.
While public input competition gains relevance the more mobile the factors are, we observe that economic integration also leads to spatial agglomeration of industries. Midelfart-Knarvik and Overman (2002), Ezcurra et al. (2006), as well as Brühlhart (2001) have shown that the regional concentration of industrial activity in Europe has been increasing over the last decades. Moreover, the rise in geographic concentration coincided with the establishment of the single European market. For the United States Ellison and Glaeser (1999) estimate that only about 20 percent of observed geographic concentration can be explained by natural comparative advantages. The remaining part has to be explained by knowledge spillover effects, labor market pooling, the circular causalities of the New Economic Geography or by the provision of industry specific infrastructure. The latter channel of industrial clustering is the one we focus on in the following.

This chapter brings together the spatial development of economic activity during a process of economic integration and the role of public input competition. We study the impact of fiscal competition via public inputs on the distribution of industry. From a positive point of view, it is shown that regional competition constitutes a prominent agglomeration mechanism. From a normative point of view, we determine whether competition leads to an efficient outcome in terms of the provision of public inputs as well as the resulting regional allocation of industry. The latter issue addresses the question of whether a federal policy reallocating public input resources across regions is necessary in order to achieve an optimal outcome.
We model public input competition in a New Economic Geography framework. This is a very useful framework to analyze the effects of public input competition on a regional level since it allows for factor mobility as well as for costly trade. As we will show, the inclusion of trade in addition to factor mobility provides important new insights. We employ an approach where trade is not determined by comparative advantages but by economies of scale and consumer preferences for differentiated varieties. Considering similar regions within federations such as the European Union or the USA this explanation of trade fits certainly better than comparative advantages. Due to the analytical tractability we use a footloose capital model (see, e.g. Martin and Rogers, 1995; Baldwin et al., 2003; Dupont and Martin, 2006). Martin and Rogers (1995) think of infrastructure as a means to reduce iceberg trade costs. In contrast, we adopt the common view that public spending directly affects the production function (see, e.g. Barro, 1990; Morrison and Schwartz, 1996). Moreover, we emphasis the strategic interaction

between regions and the cross-regional externalities that arise from local infrastructure provision when accounting for trade. For most of our analysis we consider ex-ante identical regions but in an extension we allow for differences in regional size. As opposed to the core-periphery model introduced in Chapter 2, the standard footloose capital model does not imply an agglomeration process for symmetric regions. However, once we add public input competition decreasing trade costs lead to agglomeration in the footloose capital model even for ex-ante identical regions. Hence, public input competition represents an alternative explantation for the asymmetric distribution of economic activity despite identical initial endowments.

There is an extensive literature dealing with fiscal competition in neoclassical models which is closely related to this chapter. Early contributions by Zodrow and Mieszkowski (1986) and Wilson (1986) point to an inefficiency that arises from capital mobility. It is shown that individual regions will compete for mobile capital by reducing their source tax on capital. When trying to attract mobile capital regions exert a negative externality on each others in the sense of reducing their reciprocal tax base. Therefore, tax competition yields an underprovision of public goods in the framework of Zodrow and Mieszkowski (1986) as well as Wilson (1986). However, subsequent authors criticized their result of suboptimal public good provision and show that it is driven by the assumption of a public consumption good (see Sinn, 2003 and Noiset, 1995). If public goods are productive, capital benefits from public expenditure and regions may well use such public inputs as an instrument to attract capital. As Sinn (2003) shows, a harmonization of source taxes on capital is no solution to the issue of regional tax competition since such a policy will consequently lead to an overprovision of public inputs. Source taxes on mobile factors and public inputs which benefit the productivity of mobile factors are two sides of the same coin. In the context of regional competition, the former will be reduced while the latter is being raised in order to attract mobile factors. Hence, jurisdictions will shift their expenditure from public consumption goods towards public inputs and their tax revenue from capital taxes towards the taxation of immobile factors (see, e.g., Keen and Marchand, 1997).
However, from the general premise that factor mobility and trade act as substitutes for each other this strand of literature neglects the role of regional trade. Yet, the scope for trade is an essential determinant of firms' location decisions. As was shown in Chapter 2, firms aim for locations that facilitate access to large markets. Moreover, the possibility to import and export may mitigate the willingness of workers to finance infrastructure which primarily benefits capital. In the following we will show that the inclusion of regional trade in addition to factor mobility significantly alters the results

of fiscal competition.

We derive the following main results: First of all we show that one region may decide not to provide any infrastructure and to free-ride on the other region's infrastructure if trade costs become sufficiently low. The reason is that the region benefits from importing goods at low trade costs instead of bearing the costs of infrastructure provision itself. This is the case where public input competition induces agglomeration even for ex-ante identical regions. Only when trade costs are sufficiently high, will regional competition yield a symmetric provision of public inputs and agglomeration tendencies will be eliminated.

Second, regarding the normative question of aggregate public input provision two externalities are decisive. On the one hand, infrastructure provision in one region exerts a negative externality on the other region because it attracts industry from there. On the other hand, the possibility of trade in combination with the public good characteristics of infrastructure adds a positive externality of infrastructure provision. The region that provides all the infrastructure does not take into account the cost reduction of imported manufactured goods in the foreign region. If trade costs are sufficiently low this positive externality dominates and an underprovision of infrastructure is the outcome of fiscal competition. On the contrary, if the trade costs are sizable we show that the amount of aggregate infrastructure provided in the Nash equilibrium is too high compared to a centralized provision of regional infrastructure. The reason is that imports are expensive when trade costs are high and accordingly it is important to have industry located within one's own borders such that the negative externality dominates. Hence, for high trade costs our model confirms what has been derived within the standard fiscal competition framework, namely, an oversupply of public inputs, albeit induced by a quite different mechanism. However, we show that strong agglomeration forces turn this result upside down.

Third, concerning the distribution of infrastructure across regions, low trade costs allow the regions to import industrial goods at low consumer prices. In such an integrated market a central government chooses to concentrate public infrastructure in one region. This strategy assures that all industry is located in the core region and maximizes the aggregate productivity of the industrial sector. For high trade costs, in contrast, a central government would distribute infrastructure equally across regions because the benefit from saving trade costs is higher than the loss in overall productivity. Hence, we identify the critical trade costs where agglomeration becomes efficient. This can be translated into the optimal distance between individual central places as introduced in the previous chapter.

Moreover, for a range of higher trade costs we show that the Nash equilibrium is characterized by a symmetric provision whereas a central government would place all infrastructure in one region and implement an agglomeration pattern. The reason is that an individual region waiving local industry benefits from the resulting increase in aggregate productivity only to the extent of the imported goods weighted at trade costs while the central government faces the full productivity gain. In addition, the region that loses all industry faces a higher increase in trade costs than the average federation does, since the core region saves trade costs. Therefore, individual regions start to prefer the concentration of infrastructure and the agglomeration of industry in uncoordinated equilibrium only at a lower rate of trade costs than the central government does. Note that these results arise although we abstract from local spillover-effects between firms in our model.[1]

A series of papers has addressed tax competition in the New Economic Geography, but all these analyses neglect the role of productive public expenditure. Hence, this chapter provides new insights on public input competition in such a framework.[2] Bucovetsky (2005) analyzes public input competition in a model with perfect competition and external economies of scale. He shows that both under- or overinvestment may arise as a result of regional competition depending on the extend of economies of scale in public investment. However, he does not take costly trade between the regions into account, which is central to our argument. Egger and Falkinger (2006) examine the relationship between public infrastructure competition and outsourcing in a new trade model. They conclude that public infrastructure provision may prevent international outsourcing and that regional competition may therefore result in an overprovision. In contrast to our analysis, infrastructure provision does not affect consumer prices in their model which is the source of the positive externality in our model. They think of public infrastructure provision as an instrument that reduces the fixed costs of setting up a firm, whereas we model infrastructure such that it lowers the variable costs of production. Robert-Nicoud and Sbergami (2004) link the footloose capital model with endogenous regional policy to a political economy approach. A central government which is elected by the citizens of both regions decides on the amount of subsidies payed to each region. The subsidies in turn affect the spatial allocation of industry. However, regions do not explicitly compete in their setting and no normative

[1] Such spillover effects can also lead to lower degree of agglomeration than desired from a welfare perspective as Martin and Ottaviano (1999) have shown.

[2] See Ludema and Wooton (2000); Kind et al. (2000); Andersson and Forslid (2003); Baldwin and Krugman (2004); Ottaviano and van Ypserle (2005); Borck and Pflüger (2006) or Baldwin et al. (2003) for tax competition in the NEG.

conclusions are drawn. Moreover, these subsidies reduce only the fixed costs of setting up a firm, such that the consumer prices remain again exogenous. There is clearly anecdotical evidence that public inputs affect not only set up costs, but also variable costs. For instance, improvements in public transportation systems reduce transport time and thus costs per unit.

The next section introduces the basic model and derives the impact of regional infrastructure on the long-run allocation of industry. Section 3.3 describes the externality of infrastructure provision which arises if fiscal competition between regions takes place and determines the Nash equilibria for critical values of trade freeness. Section 3.4 contrasts the Nash equilibria with the allocation of a central government taking account of the externalities. Section 3.5 extends the model to asymmetry in population size and to congestion costs and discusses the qualifications of the results. Section 3.6 summarizes the main findings and relates to policy issues.

3.2 The Model

Following Martin and Rogers (1995) we use a model where the fixed cost in the manufacturing sector is attached to an internationally mobile factor. A federation consists of two regions H and F which are symmetric in terms of preferences, technology and trade costs. There are two sectors, a manufacturing sector (M) characterized by increasing returns, monopolistic competition and iceberg trade costs, and a perfectly competitive sector labeled agriculture (A) which produces under constant returns a homogenous good that is traded without costs. This good is produced in both regions and is taken as the numéraire, i.e. its price p_A is normalized to one. Individuals consume an agricultural good C_A and a composite manufactured good C_M as in Dixit and Stiglitz (1977). The utility of consuming the agricultural and the composite manufactured good is a logarithmic quasi-linear function where μ is the expenditure share of the composite manufactured good.[3] The composite manufactured good again is given by a continuum of n differentiated varieties. In general individuals prefer to consume as many different varieties as possible. Their willingness to substitute between the quantities

[3]Since expenditure shares are exogenous in the footloose capital model, a reallocation of industry does not imply income effects. Using a quasi-linear utility function which captures only substitution effects exhibits all relevant features of the model and simplifies calculations. See Robert-Nicoud and Sbergami (2004) for an application of the footloose capital model with quasi-linear utility and Pflüger (2004) for a detailed analysis of the standard model with quasi-linear utility.

m_i of the varieties is given by the substitution elasticity $\sigma > 1$. Hence, utility is:[4]

$$U = C_A + \mu \ln C_M + \mu - \mu \ln(\mu) \ ; \ \text{where } C_M = \left[\int_0^n m_i^{\frac{\sigma-1}{\sigma}} di \right]^{\frac{\sigma}{\sigma-1}} . \tag{1}$$

Denoting personal net income by y, the budget constraint is

$$y = C_A + PC_M, \tag{2}$$

where the economy's consumer price index of the composite good P can be expressed in terms of the prices $p(i)$ of varieties i:

$$P = \left[\int_0^n p_i^{1-\sigma} di \right]^{\frac{1}{1-\sigma}} . \tag{3}$$

Utility maximization leads to demand functions and indirect utility:

$$C_A = y - \mu, \quad C_M = \mu/P, \quad m_i(p_i) = \mu \frac{p_i^{-\sigma}}{P^{1-\sigma}}, \quad V = y - \mu \ln(P). \tag{4}$$

On the supply side, there are three productive factors: inter-regionally immobile labor, mobile capital and local infrastructure. The number of immobile workers in region $j \in \{H, F\}$ is denoted by L_j, the number of capital owners by K_j, where each capital owner owns one unit of capital. The total stock of capital is given by $K^W = K_H + K_F$, while world labor endowment is given by $L^W = L_H + L_F$. Capital is invested internationally but its return is repatriated. The agricultural good is produced with labor under a linear technology. Perfect competition leads to marginal cost pricing. Furthermore the wage is equal to the marginal product of labor, i.e. one. Manufacturing firms produce with a non-homothetic technology and each firm produces one variety. A firm located in region j requires one unit of capital and, per unit of output, c_j units of labor. Total costs of a firm which produces variety i are $R_j + c_j M_i$, where R_j is the reward to capital in region j and M_i is the output of this firm.

In the Dixit-Stiglitz model of monopolistic competition mill pricing is optimal.[5] Hence,

[4] $\mu - \mu ln(\mu)$ is added to simplify the expression for indirect utility; see below.

[5] The Dixit-Stiglitz setup implies that mill prices depend only on costs in the region of origin. To the contrary, quadratic utility and linear demands (see, e.g., Ottaviano and Thisse, 2004) would imply that prices of exporters are affected by costs in the region of destination. Cost reductions in one region would translate to direct price changes of exporters in the other region.

indicating producer prices by a hat, in the region where the variety is produced $p_i = \hat{p}_i$ holds, and in the foreign region $p_i = \tau\hat{p}_i$ is fulfilled, where τ are iceberg trade costs with $\tau > 1$. Profits of firm i in region j are

$$\Pi_i^j = (\hat{p}_i - c_j)M_i - R_j. \tag{5}$$

Market clearing for a variety i produced in region j implies $M_i = m_i(\hat{p}_i)(K_j + L_j) + \tau m_i(\hat{p}_i)(K_k + L_k)$, $k \neq j$, $j, k \in \{H, F\}$. Hence, profit maximization leads to

$$\hat{p}_i = \frac{c_j\sigma}{\sigma - 1}. \tag{6}$$

Since all firms in a single region set the same price for each variety, local consumers demand the same quantity of all local varieties and we may therefore refer to regions instead of varieties. The regional price index of manufactured goods in region j follows from (3), which is given by the sum of local produced (n_j) and imported varieties (n_k).

$$P_j = \left[\hat{p}_j^{1-\sigma} n_j + (\hat{p}_k\tau)^{1-\sigma} n_k\right]^{\frac{1}{1-\sigma}}, \quad k \neq j, \quad j, k \in \{H, F\} \tag{7}$$

Variable costs c_j depend on local infrastructure in the respective region. The better the local public good supply is, the lower variable costs are. We assume that investment in local infrastructure (X_j) reduces the variable factor's input coefficient (c_j) in the production of the manufacturing sector. We abstain from an effect of infrastructure on the immobile sector because we want to focus on the part of infrastructure that is used for regional competition by attracting mobile firms. Hence, the infrastructure is specific to the mobile sector. For example one could think of IT- and telecommunication infrastructure or R&D facilities which are of negligible value for the agricultural sector. Investing in sector specific infrastructure is a rational strategy for regions experiencing increasing competition for mobile capital. Since tax revenue and the scope for expenditures is limited they will try to focus on the mobile industries and the most lucrative industries which expect a promising future. This is why we do not observe regional competition for agricultural or heavy industries. From a local politician's point of view it makes no sense to waste tax revenue for the provision of public inputs that benefit sectors which are either not very mobile or not lucrative. Pereira and Andraz (2007) analyze public infrastructure investment in Portugal between 1976-1998, which was financed to a large extent by the European Union. They can show that infrastructure investments did not only enhance labor productivity significantly but also that it shifted the countries industry mix towards more progressive sectors as for example the chemical and metal industry or the finance sector. In their analysis the seven sectors that benefited most captured about 78 percent of the benefits, but made up initially only about 21 percent of total employment.

In the model we let each region invest an amount X_j in sector specific local infrastructure which improves the productivity of labor.

$$X_j = c_j^{1-\sigma}. \tag{8}$$

The implicit assumption behind this definition is that the productivity of the variable input factor increases with local infrastructure. Moreover it may increase at a diminishing, constant or increasing rate depending on σ (diminishing rate if $\sigma > 2$), since the production function of the monopolistic competitive sector is given by:

$$Q_j = X_j^{\frac{1}{\sigma-1}} L_j, \tag{9}$$

where Q_j is total output of this sector in region j. Hence the output elasticity of infrastructure is set to $\frac{1}{\sigma-1}$, which simplifies our analysis.[6]

In order to finance infrastructure, wage and capital income of immobile residents will be taxed in each region. The total tax revenue (T_j) which is used for infrastructure investments in region j (X_j) is raised by a lump sum tax. The price of infrastructure is set to unity such that each unit of tax revenue buys one unit of infrastructure and the lump sum tax per resident is given by $t_j = \frac{T_j}{L_j+K_j} = \frac{X_j}{L_j+K_j}$. Note that this implies that taxation has no direct impact on the demand for manufactured goods, since we have assumed that there are no income effects in the markets for manufactured goods. Hence, the location of firms remains unaffected by this tax following the world income principle.

In fact our model is based on the standard result of tax competition which states, that regions will quit levying a source tax on the mobile factor but finance all expenses from a tax on immobile residents. As is shown by Sinn (2003) the costs of providing public infrastructure in a sense of a pure public good has to be covered by the immobile

[6] Expressing the input coefficient in terms of the elasticity of substitution σ is a common way to simplify New Economic Geography models (Fujita et al., 1999 p. 54, Baldwin et al., 2003, p.23). It can easily be shown that the result holds true for a general output elasticity of infrastructure, too. See Appendix A3 for a formal discussion.

factor, only.[7] In order to comprise capital income into the tax revenue we tax labor and capital income according to the residence principle as it is advocated by Bucovetsky and Wilson (1991). Moreover, since we want to focus on the effects of public input competition we refrain from distortionary taxes but think of a lump sum tax on resident income.

For now we model infrastructure as a pure public good, that is non-rivalry in its usage. Later on we also consider congestion costs in the provision of infrastructure and show how the results change if we have an impure public good.

Using the prices from (6), the definition of the price index (7), and the market clearing condition, the profits of each firm in region j from (5) are

$$\Pi_j = X_j \frac{\mu}{\sigma} \left(\frac{(K_j + L_j)}{X_j n^j + \tau^{1-\sigma} X_k n^k} + \tau^{1-\sigma} \frac{(K_k + L_k)}{X_k n^k + \tau^{1-\sigma} X_j n^j} \right) - R^j. \tag{10}$$

3.2.1 Short-Run Returns

In the monopolistic competition framework, free and instantaneous entry of firms drives pure profits to zero and the reward to capital is equal to the operating profit. Furthermore, each firm requires one unit of entrepreneurial capital (i.e., $n = K^W$), such that the number of local varieties is equal to the stock of employed capital. The share of industry or capital employed in region H can be defined as $s_n = n_H/n = n_H/K^W$, where $n = n_H + n_F$. Note that K_j represents the amount of capital, which is owned by region j, whereas n_j represent the amount of capital that is invested in region j. The share of population in region H (s_{pop}) is exogenously given by the regional factor endowments (L_H, K_H):

$$s_{pop} = \frac{K_H + L_H}{N}, \tag{11}$$

where $N \equiv K^W + L^W$ indicates worldwide population.

Using the fact that pure profits are zero, the short-run capital returns of both regions

[7]Sinn (2003) demonstrates that it is unlikely that capital taxes will be able to cover the costs of public infrastructure even if infrastructure is assumed to be an impure public good. The latter implies that public infrastructure faces congestion and that the tax rate equals a congestion charge. However, the congestion charge will under plausible circumstances not be able to cover the costs of public infrastructure provision.

result from (10) as:

$$R_H = \frac{\mu}{\sigma} X_H \left(\frac{s_{pop}}{\Delta_H} + \frac{\phi(1 - s_{pop})}{\Delta_F} \right) \frac{N}{n}, \tag{12}$$

$$R_F = \frac{\mu}{\sigma} X_F \left(\frac{\phi s_{pop}}{\Delta_H} + \frac{(1 - s_{pop})}{\Delta_F} \right) \frac{N}{n},$$

with $\Delta_H = X_H s_n + X_F \phi(1 - s_n)$ and $\Delta_F = X_F(1 - s_n) + X_H \phi s_n$,

where $\phi = \tau^{1-\sigma}$ is the degree of trade freeness with $0 < \phi \leq 1$. These short-run capital returns illustrate the two opposing effects firms consider when choosing their location. On the one hand they prefer the larger market (greater s_{pop}) and on the other hand they prefer the less crowded one (less competitors for input factors, i.e. smaller s_n). Finally, note that average short-run returns are independent of population distribution and infrastructure endowment:

$$s_n R_H + (1 - s_n) R_F = \frac{\mu N}{\sigma n} \equiv \bar{R}. \tag{13}$$

3.2.2 Long-Run Equilibrium

In the long-run equilibrium, capital owners cannot increase capital returns by relocation:

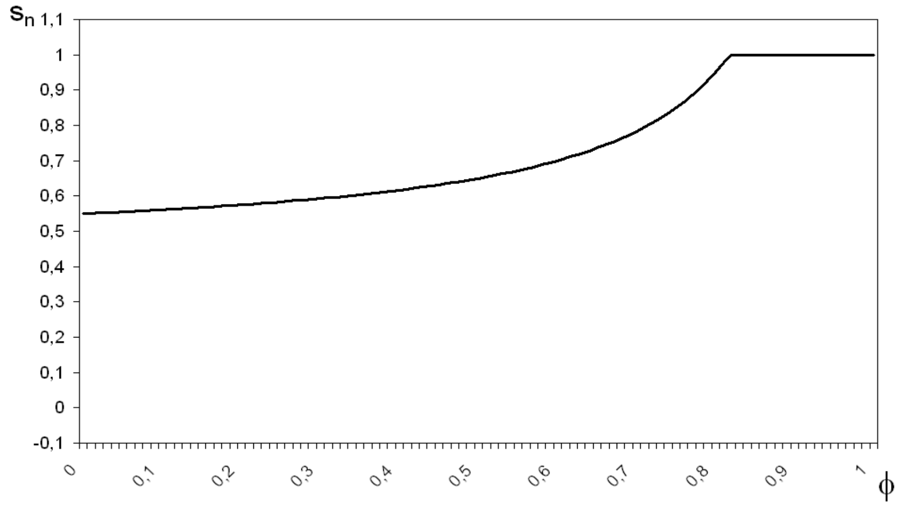
$$s_n = 0 \quad \text{if} \quad R_H - R_F|_{s_n=0} < 0 \Leftrightarrow \chi < \frac{\phi}{s_{pop} + (1 - s_{pop})\phi^2}, \tag{14}$$

$$s_n = 1 \quad \text{if} \quad R_H - R_F|_{s_n=1} > 0 \Leftrightarrow \chi > \frac{s_{pop}\phi^2 + 1 - s_{pop}}{\phi},$$

$$s_n = \frac{s_{pop}}{(1 - \phi\chi)} - \frac{\phi(1 - s_{pop})}{(\chi - \phi)} \quad \text{otherwise,}$$

where $\chi = X_H/X_F$. Either capital is completely located in the region with the larger returns or capital returns are equalized and both regions employ capital. If both regions have the same quality of infrastructure, the larger region (i.e., greater s_{pop}) will attract more and more industry during a process of falling trade costs. Sooner or later all industry will be concentrated in the larger region. This process is shown in figure 3.1, where the share of industry in region H is depicted against the freeness of trade. It is assumed that region H is larger in terms of population.

Figure 3.1: Bifurcation diagram



Having the larger share of industry implies also a lower consumer price index and accordingly a higher real income. This can be seen by rewriting the price indices (7) of regions H and F where prices (6) have been inserted:

$$P_H = \left(\frac{\sigma}{\sigma-1}\right)\left[X_H s_n + \phi X_F\left(1-s_n\right)\right]^{\frac{1}{1-\sigma}} n^{\frac{1}{1-\sigma}}, \tag{15}$$

$$P_F = \left(\frac{\sigma}{\sigma-1}\right)\left[X_F\left(1-s_n\right) + \phi X_H s_n\right]^{\frac{1}{1-\sigma}} n^{\frac{1}{1-\sigma}}.$$

Considering only the interior solution it follows from (15) that a higher share of industry in one region decreases its price index and increases the price index in the other region. However, the local share of industry increases with local infrastructure, therefore smaller regions with a lower share of expenditure can compensate their home market disadvantage by investing in infrastructure. For every combination of size asymmetry and trade freeness, there is one ratio of regional infrastructure that ensures an equal distribution of industry. Moreover, region H attracts all industry if it provides an infrastructure level X_H^{Agg}, just as well as region F does when providing X_F^{Agg}:

$$X_H^{Agg} = \frac{1-s_{pop}\left(1-\phi^2\right)}{\phi} X_F, \quad X_F^{Agg} = \frac{s_{pop}+\phi^2\left(1-s_{pop}\right)}{\phi} X_H \tag{16}$$

Taking the logarithm of price indices (15), we can express the impact of regional infrastructure in utility terms. Assuming interior solutions, that is $0 < s_n < 1$ we may further substitute for the long-run share of industry (14):

$$\ln(P_H) = \frac{1}{1-\sigma}\ln\left[\frac{X_H X_F s_{pop}(1-\phi^2)}{X_F - \phi X_H}\right] + A \quad \text{and} \tag{17}$$

$$\ln(P_F) = \frac{1}{1-\sigma}\ln\left[\frac{X_H X_F (1-s_{pop})(1-\phi^2)}{X_H - \phi X_F}\right] + A,$$

$$\text{with } A = \frac{1}{1-\sigma}\ln(n) + \ln\left(\frac{\sigma}{\sigma-1}\right).$$

These prices are only defined for $\chi > \phi$ and $1/\chi > \phi$. This means, the higher the trade freeness the lower the potential infrastructure differences. Intuitively this restriction states that consumer prices of imported varieties (i.e., the producer price plus trade costs) have to be higher than the consumer price of locally produced varieties (i.e., the producer price).[8] Infrastructure differences that imply a lower consumer price for imports than for local varieties would lead to full concentration of industry in the region with the better infrastructure, no matter what expenditure shares apply.

If all industry is concentrated in region j, the price indices are from (15):

$$\ln(P_j) = \frac{1}{1-\sigma}\ln[X_j] + A \quad \text{and} \quad \ln(P_k) = \frac{1}{1-\sigma}\ln[\phi X_j] + A, \quad k \neq j. \tag{18}$$

In the following sections it will be analyzed how the distribution of industry in figure 3.1 changes if regions are allowed to compete with infrastructure.

[8] $\chi > \phi$ implies $c_F\tau > c_H$, thus $\hat{p}_F\tau > \hat{p}_H$. $\frac{1}{\chi} > \phi$ implies $c_H\tau > c_F$, thus $\hat{p}_H\tau > \hat{p}_F$.

3.3 Decentralization

First, we analyze a situation where the two regions act independently, which we call decentralization.

3.3.1 Infrastructure and Price Index

It is crucial to the argument that we illustrate the impact of infrastructure provision on the local price index, because this is the channel through which the indirect utility of a consumer is affected. Two distinct effects of infrastructure affect regions' welfare. First, a lower input coefficient obviously allows the production of more output for given input. Differentiating (15) with respect to infrastructure and keeping the share of industry constant yields:

Direct local price index effect

$$\left.\frac{\partial \ln(P_H)}{\partial X_H}\right|_{\bar{s_n}} = \left(\frac{1}{1-\sigma}\right)\frac{1}{X_H s_n + \phi X_F (1-s_n)} s_n < 0. \tag{19}$$

Since the regional price index decreases, citizens utility increases. Second, from (12) we know that capital returns increase with the quality of local infrastructure provided. A region therefore attracts industry, when investing in infrastructure. The increasing share of local firms again lowers the local consumer price index, because less goods have to be imported in order to consume the optimal consumption bundle. This means citizens will save on trade costs. Differentiating (15) with respect to the share of industry and (14) with respect to infrastructure yields:

Indirect local price index effect

$$\frac{\partial \ln(P_H)}{\partial X_H} = \underbrace{\left(\frac{1}{1-\sigma}\right)\frac{X_H - \phi X_F}{X_H s_n + \phi X_F(1-s_n)}}_{\frac{\partial \ln(P_H)}{\partial s_n}} \underbrace{X_F \phi \left(\frac{s_{pop}}{(X_F - \phi X_H)^2} + \frac{(1-s_{pop})}{(X_H - \phi X_F)^2}\right)}_{\frac{\partial s_n}{\partial X_H}} < 0. \tag{20}$$

The first part of this effect $|\frac{\partial \ln(P_H)}{\partial s_n}|$ decreases with trade freeness ϕ, whereas the second part $|\frac{\partial s_n}{\partial X_H}|$ increases. This means for high trade freeness, it is not as important for consumers to have firms located in their home regions. However, firms get more footloose when trade costs fall, which implies that they are more easily attracted by infrastructure. This results from the fact that the home-market effect, which ties firms to the larger market, loses some of its strength. Overall the indirect effect increases with trade freeness.

For prohibitively high trade costs, that is $\phi = 0$, there is no indirect effect, because firms will choose their location only with respect to the exogenous home market size. In this case infrastructure does not have any impact on the long-run location of firms. For infinitely low trade costs, that is $\phi = 1$, the restriction described above does not allow for interior solutions. A corner solution with one region not providing any infrastructure has to apply. The amount of industry one region attracts, the other will lose. This is because the number of firms in the economy as a whole is given by the stock of capital $n = K^W = K_H + K_F$. Therefore each region exerts a negative effect on the utility of foreigners, when investing in local infrastructure. Furthermore, in uncoordinated behavior each region does not consider this indirect negative effect it has on the other region, when lowering the regional consumer price index. A negative externality evolves, which gains relevance for welfare because of monopolistic competition and trade costs.

The public good characteristics of infrastructure generates an additional, counteracting externality which becomes relevant if the trade freeness is sufficiently high. For low trade costs it might be cheaper to import manufactured goods than bearing the costs of providing enough infrastructure to have a decent share of manufacturing firms located in the respective region. Therefore, one of the two regions may decide not to provide any infrastructure at all and free-ride on the expenses of the other region. The same free-riding externality occurs if regions are very asymmetric in terms of population and infrastructure endowments. In this case it is very expensive for the smaller or less endowed region to attract industry and it may decide to better import and use the foreign infrastructure. Which of the two externalities dominates, the negative or the positive, depends crucially on the trade freeness and the initial distribution of industry as will be shown in the following.

3.3.2 Fiscal Competition Equilibria

Each region will maximize welfare of their citizens in the long-run by local infrastructure investment financed through income taxes subject to the residence principle. Thereby, each region takes infrastructure endowment in the other region as given. Since wages are equal to one and capital returns in the long-run equilibrium are the same for all capital owners, the objective of region j can be written as

$$\max_{X_j} \quad W_j \equiv L_j + K_j\bar{R} - t_j(L_j + K_j) - (K_j + L_j)\mu \ln(P_j) \quad \text{s.t.} \quad (14)$$

$$\text{where} \quad t_j(L_j + K_j) = X_j.$$

In a Nash equilibrium, both regions solve simultaneously the respective optimization problem.
In order to obtain general analytic results, we assume a symmetric distribution of population in the remaining part of this section: $s_{pop} = 1/2$. Later on, we will numerically solve for equilibria when population is asymmetrically distributed.

From (21) we get for $0 < s_n < 1$ the following first-order conditions

$$1 = \frac{\mu N}{2(\sigma - 1)} \frac{X_F}{X_H (X_F - \phi X_H)}, \quad (21)$$

$$1 = \frac{\mu N}{(\sigma - 1)} \frac{X_H}{X_F (X_H - \phi X_F)},$$

with reaction functions in the interior

$$X_{H_{1,2}} = \frac{X_F (\sigma - 1) \pm \sqrt{X_F (\sigma - 1) [X_F (\sigma - 1) - 2\mu\phi N]}}{2\phi (\sigma - 1)},$$

$$X_{F_{1,2}} = \frac{X_H (\sigma - 1) \pm \sqrt{X_H (\sigma - 1) [X_H (\sigma - 1) - 2\mu\phi N]}}{2\phi (\sigma - 1)}.$$

However, zero public good supply or “threshold” supply that ensures full agglomeration of industry might also be optimal. As a consequence, depending on trade costs, fiscal competition may either lead to interior equilibria and dispersion of industry or to corner equilibria and agglomeration, as the following proposition states.

Proposition 1 *(a) A symmetric locally stable Nash equilibrium*

$$X_H = X_F = \frac{\mu N}{2(1-\phi)(\sigma-1)} > 0 \quad \textit{with} \quad s_n = \frac{1}{2} \tag{22}$$

exists if and only if $\phi \leq 0.1748$*.*
(b) No Nash equilibrium exists if and only if $0.1748 < \phi < 0.2832$*.*
(c) Corner Nash equilibra with

$$X_H = \frac{\mu N}{2(\sigma-1)},\ X_F = 0,\ s_n = 1 \quad \textit{or} \quad X_F = \frac{\mu N}{2(\sigma-1)},\ X_H = 0,\ s_n = 0 \tag{23}$$

exist if and only if $\phi \geq 0.2832$*.*

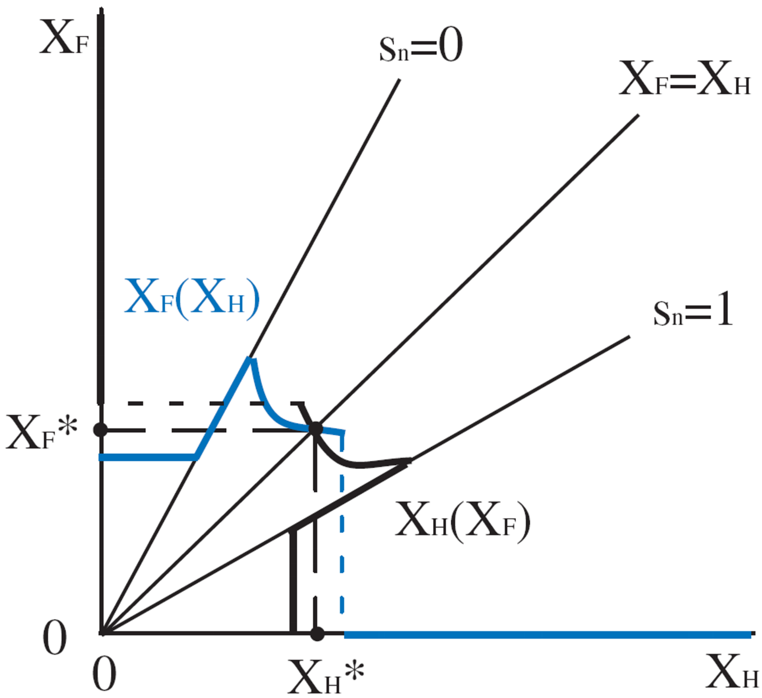
Proof. See Appendix A1. ■

Note that these critical values of trade freeness are only true for ex-ante symmetric regions ($s_{pop} = 1/2$). Section 3.5.1 discusses ex-ante asymmetric regions. In the symmetric Nash equilibrium, the provided level has the expected correlations. It increases with μ, because manufactured goods gain more weight in the individuals' preferences. The indirect effect, and accordingly the competition for industrial firms, gets the stronger the higher ϕ. Therefore, the provided level increases in trade freeness, too. The elasticity of substitution σ has a negative effect on the equilibrium provision of infrastructure, because the output elasticity of infrastructure $1/(\sigma-1)$ decreases with σ.

In this case, free-riding will not occur, because the trade costs are too high to waive local industry. In contrast, the regions prefer to invest in infrastructure in order to attract industry from the foreign region.

Corner Nash equilibria particularly require that the periphery does not prefer to deviate. It may prefer to deviate from the corner solution of zero infrastructure provision if any positive value of infrastructure provision exists that ensures a higher utility than the corner solution does. There are two possible candidates that may be beneficial for the periphery. It either provides the optimal level for an interior solution or it provides a very high level of infrastructure that attracts all industry from the core.
Corner solutions only exist if the trade freeness is sufficiently high. The intuition for this finding is that both regions try to free-ride on the expenses of the other region if the trade costs are sufficiently low, since the costs of setting up local industry exceed the costs of importing manufactured goods from the foreign region.

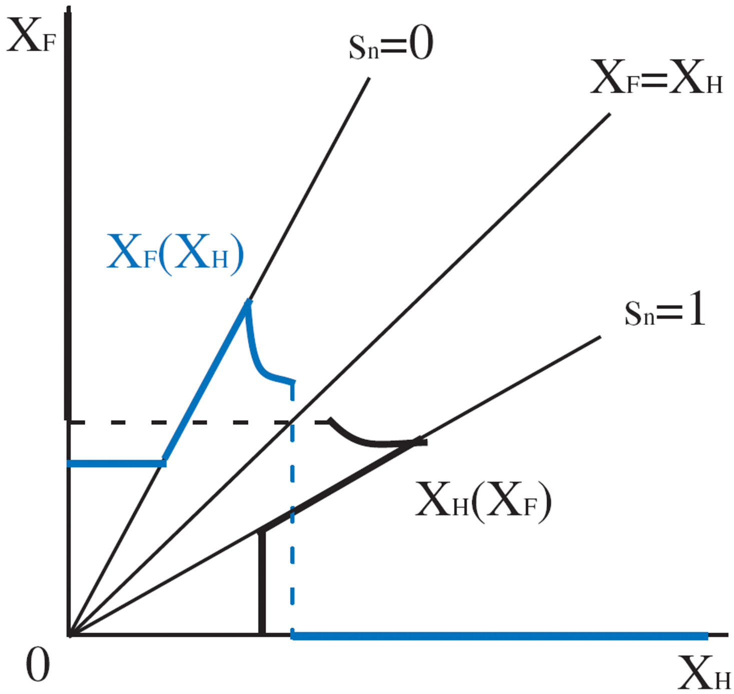
Figure 3.2: Interior Nash equilibrium at low trade freeness



In the decentralized equilibrium we may distinguish three scenarios, with respect to the trade freeness. First, if trade costs are very high, i.e., if $\phi < 0.1748$, only a stable symmetric Nash equilibrium with dispersion of industry exists. Second, for medium trade costs no Nash equilibrium in pure strategies exists at all.[9] Third, if trade freeness is high, i.e., if $\phi > 0.2832$, only corner Nash equilibria with a core-periphery pattern occur. The first case is depicted in figure 3.2, the second in figure 3.3, the third in figure 3.4. These figures show the reaction curves of both regions and – provided that they exist – their intersection points, that is the Nash equilibria. Furthermore, full-agglomeration lines where the entire industry is located in either region are shown. Not surprisingly, regions do not increase infrastructure investment further when the industry is already completely located on their territory. The periphery without any industry does not invest at all. Moreover, if the other region provides a really large quantity of public goods, attracting industry is so costly that the region refrains from investing.

[9]Since payoff functions W_j are continuous in X_H and X_F, according to the theorem of Glicksberg (1952) a mixed strategy equilibrium exists provided that the choice of infrastructure is restricted to a compact subset of $\Re_0^+$.

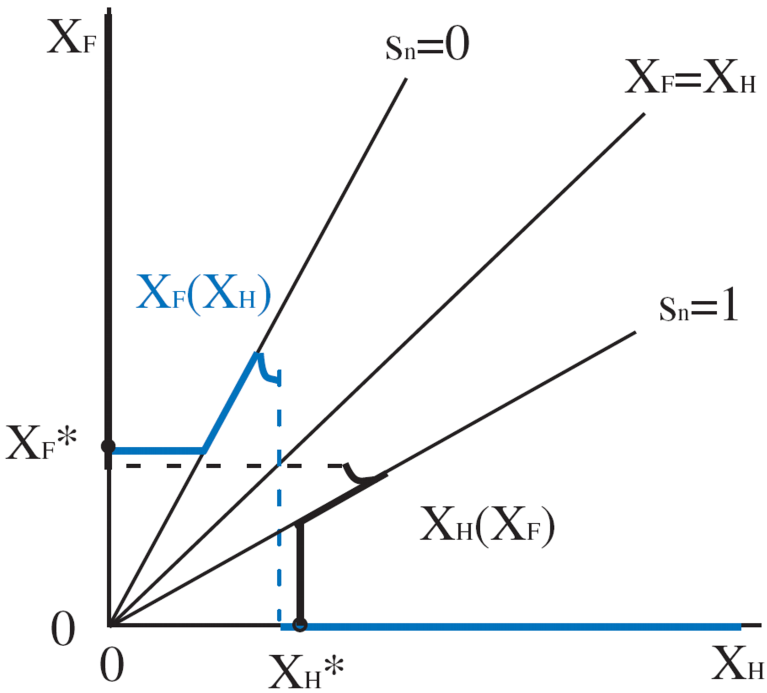
Figure 3.3: Non-existence of Nash equilibrium at medium trade freeness



The shape of the reaction curves can be easily explained. Starting at low infrastructure investment of region F, region H invests as much as necessary to attract all manufacturing firms and to become the core. When the other region increases investment further, region H must also increase investment in order to prevent capital flight. If a certain investment level of region F is achieved, this strategy becomes too costly. Region H now reduces investment, leading to industry dispersion. A further increase in region F's investment ultimately eliminates any incentive to supply infrastructure. As a consequence, region H does not supply infrastructure at all. For region F, the same reasoning applies. As can be seen from the figures, the higher the trade freeness, the smaller the area where both regions supply infrastructure goods and where industry is dispersed.

For low trade freeness (figure 3.2), reactions curves intersect once in the interior, for medium trade freeness never (figure 3.3), for high trade freeness (figure 3.4) two times at the axes. Furthermore, since in an interior equilibrium the reaction curve of region H is steeper than the reaction curve of region F, this equilibrium is locally stable, i.e. adjustment along the reaction curves ultimately leads to the equilibrium.

Figure 3.4: Corner Nash equilibria at high trade freeness



3.4 Centralization

A central government endowed with fiscal authorities may levy taxes on both regions symmetrically but distribute infrastructure asymmetrically. Furthermore, the central government internalizes the externalities that regions impose on each other. Again, workers and capital owners in both regions are taxed in order to finance infrastructure. The objective of the federal government is to maximize the sum of workers' utility in both regions:

$$\max_{X_H, X_F} \quad W \quad \equiv \quad L_H + L_F + n\bar{R} - t(L^W + K^W) \tag{24}$$

$$-(K_H + L_H)\mu \ln(P_H) - (K_F + L_F)\mu \ln(P_F) \quad \text{s.t. } (14).$$

$$\text{where} \quad t(L^W + K^W) = X_H + X_F.$$

To solve this problem, we assume again equally sized regions, i.e. $s_{pop} = 1/2$.

Proposition 2 *Symmetric distribution of infrastructure is optimal for low trade freeness, i.e., if* $\phi < 7 - 4\sqrt{3} \approx 0.0717$*. Otherwise, a corner solution with either all infrastructure in region* H *or region* F *is optimal. Independent of trade costs, the optimum aggregate infrastructure investment is*

$$X_H + X_F = \frac{\mu N}{\sigma - 1}. \tag{25}$$

Proof. See Appendix A2. ■

For high trade freeness, manufactured goods produced in one region can be consumed in both regions at low costs. Hence the federal government concentrates infrastructure in one region in order to maximize the aggregate productivity in the manufacturing sector. In this case the gain in aggregate productivity is higher than the rise of expenses on trade costs.
For low trade freeness, the federal government distributes infrastructure equally with the consequence of low aggregate productivity in order to minimize overall trade costs. The savings on trade costs due to the equalized distribution of manufacturing plants dominates the potential gain in aggregate productivity.

Comparing optimal infrastructure investment with the outcome of fiscal competition, differences occur with respect to the level and the distribution of infrastructure investment depending on the trade costs. A comparison of the critical trade costs values given in the pervious propositions yields immediately:

Proposition 3 *(a) For medium trade freeness, i.e., for* $0.1748 < \phi < 0.2832$*, the allocation of infrastructure cannot be determined consistently by decentral authorities.*
(b) A decentralized federation oversupplies public infrastructure if trade costs are high, i.e., if $\phi < 0.1748$ *(where infrastructure investment is fully equalized). For high trade freeness, i.e., for* $\phi > 0.2832$*, decentralization leads to undersupply.*
(c) When trade costs fall, a decentralized federation changes too late from full equalization of infrastructure investment and dispersed industry locations to a core-periphery pattern.

For medium trade freeness, both jurisdictions would alternately overbid and underbid each other. As a result, a decentralized federation does not even come to a solution for the allocation problem. For more extreme trade costs, decentral authorities provide infrastructural goods consistently, but never efficiently.

The central government provides always $X_H + X_F = \mu N/\left(\sigma - 1\right)$ on aggregate. However, in the decentralized case each of the two regions provides $\mu N/[2(\sigma-1)(1-\phi)]$ if the full equalization equilibrium applies. If a core-periphery pattern arises, the core region provides infrastructure $\mu N/\left[2(\sigma-1)\right]$. The intuition for those results are as follows: For low trade freeness it is crucial to have industry located nearby, therefore the negative externality from competing for industry dominates the positive externality due to the public input characteristic of infrastructure and the regions overprovide infrastructure in decentralized equilibrium. As the trade costs fall below a critical value, importing manufactured goods becomes cheap enough to waive local industry. This strategy allows the saving of the the costs of setting up infrastructure by indirectly using the foreign one. Using infrastructure without contributing to the costs of providing it represents a free-riding externality, which now leads to underprovision. Furthermore, if the industry is located in one region, the core is able to skim off agglomeration rents, and thus reduces its effort to reduce production costs.

Comparing decentralization and centralization with respect to the distribution of infrastructure an inefficiency arises for the range of trade freeness between 0.0717 and 0.2832, because corner solutions occur in the centralized economy already for lower trade freeness than in the decentralized economy. Choosing a corner solution without local industry a single region benefits from increasing aggregate productivity only to the extent weighted by the trade freeness. Therefore, the productivity gain it takes into account is smaller than the aggregate productivity gain the federation faces as a whole. Moreover, the increase in trade costs of the region that loses all industry is higher than the average increase in trade costs, since the other region saves trade costs. Both effects imply that the federal government chooses corner solutions already for a lower trade freeness than the single regions do in Nash equilibrium.

3.5 Extensions

Assuming symmetry and pure public goods greatly simplifies the analysis, but our results are much more general. They will qualitatively hold in reasonably more complex circumstances.

3.5.1 Asymmetric Regions

Asymmetry in population size does not fundamentally change the mechanisms or the results. Although it is not possible to obtain analytical results, by way of numerical analysis it can be shown that the comparison of decentral and central states leads to similar results as in the case with equally sized regions.[10] First, for low trade freeness, in a decentralized federation both regions provide public infrastructure and the aggregate level of investment is too high. However, due to the home-market effect, a disproportional high share of industry is located in the larger region even if both provide the same amount of public inputs. For high trade freeness, decentralization leads again to agglomeration and undersupply. Of course corner solutions with only one region providing infrastructure already take place for lower trade freeness than for ex-ante symmetric regions. Second, in the process of trade integration a decentralized federation changes too late from dispersion of infrastructure investment and industry locations to a core-periphery pattern. However, asymmetry is an additional source of inefficiency. For low trade freeness, competition for mobile firms strongly forces both the smaller and the larger region to supply public infrastructure. As a result, the difference in public investment in an interior Nash equilibrium is too small. A central planner would increase investment in the larger region and simultaneously reduce expenditure in the smaller region.

3.5.2 Congestion Costs

Of course, there could be congestion costs that act against the efficiency of agglomeration. One possibility to take those into account is to assume that infrastructure entails usage costs that increase with the local amount of firms (see, e.g. Sinn, 1997). This relationship can be defined by the following cost functions:

$$Cost(X_H) = \left[1 + \alpha\left(\frac{s_n}{0.5} - 1\right)\right] X_H \quad \text{and} \quad Cost(X_F) = \left[1 + \alpha\left(\frac{(1-s_n)}{0.5} - 1\right)\right] X_F,$$

where α, $0 \leq \alpha \leq 1$, represents the parameter of congestion. For $\alpha = 0$ as well as for symmetric industry distribution ($s_n = 0.5$) there is no congestion and the price of infrastructure provision remains unity as before. The higher α and the larger the asymmetry of industry distribution, the stronger the congestion. For full agglomeration

[10]For numeric simulations see Appendix A4.

the price of infrastructure in the core region becomes $1 + \alpha$. The implied utility for symmetric distribution of infrastructure is the same as in the absence of congestion. Comparing welfare with and without congestion when regions are of the same size, one can easily show that the critical value of trade freeness where agglomeration becomes efficient increases with α. Considering the Nash equilibrium congestion costs entail an additional externality. While the marginal utilities of infrastructure in both regions are not affected by the congestion costs, the marginal costs of infrastructure provision are increasing in the stock of regional infrastructure and the congestion parameter α. Hence, for low trade freeness the pressure for equalization of infrastructure levels is strengthened. However for sufficiently high trade freeness one may still expect corner solutions, i.e. full agglomeration, to arise. When deciding on whether or not to provide infrastructure a region contrasts the cost savings of not providing infrastructure to the utility loss of not having local industry and accordingly, having to import all industrial goods. Compared to the scenario without congestion costs the costs savings have risen, whereas the utility loss has not changed. Hence, corner solutions may already occur for lower trade freeness than without congestion costs. Regarding the efficiency of aggregate infrastructure provision the resulting overprovision in case of low trade freeness may be alleviated, since a region exerts not only a negative effect on the other region when investing in infrastructure by attracting industry but also a positive effect by decreasing the costs of infrastructure and accordingly the tax burden in the foreign region. The result of underprovision in the case of low trade-freeness, however, is still valid, since congestion costs do not arise in the periphery where all industry has left to the core. Obviously, the assumption of congestion costs may only act as an additional externality as long as both regions have some local industry that is as long as interior solutions arise.

3.6 Concluding Remarks

The European market becomes more and more integrated and trade costs are falling. We have shown that this phenomenon alters significantly the implications of public input competition. When trade costs are sufficiently low, it becomes favorable for regions to import industrial goods instead of providing infrastructure. Thus some regions may prefer to quit the competition for industry. Those regions become voluntarily the periphery and specialize in the agricultural sectors. Hence, the strategic behavior of individual regions which we observe to an increasing degree constitutes a prominent channel of agglomeration.

As we have shown in this chapter agglomeration becomes for a sufficiently high level of trade freeness the efficient allocation of economic activity in space. The central government internalizing externalities chooses to allocate infrastructure and accordingly manufacturing industry to a single core-region. Note that regional integration which is accompanied by rising trade freeness is certainly a desirable process since it increases aggregate welfare.[11]

These results are very much in line with the central place theory by Christaller (1933) and Lösch (1940) which states that there are different categories of central places which constitute the center of economic activity (see Chapter 2). According to the central place theory not every kind of public infrastructure should be provided at any location, but services and goods of higher order should only be provided in the major agglomerations. Our model derives an optimal distance between central places providing public infrastructure.

The critical trade freeness for which agglomeration becomes efficient can be translated into the efficient size of a market area surrounding a central place. Similarly as in Chapter 2, we can define the trade freeness as a function of the spatial distance $\phi = g(d)$ where $g'(d) < 0$. Computing the inverse function $d = g^{-1}(\phi)$ allows us to express the trade freeness derived in Section 3.4 in terms of spatial distance. However, note that the critical trade freeness for which agglomeration becomes efficient ($\phi \approx 0.0717$) is certainly not a natural constant but depends on the assumptions we made with respect to the sizes of regions, the output elasticity of infrastructure, the congestion intensity and the price of infrastructure. Hence, the true distance which ensures that agglomeration is efficient depends on these parameters as well as on the

[11] From equation (24) aggregate welfare is negatively related to the regional price-indices of both regions. The regional price-indices again decrease when trade costs fall as shown in equation (15).

specific functional relationship between distance and trade freeness which eventually remains an empirical question.
However, more importantly we have shown that regional competition will cause agglomeration to arise not before it becomes efficient. In fact the uncoordinated competition equilibrium is characterized by agglomeration only at higher trade freeness than it would be efficient. Hence, if regional policy interferes with the regional distribution of economic activity it should essentially promote regional concentration instead of dispersion.

Since externalities arise as long as trade is possible, a centralized provision of infrastructure is always welfare-improving. We have shown that for high trade costs the uncoordinated competition for local industry yields an overprovision of infrastructure while for low trade costs the free-riding externality dominates which implies an underprovision of public inputs. Hence, for low trade costs the overall level of infrastructure should be subsidized.
Thus our model implies that regional policies should subsidize infrastructure in core regions in order to achieve the efficient level of aggregate infrastructure. Moreover, from an efficiency point of view industrial concentration should already be implemented for higher trade costs than it is the outcome of regional competition. This is in sharp contrast with the German or the European Union's regional policy which aims at enhancing regional integration and at the same time subsidizes public infrastructure in peripheral regions. A trade-off between aggregate efficiency and regional cohesion arises as it was already demonstrated in other models of the New Economic Geography (see Martin, 1999).

Our model focuses on the effects of fiscal competition on infrastructure provision and industrial distribution. It neglects all effects of information asymmetries between central and decentral authorities which would make the case for decentralization. Moreover, our model features only one industrial sector with industry-specific infrastructure. The results could be translated into a world with several industrial branches where efficient agglomeration means that various industrial sectors cluster in different regions. This is a tendency we observe in federations all over the world.

Appendix

A1: Proof of Proposition 1

We analyze first candidates for interior Nash equilibria. From the first-order conditions (21) for $s_{pop} = 1/2$ and $0 < s_n < 1$ we derive the potential equilibria in the interior:

$$X_{H_1} = X_{F_1} = \frac{\mu N}{2\left(1-\phi\right)\left(\sigma-1\right)},$$

$$X_{H_2} = \frac{\mu N}{(\sigma-1)}\left(\frac{1+\phi+\sqrt{1-2\phi-3\phi^2}}{4\phi\left(\phi+1\right)}\right),\ X_{F_2} = \frac{\mu N}{(\sigma-1)}\left(\frac{1}{1+\phi+\sqrt{1-2\phi-3\phi^2}}\right),$$

$$X_{H_3} = \frac{\mu N}{(\sigma-1)}\left(\frac{1+\phi-\sqrt{1-2\phi-3\phi^2}}{4\phi\left(\phi+1\right)}\right),\ X_{F_3} = \frac{\mu N}{(\sigma-1)}\left(\frac{1}{1+\phi-\sqrt{1-2\phi-3\phi^2}}\right).$$

Note that the second and third solutions are only defined for $\phi \leq 1/3$. Calculating the second derivatives and the cross derivatives gives

$$\frac{\partial^2 W_H}{\partial X_H^2} = \left[\frac{\mu N}{2(1-\sigma)}\right]\frac{X_F\left(X_F-2\phi X_H\right)}{X_H^2\left(X_F-\phi X_H\right)^2},\qquad \frac{\partial^2 W_F}{\partial X_F^2} = \left[\frac{\mu N}{2(1-\sigma)}\right]\frac{X_H\left(X_H-2\phi X_F\right)}{X_F^2\left(X_H-\phi X_F\right)^2},$$

$$\frac{\partial^2 W_H}{\partial X_H \partial X_F} = \frac{\mu N \phi}{2\left(1-\sigma\right)\left(X_F-\phi X_H\right)^2},\qquad \frac{\partial^2 W_F}{\partial X_F \partial X_H} = \frac{\mu N \phi}{2\left(1-\sigma\right)\left(X_H-\phi X_F\right)^2}.$$

For $X_F = X_H$, the second-order condition $\partial^2 W_j/\partial X_j^2 \leq 0$ is only fulfilled for $\phi \leq 1/2$. For the asymmetric candidates, the second-order conditions $\partial^2 W_H/\partial X_H^2 \leq 0$ and $\partial^2 W_F/\partial X_F^2 \leq 0$ require $4\phi^2 + 2\phi - 1 \geq 0$, i.e. $\phi \geq 0.309017$.

The symmetric equilibrium is locally stable if the reaction curve $X_H(X_F)$ is steeper than $X_F(X_H)$, i.e., if

$$-\frac{\partial^2 W_H/\partial X_H^2}{\partial^2 W_H/(\partial X_H \partial X_F)} < -\frac{\partial^2 W_F/(\partial X_F \partial X_H)}{\partial W_F/\partial X_F^2} \Leftrightarrow \frac{2\phi-1}{\phi} < \frac{\phi}{2\phi-1}, \tag{26}$$

which is satisfied for $\phi < 1/3$.

It can be easily shown that for both asymmetric solutions there exists one region which would gain from abandoning public infrastructure investment completely. Thus, the asymmetric solutions do not present Nash equilibria.

When one single region deviates from the symmetric solution and does not provide public infrastructure, the gain for this region is

$$\frac{\mu N \{1 + (1-\phi)\ln(\phi) - (1-\phi)\ln[(1+\phi)/2]\}}{2(1-\phi)(\sigma-1)},$$

which is positive if and only if $\phi > 0.1748$. A single upwards deviation, i.e., a strong increase in investment which makes the deviating region core, pays only if $\phi > 0.4492$. Hence, already for $\phi > 0.1748$ a symmetric Nash equilibrium does not exist.

Now, we consider corner solutions. If one region, say region F, does not supply infrastructural goods, region H maximizes regional welfare W_H by choosing $X_H = \mu N/[2(\sigma-1)]$ (which fulfills the second-order conditions for H). If region F deviates by supplying infrastructure according to the reaction curve given above it gains

$$\frac{\mu N}{4\phi(\sigma-1)} \left\{ \sqrt{1-4\phi} - 1 - \phi\ln(4) - 2\phi\ln(\phi) + 2\phi\ln\left[\frac{(\phi^2-1)(\sqrt{1-4\phi}-1)}{\phi(\sqrt{1-4\phi}+1)}\right]\right\},$$

which is positive if and only if $\phi < 0.25$. If region F deviates upwards, i.e., if it supplies X_F^{Agg} which makes it the core, it gains

$$-\frac{\mu N}{4\phi(\sigma-1)}\left[1+\phi^2+2\phi\ln(2\phi) - 2\phi\ln\left(\frac{1+\phi^2}{\phi}\right)\right],$$

which is positive for $0.2203 < \phi < 0.2832$. Hence, for $\phi < 0.2832$ a corner Nash equilibrium does not exist.

As a consequence, for $0.1748 < \phi < 0.2832$ single deviations rule out any equilibrium.

A2: Proof of Proposition 2

In order to solve the optimization problem (24) for $s_{pop} = 1/2$, we determine separately interior solutions and corner solutions, compare the implied welfare levels, and, finally, choose the strategy which leads to the highest welfare level.

For infrastructure investment that implies $0 < s_n < 1$, the first-order conditions are

$$1 = \frac{\mu N}{\sigma - 1}\frac{1}{X_H}\left(2 + \frac{X_H\phi}{X_F - \phi X_H} - \frac{X_H}{X_H - \phi X_F}\right),$$
$$1 = \frac{\mu N}{\sigma - 1}\frac{1}{X_F}\left(2 + \frac{X_F\phi}{X_H - \phi X_F} - \frac{X_F}{X_F - \phi X_H}\right).$$

There are exactly three solutions for regional infrastructure investments. One solution that implies full equalization of regional infrastructure, that is a ratio $\chi_1 = 1$ and absolute levels:

$$X_{H_1} = X_{F_1} = \frac{\mu N}{2(\sigma - 1)} .$$

The trace and the determinant for this symmetric solution are

$$\text{trace}_{\chi_1} = -\frac{4(\sigma-1)[1+\phi(\phi-4)]}{\mu N(1-\phi)^2},$$
$$\det_{\chi_1} = \frac{4(\sigma-1)^2[1+\phi(\phi-6)]}{\mu^2 N^2(1-\phi)^2}.$$

Hence, for $\phi < 3 - 2\sqrt{2} \approx 0.1715$ the determinant is positive (and the trace negative), and, therefore, this symmetric solution is a maximum.

Beside this symmetric solution, there are two asymmetric solutions that imply regional infrastructure ratios $\chi_{2,3}$, with $\chi_3 < \chi_1 < \chi_2$, which are only defined for $\phi \leq 3 - 2\sqrt{2}$. However, using continuity one can easily show that those two solutions are minima.

In order to obtain candidates for corner solutions, we set without loss of generality X_F to zero. This implies a first-order condition for X_H

$$\frac{\mu N}{X_H(\sigma - 1)} - 1 = 0,$$

and, thus, $X_H = \mu N/(\sigma - 1)$ (the second-order condition is obviously fulfilled).

The difference in welfare between the corner solution and an interior symmetric allocation of infrastructure is

$$\frac{\mu N}{2(\sigma - 1)}\left[\ln(16) + \ln(\phi) - 2\ln(1+\phi)\right],$$

which is positive if and only if $\phi > 7 - 4\sqrt{3}$.

A3: Generalized Output Elasticity of Infrastructure

If we allow for a general output elasticity of infrastructure $\delta > 0$ the production function is

$$Q_j = X_j^{\delta} L_j \leftrightarrow c = X_j^{-\delta}$$

Solving for the long-run distribution of industry gives

$$s_n = \frac{s_{pop}}{(1 - \phi\chi^{\delta(\sigma-1)})} - \frac{\phi(1 - s_{pop})}{(\chi^{\delta(\sigma-1)} - \phi)} \text{ ; where } \chi = \frac{X_H}{X_F}$$

and a local price indices

$$\ln(P_H) = \frac{1}{1-\sigma}\ln\left[\frac{X_H^{\delta(\sigma-1)} X_F^{\delta(\sigma-1)} s_{pop}(1-\phi^2)}{X_F^{\delta(\sigma-1)} - \phi X_H^{\delta(\sigma-1)}}\right] + \ln\left(\frac{\sigma}{\sigma-1}\right) + \frac{1}{1-\sigma}\ln(n)$$

$$\ln(P_F) = \frac{1}{1-\sigma}\ln\left[\frac{X_H^{\delta(\sigma-1)} X_F^{\delta(\sigma-1)} (1-s_{pop})(1-\phi^2)}{X_H^{\delta(\sigma-1)} - \phi X_F^{\delta(\sigma-1)}}\right] + \ln\left(\frac{\sigma}{\sigma-1}\right) + \frac{1}{1-\sigma}\ln(n)$$

Using the indirect utility functions for the decentralized case we get first order conditions:

$$1 = \frac{\mu\delta N}{2X_H}\frac{X_F^{\delta(\sigma-1)}}{\left(X_F^{\delta(\sigma-1)} - \phi X_H^{\delta(\sigma-1)}\right)}$$

$$1 = \frac{\mu\delta N}{2X_F}\frac{X_H^{\delta(\sigma-1)}}{\left(X_H^{\delta(\sigma-1)} - \phi X_F^{\delta(\sigma-1)}\right)}$$

which imply an infrastructure provision in case of the full equalization equilibrium

$$X_H = X_F = \frac{\mu\delta N}{2(1-\phi)}$$

Determinant and trace are given by

$$\det = \frac{(1-\phi)\left(1-\phi-2\delta\phi\left(\sigma-1\right)\right)}{\mu^2 N^2 \delta^2}$$

$$tr = -\frac{2\left(1-\phi+\delta\phi\left(1-\sigma\right)\right)}{\mu N \delta}$$

The trace is negative if $\phi < \frac{1}{1+\delta(\sigma-1)}$, the determinant positive if $\phi < \frac{1}{1+2\delta(\sigma-1)}$. When setting $\delta = \frac{1}{\sigma-1}$ we get the same results as above. Infrastructure has diminishing returns if $0 < \delta < 1$. If this is the case the full equalization equilibrium is stable even for high trade freeness, whereas it gets unstable already for low trade freeness if increasing returns to infrastructure (*i.e.* $\delta > 1$) are presumed.

The federal government faces first order conditions:

$$1 = \frac{\mu\delta N}{X_H}\left[2 + \frac{\phi X_H^{\delta(\sigma-1)}}{\left(X_F^{\delta(\sigma-1)} - \phi X_H^{\delta(\sigma-1)}\right)} - \frac{X_H^{\delta(\sigma-1)}}{\left(X_H^{\delta(\sigma-1)} - \phi X_F^{\delta(\sigma-1)}\right)}\right]$$

$$1 = \frac{\mu\delta N}{X_F}\left[2 + \frac{\phi X_F^{\delta(\sigma-1)}}{\left(X_H^{\delta(\sigma-1)} - \phi X_F^{\delta(\sigma-1)}\right)} - \frac{X_F^{\delta(\sigma-1)}}{\left(X_F^{\delta(\sigma-1)} - \phi X_H^{\delta(\sigma-1)}\right)}\right]$$

which imply an infrastructure provision

$$X_H = X_F = \mu\delta N$$

Determinant and trace are given by

$$\det = \frac{1 + \phi^2 - 2\phi\left(1 + 2\delta\left(\sigma - 1\right)\right)}{\left(\phi - 1\right)^2 \mu^2 N^2 \delta^2}$$

$$tr = -\frac{2\left[1 + \phi^2 - 2\phi\left(1 + \delta\left(\sigma - 1\right)\right)\right]}{\left(\phi - 1\right)^2 \mu N \delta}$$

The trace is negative for $\phi < 1 + \delta\left(\sigma - 1\right) \pm \sqrt{\delta^2\left(1 - \sigma\right)^2 + 2\delta\left(\sigma - 1\right)}$, the determinant positive for $\phi < 1 + 2\delta\left(\sigma - 1\right) \pm 2\sqrt{\delta^2\left(1 - \sigma\right)^2 + \delta\left(\sigma - 1\right)}$. Setting $\delta = \frac{1}{\sigma-1}$ the same results as above are obtained. Moreover the lower δ the later it becomes efficient to distribute infrastructure asymmetrically. Hence, the symmetric provision of infrastructure is efficient up to a higher trade freeness for diminishing returns to infrastructure than it is for increasing returns.

A4: Simulations for Asymmetric Regions

We simulate fiscal competition with asymmetric regions, where our basic parameters are $K^W = 1, L^W = 3, \sigma = 1.5, \mu = 0.7$. Region H is only slightly larger than region F, that is $s_{pop} = 0.6$.

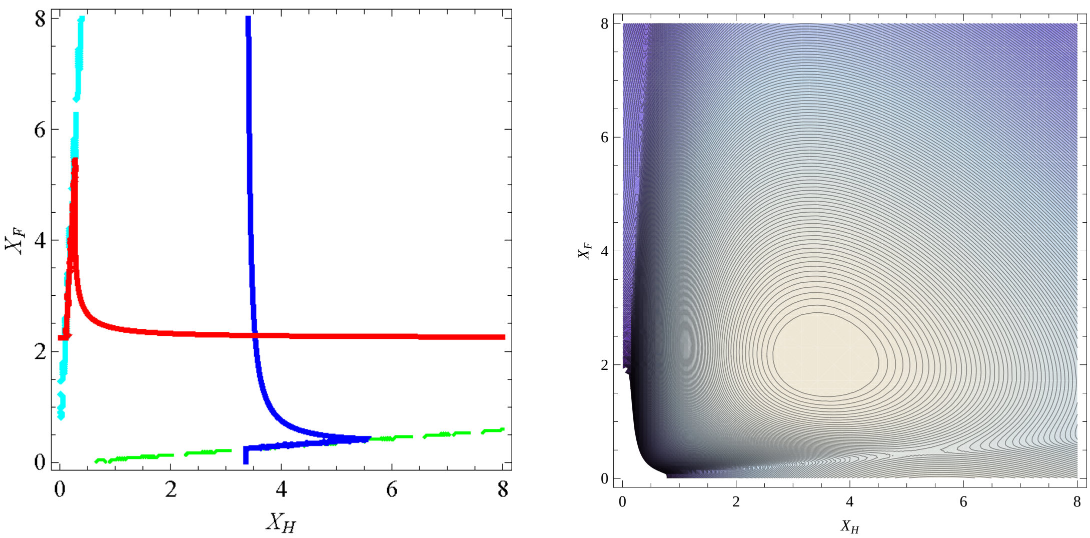
The following figures, where X_H is on the horizontal axis and X_F on the vertical axis, depict, for varying degrees of trade freeness ϕ, iso-total-welfare curves and regional reaction-curves in the interior and at "threshold" levels that ensure full agglomeration. Regional welfare is given by

$$W_j = Y_j - X_j - (K_j + L_j)\mu \ln(P_j)$$

where $Y_j = L_j + K_j\bar{R} = s_{pop}(L_j + \mu N/\sigma)$. Reaction-curves in the interior are defined by $dW_j/dX_j = 0$.

Low trade freeness: $\phi = 0.03$

Figure 3.5: Reaction functions and Welfare at low trade freeness

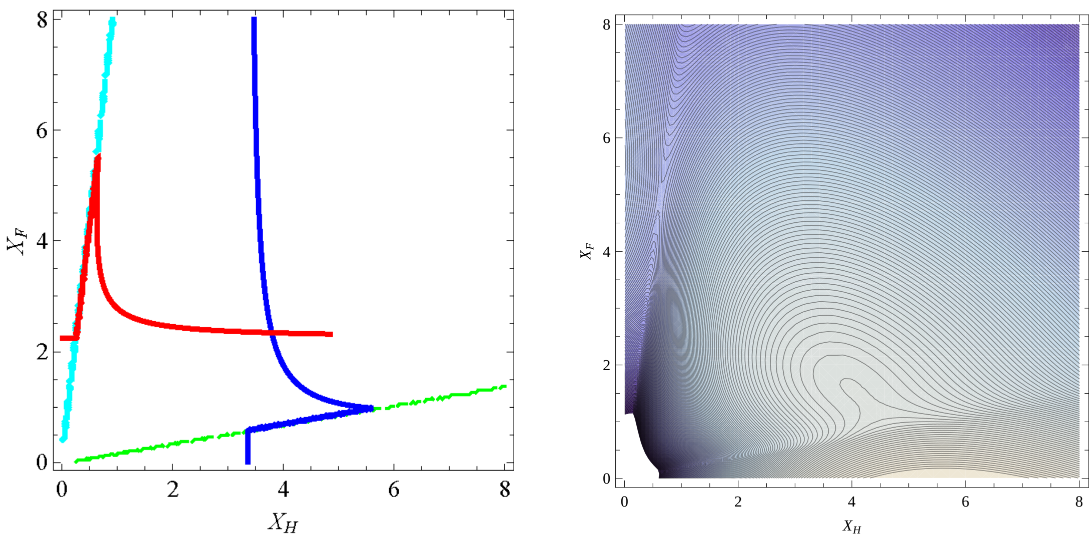


	X_H	X_F	$W_H + W_F$
Nash equilibrium	3.5229	2.2844	1.57029
Optimum	3.4949	2.1051	1.57702

The Nash equilibrium for $\phi = 0.03$ and the above parameter constellation is represented by the intersection of the reaction-curves in the left panel of figure 3.5. The red line depicts the reaction-curve of region F and the dark blue line depicts the reaction-curve of region H. Infrastructure combinations beyond the light blue and the green line imply full agglomeration in region F and H, respectively. The right panel of figure 3.5 illustrates the aggregate welfare of regions H and F. The three-dimensional contour plot assigns a dark blue shading to X_H,X_F combinations with a low level of aggregate welfare and a light beige shading to X_H,X_F combinations with a high level of aggregate welfare. The reaction-curve indicate that an interior Nash equilibrium is reached which exhibits the infrastructure levels given in the table above. The contour plot indicates that the efficient infrastructure provision requires positive amounts of X_H and X_F. The exact values are given in the table. The figures in the table show that the interior Nash equilibrium is characterized by an inefficient high level of the aggregate public infrastructure investment with a too small difference in infrastructure investments between the two regions.

Medium trade freeness: $\phi = 0.07$

Figure 3.6: Reaction functions and Welfare at medium trade freeness

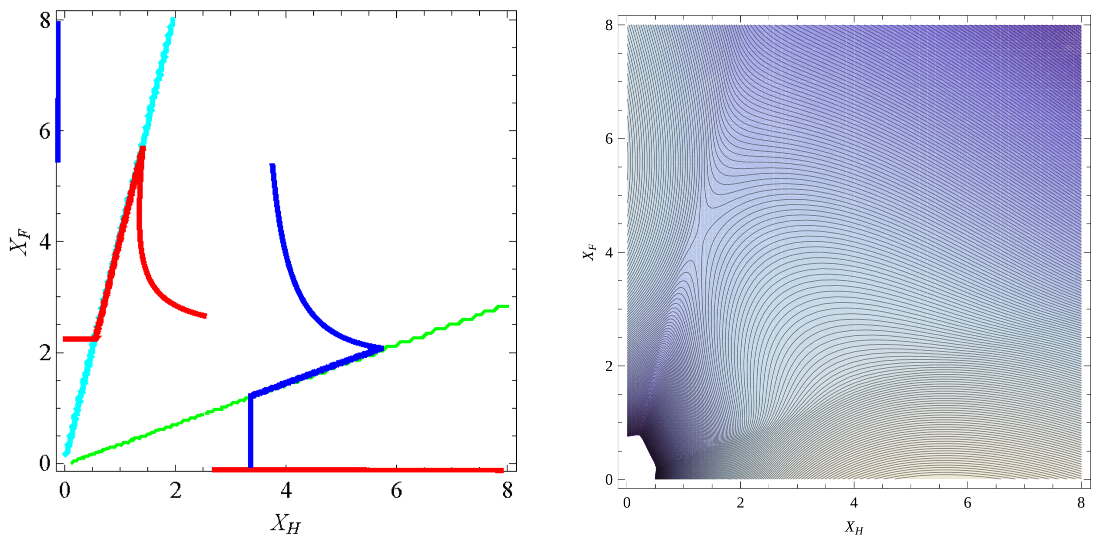


	X_H	X_F	$W_H + W_F$
Nash equilibrium	3.78931	2.34126	1.82446
Optimum	3.5168	0	2.43541

Figure 3.6 shows that the Nash equilibrium is in the interior, but full agglomeration in the larger region represents the optimum. The contour plot assigns the highest level of aggregate welfare to an infrastructure combination with $X_F = 0$. The exact figures are again given in the table. The Nash equilibrium is still characterized by an overprovision.

High trade freeness: $\phi = 0.15$

Figure 3.7: REACTION FUNCTIONS AND WELFARE AT HIGH TRADE FREENESS



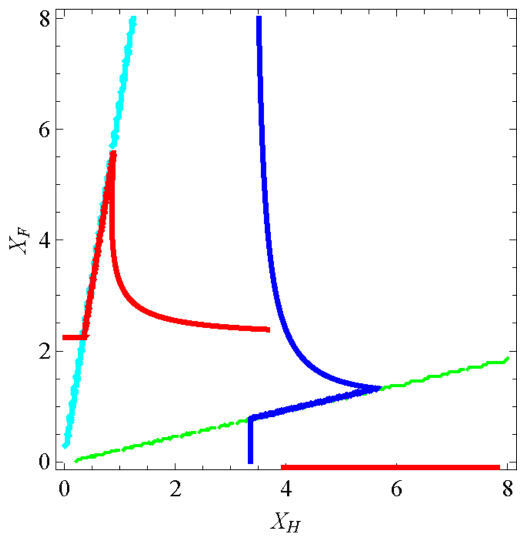
	X_H	X_F	$W_H + W_F$
Nash equilibrium	3.36	0	4.0439
Optimum	3.696	0	4.241722

For high trade freeness both the Nash equilibrium and the optimum are characterized by full agglomeration (see figure 3.7). Only region H will provide infrastructure. The contour plot assigns the highest level of aggregate welfare to an infrastructure combination with $X_F = 0$. The exact figures are again given in the table. The Nash equilibrium is characterized by underprovision, now.

Non-existence of a Nash equilibrium: $\phi = 0.095$

For a certain range of trade freeness no Nash equilibrium exists as the reactions curves in figure 3.8 show. The regions continuously over- and underbid each other.

Figure 3.8: NON EXISTENCE OF A NASH EQUILIBRIUM



Chapter 4

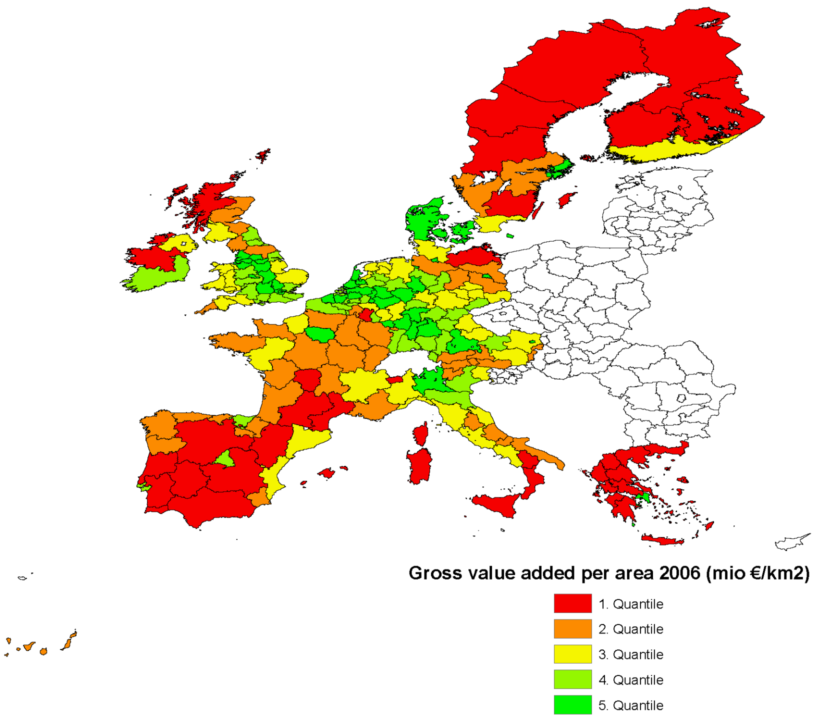
Infrastructure and the Location of Business

4.1 Introduction

This chapter provides an empirical application of the model derived in Chapter 3 using data on European regions. The effect of infrastructure on the location of business will be estimated and compared with the predictions of the theory. This is an important exercise as it allows us not only to infer the appropriateness of the theory but also gives valuable policy implications with regard to the federal provision of infrastructure. As is shown in figures 4.1 and 4.2, regional concentration is a dominant feature of economic activity in Europe. The maps illustrate the regional distribution of gross value added in the manufacturing and service sector weighted by the regions' sizes. The manufacturing sector is concentrated in northern Italy, western Germany, Denmark, the Benelux, central Britain and in the south of Ireland. The service sector's distribution is generally quite similar, but we observe differences at the coastal areas of Italy, France and Spain where the density of the service sector is quite high due to tourism while there is few manufacturing industry.

Although natural endowments like a location next to the seaside determine part of the distribution of services and manufacturing, most of the prevailing agglomeration cannot be explained by comparative advantages but originates from some source of agglomeration economies in manufacturing and service industries. As is shown in the

Figure 4.1: Regional Distribution of the Manufacturing Sector EU15

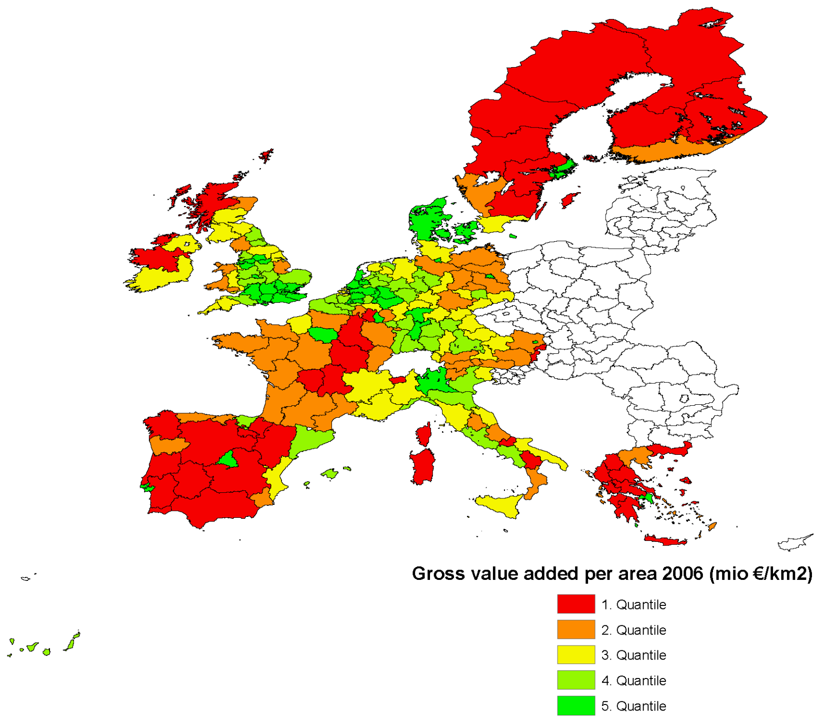


Data source: Cambridge Econometrics Regional Database.

preceding chapters, the New Economic Geography offers a theoretical approach to explain this phenomenon. A testable hypothesis, that distinguishes the New Economic Geography from the neoclassical framework, claims that improving inter-regional infrastructure may yield a higher concentration of economic activity. The preceding chapters explained why increasing the overall trade freeness (ϕ) in these models either leads continuously to a higher concentration of economic activity or suddenly creates a core-periphery pattern after exceeding a certain threshold level of trade freeness. For instance, a newly built highway which connects the economic center of two regions essentially affects the distribution of businesses between the two regions. Considering a simple two-region scenario, the region with the larger home market will benefit from the highway whereas the region with the smaller home market will lose firms to the core region.[1] Hence, the effect of inter-regional infrastructure investment depends critically on the endowments of the two regions that are affected by the investment.

[1]Note that in a neoclassical framework with decreasing returns, reducing the trade costs between a large region (with abundant capital supply and low marginal returns) and a small region (with little capital supply and high marginal returns) leads to a capital outflow from the large region to the small region.

Figure 4.2: Regional Distribution of the Service Sector EU15



Data source: Cambridge Econometrics Regional Database.

As opposed to inter-regional infrastructure which facilitates the movements of goods across the administrative boundaries of regions, intra-regional infrastructure affects only the productivity or the trade costs within the borders of a region. Hence, the latter kind of infrastructure clearly benefits the local share of economic activity. The following chapter analyzes whether we do observe such ambiguous effects of infrastructure provision and which sector fits best to the predictions of the New Economic Geography. First, we illustrate the effects infrastructure investments have in the theoretical model by deriving the relevant comparative statics. Second, we take the predictions to the data and see whether they are fulfilled.

Besides the question of whether the theory is true or not, it is important from a policy maker's point of view to know whether infrastructure affects the distribution of business at all and if this is the case, to what extent and in which way. In chapters 2 and 3 it is argued, that regional agglomeration usually is the most efficient allocation of economic activity in the New Economic Geography setting, because it allows the capture of positive agglomeration rents. Moreover, it was shown that federal governments have to intervene in the provision of infrastructure as there exist important externalities making the uncoordinated regional provision inefficient.[2] Yet, federal governments often face a political motive for reducing regional disparities and promoting the settlement of firms in lagging regions. Therefore, they use their authority over infrastructure provision for stimulating economic activity in lagging regions. The most important measure federal governments undertake in Europe in order to reduce regional disparities is the European Union's Structural Funds Program which will be topical in Chapter 5. These supranational transfers finance for the largest part regional infrastructure. About two thirds of the Structural Funds' total expenses is distributed to the so-called Objective 1 regions, which are regions lagging behind the economic development in Europe. Table 4.1 points to the significant difference between the amount of infrastructure investments in Objective 1 regions and regions which did not qualify for Objective 1 transfers. In each period, the amount of newly built and upgraded roads and railways is higher in the Structural Funds' Objective 1 regions than in the Non-Objective 1 regions. With the exception of railway growth in the first two programming periods, these differences are statistically highly significant. Hence, there is evidence that infrastructure is essentially allocated to lagging regions. The preceding chapter has argued why such a policy has to be challenged on efficiency grounds. The following chapter takes the political motive of reducing regional disparities as given and contribute to the question of whether such infrastructure investments are able to effectively reduce regional disparities at all. Moreover, it is analyzed which type of infrastructure performs best in accomplishing the federal goals.

We employ data on 204 EU15 NUTS2 regions over the period from 1988 to 2006 and estimate the effect of infrastructure on the regional share of gross value added in different sectors.[3] A challenge with this exercise is to distinguish inter- from intra-regional infrastructure. Simply considering the length of roads in the respective region will not serve this purpose, because some newly built roads will mainly reduce the trade costs within the region's borders while others significantly reduce the trade costs to neighboring regions. Therefore, we introduce a novel measure for cross-regional trade

[2]Chapter 3 shows that uncoordinated regional provision of infrastructure entails important externalities which imply overprovision of infrastructure if the trade freeness is low or underprovision if the trade freeness is high.

[3]EUROSTAT, the statistical office of the European Commission, subdivides Europe into three sub-national regional aggregates: NUTS1 (large regions with a population of 3-7 million inhabitants); NUTS2(groups of counties and unitary authorities with a population of 0.8-3 million inhabitants); and NUTS3 regions (counties of 150-800 thousand inhabitants). See Regulation (EC) No 1059/2003 of the European Parliament and of the Council of 26 May 2003 on the establishment of a common classification of territorial units for statistics (NUTS)

Table 4.1: INFRASTRUCTURE GROWTH AND EU STRUCTURAL FUNDS

	Objective 1 Regions (km)	Non-Objective 1 Regions (km)	Difference (km)	Std.Err. (km)
Roads				
1989-1993	15.702	4.108	11.594	1.562
1994-1999	10.572	6.259	4.314	1.286
2000-2006	27.835	10.070	17.765	1.649
Railways				
1989-1993	2.375	1.340	1.035	.813
1994-1999	.968	.701	.267	.332
2000-2006	14.772	9.265	5.507	1.777

Notes: Growth of roads and railways includes newly built and upgraded roads and railways. The sample consists of the EU12 NUTS3 regions for the period 1988-1993, the EU15 NUTS3 regions for the 1993-1999 period and the EU25 NUTS3 regions for the 2000-2006 period. We miss information on the four French overseas-departéments and the two autonomous Portuguese regions Madeira and Azores for all three periods. For the Dutch region Flevoland we miss information for the first period only.

freeness determined by the travel time between two regions. This will not only allow us to separate the effect of inter-regional from intra-regional infrastructure but will also reflect the quality of roads and railways as well as the traffic volume. It can be interpreted as a measure for effective infrastructure.

4.2 Literature

The related literature can be classified into two strands: a first strand which is examining the impact of public capital and in particular infrastructure capital on economic activity and a second strand which can be labeled the empirics of the New Economic Geography.

The first strand of literature has its starting point with Aschauer (1989b) who suggests that public capital had a positive impact on U.S. aggregate growth between 1949 and 1985. The paper shows that specifically core infrastructure, which is defined as expenses for nonmilitary structures such as highways, streets and water systems, plays an important role for productivity growth. A follow up paper (see Aschauer, 1989a) takes into account that public capital, in addition to stimulating private investment by increasing its productivity, may also crowd out private investment. It is shown that the net effect of public capital on private investment still remains clearly positive. This evidence is also confirmed by Munnell (1990) and Morrison and Schwartz (1996). However, other empirical work such as Evans and Karras (1994), Holtz-Eakin (1994) or Baltagi and Pinnoi (1995) come to much more pessimistic results regarding the productivity effects of public capital. After taking into account some econometric issues – for example measurement errors in the public capital stock or the inclusion of state and time specific effects – they show that the positive effect of public capital on productivity growth disappears.

From a theoretical point of view there is a clear perception that public infrastructure has a productivity enhancing effect, but it may be difficult to detect this effect empirically for several reasons. First, aggregate data often makes it difficult to disentangle single effects and to control for unobserved heterogeneity, second public investments exert cross regional spillover-effects which have to be controlled for, and third there are very different types of public spending which sometimes yield contrary effects. The first problem is addressed by a series of papers which estimate the effect of public infrastructure on the location and investment decision of firms using micro level data. Holl (2004a) and Holl (2004b) as well as Aiello et al. (2009) find robust positive effects of public infrastructure using micro level data for Spain, Portugal and Italy, respectively. The issue of regional spillovers is analyzed by Boarnet (1998) who shows that street and highway capital is significantly positive related with the output of Californian counties and that spillover effects play an important role. It is shown that infrastructure investment in one county draws away production from neighboring counties. Similar evidence is found for Germany by Borck et al. (2007). The first to distinguish empirically between inter- and intra-state public infrastructure investments are Cohen and Morrison Paul (2004). They suggest in a cost-based framework of the U.S. manufacturing sector that inter- and intra-state infrastructure are complements. In contrast, Martin and Rogers (1995) as well as Martin (1999) claim that inter-regional infrastructure provision has very different effects from intra-regional infrastructure. According to their argument, improvements in intra-regional infrastructure can reduce regional disparities, while inter-regional infrastructure investments promote further regional divergence. In the following we take up this argument and study theoretically as well as empirically whether it is true.

This brings us to the empirical literature on the New Economic Geography. We will illustrate in a simple New Economic Geography model that this discrepancy between the effects of inter- and intra-regional infrastructure investments is an implication which originates precisely from the New Economic Geography's central agglomeration force, namely the home-market effect.

Even though there has been a lot of theoretical work on the New Economic Geography in recent years, only a small number of previous studies have analyzed these theories empirically. On the one hand, this is due to the complex theoretical interrelations that are not easy to identify empirically and often predict unrealistic corner solutions where

all industry is located in one region. On the other hand, the limited availability of regional data prevents the implementation of many econometric methods. Probably the first to provide convincing evidence for the home-market effect were Davis and Weinstein (1999) and Davis and Weinstein (2003). The earlier paper investigates the role of economic geography effects in determining the production structure of Japanese regions. For eight of nineteen manufacturing sectors they find a significant impact of a variable which approximates the home-market effect. The later paper provides evidence that the home-market effect also operates on the national level (OECD countries). Redding and Venables (2004) show that the market potential is a significant and quantitatively important determinant of cross country variation in GDP per capita. They essentially test the wage equation of the New Economic Geography shown in Chapter 2 (see equation (26)) using per capita income for 101 countries in 1994. From an econometric perspective, the most reliable evidence for the home-market effect is provided by Redding and Sturm (2008). Their analysis uses the division of Germany after World War II as a natural experiment altering the market potential of West-German cities close to the border of the German Democratic Republic. The reduction in market potential makes these West-German cities located close to the border less attractive and significantly reduces their population growth rates.
The next section will derive three hypotheses of infrastructure provision using the model introduced in Chapter 3, which will subsequently be tested.

4.3 The Model

Considering the model derived in Chapter 3, the distribution of manufacturing industry is in the long run given by the condition that capital returns (R_H, R_F) have to be equalized across all regions. Solving for the share of manufacturing industry in region H (s_n) yields:

$$\begin{aligned} s_n &= 0 \quad \text{if } R_H - R_F\big|_{s_n=0} < 0 \Leftrightarrow \chi < \frac{\phi}{s_{pop} + (1-s_{pop})\phi^2}, \\ s_n &= 1 \quad \text{if } R_H - R_F\big|_{s_n=1} > 0 \Leftrightarrow \chi > \frac{s_{pop}\phi^2 + 1 - s_{pop}}{\phi}, \\ s_n &= \frac{s_{pop}}{(1-\phi\chi)} - \frac{\phi(1-s_{pop})}{(\chi-\phi)}, \quad \text{otherwise.} \end{aligned} \tag{1}$$

where s_{pop} denotes the population share in region H, ϕ the trade freeness between regions H and F and $\chi = \frac{X_H}{X_F}$ the ratio of intra-regional infrastructure in regions H and F. In order to test the model empirically, we restrain the analysis in the following to interior equilibria only. This means we require the asymmetries to be below a certain level which ensures that there are some firms operating in each region. In the following, we will analyze how different kinds of infrastructure affect the distribution of industry. We distinguish between inter-regional infrastructure which facilitates trade across regional borders (ϕ), and local infrastructure (X_i) which increases the productivity within the respective region. Moreover, we extend the model to the multi-region case, which gives important additional insights with respect to the implications of inter-regional infrastructure investment.
First, we take a look at the effect of intra-regional infrastructure on the share of manufacturing industry. Taking the partial derivative of (1) with respect to local infrastructure (X_H) yields:

$$\frac{\partial s_n}{\partial X_H} = \frac{s_{pop}\frac{\phi}{X_F}}{(1-\phi\chi)^2} + \frac{(1-s_{pop})\frac{\phi}{X_F}}{\chi - \phi} > 0. \tag{2}$$

Hence, other things equal, improving the quality of intra-regional infrastructure unambiguously increases the local share of industry.
Second, we consider the impact of inter-regional infrastructure. Such investments will reduce the trade costs between regions H and F and therefore increase the trade freeness (ϕ) between those two regions. Differentiating equation (1) with respect to the trade freeness shows the effect of inter-regional infrastructure investments on the share of manufacturing industry:

$$\frac{\partial s_n}{\partial \phi} = \frac{s_{pop}\chi}{(1-\phi\chi)^2} - \frac{(1-s_{pop})\chi}{(\chi-\phi)^2} \gtreqless 0. \tag{3}$$

Depending on the region's size (s_{pop}) as well as its relative quality of intra-regional infrastructure ($\chi = \frac{X_H}{X_F}$) the effect may be positive or negative. Assuming from now on that both regions are endowed with the same quality of local infrastructure, that is $\chi = 1$, the effect of an increase in inter-regional trade freeness becomes:

$$\frac{\partial s_n}{\partial \phi} = \frac{s_{pop} - (1-s_{pop})}{(1-\phi)^2} > 0 \text{ for } s_{pop} > 0.5. \tag{4}$$

Hence, the share of industry in region H (s_n) will only increase, if region H is larger than region F, that is $s_{pop} > 0.5$. In the case that region F is larger, an increase in trade freeness will result in a lower industry share in region H and a higher share in region F. The same thought experiment can be done for two regions that are of equal size (s_{pop}), but are endowed with different qualities of intra-regional infrastructure. In this case, a region only benefits from inter-regional infrastructure investment if it offers a higher quality of local infrastructure (X_i).

To summarize the scenario with two regions, intra-regional infrastructure unambiguously attracts firms, whereas improvements in the inter-regional infrastructure only increase a region's share of industry if the respective region is larger than its counterpart and has the better quality of local infrastructure. The reason for this finding is that regions with a larger home market and a higher productivity are more attractive for firms. If the trade freeness between the two regions is reduced, it becomes profitable for some firms to move to the larger and/or more productive region and serve the market in the smaller and/or less productive region by exports. A higher productivity reduces variable costs while the size of the market increases profits due to the home-market effect (see Chapter 2). This process takes place until, in equilibrium, profits between the two regions are again equalized.

So far we have only considered two regions, yet the presence of further competitors for the stock of firms may lead to additional effects. In a first step, we extend the model to Λ-regions which all face the same inter-regional trade freeness (ϕ). This means, shipping goods from region i to region j costs the same as shipping the goods from k to l for all $\{i, j, k, l\} \epsilon \Lambda$. Note also, that we keep the quality of local infrastructure identical in each region. In the long-run equilibrium, this leads to a share of industry in region i given by:[4]

$$s_n^i = \left[1 + \frac{\phi\Lambda}{1-\phi}\right]\left[s_{pop}^i - \frac{1}{\Lambda}\right] + \frac{1}{\Lambda}. \tag{5}$$

If region region i is exactly of average size, that is $s_{pop}^i = \overline{s}_{pop} = \frac{1}{\Lambda}\sum_{i=1}^{\Lambda}\frac{L_i+K_i}{L^W+K^W} = \frac{1}{\Lambda}$, the share of industry located in region i will be $\frac{1}{\Lambda}$. Moreover, an increase in inter-regional trade freeness affects the share of industry in region i as follows:

$$\frac{\partial s_n^i}{\partial \phi} = \frac{\Lambda s_{pop}^i - 1}{(\phi - 1)^2} > 0 \tag{6}$$

$$\text{for } \Lambda s_{pop}^i > 1 \iff L_i + K_i > \tfrac{L^W+K^W}{\Lambda}.$$

[4]See appendix A.1 for the derivation of the multi region equilibrium.

The implications with respect to inter-regional infrastructure provision are somewhat mitigated compared to the two region scenario. In the multi-region scenario, an increase in inter-regional trade freeness raises the local share of manufacturing industry, as long as the region is larger than the average region.

The assumption of trade freeness being equal between each pair of regions is quite restrictive. In a second step we relax this assumption, but for the sake of simplicity we assume that there are only $\Lambda = 3$ regions.[5] Note, that the trade freeness between two regions is still symmetric, that is, shipping a good from j to region k takes the same time and costs as shipping the good from region k to region j ($\phi_{jk} = \phi_{kj},\ \phi_{jl} = \phi_{lj},\ \phi_{kl} = \phi_{lk}$). On average, this is not a critical assumption. Suppose for now, that the trade costs between region k and l are higher than the trade costs between regions j and k as well as between j and l. The trade freeness between regions j and k is denoted by ϕ_{jk}, the trade freeness between regions j and l is ϕ_{jl} and the trade freeness between regions k and l is denoted by ϕ_{kl}. In order to keep the argument as simple as possible, let $\phi_{jk} = \phi_{jl} = \overline{\phi} > \underline{\phi} = \phi_{kl}$. Hence, the average accessibility of the regions is given by:

$$Accessibility\ \ region\ \ j = \overline{\phi}$$

$$Accessibility\ \ region\ \ k = \frac{\overline{\phi} + \underline{\phi}}{2}$$

$$Accessibility\ \ region\ \ l = \frac{\overline{\phi} + \underline{\phi}}{2}.$$

Since $\overline{\phi} > \underline{\phi}$, the accessibility of region j is the highest while region k and l have an equally lower accessibility. Calculating the equilibrium distributions of manufacturing industry for regions of identical size, that is $s^j_{pop} = s^k_{pop} = s^l_{pop}$, and local infrastructure $X_j = X_k = X_l$ yields:[6]

$$s^j_n = \frac{1}{3}\left(3 + \frac{2}{\underline{\phi} - 1} + \frac{2\underline{\phi}}{1 - 2\underline{\phi} + \overline{\phi}}\right) \tag{7}$$

$$s^k_n = s^l_n = \frac{1}{3}\left(\frac{1}{1 - \underline{\phi}} + \frac{\overline{\phi}}{2\underline{\phi} - 1 - \overline{\phi}}\right).$$

[5]The Λ region case is isomorphic but the resulting equilibrium expressions are much more complex.

[6]See appendix A.2 for the derivation of the three region equilibrium with different levels of trade freeness.

From equations (7) it can be easily seen that each region attracts one third of the total manufacturing industry, if we impose $\underline{\phi} = \overline{\phi}$. However, as long as $\underline{\phi} < \overline{\phi}$, region j will have an industry share greater than one third and regions k and l respectively an industry share smaller than one third. This is due to the term in brackets being larger than one for s_n^j, while it will be less than one for $s_n^k = s_n^l$. Intuitively, this states that with accessibility being above average a region will have a share of manufacturing industry which is above the average. Regions that are less accessible than the average, will have an industry share below the average. Moreover, improving the accessibility will lead to a more than proportional increase in the share of manufacturing industry. As an example, suppose $\underline{\phi} = 0.6$, while $\overline{\phi} = 0.65$. This means that region j is 1.04 times as accessible as regions k and l, which translates into industry shares $s_n^j \approx 0.54$ and $s_n^k = s_n^l \approx 0.23$. Thus, the share of manufacturing industry in region j is about 2.35 times the manufacturing industry share in regions k and l. The over-proportional increase takes place because accessibility leads, other things equal, to an increase in the relevant market of a firm which, via the home-market effect, increases firms' profits and yields to an inflow of new firms until the profits between the regions are equalized again. Hence, a region's overall accessibility represents an additional channel through which inter-regional infrastructure affects the local share of manufacturing industry.

Summing up the results derived from our theoretical model, we can state three implications with respect to infrastructure investment: First, intra-regional infrastructure unambiguously increases the local share of manufacturing industry. Second, inter-regional infrastructure leads to a higher accessibility and thereby raises the local share of manufacturing industry. Third, improving the inter-regional infrastructure to a region which is significantly larger or better accessible itself will have deteriorating effects on the local share of industry. Following Puga (2008), we call this potentially counterproductive effect of inter-regional infrastructure provision the hub-shadow effect. Investing in inter-regional infrastructure, through the first (accessibility) channel unambiguously benefits the local share of manufacturing, while through the second channel inter-regional infrastructure investments have ambiguous effects depending on the home-market sizes in the associated regions. It remains an empirical question whether the first or the second channel dominates when a small region improves the inter-regional infrastructure to a region with a larger home market.

4.4 Empirical Evidence

4.4.1 Empirical Specification

We base our empirical specification on a data set consisting of 204 EU15 NUTS2 regions. Using a higher level of aggregation, e.g. NUTS1 or national level, we would expect the forces working in the New Economic Geography to play a less important role. A higher level of aggregation would imply higher trade costs on average and therefore a lower level of trade freeness. As is shown in Chapter 2 as well as in Chapter 3, figure 3.1, the strength of the home-market effect increases with the overall trade freeness. Using a lower level of aggregation, e.g. NUTS3, we would expect from the same argument that the predictions of the model are even more applicable. However, it is difficult to argue why a federal government should aim at reducing disparities by infrastructure investment at the very low level of regional aggregation. The European Union's Structural Funds Program for example explicitly aims at reducing disparities at the NUTS2 level.

The model derived in Chapter 3 uses a quasi-linear utility function which neglects income effects. This was an uncritical and very helpful assumption to illustrate the theoretical results of public input competition in Chapter 3. Yet, the effects that are tested in the following do not require this assumption. Allowing for a standard Cobb-Douglas utility function with a CES subutility function over the different varieties changes the comparative statics above such that regional size is not simply reflected by its population share (s_{pop}) but by the regional share of expenditure (s_y). This is because, in equilibrium, consumers in regions with a large manufacturing sector and accordingly low trade costs for manufacturing goods may shift some expenditure from the agricultural to the manufactured good. Using the quasi-linear utility function, such a shift does not occur because a fixed share μ of the total income is spend on manufactured goods. For this reason, we employ the regional share of total EU household expenditure as an appropriate measure of the regional market-size.

Following Krugman (1991b), the New Economic Geography used to call the agglomerating sector the *manufacturing sector*. Yet, it is not clear why the predictions derived should fit the manufacturing sector better than, for instance, the service sector. Almost all firms face significant fixed costs and a consumer preference for differentiated varieties should not only be apparent for manufactured goods but in other sectors too. In fact, Fingleton and Fischer (2009) argue that the NACE-classes G to K, which are broadly defined as services, fit the monopolistic competition sector of the New Economic Geography best. Similarly, Abdel-Rahman and Fujita (1990) as well as Rivera-Batiz (1988) call the agglomerating sector the service sector. Considering the vagueness about which sector is essentially described, it is of great value to answer this open question by testing which sector fits the hypotheses derived from the model. In order to do so, we specify the NACE-classes DA-DK as the manufacturing industry sector and the classes G to K as the service sector.[7]

As a measure of the respective industry shares (s_n) we use the regional gross value added in the manufacturing/service sector divided by the sum of gross value in the manufacturing/service sector over all regions. An alternative measure would be the regional employment share of the respective sector, but we prefer gross value added over employment as it also reflects the local productivity which is directly affected by local infrastructure.

In order to account for intra-regional infrastructure, we use data on the growth of roads in the respective region. Since there is a higher potential for roads to be built in larger regions, we scale road growth by the regions' area. Hence, what we measure is the road growth per square kilometer.

Inter-regional infrastructure will be measured by the travel time between each pair of regions. In order to come as close as possible to the theoretical model, we calculate the inter-regional trade freeness between regions i and j by defining it as the inverse of the travel time between i and j ($T_{i,j}$):[8]

$$\phi_{ij} = \frac{1}{T_{ij}^{\eta}} \tag{8}$$

where η accounts for a convex distance decay function. We choose η such that the trade freeness decreases more than proportionally with travel time. This is realistic, as the travel time between two southern German cities, for instance, affects their industry share much more than the travel time to a town in northern Scandinavia.

We are particularly interested in the effects of intra-regional and inter-regional infrastructure, but according to the equilibrium industry distribution derived in equation (1) we should also control for the regional expenditure share. Since the classification of NUTS2 requires only to have a population somewhere between 800,000 and 3,000,000 inhabitants, there is significant variation with respect to the regions' population size

[7]NACE is the acronym for the statistical classification of economic activities in the European Community (Nomenclature statistique des activités économiques dans la Communauté européenne). See appendix A.3 for a list of the NACE-classes and its contents.

[8]See appendix A.4 for further details regarding the travel time variable.

which has to be controlled for. Moreover, congestion costs in the provision of public goods are very plausible, which implies that a higher population density goes in hand with a disproportionate increase in expenses for public services. This would cause a correlation between the stock of public capital and the population density. Due to those two issues, we decided to include the population density in the regression equation, in addition to the structural variables of the model.

Apparently, there are numerous other covariates which affect the regional share of the manufacturing or service sector as for example tax rates, institutions or natural endowments like a location next to a river or the seaside. Any of these covariates which are time-invariant are absorbed by the region fixed effects. Through the inclusion of time fixed effects a common time trend which could still bias the results is taken into account. Hence, the regression equation is given by:

$$g^k_{i,t} = \alpha_i + \nu_t + \boldsymbol{\gamma}\mathbf{X}_{i,t} + \delta_1 \sum_{j\neq i}^{\Lambda} \phi_{ij,t} + \delta_2 \sum_{j\neq i}^{\Lambda} \phi_{ij,t} exp_{j,t} + \delta_3 \sum_{j\neq i}^{\Lambda} \phi_{ij,t} exp_{j,t} \mathbb{1}_{exp_{j,t} > exp_{i,t}} + \mu_{i,t}$$

$$\text{where } t = \{1988, 1993, 1999, 2006\} \text{ and } k = \{manufacturing\ sector,\ service\ sector\} \quad (9)$$

where $g^k_{i,t}$ denotes region i's share of sector k at time t. The column vector $\mathbf{X}_{i,t}$ consists of the average yearly road growth over the previous years, the regional expenditure share and the population density.[9]

The sum over the trade freeness to all other regions in the sample ($\sum_{j\neq i}^{\Lambda} \phi_{ij,t}$) can be interpreted as the overall accessibility, which is not only determined by the geographic location of the region but also by the quality of inter-regional infrastructure. This variable will also reflect spillovers from infrastructure provision in other regions as $\phi_{ij,t}$ is not only affected by the infrastructure in regions i and j but also by each road located on the way from region i to region j.

Market access is measured by the trade freeness to all other regions weighted with their respective local expenditure. Intuitively, this is because it should make a difference whether one reduces the travel time to a region which is almost uninhabited or to a city which offers significant local expenditure. The latter increases the effective market access much more strongly than the former and will, according to the theory, benefit the local share of industry as long as the home market of the target region is not larger than the region's own home market. Hence, we would expect a significantly positive effect only where the expenditure of region i is greater than the expenditure

[9]The road growth variable does not only reflect newly built roads but also upgraded roads.

of target region j. This makes it necessary to include an interaction term which accounts exactly for the hub-shadow effect. It is given by $\sum_{j \neq i}^{\Lambda} \phi_{ij,t} exp_{j,t} \mathbb{1}_{exp_{j,t} > exp_{i,t}}$. The indicator ($\mathbb{1}_{exp_{j,t} > exp_{i,t}}$) equals one, if the expenditure of region i is smaller than region j's expenditure and is zero otherwise. As an example, suppose that region i invests in inter-regional infrastructure to become better connected to region j, which has a smaller home market than region i. In this case, the indicator is zero and the resulting effect is given by the change in trade freeness times δ_1 - for becoming more accessible - plus the change in ϕ_{ij} times δ_2 times the expenditure in region j - for gaining better access to the home market of region j. If, however, the home market in region j is larger than the home market in region i, the indicator becomes one and we have to add δ_3 times the change in ϕ_{ij} times the expenditure in region j to the effect described above.

In Section 4.3 we established the following hypotheses: first, regions with a high quality of local infrastructure accommodate a high share of manufacturing firms. Second, highly accessible regions, i.e. regions with a geographically central location and high quality of inter-regional transport connections, accommodate a high share of manufacturing firms and third, a region which is very close to another region with a larger home market hosts less manufacturing industry than a region with comparable inter-regional infrastructure and accessibility but without a dominant neighbor region. The last hypothesis reflects the hub-shadow effect. Regarding the empirical specification, the first hypothesis requires the coefficient of road growth per area to be positive, the second hypothesis requires δ_1 and δ_2 to be positive and the hub-shadow effect implies a negative coefficient δ_3.

4.4.2 Data and Descriptive Statistics

The data we use for estimating the specification stems from Cambridge Econometrics and the information about roads and travel times is GIS data which was provided by the Office for Regional Science, Planning and Geographic Information (RRG).

We convert the travel time matrix into a trade freeness matrix according to equation (8) and using $\eta = 1.5$. Comparing the parameters used in the literature (see Badinger and Tondl, 2003 as well as Fingleton, 2001) this value reflects a medium speed of distance decay. In Section 4.4.4, a sensitivity analysis with linear to high speed of distance decay will illustrate the robustness of our results. In order to allow a straightforward interpretation of the estimated coefficients, we normalize the variables measuring accessibility ($\sum_{j\neq i}^{\Lambda} \phi_{ij,t}$) and market access ($\sum_{j\neq i}^{\Lambda} \phi_{ij,t}exp_{j,t}$) by their means. Table 4.2 shows the descriptive statistics of our data set.

Table 4.2: DESCRIPTIVE STATISTICS

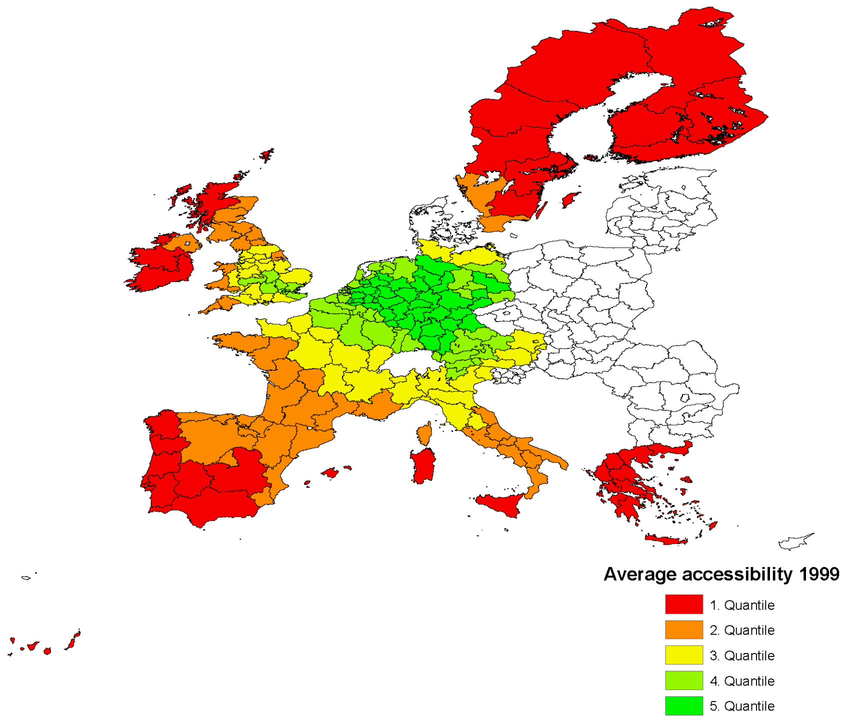
	Mean	Std. Dev.	Min	Max
	(1)	(2)	(3)	(4)
gva share service	.496	.617	.007	6.148
gva share manufacturing	.496	.546	.001	5.001
exp	.496	.467	.005	4.205
pop. density	.42	.982	.003	9.326
road	.455	.624	0	6.019
$\sum_{j\neq i}^{\Lambda} \phi_{ij}$	1	.72	.048	2.775
$\sum_{j\neq i}^{\Lambda} \phi_{ij} exp_j$	1	.747	.047	3.664
$\sum_{j\neq i}^{\Lambda} \phi_{ij,t} exp_{j,t} \mathbb{1}_{exp_{j,t}>exp_{i,t}}$	.762	.638	0	2.766

Notes: The regional shares of gross value added in the manufacturing and service sector as well as the regional share of household expenditure are measured in percent. By definition the mean of the service, industry, and expenditure share have to be equal. We define the gva share of sector k as $\frac{gva_i^k}{\sum_i^{\Lambda} gva_i^k}100$. Hence, the mean over all i is given by $\frac{1}{\Lambda}\sum_i^{\Lambda}\left(\frac{gva_i^k}{\sum_i^{\Lambda} gva_i^k}100\right) = \frac{100}{\Lambda}$, with 195 observations in the first time period and 204 over the last three time periods. This gives us $\frac{100}{4}(\frac{3}{204} + \frac{1}{195}) \approx 0.496$.

The regional share of gross value added in the manufacturing and service sector as well as the regional share of household expenditure are measured in percent. By definition, the mean over the shares is given by a constant which is approximately 0.5%. The population density is measured as population (1000s) per area (km^2). As mentioned before, road growth over the previous years in percent is scaled by area. The variable for geographic accessibility and market access are both normalized such that the means have to be one. However, the mean of the interaction between the expenditure weighted accessibility and the indicator matrix, which is equal to one only if the expenditure in region i is smaller than in region j, is obviously not equal to one, since it reflects only a part of the whole expenditure weighted accessibility, namely the part which describes the trade freeness to larger markets.

Figure 4.3 illustrates the expenditure weighted accessibility for the year 1999. It indicates a very clear core-periphery pattern which is of course mainly due to the simple geographic arrangement. Scandinavia, Portugal or Southern Italy are at the outer edge of the European circle which of course gives them the lowest average accessibility. Still when drawing a circle around the centroid of the European map (the Benelux approximately represent the centroid), we observe variance in accessibility. For instance, northern Italy and northern England have a higher expenditure weighted accessibility than eastern Spain or southern Scandinavia even though their distance from the centroid is almost the same. It is not only the immutable geography, but also the quality of infrastructure and the expenditure shares which drive this variable.

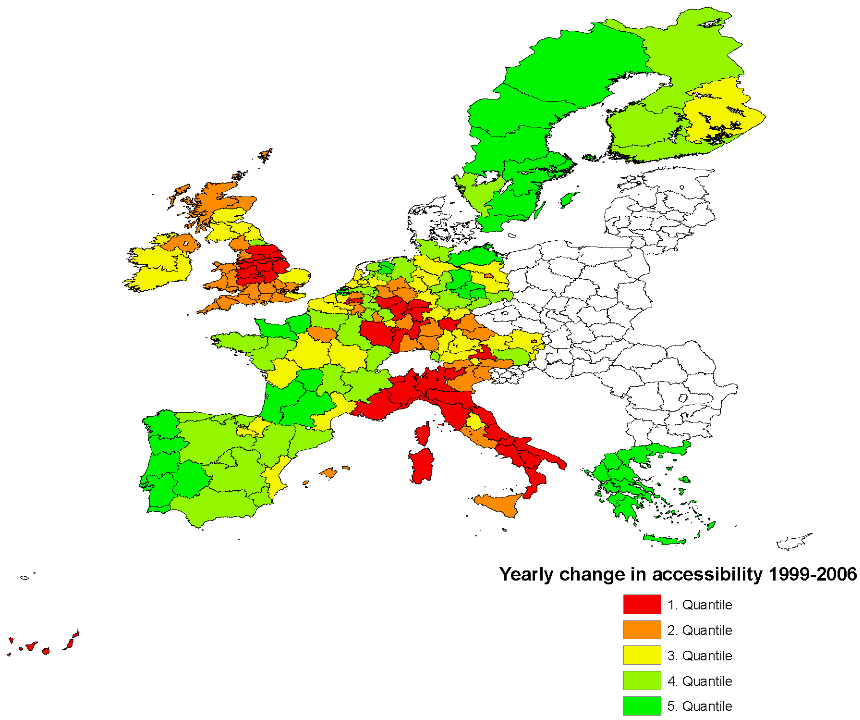
Figure 4.3: EXPENDITURE WEIGHTED ACCESSIBILITY



Note: Due to a change in the regional classification we miss compatible travel time data for Denmark. Data source: RGG and Cambridge Econometrics Regional Database.

More important for our analysis are the changes in the accessibility over time. Since we include region fixed effects in our specification (see regression equation (9)), the effects we estimate are identified from changes over time. Everything that is time-invariant – as for example the immutable geography – will be absorbed by the region fixed effects. Figure 4.4 depicts the changes in accessibility over the time from 1999-2006. The depicted accessibility measure – for the purpose of focusing on pure infrastructure changes – is not expenditure weighted but reflects the raw trade freeness. Interestingly, we observe the highest growth in accessibility in the regions at the outermost points of the European map. Portugal, Greece, Scandinavia and Eastern Germany improved their inter-regional infrastructure the most, that is they were in the highest quantile of accessibility growth. This is striking as they were rather in the lower quantiles when looking at the accessibility level in figure 4.3. Of course it is easier to improve the

Figure 4.4: Change in Accessibility



Note: Due to a change in the regional classification we miss compatible travel time data for Denmark. Data source: RGG.

accessibility from a low level of average trade freeness but still the question remains whether this prominent growth in accessibility at the outer part of Europe will relocate economic activity from the center to the investing regions. According to our theory it should have considerable effects.

4.4.3 Results

Data on the growth of roads, which includes newly built as well as upgraded roads is available from 1988-2006. We intend to use lagged values of road growth in order to eliminate the problem of reverse causality. This constrains us to estimating regression equation (9) only for the years 1993-2006, where the share of gross value added in the year 1993 is regressed on the average of newly built and upgraded roads between 1988-1993, the share of gross value added in 1999 on the average yearly road growth between 1993-1999 and lastly the share of gross value added in the year 2006 on the growth rate of roads between 1999-2006. The effects of inter-regional infrastructure can be tested for the whole sample, therefore each table includes a specification which leaves out our measure for intra-regional infrastructure in favor of using the full time period from 1988-2006.

Table 4.3: Manufacturing Share and Infrastructure

	Dependent var.: Regional share of EU15 GVA $\eta = 1.5$				
	(1)	(2)	(3)	(4)	(5)
exp	.541 (.433)	.613 (.419)	.823 (.591)	.770 (.670)	1.015 (.479)**
pop. density	-.326 (.158)**	-.346 (.156)**	-.332 (.142)**	-.340 (.136)**	-.262 (.116)**
road	.014 (.008)*	.012 (.007)*	.015 (.007)**	.014 (.007)**	
$\sum_{j\neq i}^{\Lambda} \phi_{ij}$		.738 (.291)**	.742 (.295)**	.719 (.265)***	.469 (.233)**
$\sum_{j\neq i}^{\Lambda} \phi_{ij} exp_j$			-.421 (.382)	-.321 (.566)	-.386 (.434)
$\sum_{j\neq i}^{\Lambda} \phi_{ij} exp_j \mathbb{1}_{exp_{j,t} > exp_{i,t}}$				-.165 (.367)	-.028 (.294)
Const.	.347 (.173)**	-.432 (.285)	-.121 (.407)	-.041 (.294)	.031 (.242)
Obs.	606	606	606	606	807
R^2	.065	.079	.097	.099	.166

Notes: All regressions include time and state fixed effects. The standard errors of each specification are robust against heteroscedasticity. The sample consists of 195 observations in the first time period and 204 observations over the last three time periods. We miss information on the four French overseas-départements, the three Danish NUTS2 regions and the two autonomous Portuguese regions Madeira and Azores for all time periods. For the nine East-German NUTS2 we miss information only for the first time period.

Table 4.3 presents the results for the manufacturing sector. The first four specifications are constrained to the limited sample from 1993-2006, whereas column five uses the full time span. In each of the first four columns we find a significant positive effect of intra-regional infrastructure. The second specification adds the measure of geographic accessibility which turns out to enter significantly positive too. Regarding the market-access effect and the hub-shadow effect, we do not find significant coefficients in any of the specifications.

Table 4.3 can be summarized as follows: intra-regional infrastructure as well as accessibility affect the regional share of manufacturing industry significantly and positively. One percent growth in roads per square kilometer increases the regional share of manufacturing industry by about 0.015%. The average accessibility of regions manifests in about 0.7% of the regional manufacturing share or put differently, raising the geographic accessibility by 10% above the average results in an increase in the manufacturing share of about $1.1 * 0.7 = 0.77\%$. Yet, the market-access effect and the proximity to larger markets than the own home market seem to play no role for the manufacturing sector.

Table 4.4: SERVICE SHARE AND INFRASTRUCTURE

	Dependent var.: Regional share of EU15 GVA $\eta = 1.5$				
	(1)	(2)	(3)	(4)	(5)
exp	.946 (.127)***	.979 (.119)***	.897 (.148)***	.790 (.157)***	.780 (.102)***
pop. density	.320 (.165)*	.311 (.164)*	.306 (.159)*	.288 (.155)*	.190 (.122)
road	.008 (.004)**	.007 (.003)**	.007 (.003)**	.006 (.003)**	
$\sum_{j\neq i}^{\Lambda} \phi_{ij}$		.329 (.119)***	.328 (.119)***	.283 (.103)***	.232 (.108)**
$\sum_{j\neq i}^{\Lambda} \phi_{ij} exp_j$			.164 (.108)	.367 (.172)**	.311 (.135)**
$\sum_{j\neq i}^{\Lambda} \phi_{ij} exp_j \mathbb{1}_{exp_{j,t}>exp_{i,t}}$				-.334 (.124)***	-.339 (.105)***
Const.	-.099 (.076)	-.446 (.125)***	-.568 (.175)***	-.408 (.139)***	-.251 (.138)*
Obs.	606	606	606	606	807
R^2	.63	.637	.645	.66	.644

Notes: All regressions include time and state fixed effects. The standard errors of each specification are robust against heteroscedasticity. The sample consists of 195 observations in the first time period and 204 observations over the last three time periods. We miss information on the four French overseas-départements, the three Danish NUTS2 regions and the two autonomous Portuguese regions Madeira and Azores for all time periods. For the nine East-German NUTS2 we miss information only for the first time period.

The results for the service sector, which are presented in table 4.4, are much more in line with the theory. Again, the first four specifications use the limited sample as they include our measure for intra-regional infrastructure, whereas the last column utilizes the full sample. Just as in the manufacturing sector, the effects of intra-regional infrastructure and geographic accessibility are significantly positive. Yet, both effects are less pronounced. The coefficients for intra-regional infrastructure and geographic accessibility approximately halve in magnitude. In contrast to our results for the manufacturing sector, columns (4) and (5) suggest that the service sector exhibits a significant market-access effect as well as hub-shadow effect. The magnitudes of the coefficients for the market-access effect and the hub-shadow effect are remarkably similar.[10] This means that the positive increase in market access is neutralized by the hub-shadow effect if the increase in trade freeness is aimed at a region with a larger home market.

As how to interpret these results, imagine two examples: say a region improves the trade freeness to a region with a smaller home market to ten percent above the average;

[10]In column (4) the coefficient of the market-access effect is slightly higher than the coefficient of the hub-shadow effect, while in column (5) it is the inverse. From the 95% confidence interval we cannot reject the equality of the coefficients absolute values in any case.

this results in an increase in the regional service share of about $0.28 * 1.1 = 0.308\%$ due to the better geographic accessibility plus about $0.37 * 1.1 = 0.407\%$ times the industry share of the other region. However, if the region raises the trade freeness to a region with a larger home market, the total effect will be given by the sum of the two effects already mentioned minus $0.34 * 1.1 = 0.374\%$ times the expenditure share of the other region. Hence, perfectly in line with the theory, the hub-shadow effect almost completely cancels out the market-access effect. Still a positive effect from inter-regional infrastructure investment remains through the accessibility channel.
Summarizing table 4.4, we can state that the service sector fits all hypotheses from the model. Intra-regional infrastructure turns out to significantly increase the regional share of services, just as the degree of geographic accessibility and the degree of market access do. Moreover, the hub-shadow effect appears to be highly significant and its magnitude relative to the coefficient of the market-access effect is consistent with the theory.

4.4.4 Sensitivity

The choice of the distance decay function ($\eta = 1.5$) in the tables presented above, provides some scope for discretion. A higher parameter entails a faster discounting of distant regions, or put differently, a change in the trade freeness to a distant region receives less weight than a change in trade freeness to a close region. The assumption per se is very reasonable, but the explicit size of the discount factor is not clear. Therefore we undertake a sensitivity check with linear up to strongly convex distance decay functions.

Table 4.5: Manufacturing Share and Infrastructure - Sensitivity

	$\eta = 1$	$\eta = 1.25$	$\eta = 1.5$	$\eta = 1.75$	$\eta = 2$
	(1)	(2)	(3)	(4)	(5)
exp	.681	.740	.770	.773	.755
	(.709)	(.694)	(.670)	(.639)	(.608)
pop. density	-.390	-.365	-.340	-.319	-.302
	(.153)**	(.144)**	(.136)**	(.129)**	(.125)**
road	.014	.014	.014	.015	.015
	(.007)*	(.007)*	(.007)**	(.007)**	(.008)**
$\sum_{j\neq i}^{\Lambda} \phi_{ij}$	1.334	.933	.719	.571	.437
	(.512)***	(.351)***	(.265)***	(.211)***	(.178)**
$\sum_{j\neq i}^{\Lambda} \phi_{ij} exp_j$	-.709	-.456	-.321	-.243	-.193
	(.856)	(.699)	(.566)	(.446)	(.341)
$\sum_{j\neq i}^{\Lambda} \phi_{ij} exp_j \mathbb{1}_{exp_{j,t} > exp_{i,t}}$	-.484	-.312	-.165	-.055	.013
	(.535)	(.439)	(.367)	(.303)	(.242)
Const.	.041	.018	-.041	-.064	-.028
	(.741)	(.444)	(.294)	(.232)	(.222)
Obs.	606	606	606	606	606
R^2	.114	.106	.099	.092	.087

Notes: All regressions include time and state fixed effects. The standard errors of each specification are robust against heteroscedasticity. The sample consists of 195 observations in the first time period and 204 observations over the last three time periods. We miss information on the four French overseas-départements, the three Danish NUTS2 regions and the two autonomous Portuguese regions Madeira and Azores for all time periods. For the nine East-German NUTS2 we miss information only for the first time period.

Tables 4.5 and 4.6 show the results for the manufacturing and the service sector, respectively. The first column of each table uses a linear distance decay function where $\eta = 1$ which increases up to $\eta = 2$ in the fifth column. The qualitative results turn out to be very robust with respect to the choice of η. In the manufacturing sector, both the market-access effect and the hub-shadow effect remain insignificant, while they are significant in the service sector. The magnitudes of the market-access effect and the hub-shadow effect in the service sector are such that they approximately neutralize each other independent of the chosen specification. That is, inter-regional infrastructure provision features a positive market-access effect only if the partner region has a smaller home market no matter what distance decay function is assumed. A pattern which can be observed from the sensitivity analysis with different distance decay functions is that the effects of inter-regional infrastructure become smaller and less significant the higher the parameter η. This indicates that even further distant regions exert important effects. When neglecting those distant regions to a large part, the overall effect loses power.

Table 4.6: SERVICE SHARE AND INFRASTRUCTURE - SENSITIVITY

	$\eta = 1$	$\eta = 1.25$	$\eta = 1.5$	$\eta = 1.75$	$\eta = 2$
	(1)	(2)	(3)	(4)	(5)
exp	.702	.746	.790	.831	.864
	(.165)***	(.161)***	(.157)***	(.154)***	(.152)***
pop. density	.286	.286	.288	.292	.297
	(.150)*	(.152)*	(.155)*	(.158)*	(.160)*
road	.006	.006	.006	.006	.006
	(.003)**	(.003)**	(.003)**	(.003)**	(.003)**
$\sum_{j \neq i}^{\Lambda} \phi_{ij}$	.422	.346	.283	.226	.169
	(.199)**	(.140)**	(.103)***	(.081)***	(.072)**
$\sum_{j \neq i}^{\Lambda} \phi_{ij} exp_j$	.513	.455	.367	.270	.184
	(.235)**	(.205)**	(.172)**	(.138)*	(.106)*
$\sum_{j \neq i}^{\Lambda} \phi_{ij} exp_j \mathbb{1}_{exp_{j,t} > exp_{i,t}}$	-.537	-.439	-.334	-.235	-.155
	(.173)***	(.149)***	(.124)***	(.102)**	(.081)*
Const.	-.490	-.454	-.408	-.353	-.288
	(.252)*	(.182)**	(.139)***	(.115)***	(.106)***
Obs.	606	606	606	606	606
R^2	.665	.663	.66	.655	.649

Notes: All regressions include time and state fixed effects. The standard errors of each specification are robust against heteroscedasticity. The sample consists of 195 observations in the first time period and 204 observations over the last three time periods. We miss information on the four French overseas-départements, the three Danish NUTS2 regions and the two autonomous Portuguese regions Madeira and Azores for all time periods. For the nine East-German NUTS2 we miss information only for the first time period.

4.5 Concluding Remarks

In this chapter we have shown that it makes a difference for the distribution of economic activity whether intra- or inter-regional infrastructure is provided. While the former uniformly stimulates the local share of economic activity, the latter is more effective in raising the local share of economic activity if the trade costs to a region with a smaller home market are reduced. As a result, regions with little local expenditure can be cast under the hub-shadow of a neighbor region with a major home market. This means they exhibit a lower share of economic activity than other regions of equal size and with similar accessibility but without a major home market in a neighborhood region. Due to the existence of trade and fixed costs for setting up a firm, profit-maximizing firms will decide to locate only in the hub-region with the larger home market and export their goods to the smaller neighbor market.
These theoretical findings are very important for regional policy. They show that the political goal of reducing regional disparities by providing infrastructure in lagging

regions, as it is part of the European Structural Policy as well as the German Regional Policy, will be reached more effectively by improving the intra-regional infrastructure than by improving the inter-regional infrastructure between the lagging and prosperous regions. Even though the latter policy benefits the regions by improving their average accessibility, the positive effect will be diminished by the hub-shadow effect. Another instrument which unambiguously meets the goal of reducing disparities is enlarging the home market in the lagging regions by providing direct income transfers.

The empirical analysis has shown that the service sector's distribution across European NUTS2 regions is well in line with the theoretical predictions. We find evidence for the unambiguously positive effect of intra-regional infrastructure as well as for inter-regional infrastructure improving the accessibility, altering the regional market access and featuring a significant hub-shadow effect. With respect to the manufacturing sector we could neither detect the market-access effect nor the hub-shadow effect, only the positive effect of average accessibility remains significant. This suggests that the standard New Economic Geography models essentially describe the service sector instead of the manufacturing sector. More adequate for explaining the distribution of the manufacturing sector appears to be a category of models which attributes the regional concentration of the manufacturing sector to vertical linkages between the firms (see Krugman, 1995, Baldwin et al., 2003 as well as Chapter 2). Manufacturing firms prefer to stay close to their intermediate goods suppliers and vice versa. This alters the variable which measures the relevant market size to being the demand for intermediate goods instead of the local household expenditure. All the theoretical predictions made with respect to infrastructure provision hold true in such a model, but in order to test the model we would need another measure of market size. Unfortunately, on a regional level it remains an open issue to find a variable which reliably approximates the local demand for intermediate goods.

Appendix

A1: Long-run equilibrium in the multi-region case

Considering Λ regions the market clearing condition for firm i in region j is given by:

$$M_i = m_i \widehat{p}_i (K_j + L_j) + \tau m_i \widehat{p}_i \sum_k^{\Lambda} (K_k + L_k), \ k \neq j. \tag{10}$$

Since all firms set the same price in each respective region, we refer to regions and drop the firm index i. The regional price index of manufactured goods in region j becomes in the multi-region case:

$$P_j = \left[\widehat{p}_j^{1-\sigma} n_j + \sum_k^{\Lambda} (\tau \widehat{p}_k)^{1-\sigma} n_k \right]^{\frac{1}{1-\sigma}}, \ k \neq j. \tag{11}$$

Since we refrain from asymmetries in intra-regional infrastructure, producer prices are the same in all regions. Using equation (10), (11) as well as equations (5) and (6) from Chapter 3 and the zero-profit condition, the short-run capital return in region j is:

$$R_j = \frac{\mu}{\sigma} \left(\frac{s_{pop}^j}{s_n^j + \phi(1 - s_n^j)} + \phi \sum_k^{\Lambda} \frac{s_{pop}^k}{s_n^k + \phi(1 - s_n^k)} \right) \frac{N}{n} \tag{12}$$

$$\text{where } \ k \neq j \ \text{ and } n = K^W, \ N = K^W + L^W.$$

In the long-run capital returns have to be equalized across all regions and will correspond to the average capital returns which, from equation (13) in Chapter 3 are given by $\overline{R} = \frac{\mu N}{\sigma n}$. Using this condition we can solve for the long-run distribution of manufacturing industry:

$$s_n^j = \left(s_{pop}^j - \frac{1}{\Lambda} \right) \left(1 + \frac{\phi \Lambda}{1 - \Lambda} \right) + \frac{1}{\Lambda}. \tag{13}$$

A2: Long-run equilibrium in the three-region case with different levels of trade freeness

In the case of three regions j, k, l with different trade costs between each pair of regions, the market clearing condition becomes:

$$M_i = m_i\widehat{p}_i(K_j + L_j) + \tau_{jk}m_i\widehat{p}_i(K_k + L_k) + \tau_{jl}m_i\widehat{p}_i(K_l + L_l)\sum_k^{\Lambda}(K_k + L_k) \quad (14)$$

$$\text{where } k \neq j \neq l.$$

Dropping again the firm index, we can write the regional price index of manufactured good in region j as:

$$P_j = \left[\widehat{p}_j^{1-\sigma}n_j + (\tau_{jk}\widehat{p}_k)^{1-\sigma}n_k + (\tau_{jl}\widehat{p}_l)^{1-\sigma}n_l\right]^{\frac{1}{1-\sigma}},\ k \neq j \neq l. \quad (15)$$

We assume symmetric trade costs such that $\tau_{jk} = \tau_{kj}$ for all $j \neq k$. Moreover, we impose $\phi_{jk} = \phi_{jl} = \overline{\phi} > \underline{\phi} = \phi_{kl}$. Using equations (14), (15) as well as equations (5) and (6) from Chapter 3 and the zero-profit condition, the short-run capital return in regions j, k and l are:

$$R_j = \frac{\mu}{\sigma}\left(\frac{s^j_{pop}}{s^j_n+\overline{\phi}s^k_n+\overline{\phi}s^l_n)} + \overline{\phi}\frac{s^k_{pop}}{s^k_n+\overline{\phi}s^j_n+\underline{\phi}s^l_n} + \overline{\phi}\frac{s^l_{pop}}{s^l_n+\overline{\phi}s^j_n+\underline{\phi}s^k_n}\right)\frac{N}{n} \quad (16)$$

$$R_k = \frac{\mu}{\sigma}\left(\frac{s^k_{pop}}{s^k_n+\overline{\phi}s^j_n+\underline{\phi}s^l_n)} + \overline{\phi}\frac{s^j_{pop}}{s^j_n+\overline{\phi}s^k_n+\overline{\phi}s^l_n} + \underline{\phi}\frac{s^l_{pop}}{s^l_n+\overline{\phi}s^j_n+\underline{\phi}s^k_n}\right)\frac{N}{n}$$

$$R_l = \frac{\mu}{\sigma}\left(\frac{s^l_{pop}}{s^l_n+\overline{\phi}s^j_n+\underline{\phi}s^k_n)} + \overline{\phi}\frac{s^j_{pop}}{s^j_n+\overline{\phi}s^k_n+\overline{\phi}s^l_n} + \underline{\phi}\frac{s^k_{pop}}{s^k_n+\overline{\phi}s^j_n+\underline{\phi}s^l_n}\right)\frac{N}{n}$$

$$\text{where } k \neq j \neq l \text{ and } n = K^W,\ N = K^W + L^W. \quad (17)$$

In the long run equilibrium $R_j = R_k = R_l$ has to hold which allows us to solve for the equilibrium distribution of manufacturing industry given in equation (7).

A3: NACE-classes

ESA95 code	Description	NACE code
DA	Manufacture of food products; beverages and tobacco	DA
DB	Manufacture of textiles and textile products	DB
DC	Manufacture of leather and leather products	DC
DD	Manufacture of wood and wood products	DD
DE	Manufacture of pulp, paper and paper products; publishing and printing	DE
DF	Manufacture of coke, refined petroleum products and nuclear fuel	DF
DG	Manufacture of chemicals, chemical products and man-made fibres	DG
DH	Manufacture of rubber and plastic products	DH
DI	Manufacture of other non-metallic mineral products	DI
DJ	Manufacture of basic metals and fabricated metal products	DJ
DK	Manufacture of machinery and equipment n.e.c	DK
GG	Wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods	G
HH	Hotels and restaurants	H
II	Transport, storage and communication	I
JJ	Financial intermediation	J
KK	Real estate, renting and business activities	K

A4: Deriving the inter-regional travel times on NUTS2 level

In order to calculate the travel times on NUTS2 level, we use a $N_3 \times N_3$ travel-time matrix $\mathbf{T}$, whose elements are defined as travel times between all pairs of NUTS3 regions (Note that there are $i_3 = 1, ..., N_3$ NUTS3 regions and $i_2 = 1, ..., N_2$ NUTS2 regions). Data on travel times is available for the years 1998, 1993, 1999, and 2006. The main diagonal elements of $\mathbf{T}$ are zero. The travel-time matrix $\mathbf{T}$ is translated into a $N_3 \times N_3$ trade-freeness matrix Φ_3 according to equation (8). Define a $N_3 \times N_2$ indicator matrix $\mathbf{Z}_{32} = [z_{i_3,i_2}]$, which indicates whether a NUTS3 region i_3 lies within the respective NUTS2 region i_2 or not, i.e. $z_{i_3,i_2} = 1$ if $i_3 \epsilon \, i_2$ and zero otherwise. Then the NUTS2 trade-freeness matrix Φ_2 is given by

$$\Phi_2 = [\mathbf{Z}_{32}^{'} \Phi_3 \mathbf{Z}_{32}] \circ \mathbf{W} \tag{18}$$

where $\circ$ denotes multiplication element-by-element and $\mathbf{W} = [w_{i_2}]$ is a vector which reflects the number of NUTS3 regions in the respective NUTS2 region. This weighting is necessary as the elements of Φ_2 would otherwise increase in the number of NUTS3 regions a NUTS2 region contains. The elements w_{i_2} of vector $\mathbf{W}$ are defined as $\frac{1}{N_3^{i_2}(N_3 - N_3^{i_2})}$ where $N_3^{i_2}$ denotes the number of NUTS3 regions in NUTS2 region i_2 and N_3 is the total number of NUTS3 regions. The $N_2 \times N_2$ matrix Φ_2 is calculated for the years 1998, 1993, 1999, and 2006 and consists of the elements $\phi_{ij,t}$ which are used in equation (9).

Chapter 5

The Effects of EU Structural Funds on Regional Performance*

5.1 Introduction

The preceding chapters have analyzed why regional concentration of economic activity evolves in the first place and whether this outcome is efficient. Moreover, the role of federal infrastructure investments in affecting the regional distribution of industry was critically evaluated. As pointed out in Chapter 3, agglomeration represents for sufficiently high trade freeness the efficient allocation of economic activity. Nevertheless, many federations – national or supra-national in scope – rely on regional transfer schemes which allow for transfers from rich core-regions to the lagging periphery. Examples of such federations are the German States (Länder) on the national level or the European Union (EU) on the supra-national level. The most important aim of the aforementioned transfers is to establish equalization – at least partially – of fiscal capacity and per-capita income among the participating jurisdictions (see Ma, 1997). In comparison to other federations, the magnitude of equalization transfers is particularly large within the EU. Moreover, the European regional transfer scheme is intended to be not only redistributive in terms of fiscal capacity and per-capita income but explicitly aims at attracting economic activity to peripheral regions.

This chapter analyzes whether the EU's regional policy was successful in equalizing per-capita income and implementing a more even distribution of economic activity. Moreover, it compares the achieved effects to its costs.

*This chapter is based on Becker, Egger, and von Ehrlich (2009)

The lion's share of fiscal equalization transfers at the level of the EU is spent under the auspices of the Structural Funds Program. Starting in 1988, this program distinguishes between transfers under three mutually exclusive schemes: Objective 1, Objective 2, and Objective 3.

The goal of this chapter is to assess the causal effect of Objective 1 status on per-capita GDP and employment growth of treated regions in the EU, using a host of micro-econometric techniques for program evaluation. The analysis sheds light on the effectiveness of the Objective 1 scheme (i.e., whether it causes treated regions to grow faster than control regions) and its net benefits (i.e., whether the growth induced justifies the costs incurred).

We confine our analysis to Objective 1 treatment for three reasons. First, Objective 1 funding has the explicit aim of fostering GDP-per-capita growth in regions that are lagging behind the EU average and of promoting aggregate growth in the EU (European Commission, 2001). Second, Objective 1 expenditures form the largest part of the overall Structural Funds Program budget. They account for more than two thirds of the program's total budget: 70% in the 1988-93 period, 68% in the 1994-99 period and 72% in the 2000-06 period (see European Commission, 1997, p. 154f., and European Commission, 2007, p. 202). Third, the Objective 1 scheme has been largely unchanged over all three programming periods of its existence.[1]

A region qualifies for Objective 1 transfers if its GDP per capita in purchasing power parity terms (PPP) is less than 75% of the EU average. For the programming periods 1989-93, 1994-99, and 2000-06, the European Commission computed the relevant threshold of GDP per capita in PPP terms based on the figures for the last three years of data available at the time when the Commission's regulations came out.

To further understand the Objective 1 scheme, it is useful to recall the classification system of regional units in the EU. Eurostat, the statistical office of the European Commission, distinguishes between three sub-national regional aggregates: NUTS1 (large regions with a population of 3-7 million inhabitants); NUTS2 (groups of counties

[1] Objective 2 covers regions that face socioeconomic problems which are mainly defined by high unemployment rates. More precisely, regions must satisfy three criteria to be eligible for Objective 2 transfers: first, an unemployment rate above the Community average; second, a higher percentage of jobs in the industrial sector than the Community average; and, third, a decline in industrial employment. Objective 3 deals with the promotion of human capital. The main goal is the support of the adaptation and modernization of education, training and employment policies in regions. Objectives 2 and 3 were modified slightly over the programming periods considered here. In 1989-93 and 1994-99 three additional objectives of minor importance existed which were abolished in 2000-06. For the new programming period 2007-13 the three objectives have been renamed Convergence Objective, Regional Competitiveness and Employment Objective, and European Territorial Co-operation Objective.

and unitary authorities with a population of 0.8-3 million inhabitants); and NUTS3 regions (counties of 150-800 thousand inhabitants).[2]

With very few exceptions, transfer eligibility is determined at the NUTS2 level and in advance for a whole programming period of several years.[3] For instance, in the 1994-99 programming period, the European Commission provided Objective 1 transfers to 64 out of 215 NUTS2 regions in the EU15 area. A graphical illustration of the regions receiving Objective 1 funds ("treated regions") across the three most recent budgetary periods is provided in Figure 5.1.

The amounts that are paid are quite significant for the recipient regions. In the 1994-99 programming period the 64 Objective 1 treated NUTS2 regions received on average transfers in the order of 1.8 percent of their GDP (see European Commission, 1997, 2007). Table 5.1 provides information on the transfers' magnitude for the three most recent programming periods. The number of Objective 1 regions as well as the total amount of transfers steadily increased over the three programming periods. As a fraction of the average Objective 1 regions GDP the transfers remained relatively constant between 1.4 and 1.8 percent.

A number of questions relating to these expenses are of obvious interest to both policy makers and economists. To which extent do economic outcomes in the recipient regions actually respond to such re-distributional transfers? This calls for an evaluation of the overall (causal) impact of transfers. Moreover, one could ask about the net benefits of transfers: Does the response in economic outcomes in the treated regions justify the size of the program and, in particular, its costs to the untreated jurisdictions? The related literature provides no convincing answer to these questions.

[2] NUTS is the acronym for *N*omenclature des *U*nités *T*erritoriales *S*tatistiques coined by Eurostat. The highest level of regional aggregation (NUTS1) corresponds to Germany's *Bundesländer*, France's *Zones d'Études et d'Aménagement du Territoire*, the United Kingdom's *Regions of England/Scotland/Wales* or Spain's *Grupos de Comunidades Autónomas.* At the other end of the NUTS classification scheme, NUTS3 regions correspond to *Landkreise* in Germany, to *Départements* in France, to *Unitary Authorities* in the UK or to *Comunidades Autónomas* in Spain.

[3] Owing to their territorial adjacency to Belgium's Objective 1 region Hainaut, the three French préfectures Valenciennes, Douai, and Avesnes (within the NUTS3 region Nord) received Objective 1 status in the 1994-99 programming period even though their NUTS2 mother region Nord-Pas-de-Calais did not qualify. The Austrian region Burgenland as the single Objective 1 region of the 1995 accession countries (Austria, Finland, and Sweden) did only receive Objective 1 funds from 1995 onwards. Similarly, the Objective 1 regions of the 2004 accession countries (Cyprus, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, and Slovak Republic, and Slovenia) did only receive funds from 2004 onwards.

Figure 5.1: OBJECTIVE 1 REGIONS

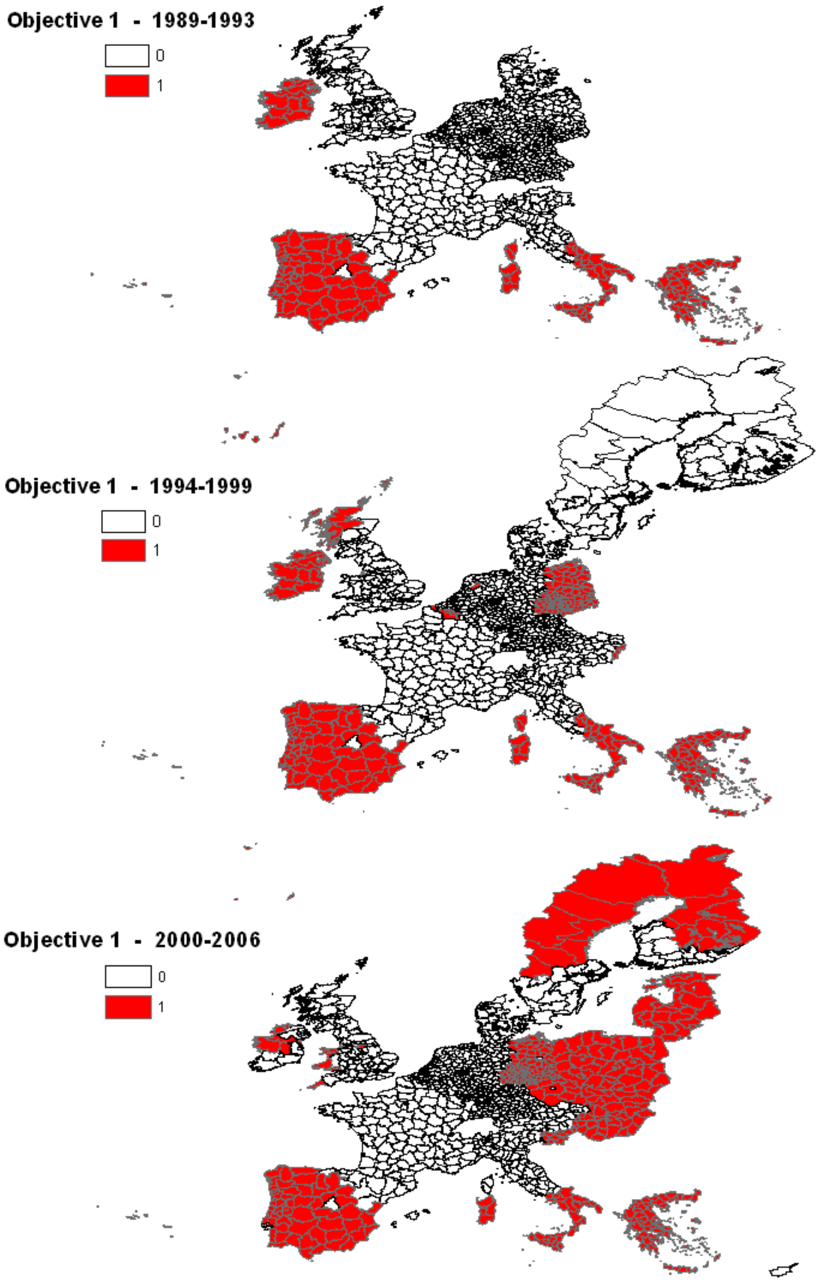


Table 5.1: Objective 1 regions

	1989-1993	1994-1999	2000-2006
	(1)	(2)	(3)
NUTS2			
Total Number of NUTS2 Regions	193	215	285
Number of Obj.1 NUTS2 Regions	58	64	129
NUTS3			
Total Number of NUTS3 Regions	1015	1091	1213
Number of Obj.1 NUTS3 Regions	286	309	417
Overall yearly funds (Mio. Euro)	8763.600	15661.670	23144.020
Overall yearly funds (Mio. Euro PPP)	10278.840	17479.010	26480.150
Yearly funds as fraction of Obj. 1 NUTS2 region GDP	.014	.018	.017
Yearly funds per inhabitant of Obj. 1 NUTS2 region (Euro PPP)	125	193	229

Notes: Data on EU Structural Funds stem from the European Commission (1997, pp. 154-155, and 2007, p. 202). To obtain average yearly funds we divide period-specific figures by the number of years the respective programming period lasted. We calculate the funds in PPP terms by weighting the funds each single country received in the respective programming period with the country's Purchasing Power Parity Index of the programming period's initial year. Funds per GDP and funds per inhabitant are calculated as the average yearly funds divided by regional GDP and regional population, respectively, prior to the programming period. This is 1988 and 1989 for the EU12 and the German New Länder, respectively, in the first period, 1993 for the EU12 regions in the second period but 1994 for the countries joining in 1995 (Austria, Finland, and Sweden), and 1999 for the EU15 in the third period but 2003 for the accession countries of 2004. Moreover, we adjust for the number of years the respective countries actually received funds. This is 5 years and 4 years for the EU12 and the German New Länder, respectively, in the first period, 6 years and 5 years for the EU12 and the new members of 1995, respectively, in the second period, and 7 years for the EU15 but 3 years for the new accession countries of 2004.

A small number of previous studies looked into the impact of re-distributional regional policies on economic outcomes (see Section 5.2 for a detailed discussion of the literature). Most of that research focused on the impact of the EU's Structural Funds Program. Yet, essentially all existing work on that topic uses fairly aggregated regional data at the NUTS1 level. Whereas some papers even used NUTS2-level data, they did not exploit important features of the design of the program. This might be problematic because, by design of the program, regions which are eligible for transfer payments under Objective 1 ("poor regions") differ systematically on average from non-eligible ones ("rich regions"). Furthermore, with regard to transfers under the auspices of the Structural Funds Program, most papers use cross-sectional data. Hence, the level of aggregation, the cross-sectional nature of the data employed, and the type of empirical methods applied in previous work rendered identification of the causal effect of the program difficult if not impossible.

We compile data on 285 NUTS2 and 1,213 NUTS3 regions in Europe for three programming periods – 1989-93, 1994-99, and 2000-06 – to assess the causal effect of transfers through the EU's Structural Funds Program on economic outcomes such as average annual growth of GDP per capita and employment growth of treated versus untreated regions. Ideally, in an experimental setting, we would randomly assign regions to a treatment and control group, i.e., give structural funds to some randomly selected regions and compare their economic outcomes to those of randomly selected control regions. While such an ideal experiment is not possible, the EU criteria for assigning Objective 1 status have quasi-experimental features. The 75% threshold at the NUTS2 level gives rise to a regression-discontinuity design (RDD) whereby regions very close to that threshold are likely to be very similar ex ante, but those below the 75% threshold qualify for Objective 1 funds, whereas those above do not. This rule is strictly applied in the vast majority of cases.[4]

In this chapter a *sharp* RDD (considering only compliers with the 75% rule) as well as a *fuzzy* RDD (considering compliers and non-compliers with the 75% rule) is pursued.[5] We analyze causal effects on growth of per-capita GDP at purchasing power parity (PPP) at NUTS2 level for most of the chapter, but also at NUTS3 level since part of the fuzziness in the design is brought about by exceptions below the NUTS2 level. Similarly, possible effects on employment are considered. Moreover, the analysis includes dose-response-function estimates of the amount of Objective 1 transfers at the NUTS3 level for the central programming period (1994-99) for which we have data on the actual level of Objective 1 transfers. The latter allows us to estimate the effects of a continuous-treatment intensity instead of a binary treatment. Finally, a back-of-the-envelope calculation of the net benefits of the program is presented.

Overall, we identify positive causal effects of Objective 1 treatment on the growth of per-capita income at PPP with both discrete and continuous treatments. The positive effect is qualitatively robust to estimation methods applied. In the preferred specification and procedure, we estimate a differential impact of Objective 1 program participation on the growth of GDP per capita at PPP of about 2.0 percentage points within the same programming period. No such effects can be found for employment. A back-of-the envelope calculation suggests that – on average – the funds spent on Objective 1 have a return which is about 1.21 times their direct costs in terms of GDP. Hence, the program seems effective and generates benefits in the recipient regions which exceed the direct costs to the EU budget. The results indicate that the EU finances productive investments such that the transfers are not a mere redistribution of

[4]For 628 out of the 674 NUTS2 observations across all periods, Objective 1 status complies with the formal 75%-rule. See below for further details on exceptions from the rule.

[5]Eligibility is synonymous with actual treatment under sharp RDD but not under fuzzy RDD.

income but exhibit multiplicative effects. However, the analysis does not find evidence for employment effects. Hence, it remains questionable whether firms are effectively attracted to lagging regions by Objective 1 transfers.

The chapter is organized as follows. The next section provides a discussion of the state of the literature on the evaluation of the EU's Structural Funds Program. Section 5.3 presents our data and shows descriptives on treated (i.e., Objective 1) and untreated (i.e., non-Objective 1) NUTS2 and NUTS3 regions. Section 5.4 summarizes the findings about the (causal) effects of Objective 1 treatment on the growth of GDP per capita when using the aforementioned quasi-experimental design. Section 5.5 provides sensitivity checks and extensions and delivers a back-of-the-envelope calculation of the net benefits of the European Union's Objective 1 Program based on the preferred estimates of the treatment effect. The last section concludes with a summary of the most important findings.

5.2 Effects of the Structural Funds Program: State of the Debate

The interest in effects of the EU's structural policy roots in empirical work on regional growth and convergence. Sala-i-Martin (1996) started the debate by diagnosing a failure of the EU's structural policy based on cross-sectional regressions showing that the regional growth and convergence pattern in the EU was not different from the one in other federations which lack such an extensive cohesion program. Obviously, such a conclusion requires comparability of federations and their regions in all other respects, which is not necessarily the case. However, Boldrin and Canova (2001) came to similar conclusions when focusing on regional growth within the EU and comparing recipient and non-recipient regions. Yet, both papers looked at the combined Structural Funds Program and not at the Objective 1 scheme, which primarily aims at closing the gap in per-capita income. Furthermore, they used fairly aggregated NUTS1 and NUTS2 data, since data at the NUTS3 level was not available at the time.

This evidence is different from the findings of Midelfart-Knarvik and Overman (2002) who identify a positive impact of the Structural Funds Program on industry location and agglomeration at the national level. [6] Similarly, Beugelsdijk and Eijffinger (2005) and Ederveen et al. (2006) took a national perspective and found a positive relationship between Structural Funds Program spending and GDP-per-capita growth (at least, in countries with favorable institutions). At the sub-national (NUTS1 or NUTS2) level, Cappelen et al. (2003) as well as Ederveen et al. (2002) detect a significant positive impact of structural funds on regional growth while Dall'erba and Le Gallo (2008) do not support this conclusion. Eggert et al. (2007) conclude that the Structural Funds support regional convergence in Germany but may exert a negative effect on aggregate growth.

However, as argued in the introduction, there is a number of potential problems with evaluations in earlier work which mostly relates to the limited availability of sufficient data in the cross-sectional as well as the time dimensions, and to the methods applied.[7] With much more data at hand now, we may revisit earlier conclusions and exploit information contained in the design of the program by means of a regression-discontinuity design.[8]

5.3 Data and Descriptive Statistics

5.3.1 Data Sources

For the empirical analysis, data from several sources are linked. The main outcome variable of interest is average annual growth of GDP per capita at purchasing power parity (PPP) during a programming period. In an extension, we look at average annual employment growth as an alternative outcome. Data for these variables at the

[6] However, they find that the funds seem to stimulate economic activity counter to the comparative advantage of the recipient countries.

[7] In many of the previous studies, the number of observations and, hence, the number of treated and untreated regions, is so small that it almost precludes the use of modern techniques for program evaluation, such as our regression-discontinuity design.

[8] A related approach of identifying causal effects of regional policy for one selected EU country is conducted in Criscuolo et al. (2007). They use micro level data on firms in the United Kingdom (UK) and employ a quasi-experimental framework to identify the causal effects of the UK's *Regional Selective Assistance* program on firm performance. They generate an instrument for recipient status of state aid by exploiting changes in the area-specific eligibility criteria. The eligibility criteria in the UK are determined by the European Commission's guidelines for regional development policies which also underly the Structural Funds Program. The revision of regional eligibility for structural funds before each programming period also determines the provision of *Regional Selective Assistance* to firms in the UK and may therefore be used as an exogenous instrument. The authors find a significant positive effect of state aid on investment as well as on employment.

NUTS2 and NUTS3 regional levels are taken from Cambridge Econometrics' Regional Database. Data on Objective 1 treatment in the Structural Funds Program at various levels of regional aggregation were collected from documents of the European Commission concerning structural funds.[9] In one of the sensitivity checks, we exploit information about the amount of funds paid instead of using a binary treatment indicator. The corresponding information has been provided by ESPON (European Spatial Planning Observation Network). However, data on funds paid at the regional level are only available for the programming period 1994-99.

In part of the analysis, we use information on sectoral employment,[10] population, and investment as control variables at the level of NUTS2 and NUTS3 regions from Cambridge Econometrics' Regional Database. Moreover, some of the sensitivity checks involve data on the geographical size and location of regions from the Geographic Information System of the European Commission (GISCO) to guard against a possible bias of the Objective 1 treatment effect associated with spillovers across regional borders. Finally, some of the empirical models in the sensitivity analysis involve a measure of countries' voting power in the EU Council (measured by the Shapley and Shubik, 1954 index). Those are taken from Felsenthal and Machover (1998) for the years until 2004 (for EU12 and EU15), and from Widgrén (2009) for the current voting scheme in the EU27 under the rules of the Treaty of Nice.

5.3.2 Descriptive Statistics

It is instructive to consider the variation in GDP per capita across NUTS2 jurisdictions in the EU because this gives us a picture of the prevailing disparities in Europe. This is done in Table 5.2 for the year 1999 (i.e., the year prior to the last available programming period, 2000-06).

[9] For each programming period, eligibility was determined by the European Commission one year in advance of the start of the programming period on the basis of the figures for the last three years of data that were available at the time. Concerning the first programming period 1989-93, see Council Regulation number 2052/88 and – regarding the New German Länder – see Official Journal L 114, 07/05/1991. The NUTS2 regions covered by Objective 1 in 1994-99 are listed in Council Regulation 2081/93 and – regarding the new member states Austria, Finland, and Sweden – in the Official Journal L 001, 01/01/1995. For the last programming period 2000-06, data stem from Council Regulation 502/1999 and – for the new member countries of 2004 – from the Official Journal L 236, 23/09/2003. All the regulations are available on the database for European Law, EUR-Lex.

[10] Sectoral employment is used to compute sector shares of agriculture and services to establish comparability of recipient and non-recipient regions in the instrumental variables estimation under fuzzy RDD.

Table 5.2: Disparities in the EU25 1999 (GDP per capita PPP)

	Country Avg. (Euro PPP)	Country Max (Euro PPP)	Country Min (Euro PPP)	Country Avg. rel. to EU25	Country Max rel. to EU25	Country Min rel. to EU25
Austria	18855.38	29546.84	13446.46	1.02	1.59	.72
Belgium	18466.26	43347.16	14331.10	.99	2.34	.77
Cyprus	14861.88	14861.88	14861.88	.80	.80	.80
Czech Republic	11411.80	23708.24	9554.07	.61	1.28	.51
Germany	19929.09	35739.29	12738.76	1.07	1.93	.69
Denmark	22634.88	27954.49	17869.64	1.22	1.51	.96
Estonia	6252.50	10644.65	4636.73	.34	.57	.25
Spain	16005.10	22823.61	11146.41	.86	1.23	.60
Finland	20302.39	28662.20	15392.66	1.09	1.54	.83
France	19790.04	32908.45	16100.37	1.07	1.77	.87
Greece	12530.61	16631.15	9377.14	.68	.90	.51
Hungary	8598.66	14861.88	6192.45	.46	.80	.33
Ireland	21651.46	24769.80	16454.23	1.17	1.33	.89
Italy	21184.88	29900.69	12915.68	1.14	1.61	.70
Lithuania	6243.72	9153.68	4171.41	.34	.49	.22
Luxembourg	40693.25	40693.25	40693.25	2.19	2.19	2.19
Latvia	5296.85	10829.71	3191.77	.29	.58	.17
Malta	14508.03	14508.03	14508.03	.78	.78	.78
Netherland	22107.05	29016.05	16808.08	1.19	1.56	.91
Poland	8382.42	13092.61	6015.52	.45	.71	.32
Portugal	13250.58	21408.19	12207.97	.71	1.15	.66
Sweden	19942.22	30431.47	18754.28	1.07	1.64	1.01
Slovenia	12438.66	19182.09	9761.78	.67	1.03	.53
Slovak Republic	8824.24	18931.21	6546.31	.48	1.02	.35
United Kingdom	19392.81	49362.68	12384.90	1.04	2.66	.67

Notes: The table shows average, maximum, and minimum GDP per capita at PPP within each country for NUTS2 regions.

The number of countries considered in the table is 25. Between 1986 and 1995, the EU consisted of 12 economies as included in the programming period 1989-93. Countries that joined the EU in 1995 (Austria, Finland, and Sweden) were included in the EU regulations for the programming period 1994-99. Similarly, the Eastern Enlargement of the European Union (in 2004) by 10 economies[11] was incorporated in the programming period 2000-06. Table 5.2 sheds light on the variation of GDP per capita across NUTS2 regions within individual countries and relative to the average GDP of EU25 countries, using data from the year 1999. The per-capita GDP of the poorest NUTS2 region located in Latvia is only 17 percent of the EU25 average while the richest NUTS2 region, *Inner London*, has a per-capita GDP of 2.66 times the average. Even within individual countries as for example Germany the disparities range from about 69 percent to 193 percent of the EU average.

We may summarize insights from that exercise as follows. There is considerable variation in GDP per capita both between and within EU countries. The former obviously

[11] Cyprus, Malta, and 8 Central and Eastern European countries: Czech Republic, Estonia, Hungary, Lithuania, Latvia, Poland, Slovenia, Slovak Republic.

strongly increased after the EU's enlargement in 2004. Some countries host NUTS2 regions above and below the 75% threshold.

According to the 75%-rule, all NUTS2 regions in a country would be eligible for Objective 1 transfers if the maximum GDP per capita across all regions were smaller than 75% of the EU25 average (see the fifth data column of Table 5.2). Suppose 1999 would have been the decisive year for the determination of Objective 1 eligibility for all regions in the EU25 countries. In this case, the Baltic countries (Estonia, Latvia, and Lithuania) as well as Poland would have been eligible in total.[12] Instead, none of the NUTS2 regions in a country would be eligible for Objective 1 transfers if the minimum GDP per capita in a region were higher than 75% of the EU25 average. This is the case for Luxembourg, Cyprus, and Malta (all of them cases of small countries consisting of only one NUTS2 region) as well as for Belgium, Denmark, Finland, France, Ireland, the Netherlands, and Sweden. However, the actual eligibility criterion applied for the 2000-06 programming period was somewhat different from that: the NUTS2 average GDP per capita over the years 1994-96 relative to the Community average was used for the EU15 countries while the average over the years 1997-99 was applied for the accession countries of 2004.

Table 5.3: Objective 1 recipient vs. non-recipient regions

	Mean recipient	Mean non-recipient	Difference col.(1)-col.(2)	Std. Err. of col.(3)
	(1)	(2)	(3)	(4)
1989-93 EU12				
GDP per capita 1988	8586.20	13634.19	-5047.99	478.23
No. of observations	52	134		
1994-99 EU15				
GDP per capita 1993	10795.99	16298.13	-5502.14	536.56
No. of observations	58	151		
2000-06 EU25				
GDP per capita 1999	11157.73	21251.68	-10093.94	556.27
No. of observations	123	156		

Notes: The table shows differences in GDP per capita (PPP) of recipient and non-recipient regions at the NUTS2 level. We miss information on the four French overseas-départements and the two autonomous Portuguese regions Madeira and Azores for all three programming periods. For the Dutch region Flevoland we miss information for the first period only. Regarding the East-German NUTS3 regions we calculated GDP per capita for the years 1989 and 1990 using information from the German Democratic Republic's statistical yearbook.

[12] Of course, actual Objective 1 transfer eligibility of the Baltic countries as well as Poland became only relevant after their EU membership in 2004.

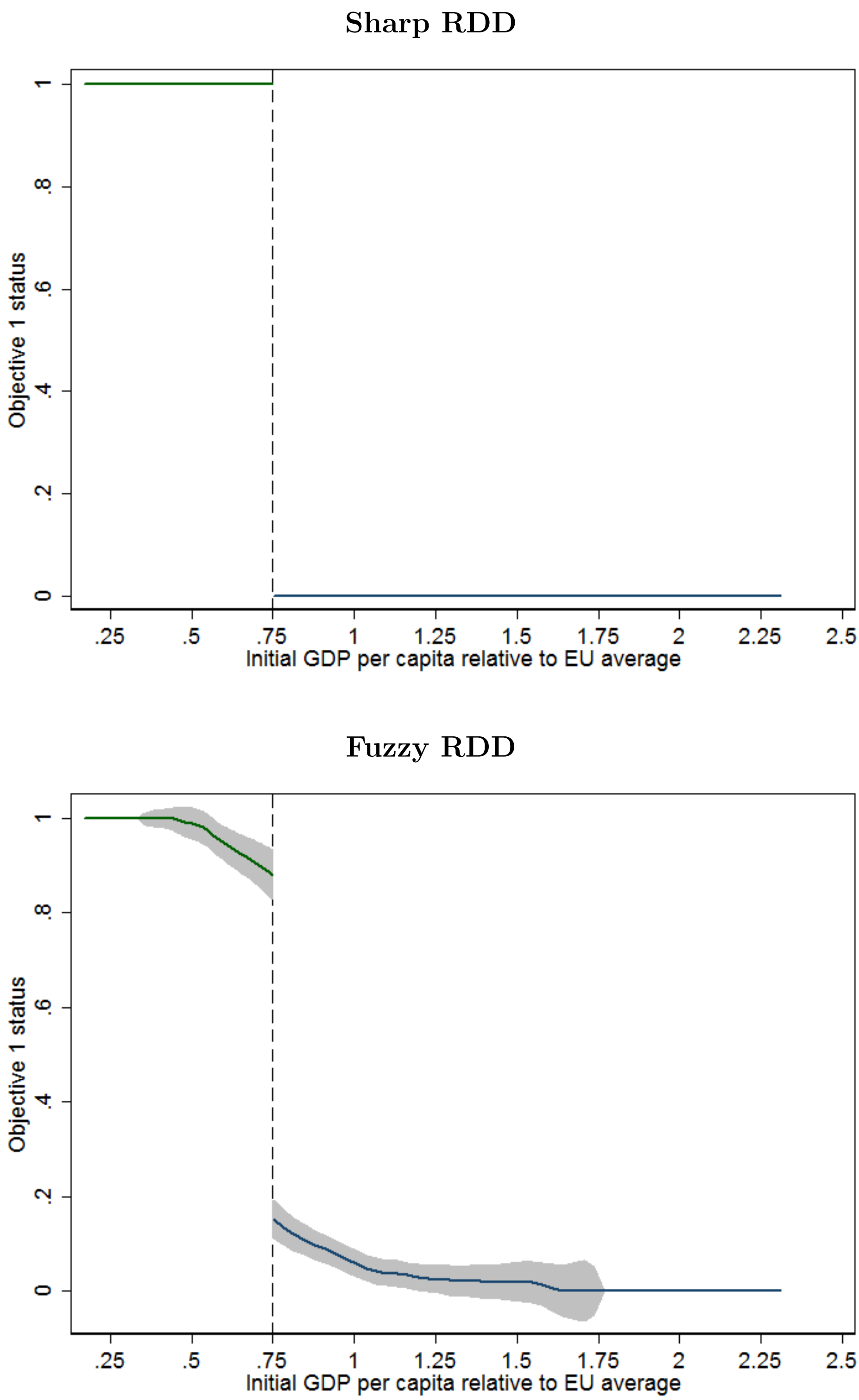
Since a region's initial GDP per capita is the only official criterium for Objective 1 status, Table 5.3 compares treated and non-treated regions with respect to the difference in their GDP per capita. The prime target of Objective 1 transfers is the reduction of this gap. Not surprisingly, the average difference in per-capita GDP between Objective 1 and non-Objective 1 regions in column 3 increases as further countries join the EU over the course of the three programming periods. In 1988, for the EU12, the average NUTS2 recipient region had a per-capita GDP that was 63 percent of the average non-recipient region. In 1999, for the EU25, the average recipient region had a per-capita GDP that was 53 percent of the average non-recipient region.

Given the ex ante differences between Objective 1 and non-Objective 1 regions, an unconditional comparison of their economic performance seems like comparing apples to oranges. The main problem is that real per-capita GDP determines not only the probability of Objective 1 treatment but – according to the convergence hypothesis – also the growth of real per-capita income (i.e., the outcome). Hence, the challenge is to disentangle the impact of initial levels of real per capita income on growth per se from the discontinuity related to and associated with Objective 1 at a level of per-capita income which is less than 75% of the EU average.

Figure 5.2 illustrates graphically how the probability of Objective 1 treatment relates to region-specific per-capita GDP at PPP prior to a programming period. There, we plot local polynomial functions of NUTS2-level real per-capita income prior to each programming period against Objective 1 treatment (a binary variable) of NUTS2 regions during that period, one to the left of the 75%-threshold and one to the right of it.

In the upper panel, a sharp regression-discontinuity design (see Cameron and Trivedi, 2005, Imbens and Lemieux, 2008, as well as Angrist and Pischke, 2009, for a general discussion of RDD) is enforced. For this, we remove regions where Objective 1 status does not comply with eligibility status: some of these regions got treated even though they were too rich to be eligible and others were not treated even though they were poor enough to be eligible (see Table 5.4 for the number of regions by programming period and compliance status). With the sharp RDD, all regions below the per-capita income threshold get Objective 1 treatment with probability one and all regions above the threshold get treatment with probability zero so that the local polynomial functions are straight horizontal lines.

Figure 5.2: Objective 1 status and the 75% GDP threshold



Note: The sharp RDD is based on those NUTS2 regions (628 out of 674) where eligibility according to the 75% rule coincides with actual Objective 1 status. The fuzzy RDD uses all NUTS2 regions, including those that constitute exceptions from the 75% rule. See main text for details.

Local polynomial smooth along with confidence intervals; based on Epanechnikov kernel with rule-of-thumb bandwidth.

Table 5.4: Eligibility and actual treatment under Objective 1 according to 75% GDP per capita threshold

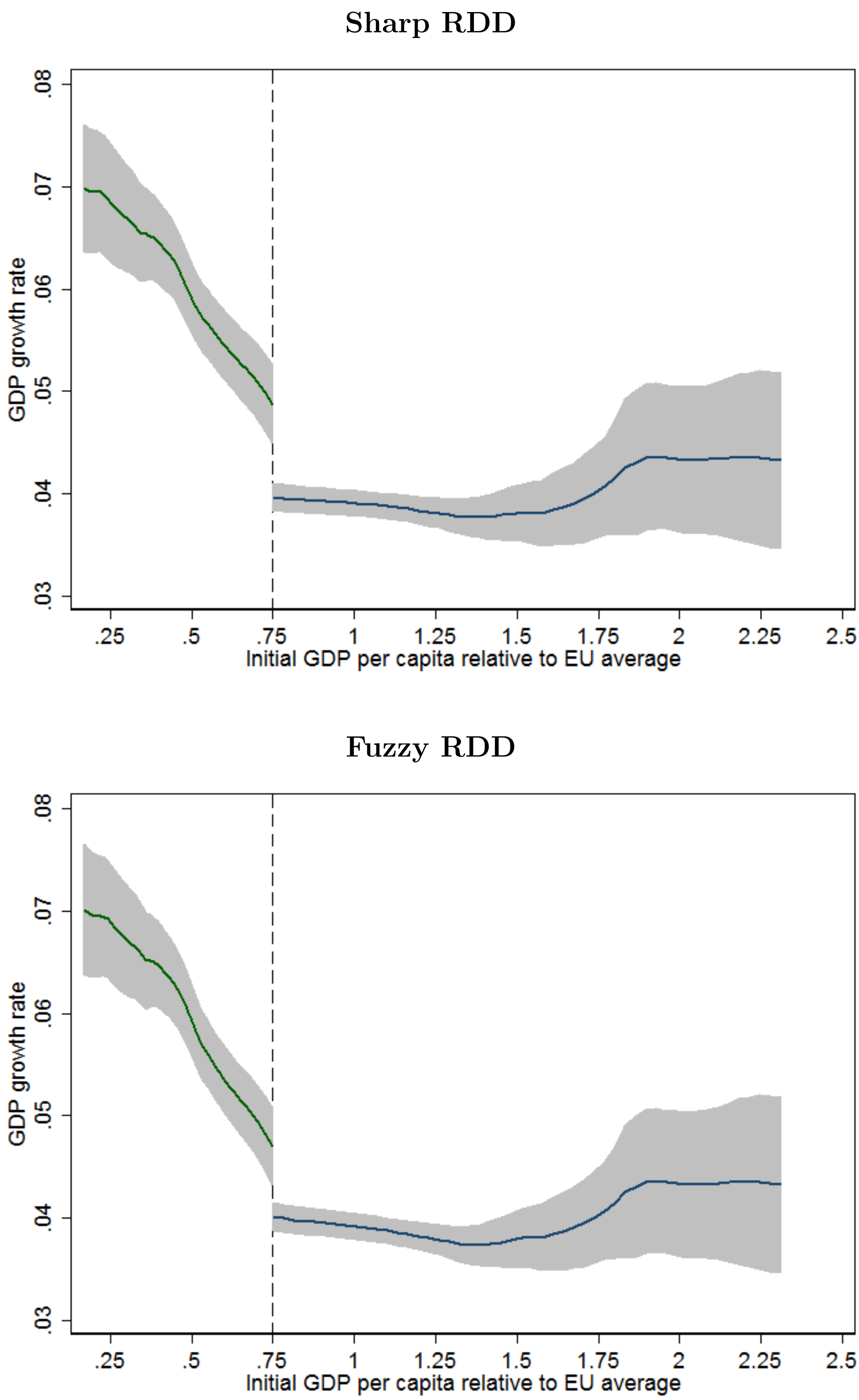
	Recipients NUTS2	Non-recipients NUTS2	Recipients NUTS3	Non-recipients NUTS3
1989-93 EU12				
Eligible	43	4	246	98
Non Eligible	9	130	34	631
1994-99 EU15				
Eligible	44	3	260	108
Non Eligible	14	148	43	674
2000-06 EU25				
Eligible	111	4	345	95
Non Eligible	12	152	66	701

Notes: Eligible regions are characterized by a GDP per capita of less than 75% of EU average in the qualifying years of each programming period (typically a 3-year average over the years preceding the start of a new programming period; see main text for details). Recipient regions are those that did effectively receive Objective 1 status. We miss information on the four French overseas-départements and the two autonomous Portuguese regions Madeira and Azores for all three programming periods. For the Dutch region Flevoland we miss information for the first period only. Regarding the East-German NUTS3 regions we calculated GDP per capita growth for the years 1989 and 1990 using information from the German Democratic Republic's statistical yearbook.

Table 5.4 provides information on eligibility and actual Objective 1 treatment for NUTS2 and NUTS3 regions. Across the three programming periods 1989-93, 1994-99, and 2000-06, the number of observations which conform to the sharp RDD is 628 (out of 674) for NUTS2 regions and 3,142 (out of 3,300) for NUTS3 regions. Only about 5% of all regions are thus exceptions from the rule.

The lower panel of figure 5.2 uses all NUTS2 regions, including exceptions from the 75% rule, which gives rise to a fuzzy regression-discontinuity design. The main feature of a fuzzy RDD is that the size of the discontinuity in treatment probability is smaller than unity. Graphically this means that the jump at the 75%-threshold in the lower panel of figure 5.2 is – in contrast to the upper panel (sharp RDD) – less than unity. As we approach the 75%-threshold from below, some regions that would be formally eligible do not obtain Objective 1 status. Hence, the probability of Objective 1 status is smaller than unity. For instance, the UK did not deliver GDP data at the NUTS2-level at the time Objective 1 status was determined in the programming period 1989-93. Only ex post, when the data became available, it turned out that some British NUTS2 regions should have been eligible for Objective 1 funds. As we approach the 75%-threshold from above, the probability of Objective 1 status exceeds zero, witnessing cases where governments negotiated exceptions from the 75% rule for regions which were too rich to be formally eligible.

Figure 5.3: Growth and the 75% GDP threshold



Note: The sharp RDD is based on those NUTS2 regions (628 out of 674) where eligibility according to the 75% rule coincides with actual Objective 1 status. The fuzzy RDD uses all NUTS2 regions, including those that constitute exceptions from the 75% rule. See main text for details.

Local polynomial smooth along with confidence intervals; based on Epanechnikov kernel with rule-of-thumb bandwidth.

To illustrate the effect of the discontinuity in Objective 1 treatment on economic outcomes, we plot local polynomial functions of per-capita GDP at PPP prior to each programming period against average annual growth of per-capita GDP at PPP during that period. Identification of a causal effect of Objective 1 treatment on growth by means of RDD requires that there is a discontinuity at the threshold in Figures 5.2 and 5.3.

The results in the figures are promising in that regard, since the discontinuity in Figure 5.2 is obvious and the confidence intervals of the local polynomial functions to the right and the left of the 75% per-capita income threshold in Figure 5.3 are non-overlapping. However, the lower panel of Figure 5.2 suggests that the design is fuzzy which requires instrumental variable estimation for consistent estimation of the Objective 1 treatment effect (see Wooldridge, 2002, and Angrist and Pischke, 2009). Accordingly, we proceed with regression analysis in the next section to identify the treatment effect by means of instrumental variables estimation.

Before we proceed to the regression analysis, Table 5.5 displays descriptive statistics of the variables entering in our regressions.

Table 5.5: Descriptive Statistics

	Mean	Std. Dev.	Min	Max
	(1)	(2)	(3)	(4)
GDP per capita growth (NUTS2)	.042	.018	-.008	.131
GDP per capita growth (NUTS3)	.041	.022	-.039	.251
Employment growth (NUTS2)	.005	.014	-.062	.079
Employment growth (NUTS3)	.005	.022	-.162	.273
Objective 1 (NUTS2)	.306	.461	0	1
Objective 1 (NUTS3)	.305	.46	0	1
Avg. GDP per capita threshold years	12927.27	4562.467	3343.816	37835.19
Employment share	.427	.119	.017	1.634
Agricultural share	.082	.102	0	.849
Service share	.613	.122	.031	.997
Population growth	.003	.008	-.058	.078
Population density	.492	1.065	.002	20.381
Shapley-Shubik index (v2)	.106	.04	0	.134

Notes: We miss information on the four French overseas-départements and the two autonomous Portuguese regions Madeira and Azores for all three programming periods. For the Dutch region Flevoland we miss information for the first period only. Regarding the East-German NUTS3 regions we calculated GDP per capita growth for the years 1989 and 1990 using information from the German Democratic Republic's statistical yearbook.

In particular, this table provides information on four moments of the distribution of all variables in use, namely the mean, the standard deviation, the minimum, and the maximum. The variables are GDP per capita growth in three programming periods at the NUTS2 level, GDP per capita growth in three programming periods at the NUTS3 level, employment growth in three programming periods at the NUTS2 level, employment growth in three programming periods at the NUTS3 level, the Objective 1 treatment indicator in three programming periods at the NUTS2 level, the Objective 1 treatment indicator in three programming periods at the NUTS3 level, average GDP per capita in threshold years (i.e., prior to the respective three programming periods), the employment share, the agricultural share (in total employment), the service share (in total employment), the population growth rate, the population density (in 1,000 inhabitants per square km), and the Shapley-Shubik index of voting power in the EU Council.

5.4 Regression Analysis

We seek to estimate the causal effect of Objective 1 status on regional economic performance by means of regression analysis. In the following regression models for sharp and fuzzy RDD will be employed. However, to begin with we introduce such models in formal accounts and then discuss the associated results with the data at hand.

5.4.1 The Regression-Discontinuity Design (RDD)

Think of a NUTS2 region A with a GDP per capita of 74.99% and a NUTS2 region B with a GDP per capita of 75.01%, one eligible, one not. These two regions are certainly more comparable than regions far away from the threshold. In our context, the reason is simply that – on average – regions which are far below the steady state of per-capita income grow faster than regions which are closer to the steady state. The crucial question is whether the discontinuity at the threshold associated with Objective 1 status is discernible from a polynomial function of per-capita income of reasonable order.

In case of a sharp RDD, one could stop at displaying the upper panel of Figure 5.3. Formally, a sharp RDD with data across regions and periods can be estimated by means of the regression of the following form (see Angrist and Pischke, 2009):

$$Growth_{it} = \theta + \rho Treat_{it} + \mathbf{x}_{it}\boldsymbol{\pi} + \lambda_i + \mu_{it}, \tag{1}$$

where $Growth_{it}$ represents average annual growth of region i's real per-capita income in PPP terms during programming period t, $Treat_{it}$ is a binary indicator variable for Objective 1 treatment which is unity in case of treatment of region i in programming period t and zero else, and $\mathbf{x}_{it}$ is a row-vector of control variables. In its purest form, $\mathbf{x}_{it}$ only consists of $f(GDP_{it}/capita_{it})$, a polynomial function of p^{th} order of per-capita income prior to programming period t (see above for details on the years which the European Commission considered to determine Objective 1 treatment status based on real per-capita income).[13] Furthermore, θ is a constant, ρ is the average treatment effect of Objective 1 treatment status, $\boldsymbol{\pi}$ represents a column vector of parameters on $\mathbf{x}_{it}$, λ_i is a region-specific effect that may be random or fixed,[14] and μ_{it} is a possibly heteroskedastic disturbance term.

The lower panel of Figure 5.2 illustrates that, with partial non-compliance, the 75%-rule gives rise to a fuzzy regression-discontinuity design (fuzzy RDD) that requires instrumental variables estimation.

With the fuzzy RDD, $Treat_{it}$ is not exogenous anymore in (1). Formally, $Treat_{it}$ in (1) may be instrumented by a first stage regression of the form

$$Treat_{it} = \alpha + \beta Rule_{it} + \mathbf{x}_{it}\boldsymbol{\gamma} + \epsilon_{it} \tag{2}$$

or

$$P(Treat_{it} = 1) = f(\delta + \zeta Rule_{it} + \mathbf{x}_{it}\boldsymbol{\eta} + \nu_{it}) \tag{3}$$

where i refers to a region and t to a programming period (1989-93; 1994-99; 2000-06). $Treat_{it}$ and $\mathbf{x}_{it}$ are defined as with the sharp RDD. $Rule_{it}$ is a binary indicator variable for Objective 1 eligibility which is unity in case of eligibility of region i in period t according to the 75% rule and zero else. α, β, δ, and ζ are unknown parameters, and $\boldsymbol{\gamma}$ and $\boldsymbol{\eta}$ are column vectors of unknown parameters. ϵ_{it} and ν_{it} are the disturbance terms of the two models. While some distributional assumption has to be made for

[13] The polynomial function is defined as $f(GDP_{it}/capita_{it}) = \pi_1(GDP_{it}/capita_{it})^1 + ... + \pi_p(GDP_{it}/capita_{it})^p$, where $\pi_1, ..., \pi_p$ are parameters to be estimated. In our empirical models, we will use fourth-order polynomials so that $p = 4$.

[14] With fixed λ_i, we may include region-specific indicator variables or, where more convenient, averages of all right-hand-side variables across the programming periods t as in Mundlak (1978).

Table 5.6: RDD NUTS2 - Objective 1 and GDP/capita growth

	sharp RDD		fuzzy RDD			
	Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
			linear 1st		nonlinear 1st	
	(1)	(2)	(3)	(4)	(5)	(6)
Objective 1	.016	.012	.018	.015	.019	.020
	(.002)***	(.005)**	(.002)***	(.004)***	(.002)***	(.004)***
Const.	.073	.033	.070	.090	.068	.088
	(.016)***	(.027)	(.016)***	(.023)***	(.016)***	(.023)***
Obs.	628	628	674	674	674	674
Regions	263	263	279	279	279	279
R^2	.220	.023	.172	.200	.166	.180

Notes: All regressions include a fourth-order polynomial of the average GDP per capita at the NUTS2 level over the years decisive for the eligibility threshold. In all pooled OLS regressions the standard errors are robust to clustering at the NUTS2 level. The first two columns report sharp RDD specifications while columns three to six show the fuzzy RDD specifications using a two-stage instrumental variables procedure. The first stage is estimated linearly in specifications (3) and (4) whereas specifications (5) and (6) build on a nonlinear first stage. In each case we run the pooled OLS regressions as well as the fixed effects regressions. The sample consists of the EU12 NUTS2 regions for the first period, the EU15 NUTS2 regions for the second period, and the EU25 NUTS2 regions for the third programming period. We miss information on the four French overseas-départements and the two autonomous Portuguese regions Madeira and Azores for all three periods. For the Dutch region Flevoland we miss information for the first period only. Regarding the East-German NUTS2 regions we calculated GDP per capita growth for the years 1989 and 1990 using information from the German Democratic Republic's statistical yearbook.

the nonlinear probability model, this is not the case for the linear probability model. However, ϵ_{it} will be generally heteroskedastic and we correct for this in the estimation of the variance-covariance matrix.

To identify the causal effect of Objective 1 status, $Treat_{it}$, on annual growth of real per-capita income of region i during programming period t, $Growth_{it}$, we need to instrument $Treat_{it}$ in the first stage either as in the linear probability model (2) or by $\widehat{P}(Treat_{it} = 1)$ obtained from (3). As long as we control for $\mathbf{x}_{it}$ in the second stage, either approach represents a just-identified two-stage model.

Tables 5.6 and 5.7 summarize our findings for six different models each. Three of the six models are estimated by pooled OLS and three of them include fixed effects at the respective regional level (NUTS2 in Table 5.6 and NUTS3 in Tables 5.7).[15] For both pooled OLS and fixed effects estimates, there are three models in each table: one which (inappropriately) assumes a sharp RDD with $Treat_{it}$ exogenous in the second-stage model in (1);[16] one which assumes a linear first stage model with fuzzy RDD;

[15]We always cluster standard errors at the level of NUTS2 regions.

[16]As explained above, under the assumption of a sharp RDD, we exclude observations that do not comply with the 75% per-capita income rule. There are 46 NUTS2 observations across the three programming periods which do not comply with the 75% per-capita income rule. Therefore, the number of observations is smaller in the models assuming a sharp RDD than in those assuming a fuzzy RDD.

Table 5.7: FUZZY RDD NUTS3 - OBJECTIVE 1 AND GDP/CAPITA GROWTH

	sharp RDD		fuzzy RDD			
	Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
			linear 1st		nonlinear 1st	
	(1)	(2)	(3)	(4)	(5)	(6)
Objective 1	.009	.011	.010	.009	.011	.017
	(.002)***	(.004)***	(.002)***	(.004)**	(.002)***	(.004)***
Const.	.111	.111	.109	.109	.106	.093
	(.023)***	(.014)***	(.023)***	(.029)***	(.023)***	(.028)***
Obs.	3142	3142	3300	3300	3300	3300
Regions	1150	1150	1207	1207	1207	1207
R^2	.158	.049	.144	.151	.143	.142

Notes: All regressions include a fourth-order polynomial of the average GDP per capita at the NUTS2 level over the years decisive for the eligibility threshold. The first two columns report sharp RDD specifications while columns three to six show the fuzzy RDD specifications using a two-stage instrumental variables procedure. The first stage is estimated linearly in specifications (3) and (4) whereas specifications (5) and (6) build on a nonlinear first stage. In each case we run the pooled OLS regressions as well as the fixed effects regressions. The standard errors are robust to clustering at the NUTS2 level. The sample consists of the EU12 NUTS3 regions for the first period, the EU15 NUTS3 regions for the second period, and the EU25 NUTS3 regions for the third programming period. We miss information on the four French overseas-départements and the two autonomous Portuguese regions Madeira and Azores for all three periods. For the Dutch region Flevoland we miss information for the first period only. Regarding the East-German NUTS3 regions we calculated GDP per capita growth for the years 1989 and 1990 using information from the German Democratic Republic's statistical yearbook.

and one which assumes a nonlinear (probit) first stage model with fuzzy RDD.[17] All of the models include a fourth-order polynomial function of real per-capita GDP at the level of the corresponding NUTS2 region in the relevant years prior to programming period t.[18]

The models in Table 5.8 do not only include the above-mentioned fourth-order polynomial function in $\mathbf{x}_{it}$ but also other covariates which vary across programming periods and countries (voting share in the EU Council as measured by the Shapley-Shubik index of the number of votes of a country in the respective programming period) or across programming periods and NUTS3 regions (Employment share in total population; Share of agriculture in total employment; Share of services in total employment; Population growth; Population density). These variables reflect the bargaining power of countries at the level of the European Union as well as structural characteristics, which should be associated with the probability that a NUTS3 region obtains Objective 1 transfers from the corresponding NUTS2 recipient region. Specifications (5) and (6) in Table 5.8 not only control for the NUTS3 level covariates but also include

[17]With nonlinear first stage models, we generally apply a Mundlak-Chamberlainian approach under fixed effects estimation, since a simple within-transformation of the data would be inappropriate (see Wooldridge, 1995). We estimate the nonlinear probability models for each programming period separately and then include time averages $\bar{\mathbf{x}}_i$ of the covariates $\mathbf{x}_{it}$ in the second stage model.

[18]Irrespective of whether we use NUTS2 or NUTS3 data, the polynomial function involves always NUTS2 real per-capita GDP. The reason is that the rule generally applies to NUTS2 regions.

Table 5.8: Fuzzy RDD NUTS3 with controls - Objective 1 and GDP/capita growth

	Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
	(1)	(2)	(3)	(4)	(5)	(6)
Objective 1	.012 (.002)***	.017 (.004)***	.010 (.002)***	.014 (.004)***	.015 (.003)***	.017 (.003)***
Empl. share	.007 (.006)	-.063 (.024)***	.008 (.006)	-.063 (.024)***	.009 (.006)	-.061 (.026)**
Agri. share	-.030 (.006)***	-.031 (.028)	-.040 (.006)***	-.028 (.028)	-.057 (.009)***	-.046 (.026)*
Serv. share	.007 (.005)	.054 (.017)***	.005 (.005)	.057 (.017)***	.002 (.005)	.050 (.017)***
Pop. growth	-.197 (.087)**	-.273 (.127)**	-.196 (.089)**	-.274 (.127)**	-.334 (.085)***	-.352 (.110)***
Pop. density	-.0002 (.0005)	.073 (.028)***	-.0001 (.0004)	.071 (.028)**	-.001 (.0006)**	.059 (.028)**
Voting power in EU council			-.077 (.014)***	.067 (.063)		
Country-time dummies	no	no	no	no	yes	yes
Const.	.100 (.020)***	.075 (.025)***	.110 (.020)***	.074 (.023)***	.122 (.019)***	.126 (.037)***
Obs.	3300	3300	3300	3300	3300	3300
Regions	1207	1207	1207	1207	1207	1207
R^2	.175	.202	.193	.226	.314	.331

Notes: All regressions include a fourth-order polynomial of the average GDP per capita at the NUTS2 level over the years decisive for the eligibility threshold. In each column we estimate fuzzy RDD specifications using a two-stage instrumental variables procedure where the first stage is estimated nonlinearly. In each case we run the pooled OLS regressions as well as the fixed effects regressions. The standard errors are robust to clustering at the NUTS2 level. The sample consists of the EU12 NUTS3 regions for the first period, the EU15 NUTS3 regions for the second period, and the EU25 NUTS3 regions for the third programming period. We miss information on the four French overseas-départements and the two autonomous Portuguese regions Madeira and Azores for all three periods. For the Dutch region Flevoland we miss information for the first period only. Regarding the East-German NUTS3 regions we calculated GDP per capita growth for the years 1989 and 1990 using information from the German Democratic Republic's statistical yearbook.

country-year dummies. The latter is supposed to capture any country-specific effect that might impact a country's bargaining power in any of the periods. It is thus even more general than the inclusion of the Shapley-Shubik of voting power in the EU Council which we use in columns (3) and (4). Considering the similar results for linear and nonlinear first stage specifications in Tables 5.6 and 5.7, we report only the results for the nonlinear specifications in Table 5.8.

Tables 5.6, 5.7, and 5.8 can be summarized as follows. First, Objective 1 treatment enters positively and is significantly different from zero in all models estimated. The difference between pooled OLS and fixed effects in the second stage is fairly small. The confidence intervals around the Objective 1 treatment indicators under pooled OLS and fixed effects estimation are overlapping with conventional confidence bands. This indicates that the bias from omitted time-invariant variables is small with the RDD in our application. The difference between the fuzzy RDD models with a linear versus a nonlinear probability model in the first stage is negligible. The relevant results of each of the respective models (1)-(6) in Tables 5.6 and 5.7 are not statistically different across tables. Including controls neither affects the significance nor the magnitude of the coefficients substantially. This indicates that the basic assumptions about RDD are met with the data at hand (i.e., it is enough to include a polynomial function $f(GDP_{it}/capita_{it})$ for consistent estimation of the treatment effect; see Angrist and Pischke, 2009). The latter is true also irrespective of whether we use NUTS2 controls or NUTS3 controls with outcomes at the NUTS3 level. To see this, compare the results from columns (5) and (6) in Table 5.8 with the corresponding columns in Tables 5.6 and 5.7. Overall, we identify a positive effect of Objective 1 treatment that is significantly different from zero and robust in various regards investigated in Tables 5.6, 5.7, and 5.8.

5.5 Sensitivity Checks, Evaluation, and Extensions

In the following we use the positive average treatment effects of Objective 1 transfers in models (5) and (6) reported in Tables 5.6 and 5.8 as our reference estimates. In this section, we proceed along several lines. First, we check the sensitivity of the estimates with regard to using an alternative approach to estimating binary treatment effects. Second, we check for and possibly avoid a potential bias of cross-border spillovers from Objective 1 treatment on regions in the control group. Third, we investigate possible effects on employment to shed light on a possible impact on an outcome other than income. Fourth, we assess the net benefits based on the preferred point estimates by means of a back-of-the envelope calculation. Finally, we exploit information about the extent of Objective 1 transfers in one of the three programming periods and provide dose-response estimates of the role of Objective 1 transfers on per-capita income growth.

Switching regression estimates of Objective 1 treatment effects To prove the robustness of our results with respect to the empirical estimation strategy, we may use a switching regression model as described in Wooldridge (2002, Procedure 18.4) instead of RDD.[19] In this model the treatment effect is identified from regions changing their Objective 1 status over the periods under consideration. Figure 5.4 illustrates the NUTS2 regions that either newly obtained Objective 1 status or lost Objective 1

[19]In fact, Wooldridge (2002, Procedure 18.4) describes the cross-sectional case. The extension to the case of panel data is similar to Wooldridge's (1995) panel estimator for the sample selection model.

status in the 1994-1999 and 2000-2006 programming periods, respectively. The model starts from the following system of equations:

$$Growth_{1,it} = \mathbf{x}_{it}\boldsymbol{\upsilon}_1 + \psi_{1,it} \tag{4}$$

$$Growth_{0,it} = \mathbf{x}_{it}\boldsymbol{\upsilon}_0 + \psi_{0,it} \tag{5}$$

where $Growth_{1,it}$ is the outcome (average annual growth during programming period t) in case of Objective 1 treatment and $Growth_{0,it}$ is the outcome in case of non-treatment. After stacking the two processes, the outcome $Growth_{it}$ can be written as

$$Growth_{it} = \rho_0 + (\rho_1 - \rho_0)Treat_{it} + \mathbf{x}_{it}\boldsymbol{\upsilon} + \chi_{it} \tag{6}$$

with $\chi_{it} = Treat_{it}\psi_{1,it} + (1 - Treat_{it})\psi_{0,it}$. If $E[\chi_{it}] \neq 0$, as is the case if eligibility and actual treatment do not coincide, an OLS estimate of equation (6) will give a biased estimate of the average treatment effect $\rho = \rho_1 - \rho_0$.

One can however compute two (period-specific) selection terms, one for selection into treatment and one for selection into non-treatment, and add those to the outcome equation. Wooldridge (2002, Procedure 18.4) proposes to estimate (period-specific) probits of $Treat_{it}$ on $Rule_{it}$ and $\mathbf{x}_{it}$ as in equation (3) above.[20] One then computes predicted probabilities, $\hat{\Phi}_{it}(\hat{\delta} + \hat{\zeta}Rule_{it} + \mathbf{x}_{it}\hat{\boldsymbol{\eta}})$, along with $\hat{\phi}_{it} = \phi(\hat{\delta} + \hat{\zeta}Rule_{it} + \mathbf{x}_{it}\hat{\boldsymbol{\eta}})$, where ϕ and Φ are the normal density and the normal cumulative distribution function, respectively.

We then run pooled OLS regressions of $Growth_{it}$ on $Treat_{it}$, $\mathbf{x}_{it}$, $Treat_{it}(\hat{\phi}_{it}/\hat{\Phi}_{it})$ and $(1-Treat_{it})[\hat{\phi}_{it}/(1-\hat{\Phi}_{it})]$ and – in the fixed effects version – the Mundlak-Chamberlain terms $\bar{\mathbf{x}}_i$.[21]

Table 5.9 displays the results from this estimation. In columns (1) and (2), results are shown for the sample of NUTS2 regions. The point estimates on the Objective 1 indicator are 0.016 and 0.014, respectively, and are very close to the results in columns (3)-(6) of Table 5.6. We also obtain estimates for the sample of NUTS3 regions in columns (3) and (4). In the covariate vector $\mathbf{x}_{it}$ we use the same variables as in columns (3) and (4) of Table 5.8. Again, the switching regression results are very similar to the

[20] As in Section 5.4.1, one may apply a Mundlak-Chamberlainian approach as in Wooldridge (1995), i.e., estimate the nonlinear probability models for each programming period separately and then include time averages $\bar{\mathbf{x}}_i$ of the covariates $\mathbf{x}_{it}$ in the second stage model in addition to $\mathbf{x}_{it}$.

[21] We have skipped the arguments of $\hat{\phi}_{it}$ and $(1 - \hat{\Phi}_{it})$ for the sake of brevity.

fuzzy RDD results.

In summary, the results from the switching regression model are very close to the RDD. This indicates that our results do not depend on the assumptions adopted in RDD as compared to switching regression analysis.

Figure 5.4: Regions changing Objective 1 status

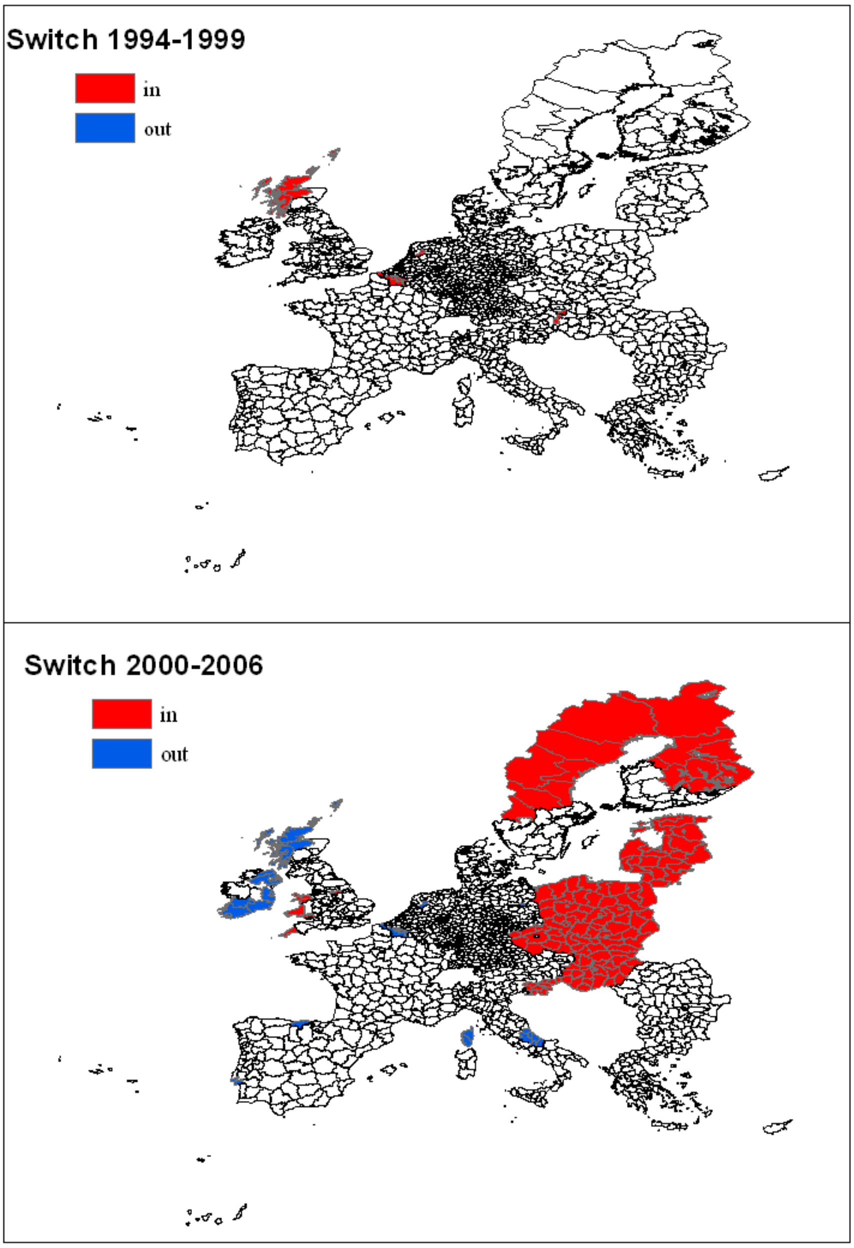


Table 5.9: SWITCHING REGRESSION MODELS - OBJECTIVE 1 AND GDP/CAPITA GROWTH

	NUTS2		NUTS3	
	Pooled OLS	FE	Pooled OLS	FE
	(1)	(2)	(3)	(4)
Objective 1	.016 (.002)***	.014 (.003)***	.013 (.002)***	.013 (.002)***
Empl. share			.008 (.004)**	-.059 (.020)***
Agri. share			-.055 (.006)***	-.040 (.024)
Serv. share			.004 (.004)	.050 (.021)**
Pop. growth			-.340 (.062)***	-.360 (.107)***
Pop. density			-.001 (.0005)***	.059 (.018)***
Country-time dummies			yes	yes
Selection correction terms[a]	yes	yes	yes	yes
Const.	.070 (.016)***	.089 (.023)***	.122 (.013)***	.114 (.029)***
Obs.	674	674	3300	3300
Regions	279	279	1207	1207
R^2	.215	.228	.324	.340

Notes: All regressions include a fourth-order polynomial of the average GDP per capita at the NUTS2 level over the years decisive for the eligibility threshold. In each column we estimate fuzzy RDD specifications using a two-stage instrumental variables (IV) procedure where the first stage is estimated nonlinearly. The first two columns report NUTS2 specifications while columns (3) and (4) show NUTS3 level regressions. In each case we run the pooled OLS regressions as well as the fixed effects regressions. The standard errors are robust to clustering at the NUTS2 level in the NUTS3 regressions as well as in the NUTS2 pooled OLS estimations. We take the sample of EU12 NUTS2 (and NUTS3) regions for the first period, EU15 NUTS2 (and NUTS3) regions for the second period, and EU25 NUTS2 (and NUTS3) regions for the third programming period. We miss information on the four French overseas-départements and the two autonomous Portuguese regions Madeira and Azores for all three periods. For the Dutch region Flevoland we miss information for the first period only. Regarding the East-German NUTS3 regions we calculated GDP per capita growth for the years 1989 and 1990 using information from the German Democratic Republic's statistical yearbook. [a] Following Wooldridge (2002, p.631), we compute inverse Mills ratios for treated and control regions as regressors in the second-stage model; see main text for details.

Avoiding a possible downward bias of the treatment effects from cross-border spillovers One concern with the estimates in Tables 5.6, 5.7 and 5.8 is that Objective 1 transfers may be used to finance public infrastructure, generating not only local effects on the treated regions but also spillovers to neighboring regions as was argued in Chapter 4. The latter would violate what is called the *stable unit*

Table 5.10: Spatial Exclusion Mechanism RDD NUTS2 - Objective 1 and GDP/capita growth

	100km	150km	200km
	(1)	(2)	(3)
Objective 1	.017	.020	.020
	(.003)***	(.004)***	(.004)***
Const.	.084	.084	.081
	(.023)***	(.022)***	(.022)***
Obs.	625	586	544
Regions	261	226	213
R^2	.188	.190	.192

Notes: All regressions include a fourth-order polynomial of the average GDP per capita at the NUTS2 level over the years decisive for the eligibility threshold. In each column, we estimate fuzzy RDD specifications using a two-stage instrumental variables procedure where the first stage is estimated nonlinearly. All specifications include region fixed effects. The standard errors are robust to clustering at the NUTS2 level. The distance matrices are used such that only those non-Objective 1 regions are included in the control group that are more than 100 km away from the next Objective 1 region in the first column, 150km in the second column, and 200km in the third column. The sample consists of the EU12 NUTS2 regions for the first period, the EU15 NUTS2 regions for the second period, and the EU25 NUTS2 regions for the third programming period. We miss information on the four French overseas-départements and the two autonomous Portuguese regions Madeira and Azores for all three periods. For the Dutch region Flevoland we miss information for the first period only. Regarding the East-German NUTS3 regions we calculated GDP per capita growth for the years 1989 and 1990 using information from the German Democratic Republic's statistical yearbook.

treatment value assumption and lead to downward-biased estimates of the average Objective 1 treatment effect, unless spillovers are captured by the polynomial function of NUTS2 per-capita income. The reason is that positive spillovers affect per-capita income growth of untreated (control) neighboring units positively which reduces the difference between growth rates of the treated and the untreated regions.

Provided that the spillovers are of medium reach (e.g., do not exceed a distance of 100-200 kilometers), such a bias can be avoided by the following procedure. Let us assume that Objective 1 transfers mainly affect regions within a radius of 100, 150, or 200 kilometers but not more distant ones. We may then exclude all untreated regions that are located within such a radius from any treated region. The corresponding estimates, especially the ones using a radius of 200 kilometers, would then be free of a bias from spillovers.[22]

Tables 5.10 and 5.11 summarize the results from fixed region effects estimation involving a nonlinear probability model in the first stage. The estimates are obtained in the restricted samples excluding untreated observations within the mentioned radius around treated observations. Obviously, the number of observations declines with a larger radius applied around Objective 1 treated regions. With a radius of 100 kilome-

[22] Admittedly, we ignore Objective 1-induced (second or higher order) spillover effects from untreated regions to other untreated regions, but these should be negligible with spillovers of reasonable magnitude.

Table 5.11: Spatial Exclusion Mechanism Fuzzy RDD NUTS3 - Objective 1 and GDP/capita growth

	100km	150km	200km
	(1)	(2)	(3)
Objective 1	.007	.019	.019
	(.003)***	(.004)***	(.004)***
Empl. share	-.067	-.042	-.037
	(.028)**	(.023)*	(.022)*
Agri. share	-.008	-.082	-.080
	(.020)	(.030)***	(.031)***
Serv. share	.037	.047	.050
	(.017)**	(.020)**	(.021)**
Pop. growth	-.388	-.424	-.443
	(.124)***	(.124)***	(.134)***
Pop. density	.003	.066	.076
	(.010)	(.035)*	(.036)**
Country-time dummies	yes	yes	yes
Const.	.005	.128	.127
	(.031)	(.041)***	(.042)***
Obs.	2861	2567	2420
Regions	1040	876	828
R^2	.362	.325	.324

Notes: All regressions include a fourth-order polynomial of the average GDP per capita on NUTS2 level over the years decisive for the eligibility threshold. In each column, we estimate fuzzy RDD specifications using a two-stage instrumental variables procedure where the first stage is estimated nonlinearly. All specifications include region fixed effects. The standard errors are robust to clustering at the NUTS2 level. The distance matrices are used such that only those non-Objective 1 regions are included in the control group that are more than 100 km away from the next Objective 1 region in the first column, 150km in the second column, and 200km in the third column. The sample consists of the EU12 NUTS3 regions for the first period, the EU15 NUTS3 regions for the second and the EU25 NUTS3 regions for the third programming period. We miss information on the four French overseas-départements and the two autonomous Portuguese regions Madeira and Azores for all three periods. For the Dutch region Flevoland we miss information for the first period only. Regarding the East-German NUTS3 regions we calculated GDP per capita growth for the years 1989 and 1990 using information from the German Democratic Republic's statistical yearbook.

ters the number of NUTS2 observations drops from 674 in Table 5.6 to 625 observations in Table 5.10. With a radius of 200 kilometers, the number of observations drops to 544. Similarly, the number of NUTS3 observations drops from 3,300 in Table 5.7 to 2,861 in Table 5.11 with a radius of 100 kilometers and to 2,420 with a radius of 200 kilometers.

A comparison of the results between the fixed effects two-stage least squares estimates with a nonlinear first stage in Tables 5.6 and 5.10 suggests that, at the NUTS2 level, the downward bias of the average Objective 1 treatment effect on per-capita income growth due to cross-regional spillovers in Table 5.6 is quite small: the coefficient estimate for a radius of 200 kilometers in Table 5.10 amounts to 0.020 which is identical to its benchmark in Table 5.6. Similar conclusions apply for the case with NUTS3 regions. The corresponding estimate at the NUTS3 level is 0.011 with a radius of 200 kilometers in Table 5.11 while it was 0.007 in Table 5.8. We conclude that there are only modest cross-region spillovers.

Table 5.12: RDD NUTS2 - Objective 1 and Employment growth

	sharp RDD		fuzzy RDD			
	Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
			linear 1st		nonlinear 1st	
	(1)	(2)	(3)	(4)	(5)	(6)
Objective 1	.001	.009	.001	-.005	.002	-.003
	(.002)	(.005)*	(.002)	(.003)*	(.002)	(.003)
Const.	-.005	-.071	-.009	.024	-.011	.023
	(.018)	(.020)***	(.018)	(.026)	(.018)	(.026)
Obs.	628	628	674	674	674	674
Regions	263	263	279	279	279	279
R^2	.037	.136	.031	.061	.031	.065

Notes: All regressions include a fourth-order polynomial of the average GDP per capita at the NUTS2 level over the years decisive for the eligibility threshold. In all pooled OLS regressions the standard errors are robust to clustering at the NUTS2 level. The first two columns report sharp RDD specifications while columns three to six show the fuzzy RDD specifications using a two-stage instrumental variables procedure. The first stage is estimated linearly in specifications three and four whereas specifications five and six build on a nonlinear first stage. In each case we run the pooled OLS regressions as well as the fixed effects regressions. The sample consists of the EU12 NUTS2 regions for the first period, the EU15 NUTS2 regions for the second period, and the EU25 NUTS2 regions for the third programming period. We miss information on the four French overseas-départements and the two autonomous Portuguese regions Madeira and Azores for all three periods. For the Dutch region Flevoland we miss information for the first period only. Regarding the East-German NUTS3 regions we calculated employment growth for the first period by relying on information for the years 1991-1993 only.

Objective 1 treatment effects on regional employment growth Particularly because a fraction of the funds is used as employment subsidies, there might also be a positive effect of Objective 1 treatment on employment growth. Moreover, infrastructure investments are thought to attract firms (see chapters 3 and 4) such that regional employment should be positively affected. We address this question by means of Tables 5.12 and 5.13 which are identical to Tables 5.6 and 5.8, except that they use average annual employment growth instead of per-capita income growth during a programming period as the dependent variable.

The treatment effect estimates in these tables indicate that significant employment growth effects of Objective 1 cannot be identified. This result is obtained irrespective of whether we use outcomes at the NUTS2 or the NUTS3 regional level; see the preferred model (6) in Table 5.12 or models (2), (4), or (6) in Table 5.13. Even though employment growth is not the main target of Objective 1 transfers, it should be achieved indirectly. Yet, together with the results on per-capita income growth in Tables 5.6 and 5.8, our findings indicate that income growth effects must be triggered mainly by investment growth and productivity growth. Employment seems to remain unaffected by Objective 1 transfers. This sheds a rather negative light on the part of Objective 1 transfers that are used for employment subsidies, these seem to be ineffective.

Table 5.13: Fuzzy RDD NUTS3 with controls - Objective 1 and Employment growth

	Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
	(1)	(2)	(3)	(4)	(5)	(6)
Objective 1	.0006	-.001	-.002	-.004	-.008	-.003
	(.002)	(.003)	(.002)	(.003)	(.003)***	(.002)
Empl. share	-.045	-.241	-.044	-.241	-.045	-.243
	(.011)***	(.067)***	(.011)***	(.067)***	(.013)***	(.068)***
Agri. share	.015	.058	.003	.057	.022	.069
	(.012)	(.038)	(.012)	(.038)	(.016)	(.037)*
Serv. share	.025	.017	.022	.016	.029	-.005
	(.008)***	(.023)	(.008)***	(.023)	(.011)***	(.029)
Pop. growth	.918	.998	.918	.997	.888	.761
	(.096)***	(.200)***	(.095)***	(.200)***	(.099)***	(.195)***
Pop. density	.0008	-.003	.001	-.002	.0008	-.024
	(.0006)	(.011)	(.0007)	(.011)	(.0006)	(.012)**
Voting power in EU council			-.093	-.044		
			(.014)***	(.038)		
Country-time dummies	no	no	no	no	yes	yes
Const.	-.027	.014	-.013	.012	-.028	-.063
	(.019)	(.028)	(.018)	(.025)	(.023)	(.030)**
Obs.	3300	3300	3300	3300	3300	3300
Regions	1207	1207	1207	1207	1207	1207
R^2	.213	.28	.237	.304	.297	.372

Notes: All regressions include a fourth-order polynomial of the average GDP per capita at the NUTS2 level over the years decisive for the eligibility threshold. In each column we estimate fuzzy RDD specifications using a two-stage instrumental variables procedure where the first stage is estimated nonlinearly. In each case we run the pooled OLS regressions as well as the fixed effects regressions. The standard errors are robust to clustering at the NUTS2 level. The sample consists of the EU12 NUTS3 regions for the first period, the EU15 NUTS3 regions for the second period, and the EU25 NUTS3 regions for the third programming period. We miss information on the four French overseas-départements and the two autonomous Portuguese regions Madeira and Azores for all three periods. For the Dutch region Flevoland we miss information for the first period only. Regarding the East-German NUTS3 regions we calculated employment growth for the first period by relying on information for the years 1991-1993 only.

Assessing the net benefits of Objective 1 treatment With the estimates at hand, we may easily infer whether the Objective 1 transfers are on average cost-effective or not, when requiring positive net benefits within a programming period. In Table 5.6 the average estimate on the Objective 1 effect across columns (3) through (6) is 0.018. However, there is evidence of moderate cross-regional spillovers so that we refer to the point estimate of 0.020 in column (3) of Table 5.10 as the preferable estimate. Accordingly, Objective 1 treatment led to an average growth advantage of real GDP per capita of approximately 2.0 percentage points in recipient regions. The level of GDP per capita and GDP (at PPP) in the average treated NUTS2 region and year amounted to $11,074$ Euro and $16,000$ million Euros, respectively.[23] The average Objective 1 NUTS2 region's population changed only slightly over the average period

[23]Taking GDP and GDP per capita prior to each single programming period, i.e., in 1988 for the EU12 in the first period and 1989 for the German New Länder, 1993 for the EU12 and 1994 for Austria, Finland, and Sweden in the second period, 1999 for the EU15 and 2003 for the new accession countries in the third programming period.

with a yearly growth rate of 0.1 percent. Hence, Objective 1 treatment caused absolute GDP to change by about the same rate as per-capita GDP, namely by 2.0% or 320 million Euros (at PPP) per year in the average treated region and programming period. Aggregating this effect up for all treated regions in the average programming period results in a treatment effect of 23.06 billion Euros (at PPP) per year within the EU as a whole.[24] The total cost of the Objective 1 program was 18.98 billion Euro (at PPP) per year in the average programming period (see Table 5.1). Then, we may conclude that the Objective 1 program induces a net effect of 4.08 billion Euros (at PPP) per year or 21 percent of the expenses per year in the EU as a whole. In other words, every Euro spent on Objective 1 transfers leads to 1.21 EUR of additional GDP. However, this simple cost benefit calculation may not be interpreted as a proof of efficiency as it neglects any inefficiencies resulting from tax collection as well as the inefficiencies that result from the potential distortion of the optimal spatial allocation of economic activity (see Chapters 2 and 3).[25] Yet, the cost benefit calculation demonstrates that the transfers are not wasteful and make more than a pure redistribution of income. The public investments seem to have productivity enhancing effects in the recipient regions.

Estimating the effect of marginal changes in Objective 1 transfers: dose-response function estimation The amount of Objective 1 transfers differs across recipient regions. From our estimates one should not expect that any treatment intensity leads to a positive net effect. Obviously, there has to be an optimal level of transfers which yields the most favorable cost-benefit ratio. If the transfers cross a certain level the direct costs should exceed the induced productivity effects. The *Dose-response function* estimation allows us to assess how Objective 1 regions respond to variations in the treatment intensity.[26] The derivative of the estimated Dose-response function provides the marginal effects of a change in treatment intensity.

[24]There were 58 treated NUTS2 regions in the first period of which the 11 New Länder regions received funds only for 4 years. In the second period, there were 64 treated NUTS2 regions of which the Austrian NUTS2 region Burgenland received funds only in 5 out of 6 years, and in the third period there were 129 treated NUTS2 regions of which the 67 Objective 1 regions in the new accession countries received funds only from 2004 onwards, that is for 3 rather than 7 years. This makes a total of 1297 region-year observations of Objective 1 treatment or, on average, 72.06 regions per year receiving treatment over the three periods.

[25]A crucial assumption for this cost assessment is that the associated collection of taxes did not distort economic activity in net paying regions. Moreover, we abstract from administrative costs associated with the collection of these taxes. Hence, we assume that one Euro of Objective 1 transfers is identical to one Euro of costs.

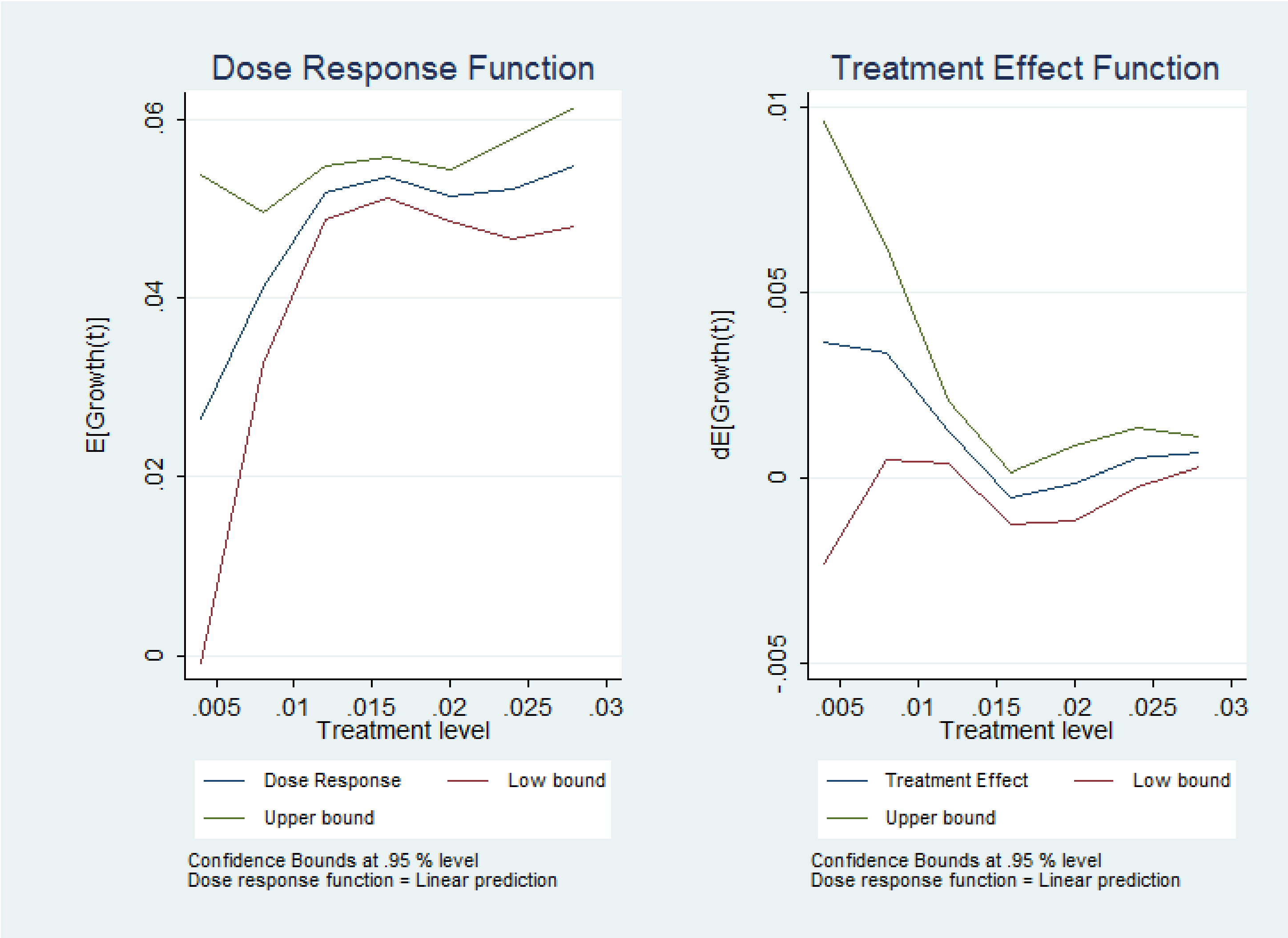
[26]The latter could not be done throughout our analysis, because of the limited availability of continuous information about Objective 1 transfers.

In this final assessment, we consider continuous Objective 1 transfers and estimate their effects on economic growth. Binary indicators of Objective 1 recipient status may conceal varying effects of different magnitudes of EU transfers. We consider Objective 1 funds as a fraction of GDP as the continuous treatment variable. However, there are not enough data points to estimate a dose-response function at the NUTS2 level, and information on continuous transfers at the NUTS3 level of aggregation is only available for the programming period 1994-99. So, unlike with binary Objective 1 transfer recipient status, we have to focus on the analysis at the NUTS3 level but may not entertain the advantages of region-fixed effects estimation.

Hirano and Imbens (2004) derive an extension of the Rosenbaum and Rubin (1983) propensity-score method to estimate the average local effects of continuous treatments. The estimates from generalized propensity score matching are typically reported with a dose-response function. In our case, the dose-response function depicts the growth rate over a programming period as a function of EU transfers relative to GDP. Figure 5.5 plots the dose-response function (left panel) as well as the treatment effect function (the derivative of the dose-response function) for the central programming period, 1994-99.

The dose-response function indicates that at low values of the treatment (Objective 1 transfer volumes that are small relative to a region's GDP) regions grow at around 3 percent per year, similar to regions without Objective 1 status. An increase in Objective 1 transfers leads to a more-than-proportionate increase in average annual growth of per-capita GDP at PPP growth until a level of transfers of 1.25% of the recipient region's GDP is reached. Beyond that value, the marginal effect of further transfers declines. That is also evidenced by the treatment effect function where confidence bands include a zero additional growth effect for transfer-to-GDP ratios above 1.25 percent. These results point to the existence of leakage effects if transfers exceed a certain value. Hence, receiving more transfers leads to higher GDP per capita growth but at a diminishing rate. After crossing a certain treatment intensity leakage effects dominate and the effect of further transfers is zero.

Figure 5.5: Dose-response estimates of Objective 1 transfers



Note: The figure plots the dose-response function (left panel) as well as the treatment effect function (the derivative of the dose-response function) for the central programming period, 1994-99. Objective 1 transfers are computed as a fraction of beginning of period GDP. See main text for details.

5.6 Concluding Remarks

This chapter considers the estimation of causal effects of the European Union's (EU) Objective 1 transfers on economic growth. Objective 1 funds aim at facilitating convergence and cohesion within the EU and constitute the major part of the EU's Structural Funds Program. They target fairly large, sub-national regional aggregates – referred to as NUTS2 regions – to foster growth in regions, whose per-capita GDP in purchasing power parity is lower than 75% of the EU's average per-capita income.

The 75%-rule gives rise to a regression-discontinuity design that exploits the jump in the probability of Objective 1 recipience. In the vast majority of cases (93% of the observations at NUTS2 level), the 75%-rule is strictly applied. Only 7% of our observations do not comply with the assignment mechanism. These are regions which either obtained Objective 1 funds although they were not eligible according to the rule or they did not receive funds although they were eligible. When non-complying regions are excluded, we have a sharp regression-discontinuity design. Including non-compliers leads to a fuzzy regression-discontinuity design. We perform regressions both at the NUTS2 level (the level of aggregation at which Objective 1 status is generally decided) and at the NUTS3 level for sharp and fuzzy regression-discontinuity designs. In some specifications, further covariates are included in order to control for selection into treatment. Amongst others, we include a country's voting power in the EU council to account for non-compliance with the 75%-rule.

The results can be summarized as follows. On average, Objective 1 status raises real GDP per capita growth by about 2.0% within the same programming period. Second, different from the positive effects on per-capita GDP, we do not find significant employment effects during the period in which transfers are allocated. There may be various reasons for that. One reason could be that Objective 1 transfers mainly stimulate investment. Another reason could be that the creation of jobs takes longer than the duration of a programming period of five to seven years.

Several robustness checks are performed. First, we deal with possible spillovers of Objective 1 funds on neighboring regions by estimating separate regressions in which we exclude control regions adjacent to treated regions. Second, we use a switching regression model as an alternative model of selection into treatment. The results are robust to these changes in estimation sample and estimation method. Finally, we shed light on the role of the magnitude of funds received for economic outcomes. With regard to the latter, we find that a marginal increase in Objective 1 transfers at low levels of funding leads to a disproportionate increase in growth of per-capita GDP at PPP. However, beyond a certain level, further increases in Objective 1 funding do not lead to further growth effects.

A simple back-of-the envelope calculation of the net benefits of Objective 1 transfers suggests the following. According to conservative benchmark estimates, every Euro spent on Objective 1 transfers leads to 1.21 Euros of additional GDP. The latter seems to be associated with investment (e.g., in infrastructure) and, eventually, productivity gains but not with the creation of new jobs. From this, we may conclude that, on average, Objective 1 transfers under the EU's Structural Funds Program are effective and – in net terms – not wasteful. Hence, the EU's Structural Funds are not only effective in reducing regional income disparities but finance on average productivity enhancing investments in the recipient regions. However, as argued in the preceding chapters it remains doubtful whether it is efficient to intervene in the regional distribution of economic activity by reducing regional disparities at all.

Chapter 6

Summary and Policy Implications

The preceding chapters provided a comprehensive theory of economic geography by combining traditional approaches with micro-foundations of agglomeration economies. Four different mechanisms, which underly economic activity's tendency to concentrate in space, were subject of the analysis. Each of the agglomeration mechanisms is characterized by the interplay of agglomeration economies and dispersion forces. Agglomeration economies emerge from knowledge spillovers, uncertainty, the provision of public goods as well as from the interaction of imperfect competition and increasing returns to scale. The spatial scarcity of production factors or consumption goods typically represents the major force that countervails agglomeration economies. The relative strength of agglomeration economies and dispersion forces, which characterizes the equilibrium distribution of economic activity, depends on a set of parameters derived from the theoretical analysis. In a framework with multiple agglomerations, these parameters determine the distance between individual agglomerations or – in the terminology of the traditional theories – the distance between individual central places.

The strength of agglomeration economies and dispersion forces vary among industries such that different industries provide their products and services at a different frequency in space. Industries with weak agglomeration economies but strong dispersion forces provide their goods and services at various places while industries that feature strong agglomeration economies and weak dispersion forces are located only at a few higher-order central places. Depending on which industry one considers some agglomeration economies are more relevant than others. This is illustrated in Chapter 4 which provides empirical evidence that the New Economic Geography's agglomeration mechanism particularly fits the spatial distribution of the service sector in Europe.

The ongoing integration of the European Single Market reduces the overall level of trade and transportation costs. On the one hand this benefits aggregate welfare, but on the other hand this process is associated with a significant reorganization of economic activity.[1] The models in Chapters 2 and 3 illustrated that a lower level of trade costs yields stronger agglomeration tendencies and a more uneven distribution of economic activity in space. The latter represents a challenge for European policy which regards regional cohesion as a principal aim. As pointed out in Chapter 5 the European Union's Structural Funds Program – which represents the most important instrument of European regional policy – effectively contributes to the objective of reducing regional disparities. However, this policy of spatial equalization goes at the expense of aggregate efficiency. The theoretical analyses in Chapters 2 and 3 argued that there are efficiency gains from regional concentration of economic activity. The agglomeration economies introduced in the previous chapters allow for gains from variety, sharing risk, realizing knowledge spillovers, and the efficient usage of public goods. Hence, similarly to the distribution of personal income, a trade-off between equality and efficiency arises in a spatial dimension.

In addition to the efficiency analysis, the effectiveness of regional policy was discussed theoretically. Relying on a New Economic Geography model, it was argued that regional policy should focus on providing public infrastructure within peripheral regions in order to achieve a competitive level of productivity before improving the connections to core regions (see Chapter 4). Germany serves as an example for persistent disparities in spite of great efforts made by regional policy. This may partly be attributed to misguided regional policy. However, the main reason for this development was the suppression of a major dispersion force. After reunification, the employers and unions agreed upon raising the eastern wages rapidly to the western level (see Sinn and Sinn, 1991). Thereby, the spatial equilibrium was distorted insofar as the gap between factor prices in western and eastern regions was significantly reduced. Accordingly, the balance between agglomeration economies and dispersion forces was shifted in favor of West German core-regions which offered the whole range of agglomeration benefits. The efforts made by regional policy to increase productivity in East German regions through public investments and subsidies were not powerful enough to compensate for the disabled dispersion force.

[1] The interrelation between trade costs, welfare and agglomeration is analyzed in Chapter 3. From equation (24) aggregate welfare is negatively related to the regional price-indices of both regions. The regional price-indices again decrease when trade costs fall as shown in Chapter 3, equation (15). However, a lower level of trade cost increases regional disparities as shown in figure 3.1.

From an efficiency point of view, the theoretical analysis of the previous chapters suggests that regional policy should abstain from intervening in the spatial distribution of economic activity in the sense of attracting firms to peripheral regions. This may lead to a higher degree of agglomeration, but for mobile citizens, regional concentration of economic activity does not necessarily imply regional income disparities. However, migration is costly and not all citizens are mobile. Moreover, peripheral regions may lack the fiscal capacity to provide their remaining citizens with an adequate level of public goods and services. Therefore, regional transfers in terms of a fiscal equalization scheme are necessary in order to comply with a harmonization of individual prospects. Yet, any equalization scheme should aim at minimizing distortions of the spatial distribution of economic activity such that benefits from agglomeration economies are realized.

Bibliography

Abdel-Rahman, H. and M. Fujita (1990). Product variety, Marshallian externalities, and city sizes. *Journal of Regional Science 30*, 165–183.

Aiello, F., A. Iona, and L. Leonida (2009). Regional infrastructure and firm investment: Theory and empirical evidence for Italy. *Empirical Economics*. forthcoming.

Alonso, W. (1964). *Location and Land Use*. Cambridge, MA: Harvard Universiy Press.

Andersson, F. and R. Forslid (2003). Tax competition and economic geography. *Journal of Public Economic Theory 5*, 279–303.

Angrist, J. D. and J.-S. Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

Aschauer, D. A. (1989a). Does public capital crowd out private capital? *Journal of Monetary Economics 24*, 171–188.

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

Badinger, H. and G. Tondl (2003). Trade, human capital and innovation: The engines of European regional growth in the 1990s. In B. Fingleton (Ed.), *European Regional Growth*, pp. 215–240. Berlin: Springer.

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

Baldwin, R. E. and P. Krugman (2004). Agglomeration, integration and tax harmonisation. *European Economic Review 48*, 1–48.

Baltagi, B. H. and N. Pinnoi (1995). Public capital stock and state productivity growth: Further evidence from an error components model. *Empirical Economics 20*, 351–359.

Barro, R. J. (1990). Government spending in a simple model of endogenous growth. *Journal of Political Economy 98*, 103–125.

Becker, S. O., P. H. Egger, and M. von Ehrlich (2009). Going NUTS: The effect of EU Structural Funds on regional performance. mimeo, University of Munich.

Berlin Institut (2009). Demografischer Wandel – Ein Politikvorschlag unter Besonderer Berücksichtigung der Neuen Länder. Berlin Institut für Bevölkerung und Entwicklung, Berlin.

Beugelsdijk, M. and S. C. Eijffinger (2005). The effectiveness of Structural Policy in the European Union: An empirical analysis for the EU-15 in 1995-2001. *Journal of Common Market Studies 43*, 37–51.

BMVBS (2006). Perspektiven der Raumentwicklung in Deutschland. Bundesministerium für Verkehr, Bau und Stadtentwicklung (Federal Ministry of Transport, Building and Urban Affairs), Berlin.

Bénassy-Quéré, A., N. Gobalraja, and A. Trannoy (2007). Tax and public input competition. *Economic Policy 50*, 385–430.

Boarnet, M. G. (1998). Spillovers and the locational effects of public infrastructure. *Journal of Regional Science 38*, 381–400.

Boldrin, M. and F. Canova (2001). Europe's regions - income disparities and regional policies. *Economic Policy 32*, 207–253.

Borck, R., M. Caliendo, and V. Steiner (2007). Fiscal competition and the composition of public spending: Theory and evidence. *Finanzarchiv (Public Finance Analysis) 63*, 264–277.

Borck, R. and M. Pflüger (2006). Agglomeration and tax competition. *European Economic Review 50*, 647–668.

Brühlhart, M. (2001). Evolving geographic concentration of European manufacturing industries. *Weltwirschaftliches Archiv (Review of World Economics) 137*, 215–243.

Bucovetsky, S. (2005). Public input competition. *Journal of Public Economics 89*, 1763–1787.

Bucovetsky, S. and J. D. Wilson (1991). Tax competition with two tax instruments. *Regional Science and Urban Economics 21*, 333–350.

Cameron, A. C. and P. K. Trivedi (2005). *Microeconometrics: Methods and Applications.* New York, NY: Cambridge University Press.

Capello, R. (2007). *Regional Economics.* New York, NY: Routledge.

Cappelen, A., F. Castellacci, J. Fagerberg, and B. Verspagen (2003). The impact of EU regional support on growth and convergence in the European Union. *Journal of Common Market Studies 41*, 621–644.

Christaller, W. (1933). *Die Zentralen Orte Süddeutschlands.* Darmstadt: Wissenschafliche Buchgesellschaft.

Cohen, J. P. and C. J. Morrison Paul (2004). Public infrastructure investment, interstate spatial spillovers and manufacturing costs. *Review of Economics and Statistics 86*, 551–560.

Criscuolo, C., R. Martin, H. G. Overman, and J. van Reenen (2007). The effect of industrial policy on corporate performance: Evidence from panel data. *Working Paper LSE.*

Dall'erba, S. and J. Le Gallo (2008). Regional convergence and the impact of Structural Funds over 1989-1999: A spatial econometric analysis. *Papers in Regional Science 87*, 219–244.

Davis, D. R. and D. E. Weinstein (1999). Economic geography and regional production structure: An empirical investigation. *European Economic Review 43*, 379–407.

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

Dixit, A. and J. E. Stiglitz (1977). Monopolistic competition and optimum product diversity. *American Economic Review 87*, 297–308.

Dupont, V. and P. Martin (2006). Subsidies to poor regions and inequalities: Some unpleasant arithmetic. *Journal of Economic Geography 6*, 223–240.

Duranton, G. and D. Puga (2001). Nursery cities: Urban diversity, process innovation, and the life cycle of products. *American Economic Review 91*, 1454–1477.

Duranton, G. and D. Puga (2004). Micro-foundations of urban agglomeration economies. In J. V. Henderson and J.-F. Thisse (Eds.), *Handbook of Regional and Urban Economics*, Volume 4, pp. 2063–2117. Amsterdam: Elsevier.

Ederveen, S., H. L. de Groot, and R. Nahuis (2006). Fertile soil for Structural Funds? A panel data analysis of the conditional effectiveness of European cohesion policy. *Kyklos 59*, 17–42.

Ederveen, S., J. Gorter, R. de Mooij, and R. Nahuis (2002). *Funds and Games: The Economics of European Cohesion Policy*. Amsterdam: CPB & Koninklijke De Swart.

Egger, H. and J. Falkinger (2006). The role of public infrastructure and subsidies for fim location and international outsourcing. *European Economic Review 50*, 1993–2015.

Eggert, W., R. Fenge, G. König, and M. von Ehrlich (2007). Konvergenz- und Wachstumseffekte der Europäischen Regionalpolitik in Deutschland. *Perspektiven der Wirtschaftspolitik 8*, 130–146.

Ellison, G. and E. L. Glaeser (1999). The geographic concentration of industry: Does natural advantage explain agglomeration. *American Economic Review 89*, 311–316.

European Commission (1997). Regional development studies - the impact of structural policies on economic and social cohesion in the union 1989-99. Luxembourg: Office for Official Publications of the European Commission.

European Commission (2001). Unity, solidarity, diversity for europe, its people and its territory - second report on economic and social cohesion. Luxembourg: Office for Official Publications of the European Commission.

European Commission (2007). Annex to the 18th annual report on implementation of the Structural Funds 2006. Brussels: Commission of the European Communities.

European Commission (2008). Working for the regions – eu regional policy 2007-2013. Luxembourg: Office for Official Publications of the European Communities.

European Commission (2009). EU budget – Sustainable development and innovation at the core of the EU budget. Luxembourg: Office for Official Publications of the European Communities.

Evans, P. and G. Karras (1994). Is government capital productive? evidence from a panel of seven countries. *Journal of Macroeconomics 16*, 271–279.

Ezcurra, R., P. Pascual, and M. Rapun (2006). The dynamics of industrial concentration in the regions of the European Union. *Growth and Change 37*, 200–229.

Felsenthal, D. and M. Machover (1998). The weighted voting rule in the EU's council of ministers, 1958-95: Intentions and outcomes. *Electoral Studies 16*, 33–47.

Fenge, R., M. von Ehrlich, and M. Wrede (2009). Public input competition and agglomeration. *Regional Science and Urban Economics 39*, 621–631.

Fingleton, B. (2001). Theoretical economic geography and spatial econometrics: Dynamic perspectives. *Journal of Economic Geography 1*, 201–225.

Fingleton, B. and M. M. Fischer (2009). Neoclassical theory versus new economic geography: Competing explanations of cross-regional variation in economic development. *Annals of Regional Science*. forthcoming.

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

Gabe, T. M. and K. P. Bell (2004). Tradeoffs between local taxes and government spending as determinant of business location. *Journal of Regional Science 44*, 21–41.

Glicksberg, I. L. (1952). A further generalization of the Kakutani fixed point theorem with application to Nash equilibrium points. *Proceedings of the American Mathematical Society 3*, 170–174.

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

Hirano, K. and G. W. Imbens (2004). The propensity score with continuous treatments. In A. Gelman and X.-L. Meng (Eds.), *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*, pp. 73–83. Wiley Series in Probability and Statistics, Chichester: Wiley.

Holl, A. (2004a). Manufacturing location and impacts of road transport infrastructure: Empirical evidence from Spain. *Regional Science and Urban Economics 34*, 341–363.

Holl, A. (2004b). Transport infrastructure, agglomeration economies, and firm birth: Empirical evidence from Portugal. *Journal of Regional Science 44*, 693–712.

Holtz-Eakin, D. (1994). Public-sector capital and the productivity puzzle. *Review of Economics and Statistics 76*, 12–21.

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

Imbens, G. W. and T. Lemieux (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics 142*, 615–635.

Keen, M. and M. Marchand (1997). Fiscal competition and the pattern of public spending. *Journal of Public Economics 66*, 33–53.

Kind, H. J., K. H. Knarvik, and G. Schjelderup (2000). Competing for capital in a lumpy world. *Journal of Public Economics 78*, 253–274.

Krugman, P. (1991a). *Geography and Trade.* Cambridge, MA: MIT Press.

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

Krugman, P. (1995). Globalization and the inequality of nations. *Quarterly Journal of Economics 60*, 857–880.

Lösch, A. (1940). *Die Räumliche Ordnung der Wirtschaft.* Jena: Gustav Fischer Verlag.

Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics 22*, 3–42.

Ludema, R. D. and I. Wooton (2000). Economic geography and the fiscal effects of regional integration. *Journal of International Economics 52*, 331–357.

Ma, J. (1997). Intergovernmental fiscal transfers in nine countries: Lessons for developing countries. *World Bank Policy Research Working Paper No. 1822.*

Marshall, A. (1890). *Principles of Economics.* London: Macmillan and Co.

Martin, P. (1999). Public policies, regional inequalities and growth. *Journal of Public Economics 73*, 85–105.

Martin, P. and G. I. P. Ottaviano (1999). Growing locations: Industry location in a model of endogenous growth. *European Economic Review 43*, 281–302.

Martin, P. and C. A. Rogers (1995). Industrial location and public infrastructure. *Journal of International Economics 39*, 335–351.

Midelfart-Knarvik, K. H. and H. G. Overman (2002). Delocation and European integration- is structural spending justified? *Economic Policy 17*, 323–359.

Morrison, C. J. and A. E. Schwartz (1996). State infrastructure and productive performance. *American Economic Review 86*, 1095–1111.

Moses, L. N. (1958). Location and the theory of production. *Quarterly Journal of Economics 78*, 259–72.

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica 46*, 69–85.

Munnell, A. H. (1990). How does public infrastructure affect regional economic performance? *New England Economic Review Sept/Oct*, 11–33.

Nelson, R. and E. Phelps (1966). Investment in humans, technological diffusion and economic growth. *American Economic Review 56*, 69–75.

Noiset, L. (1995). Pigou, Tiebout, property taxation, and the underprovision of local public goods: Comment. *Journal of Urban Economics 38*, 312–316.

Ottaviano, G. I. P. and J.-F. Thisse (2004). Agglomeration and economic geography. In J. V. Henderson and J. Thisse (Eds.), *Handbook of Regional and Urban Economics*, Volume 4, pp. 2563–2608. Amsterdam, NL: Elsevier.

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

Overman, H. G. and D. Puga (2009). Labour pooling as a source of agglomeration: An empirical investigation. *CEPR Discussion Paper Nr.7174*.

Pereira, A. M. and J. M. Andraz (2007). Public investment in transportation infrastructures and industry performance in Portugal. *Journal of Economic Development 32*, 1–20.

Pflüger, M. (2004). A simple, analytically solvable chamberlinian agglomeration model. *Regional Science and Urban Economics 34*, 565–573.

Polanyi, M. (1958). *Personal Knowledge*. Chicago: The University of Chicago Press.

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

Puga, D. (2008). Agglomeration and cross-border infrastructure. *European Investment Bank Papers 13*, 102–124.

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

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

Rivera-Batiz, F. (1988). Increasing returns, monopolistic competition, and agglomeration economies in consumption and production. *Regional Science and Urban Economics 18*, 125–153.

Robert-Nicoud, P. and F. Sbergami (2004). Home-market vs. vote-market effect: Location equilibrium in a probabilistic voting model. *European Economic Review 48*, 155–179.

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy 94*, 1002–1037.

Rosenbaum, P. R. and D. B. Rubin (1983). Central role of the propensity score in observational studies for causal effects. *Biometrika 70*, 41–55.

Sala-i-Martin, X. (1996). Regional cohesion: Evidence and theories of regional growth and convergence. *European Economic Review 40*, 1325–1352.

Salop, S. C. (1979). Monopolistic competition with outside goods. *Bell Journal of Economics 10*, 141–156.

Shapley, L. S. and M. Shubik (1954). A method for evaluating the distribution of power in a committee system. *American Political Science Review 48*.

Sinn, G. and H.-W. Sinn (1991). *Kaltstart: Volkswirtschaftliche Aspekte der deutschen Vereinigung*. Tübingen: J.C.B. Mohr.

Sinn, H.-W. (1997). The selection principle and market failure in systems competition. *Journal of Public Economics 66*, 247–274.

Sinn, H.-W. (2003). *The New Systems Competition*. Malden, MA: Blackwell Publishing.

von Böventer, E. (1962). Die Struktur der Landschaft - Versuch einer Synthese und Weiterentwicklung der Modelle J.H. von Thünens, W. Christallers und A. Löschs. In E. Schneider (Ed.), *Schriften des Vereins für Socialpolitik*, pp. 77–133. Berlin: Verlag von Duncker & Humblot.

von Böventer, E. (1969). Walter Christaller's central places and peripheral areas: The central place theory in retrospect. *Journal of Regional Science 9*, 117–124.

von Thünen, J. H. (1826). *Der Isolierte Staat in Beziehung auf Landschaft und Nationalökonomie*. Hamburg: Puthes.

Weber, A. (1909). *Reine Theorie des Standorts*. Tübingen: Mohr.

Widgrén, M. T. (2009). The impact of council voting rules on EU decision-making. *CESifo Economic Studies 55*, 30–56.

Wilson, J. D. (1986). A theory of interregional tax competition. *Journal of Urban Economics 19*, 296–315.

Wooldridge, J. M. (1995). Selection corrections for panel data models under conditional mean independence assumptions. *Journal of Econometrics 68*, 115–132.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Zodrow, G. R. and P. Mieszkowski (1986). Pigou, Tiebout, property taxation, and the underprovision of local public goods. *Journal of Urban Economics 19*, 356–370.

Curriculum Vitae

Since December 2006	Research and Teaching Assistant Chair of Prof. Dr. Dr. h.c. Hans-Werner Sinn Ph.D. Program in Economics Munich Graduate School of Economics Ludwig-Maximilians-Universität, Munich
March 2006 - December 2006	Junior Researcher and Doctoral Candidate Ifo Institute for Economic Research, Munich
August 2004 - March 2005	Visiting Student in Economics Columbia University, New York
October 2001 - June 2006	Studies in Economics (Dipl.-Volksw.) Ludwig-Maximilians-Universität, Munich
June 2001	Abitur Allgäu Gymnasium, Kempten
November 21, 1980	Born in Heidelberg, Germany