HWWA HAMBURG

DISCUSSION PAPER

New Economic Geography in Germany: Testing the Helpman-Hanson Model

Steven Brakman
Harry Garretsen
Marc Schramm

HWWA DISCUSSION PAPER

172

Hamburgisches Welt-Wirtschafts-Archiv (HWWA)
Hamburg Institute of International Economics
2002
ISSN 1616-4814

The HWWA is a member of:

- Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz (WGL)
- Arbeitsgemeinschaft deutscher wirtschaftswissenschaftlicher Forschungsinstitute (ARGE)
- Association d’Instituts Européens de Conjoncture Economique (AIECE)

New Economic Geography in Germany: Testing the Helpman-Hanson Model

Steven Brakman
Harry Garretsen
Marc Schramm

A first draft of this paper was presented by Harry Garretsen at a joint seminar of the Dept. of Economics, Hamburg University and the HWWA at October 30th, 2001 in Hamburg. We would like to thank the participants of the seminar and also Michael Roos for their comments. In particular we would like to thank Michael Funke for the invitation and Konrad Lammers for the opportunity to include this paper in the HWWA Discussion Paper series.

This discussion paper is assigned to the HWWA's research programme "European Integration and Spatial Development".

HWWA DISCUSSION PAPER

Edited by the Department
EUROPEAN INTEGRATION
Head: Dr. Konrad Lammers

Hamburgisches Welt-Wirtschafts-Archiv (HWWA)
Hamburg Institute of International Economics
Öffentlichkeitsarbeit
Neuer Jungfernstieg 21 - 20347 Hamburg
Telefon: 040/428 34 355
Telefax: 040/428 34 451
E-mail: hwwa@hwwa.de
Internet: http://www.hwwa.de/

Steven Brakman
Groningen University
International Economics Department
P.O. Box 800, 9700 AV Groningen, Netherlands
Phone : +31 50 363 3746
Fax : +31 50 363 7337
E-mail: s.brakman@eco.rug.nl

Harry Garretsen
Nijmegen School of Management
University of Nijmegen
PO Box 9108, 6500 HK Nijmegen, Netherlands
Phone: 00-31-24-3615889
E-mail: H.Garretsen@nsm.kun.nl

Marc Schramm
University of Nijmegen
Centrum voor Duitsland-Studies
Huize Heyendael, Geert Grooteplein 9
Postbus 9103, 6500 HD Nijmegen, Netherlands
Phone: 024-3611137
Fax: 024-3613006
E-mail: m.schramm@cds.kun.nl

CONTENTS

Abstract

In this paper we find evidence that the new economic geography approach is able to describe and explain the spatial characteristics of an economy, in our case the German economy. Using German district data we estimate the structural parameters of a new economic geography model as developed by *Helpman* (1998) and *Hanson* (1998) and we find confirmation for a spatial wage structure. The advantage of the Helpman-Hanson model is that it incorporates the fact that agglomeration of economic activity increases the prices of local (non-tradable) services, like housing. This model thereby provides an intuitively appealing spreading force that allows for less extreme agglomeration patterns than predicted by the bulk of new economic geography models. Based on different estimation strategies and taking a number of features of the re-unified German economy into account, we do not only test for the spatial distribution of wages but also for the spatial structure with respect to German unemployment, employment and land prices.

JEL-Code: R10, R12, R23

Keywords: economic geography, empirical estimation, Germany

1 INTRODUCTION

Initiated by *Krugman* (1991) there has been a renewed interest in mainstream economics in recent years for the question how the spatial distribution of economic activity comes about. The literature on the so called new economic geography or geographical economics, shows how modern trade and growth theory can be used to give a sound theoretical foundation for the location of economic activity across space.[1] The seminal book by *Fujita*, *Krugman* and *Venables* (1999) develops and summarizes the main elements of the new economic geography approach. The emphasis in this book is strongly on theory and empirical research into the new economic geography is hardly discussed at all. As already observed by *Krugman* (1998, p. 172) in his survey of the new economic geography, this is no coincidence since there is still a lack of direct testing of the empirical implications of the new economic geography models. In his review of *Fujita*, *Krugman* and *Venables* (1999), *Neary* (2001) reaches a similar conclusion. In order to make progress an empirical validation of the main theoretical insights is called for. The reasons that the empirical research lags behind is that the new economic geography models are characterized by non-linearities and multiple equilibria which makes empirical validation relatively difficult.

To date, there is a substantial amount of empirical research that shows that location matters, but there are indeed still relatively few attempts to specifically test for the relevance of the structural parameters of new economic geography models (see the survey by *Overman*, *Redding* and *Venables* (2001)). A notable exception is the work by *Gordon Hanson* (1998, 1999). Hanson uses a new economic geography model developed by *Helpman* (1998) and then directly tests for the significance of the model parameters. Based on US county-data he finds confirmation for his version of the Helpman model. In this paper we apply the Helpman-Hanson model to the case of Germany. The goal of the paper is twofold.

First, we want to establish whether the Helpman-Hanson model holds for Germany, that is to say we want to know whether the key model parameters are significant or not. The main equation to be estimated will be a nominal wage equation, central to this equation

1 Elsewhere, see in particular *Brakman*, *Garretsen* and *van Marrewijk* (2001), we have argued that is more accurate to use the phrase “geographical economics” instead of “new economic geography” because the approach basically aims at getting more geography into economics rather than the other way around, but we stick here to the latter to avoid confusion.

is the idea nominal wages will be higher in those regions that have easy access to economic centers because for those regions demand linkages are relatively strong.

Second, we want to extend the analysis by Hanson by taking on board several features of the German economy that set Germany apart from the case of the USA and analyze their empirical implications. Apart from wages we will incorporate "spatial" features of other variables. The main eographical unit of analysis is the German city-district (*Stadtkreis).*

Even though the case of post-reunification Germany is thought to be well-suited for a new economic geography approach (see *Brakman* and *Garretsen* (1993) for an early qualitative attempt), the goal of the present paper is *not* to analyze whether or not our new economic geography model is the "best" model to analyze Germany after the fall of the Berlin Wall in 1989. In a similar vein, we do also not test the Helpman-Hanson model against possible alternative explanations of the regional distribution of economic activity in Germany. Our goal is more limited, we want to assess the empirical relevance of a particular new economic geography model for Germany.[2] By doing so we will take a number of characteristics of the German economy into account, which are of a 'geographical' nature. Basically this is the rationale for choosing Germany as an example to investigate the relevance of the New Economic Geography approach. In recent history Germany experienced the "rise and fall" of the Berlin Wall which from the point of view of the new economic geography creates a unique testing ground. In our paper we take the approach recommended by *Hanson* (2000) as a starting point. He concludes his survey of the empirical literature of spatial agglomeration by stating that the well-documented correlation of regional demand linkages with higher wages "*would benefit from exploiting the well-specified structural relationships identified by theory as a basis for empirical work*" (*Hanson*, 2000, p. 28).

The paper is organized as follows. In section 2 we briefly provide some data on the level of the German states (*Bundesländer)* to support the idea that geography might matter in Germany. In section 3 we first discuss the main elements of the theoretical model and focus on the derivation of the empirical specification of the wage equation that is our basic equation in the subsequent part of the paper. This wage equation is our vehicle to test for the presence of regional demand linkages that are central to the core new economic geography model in the underpinning of the spatial agglomeration of

2 For a similar attempt see *Roos* (2001)

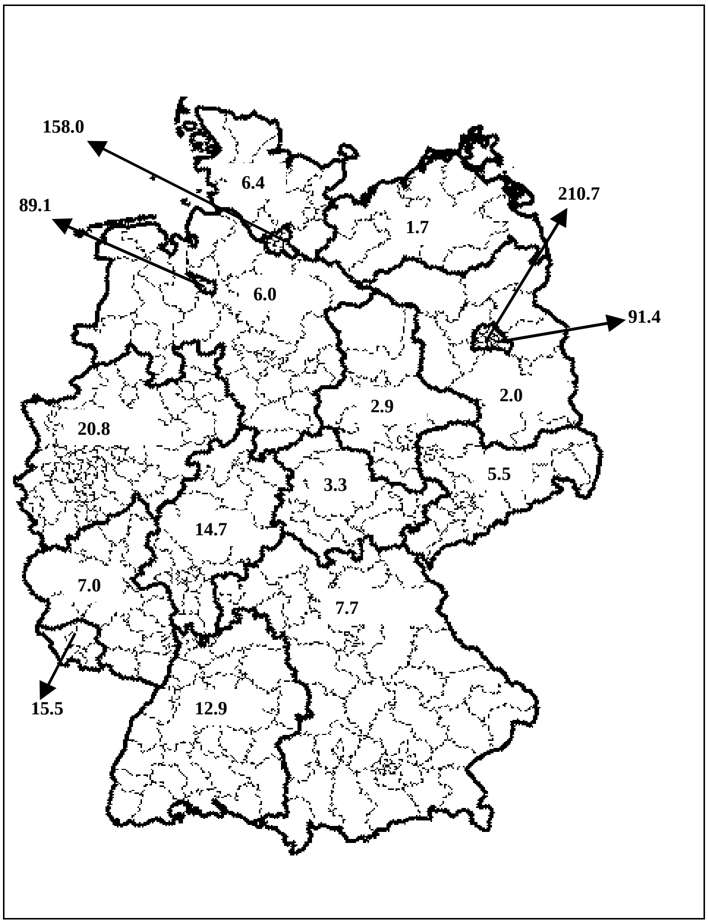
economic activity. Section 4 discusses some of the problems associated with estimating our wage equation and then gives the main estimation results for the basic wage equation and also supplies three alternative estimation strategies. In section 5 we address the role of two features of the German economy that might have a bearing on our results: the role of transfers (as proxied by the difference between regional GDP and regional income tax base) and the alleged inflexibility of the German labor market. With respect to the latter we will provide estimation results on the spatial characteristics of additional variables (besides wages) notably regional unemployment, employment and land prices. We also discuss the limitations of our approach and some possibilities for future research. Section 6 concludes the paper. Our main conclusion will be that the Helpman-Hanson model performs rather well for the case of Germany and we thereby find support for the empirical relevance of the new economic geography approach.

2 GEOGRAPHY AND GERMANY

In this section we briefly present some data in order to illustrate the spatial distribution of some key variables across Germany. A quick look at the Maps 1-3 below immediately shows that there are indeed geographical or spatial differences within Germany with respect to the economic variables that are at the heart of our model. Take, for instance, a look at Map 1 which gives GDP per km^2 for the German states (*Bundesländer)*. This map shows that geographical differences with respect to GDP are quite large (and even more skewed than GDP per capita, not shown here). The map also indicates that GDP per km^2 is higher in the former West Germany and this is not only true for smaller city-states like Bremen, Hamburg and West-Berlin.[3]

3 Maps 1-3 are based on information of 441 districts (*Kreise*), which we aggregated to 16 states (*Bundesländer*), to avoid information overload in the maps. The solid lines indicate the states, the dashed lines the districts.

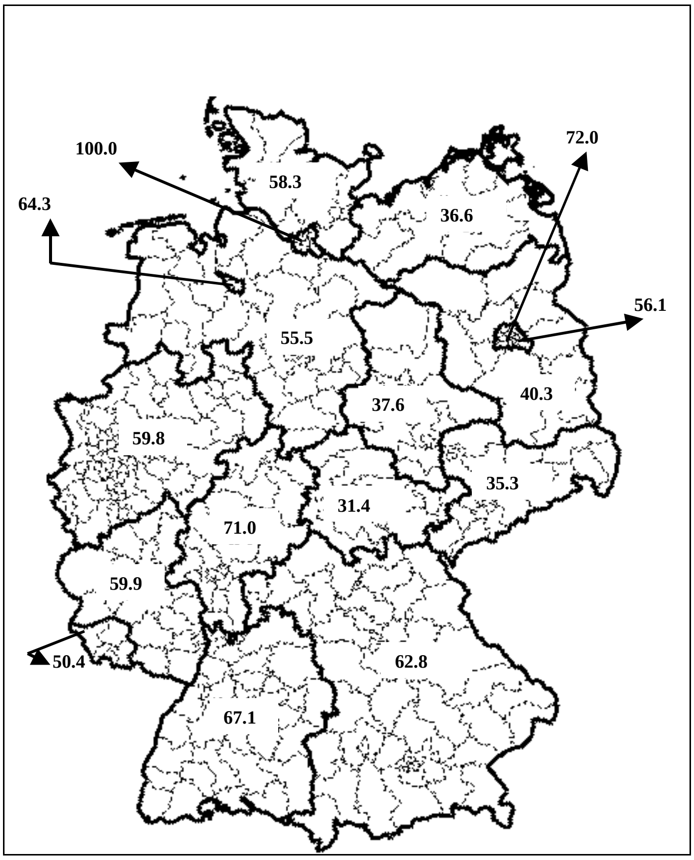
Map 1: GDP per km^2 in Millions of DM (1994)



Source: Federal Statistical Office, Wiesbaden.

Central in this paper is the Helpman-Hanson model. The key equation in this model, as will be explained in the next section, describes the spatial nature of (nominal) wages. Map 2 indicates that not only hourly (manufacturing) wages differ remarkably between states, but the map also suggests that in the eastern part of Germany wages are on average lower than in western Germany. The dividing line between high and low wages to some extent identifies the former border between East and West Germany. In the Helpman-Hanson model wages in a region are higher if that region is part of or close to a large market, proxied by GDP. This is in line with Maps 1 and 2 because these two maps suggest a positive correlation between nominal wages and gdp (per km^2).

Map 2: Average Hourly Wage in the Manufacturing Sector, 1995

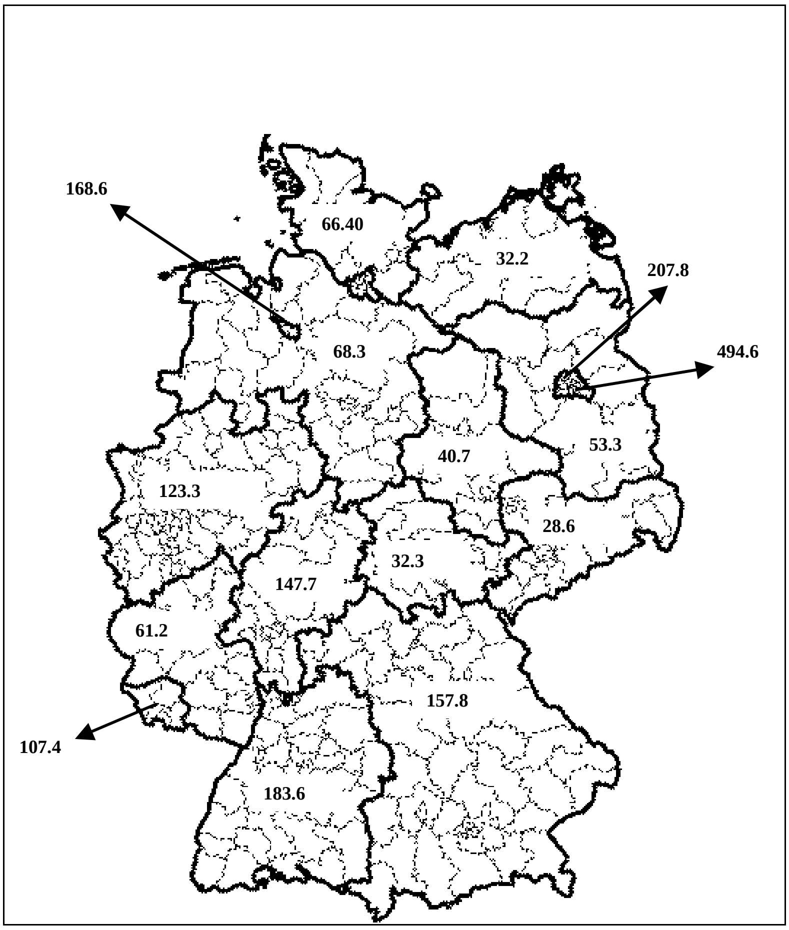


Source: Federal Statistical Office, Wiesbaden

Agglomeration in new economic geography models is, as geographers have known for a long time, the result of the combination of agglomerating and spreading forces. An important agglomerating force is for instance the size of the market (see Map 2). Among the spreading forces are the demand from immobile workers in peripheral regions, but also negative feedbacks in the core-regions such as congestion or the relatively high cost of housing and other local goods. An indication for the presence of these spreading forces in core regions are, for example, land prices. As Map 3 indicates, land prices in eastern Germany seem on average lower than in western Germany, but a possible dividing line between eastern and western Germany is less clear-cut than with respect to regional wages. Land prices can be looked upon as a proxy for housing prices and as will become clear in Section 3 housing prices are the spreading force in the Helpman-

Hanson model. Hence, Maps 1-3 give a first indication of the spatial distribution of the three key variables in the theoretical model, nonimal wages, the size of the market (gdp) and housing prices (here proxied by land prices). Taken together these maps suggest that there is no random distribution of economic activity across Germany and that high wages go along with high gdp and high land prices. A look at the *Kreise* data on which the Maps are based confirms this conclusion. The highest (lowest) values for the three variables are invariably observed in western (eastern) German *Kreise.*

Map 3: Prices of Land per m^2 (1995)



Source: Federal Statistical Office, Wiesbaden

3 THEORETICAL MODEL AND DATA

3.1 The rationale for the Helpman-Hanson Model

The benchmark model of the new economic geography, developed by *Krugman* (1991), is in general not suited for empirical validation, because it produces, in the long-run, for an intermediate range of trade costs only one, or at most a very few (equally sized) locations with manufacturing economic activity. This is clearly not in accordance with the facts about the spatial distribution of manufacturing activity for the US or any other industrialized country. Furthermore, it lacks some of the spatial characteristics of agglomerations, which have been found to be very relevant empirically, most importantly the tendency of prices of local (non-tradable) goods to be higher in agglomerations (see for example the survey by *Anas*, et al., 1998, and our Map 3 for that matter).

How can one arrive at a model that is better suited for empirical testing, that is to say a model that is less biased in favour of (complete) agglomeration? *Krugman* and *Venables* (1995) offer a useful starting-point. They assume, in contrast with *Krugman* (1991), no labour migration between regions, so when a sector expands the labour supply must come from other sectors in that region. Cumulative causation in this model comes from input-output linkages between firms, which are now assumed to use each other output as an intermediate input. Firms benefit from being close to each other by not paying transport cost on intermediate factors of production. In agriculture, only labor is used, with constant returns to scale and it can be costlessly traded. The latter assumption assures that as long as both regions produce both goods the wage rate equals unity (by choice of units). Typically this model produces two types of equilibria (see also *Fujita*, *Krugman* and *Venables*, 1999, chapter 14). For high trade costs of manufactures, a symmetric equilibrium, and for low trade costs a core-periphery solution (for intermediate transportation costs, asymmetric but unstable equilibria are possible). So, without complete specialization this model produces in a qualitative sense still the same type of equilbria as the *Krugman* (1991) model. *Krugman* and *Venables* (1996) extend this model by assuming two manufacturing sectors, each of which sells and buys more to firms in the same sector than to firms of the other sectors. Complete agglomeration is now less likely, because favorable cost and demand linkages benefit firms in the same sector while competition in product and labor markets harm all firms in all sectors equally. For low trade costs this results in regions to become specialized in one sector only.

It is only a small step to make this Krugman-Venables model more in line with the stylized facts; simply assume that the production function in agriculture is increasing (in labor) and concave (see *Puga*, 1999 and *Fujita*, *Krugman* and *Venables*, 1999, p. 244). This introduces an extra spreading force into the model. Complete agglomeration is now less likely, as agglomeration drives up wages in the core region, making it attractive for firms to re-locate to a peripheral region where labor costs are lower. If one plots the share of industry in a specific region against trade costs this typically results in a Ώ-type of relationship between the share of industry in each region and trade cost (see in particular *Puga*, 1999, Figure 6 or *Puga*, 2001, Figure 8, and also *Fujita*, *Krugman* and *Venables*,1999, Figure 14.8). For high trade costs, there is (equal) spreading of industrial activity, for intermediate levels of trade costs full as well as partial agglomeration results, and for low trade costs there is a return to spreading. Given the observation that full agglomeration is not in accordance with the facts, new economic geography models based on forward and backward linkages and with no interregional labor mobility seem therefore useful models for empirical testing. Unfortunately, however, direct testing of these models is rather cumbersome because it requires detailed information on input-output linkages between firms on a regional level (the importance of which is clearly illustrated by *Krugman* and *Venables*, 1996).

The reasons stated above are the main arguments why the model developed by *Helpman* (1998), with its empirical applications by *Hanson* (1998, 1999), is a useful alternative for empirical research. It combines the "best of the two worlds" since it shares with *Krugman* (1991) its emphasis on demand linkages (which are more easy to test for than input-output linkages). But at the same time through the inclusion of a non-tradable consumption good (i.e housing), the model is capable of producing similar equilibria as the aforementioned models based on input-output linkages and immobile factors of production.[4] The price of housing in the *Helpman* (1998) model which increases with agglomeration, serves as an analogous spreading force as the rising wages in *Puga* (1999). In fact, it can be shown that in terms of equilibrium outcomes the Helpman model yields similar results as, what has been dubbed, the second core model of new economic geography where there is no interregional labor mobility and the possibility of

4 For the differences between *Krugman* (1991) and *Helpman* (1998), see *Helpman* (1998, pp. 49-53). For a very useful general framework to understand the different implications of models with and without interregional labor mobility see *Puga* (1999, 2001). For the observation that the Helpman model is at home in the class of models that display the above mentioned Ω-relationship see *Puga* (1999, p. 324), *Puga* (2001, p. 16). *Ottaviano* and *Thisse* (2001, p. 175) also note that this relationship applies to *Helpman* (1998).

agglomeration arises through intricate input-output linkages between firms (*Venables*, 1996, *Krugman* and *Venables*, 1995, 1996, *Puga*, 1999).

3.2 The Helpman-Hanson Model

We briefly discuss the theoretical approach in *Hanson* (1998, 1999) and focus on the equilibrium conditions because these are needed to arrive at the basic wage equation that will be estimated.[5] With one notable exception (the inclusion of a non-tradable good (housing)) the micro-foundation for the behavior of the individual consumers and producers is the same as in the seminal *Krugman* (1991) model. Here, we only discuss the resulting equilibrium conditions and for the full model specification we refer to *Hanson* (1998, 1999). In the model consumers derive utility from consuming a manufacturing good, which is tradable albeit at a cost, and from housing which is a non-trabable good between regions. The manufacturing good consists of many varieties and each firm offers one variety and this is modeled with well-known Dixit-Stiglitz formulation of monopolistic competition. The only factor input in the model is labor and labor is needed to produce the manufacturing good and labor can move between regions in the long run. In this set-up of the model the perfectly competitive housing sector serves as the spreading force, because housing (a non-tradable good) is relatively more expensive in the centers of production where demand for housing is high. As we will see below apart from the inclusion of a homogenous non-tradable good (housing) at the expense of a homogenous tradable good (agriculture), there are no fundamental differences between *Krugman* (1991) and *Helpman* (1998). In particular in both models agglomeration is driven by demand linkages and the interregional mobility of labor.

This extension of core model thus allows for a richer menu of equilibrium spatial distributions of economic activity then the core model. As trade or transportation costs fall agglomeration remains a possible outcome but now also (renewed) spreading and partial agglomeration are feasible. Partial agglomeration means that all regions have at least some industry. Notwithstanding the different implications of *Helpman* (1998) compared to *Krugman* (1991) the equilibrium conditions (five in total) are very similar to the core model, in particular the equilibrium wage equation, which is central to the empirical analysis, is identical to the (normalized) equilibrium wage equation in *Krugman* (1991):

---

5 For an in-depth analysis of the core model see *Fujita*, *Krugman* and *Venables* (1999, chapters 4 and 5) or *Brakman*, *Garretsen* and *van Marrewijk* (2001, chapters 3 and 4).

(1) $W_r = \left[\sum_s Y_s I_s^{\varepsilon-1} T^{D_{rs}(1-\varepsilon)}\right]^{1/\varepsilon}$

(2) $I_r = \left[\sum_s \lambda_s \left(T^{D_{rs}}\right)^{1-\varepsilon} W_s^{1-\varepsilon}\right]^{1/(1-\varepsilon)}$

(3) $Y_r = \lambda_r L W_r$

In which in equation (1) W_r is the region's *r* (nominal) wage rate, *Y* is income, *I* is the price index for manufactured goods, ε is the elasticity of substitution for manufactured goods. *T* is the transport cost parameter, and $T_{rs} = T^{D_{rs}}$, where D_{rs} is the distance between locations *r* and *s*. Transport costs *T* are defined as the number of manufactured goods that have to be shipped in order to ensure that one unit arrives over one unit of distance. Given the elasticity of substitution ε, it can directly be seen from equation (1) that for every region wages are higher when demand in surrounding markets (Y_s) is higher (including its own market), when access to those markets is better (lower transport costs *T*). Also regional wages are higher when there is less competition for the varieties the region wants to sell in those markets (this is the extent of competition effect, measured by the price index I_s).

Equation (2) gives the equilibrium price index for region *r*, where this price index is higher if a region has to import a relatively larger part of its manufactured goods from more distant regions. Note that the price index *I* depends on the wages *W.* Equation (3) simply states income in region *r,* Y_r, has to equal the labor income earned in that region, where λ_r is region *r*'s share of the total manufacturing labor force *L*.

The main aim of our empirical research is to find out whether or not a spatial wage structure, that is a spatial distribution of wages in line with equation (1), exists for Germany. Equation (1) cannot be directly estimated as there are typically no time series of local price indices for manufactures (where local refers to the US county level in Hanson's study and to the city-district level in our case). And, even more problematic (see equation (2)), the price index *I* is endogenous, and inter alia depends on each of the local wage rates, which makes a reduced form of equations (1) and (2) extremely lengthy and complex. These problems have somehow to be solved in order to estimate a spatial wage structure for Germany.

Hanson uses the following estimation strategy based on the remaining two equilibrium conditions. In order to arrive at a wage equation that can actually be estimated he rewrites the price index in exogenous variables which can actually be observed for his sample of US counties.

First, he uses:

$$(4) \quad P_r H_r = (1-\delta) Y_r$$

Equation (4) states that the value of the fixed stock of housing equals the share of income spent on housing, where P_r is the price of housing in region r, H_r is the *fixed* stock of housing in region r and $(1\text{-}\delta)$ is the share of income spent on housing and δ is thus the share of income spent on manufactures.[6]

Second, real wage equalization between regions is assumed:

$$(5) \quad W_r \Big/ P_r^{1-\delta} I_r^{\delta} = W_s \Big/ P_s^{1-\delta} I_s^{\delta}$$

Equation (5) is quite important. It is assumed that the economy has reached a long-run equilibrium in which real wages are identical. This implies that labor has no incentive to migrate (interregional labor mobility is solely a function of interregional real wage differences).[7] The assumption of interregional labor mobility and the notion that agglomeration leads to interregional wage differences are not undisputed for a country like Germany with an allegedly "rigid" labor market, see in particular *Puga* (2001, p. 18) for implications of low labor mobility and no interregional wage differences from a new economic geography perspective. We return to this issue in section 5.

The importance of a non-tradable housing sector as a spreading force is implied by (5). A higher income Y_s implies, *ceteris paribus*, higher wages in region *r,* see equation (1), but it also, given the stock of housing, puts an upward pressure on housing prices P_r, equation (4). Combining (4) and (5) allows us to rewrite the price index in terms of the housing stock, income and nominal wages. The equilibrium condition for the housing market can be written as $P_r=(1-\delta)Y_r/H_r$ and this expression for P_r is then substituted into equation (5) which defines the price index I_r in terms of W_r, Y_r and H_r. Substituting this

6 Note, that direct observation of a housing price index could serve a similar purpose. We will return to this in section 4.

7 *Overman*, *Redding* and *Venables* (2001, p. 17) discuss how the model used by Hanson can be seen as a specific version of a more general new economic geography model.

in (1) results in a wage equation which can be estimated. This will also be the benchmark wage equation in our empirical analysis.

$$(6) \qquad \log(W_r) = k_0 + \varepsilon^{-1} \log\left(\sum_s Y_s^{\varepsilon+(1-\varepsilon)/\delta} H_s^{(1-\delta)(\varepsilon-1)/\delta} W_s^{(\varepsilon-1)/\delta} T^{(1-\varepsilon)D_{rs}}\right) + err_r$$

Where k_0 is a parameter and err_r is the error term. Equation (6) includes the three central structural parameters of the model, namely share of income spent on manufactures, δ, the substitution elasticity, ε and the transport costs, *T.* Given the availability of data on wages, income, the housing stock, and a proxy for distance, equation (6) can be estimated. The dependent variable is the wage rate measured at the US county level and Hanson finds strong confirmation for underlying model to the extent that the three structural parameters are significant and have the expected sign which, in terms of equation (6), means that that there is a spatial wage structure. In section 4 we will begin our empirical inquiry of the German case by estimating equation (6) for our sample of German city districts.

3.3 Data and Estimation Issues

Before we turn to the estimation results a few words on the construction of our data set are in order. Germany is administratively divided into about 441 districts (*Kreise*). Of these districts a total of 119 districts are so called city-districts (*kreisfreie Stadt*), in which the district corresponds with a city. 114 of these city districts are included in the sample. We use district statistics provided by the regional statistical offices in Germany. The data set contains local variables, like the value added of all sectors in that district (GDP), the wage bill and the number of hours of labor in firms with 20 or more employees in the mining and manufacturing sector. Combining the latter two variables gives the regional wage W_r, which is measured as the average hourly wage in the manufacturing and mining sector. Since we also want to analyze the cities' *Hinterland* we also included 37 aggregated (country) districts, constructed from a larger sample of 322 country districts.[8] The total number of districts in our sample is thus 151, namely 114 city districts and 37 country districts. Transport costs are, of course, a crucial variable.

8 From a total of 441 districts we subtract the 119 city-districts and this gives us 322 country districts. Many of these 322 country districts are very small. In order to arrive at a geographical unit that is more in line with that of the city-district we decided not to use the 322 corresponding *Kreise* but to use a larger geographical unit of analysis the so called *Bezirke* and this reduces the 322 districts to the 37 country districts, Furthermore, this simplifies the distance matrix considerably.

We do not use the geodesic distance between districts, because this measure does not distinguish between highways and secondary roads. Instead, distance is measured by the average number of minutes of travel by car it take to get from city district A to city district B. The data are obtained from the Route Planner 2000 (Europe, And Publishers, Rotterdam). For the data on the housing stock H_r, required to estimate equation (6), we use the number of rooms in residential dwellings per district. In some of our estimations we also include one or more of the following regional variables, unemployment, employment, income (personal income tax base) and land prices *(Baulandpreise)*.

Since we only have one observation for each variable per district for the average hourly wage and for GDP (for 1995 and 1994 respectively) we have to estimate the wage equation in levels and we therefore also restrict ourselves to cross-section estimations. The estimation of an equation like equation (6) raises several estimation issues. First of all, there is the issue of the endogeneity of particular right hand side variables like Y_s. In our case this problem is somewhat reduced by the fact that wage data are for 1995 and GDP data are for 1994 (and thus precede the wage data). At any rate this still leaves, however, the local wage rate itself as endogenous variable (see W_s in equation (6)). To check for this we have experimented with instrumental variables (IV) in our estimation. As always, it is difficult to find good instruments and we (inter alia) used the size of districts, the size of the district' s population and the population density as instruments. The main conclusion is that these IV-estimations do not lead to different results, so we do not report them below. If we would have been able to use multiple years of observation for each variable, the time-series element of the data would have allowed us, as in Hanson's work, to estimate in first differences and it would then also be worthwhile to experiment with different geographical units of analysis. Estimation in first differences would allow us to deal with time-invariant, district-specific effects that may have a bearing on district-wages. This is not possible in our cross-section setting. With respect to the geographical unit of analysis, in our estimations the left and right hand side variables are both measured at the district level. In *Hanson* (1998, 1999) or *Roos* (2001) the latter are typically measured a higher level of aggregation (e.g. the US state and *Bundesland* level*)* so as to make it less likely that a shock to district wages W_j has an impact on Y_s or W_s. On the other hand, less geographical aggregation of the data makes it less likely that location-specific shocks (via the error-term in equation (6)) have an impact on the independent variables.

Related with this last observation is another estimation issue, namely that the variance of the error-term varies systematically across the various districts. To address the issue of heteroscedasticity we applied the Glejser-test and used weighted least squares (WLS). Therefore, we estimated (via non-linear least squares, NLS) equation (6) or any other of our specifications and we then we regressed the (absolute of the) resulting residuals on the right hand side variables. A significant impact of these variables on the residuals indicates heteroscedasticity and for every specification it turned out that this is indeed something that has to be taken into account. To deal with this we therefore used weighted least squares (WLS) estimations where the weights are for each specification taken from the estimation results from regressing the absolute residuals (from the "unweighted" NLS estimation) on the right hand side variables.

4 BASIC ESTIMATION RESULTS FOR GERMANY

4.1 Estimating the Benchmark Wage Equation

We now turn to the attempt to estimate the structural parameters using the wage equation (6) for Germany. In doing so, we will not only be able to estimate the structural parameters δ, ε and *T* (and to establish the existence of a spatial wage structure) but we can also verify the so-called no-black hole condition, which gives an indication for the convergence prospects in Germany. In section 4 we first estimate equation (6) and then discuss three alternative estimation strategies.

Table 1a gives the estimation results for the estimation of equation (6). We also included a dummy variable for East German districts and a dummy variable for country districts.[9] The dummy for East German districts is motivated by the fact that wages (and labor productivity) in East Germany are lower than in West Germany. As the inclusion of these two dummies turned out to be immaterial for the conclusions with respect to the structural parameters they are not reported here but we will return to them in subsequent estimations.[10]

9 Inspection of the wage data revealed that there is one very large outlier, the district of Erlangen in Bavaria, which has by far the highest wage so we included dummy for this district as well.

10 In our estimations we consider Germany to be a closed economy, elsewhere (see *Brakman*, *Garretsen* and *Schramm*, 2000) we have checked whether the inclusion Germany's main trading partners would influence te outcomes but this was not the case. We did not control for fixed regional endowments as

Table 1a: Estimating the Structural Parameters for Germany

	Coefficient	Standard error	t-statistic
δ	2.449	1.179	2.076
ε	3.893	0.473	8.226
Log(T)	0.009	0.001	9.162
Adj. R^2 = 0.95; number of observations = 150; weighted least squares			
Implied values:			
ε/(ε-1)	1.343	ε(1-δ)	-5.46

All three structural parameters are found to be significant and they also have the correct sign thereby validating the Helpman-Hanson model. The substitution elasticity ε is significant and the coefficient implies a profit margin of slightly above 30% (given that ε/(ε-1) is the mark-up), which is fairly reasonable, although higher than found for the US by *Hanson* (1998, 1999). Note that the value ε(1-δ) is used to determine whether a reduction of transport costs affects spatial agglomeration of economic activity: the so called no black hole condition for the *Helpman* (1998) model holds if ε(1-δ) <1 (see below).[11]

The coefficient for δ is, however, (implausibly) large because it indicates that Germans do not spend any part of their income on housing (see equation (2)). The high value is in accordance with the findings of Hanson, who also finds that δ is large for the USA (above 0.9 and in some cases also not significantly different from 1). Finally, the transport cost parameter has the expected sign and is highly significant. All in all, the estimation results provide support for the idea of a spatial nominal wage structure, to see this substitute the estimated coefficients into wage equation (6) and one can see how *ceteris paribus* the presence of nearby large markets (hence low T and high Y) increases wages in district j. Given the fact that we find that δ is clearly not significantly lower than *1*, H_s does not exert an impact on wages, but W_s does.[12]

f.i. climate. *Hanson* (1999) does control for these endowments in his study for the USA but for a relatively small country like Germany these kind of differences are thought not to be relevant.

11 In *Krugman* (1991) the no black hole condition is met if *ε(1-δ)>1*. *Helpman* (1998) shows how this difference is ultimately due to the fact that the spreading force in the Krugman model is a homogeneous tradable good (the agricultural good) whereas in the Helpman model it is a homogeneous non-tradable good (housing which is in fixed supply) is responsible for this difference.

12 Restricting δ to actual values of the share of income spent on non-tradable services (or non-tradable housing services) has virtually no impact on the estimated size and significance of the transport costs *T*, or on the explanatory power of the estimated equation, which is still able to explain 46% of the

Furthermore, Table 1a enables us to see whether or not the *no black hole condition* is met. It is indeed the case that $\varepsilon(1-\delta)<1$, although not significantly (except for the case in which δ is fixed, see however footnote 13). This implies that transport costs has an impact on the degree of agglomeration, that is to say agglomeration is not inevitable if transport costs can be sufficiently reduced. For Germany this seems to indicate that a lowering of transport costs might lead to more even spreading of economic activity, which is good news for the peripheral districts, the bulk of which is located in Eastern Germany. In the Helpman-Hanson model if $\varepsilon(1-\delta)>1$, this means that a region's share of manufacturing production is a function of its (fixed) relative housing stock only (*Helpman*, 1998, p. 40).

The estimation of wage equation (6) provides some empirical support for the new economic geography approach, here the Helpman-Hanson model, and our estimations for Germany lead to similar conclusions as Hanson's estimation for the USA. At the same time it is, however, clear, that the economy of post-reunification Germany differs in a number of important respects from the US economy to the effect that the our estimation results might be improved upon if we take more "German features" on board. This is the subject of the next section, where we will analyze the effects of changes in the basic wage specification as given by equation (6). Before we address these German features, we first turn to three alternative strategies to estimate wage equation (6). The first strategy is to use land prices as a proxy for the housing prices which makes equation (4) redundant. The second and third strategy deal with the dismissal of the assumption of real wage equalization (recall equation (3)) and this will be discussed in the next subsection.

We do not have data on housing prices but instead we use land prices (*Baulandpreise)* as a proxy. From Figure 3 in section 2 we already know that, at least at the state level, land prices, are much higher in states with a higher GDP. An estimation of wage equation (6) with these price data provides a more direct test of the Helpman-Hanson model, because the influence of agglomeration on prices of local non-tradables is driving the spreading force in the Helpman-Hanson model. Table 1b gives the estimation results for 146 districts,[13] (we only have data on land prices for a subset of our districts). The

variance in wages, as compared to 48% in the unrestricted specification. A likelihood-test indicates that the restricted model has to be rejected as being inferior compared to the unrestricted model.

13 For 1 East German city district and 5 West German city districts there are no data on land prices. So they are excluded, except for Hamburg, which is also a (city) state.

Table 1b: Estimating the Wage Equation with Land Prices

	Coefficient	Standard error	t-statistic
δ	0.577	0.0193	29.76
ε	3.822	0.2866	13.33
Log(T)	0.0078	0.0010	7.389
Adj. R^2 = 0.98; number of observations = 146; weighted least squares			
Implied values:			
ε/(ε-1)	1.3	ε(1-δ)	1.63

three structural parameters are again clearly significant. The main differences with Table1a is that the δ-coefficient is now found to be lower. The latter is especially relevant since now we find that δ is significantly smaller than 1 which indicates that a significant part of income (1-0.57) is indeed spent on housing and that the housing sector can indeed act a spreading force (the actual share of income spent on manufactures in Germany was 0.68 which corresponds with our estimated share of 0.57). The other main difference with table 1a is that the "no black hole condition" is no longer met (ε(1-δ)=1.63>1), which would imply that the spatial distribution of economic activity (and hence of district wages) only depends on the (fixed) distribution of the housing stock and that it would not depend on the level of transportation costs at all.[14]

4.2 No Real Wage Equalization

The assumption of real wage equalization boils down to imposing a long-run equilibrium and this (implicitly) implies a sufficient degree of labor mobility and wage flexibility. In general, the requirement that interregional real wages are equal by assumption is not very appealing because it always assumes that the economy is in a long-run equilibrium. Furthermore, specifically in the German case this assumption seems at odds with the stylized fact that (real) wages differed between eastern and western German regions at the start of the reunification process. Our second and third alternative estima-

14 The standard deviation of the estimation of ε(1-δ) indicates that the estimated value (1.63) is significantly greater than 1 (standard deviatio 0.174).

tion strategies are to estimate a wage equation and the structural parameters *without* invoking real wage equalization.

The second strategy is simply to re-estimate equation (6) with land prices as our proxy for P_s (see Table 1b) and by adding the possibility for a real wage differential *between* (but not within) East and West Germany (see **appendix 1** for a derivation of the resulting wage equation). The coefficient ϕ captures the east-west German real wage differential and is constructed in such a way that, for instance because of trade union preferences, $\phi>1$ indicates that real wages in western Germany are higher. Table 2 shows that indeed $\phi>1$. For the other coefficients the estimation results in Table 2 are very much in line with the ones reported in table 1b.[15] It is by no means obvious that *real* wages should be higher in western Germany. There is some evidence (see for instance Sinn, 2000) there is evidence that following the reunification there has been a process of real wage equalization between the former FRG and GDR. Nominal wages are higher in western Germany but the data show (see also our Maps 1 and 3) that housing rents(!) and other local prices are also considerably higher on average in the western part of Germany thereby fostering, like the Helpman-Hanson model predicts, a tendency towards real wage equalization.[16] The latter might be true but our estimation results that (at least in the mid 1990s) this equalization was still a long way off.

Table 2: Estimating the Wage Equation with Incomplete Real Wage Equalization between East and West Germany

	Coefficient	Standard error	t-statistic
δ	0.6713	0.02756	24.355
ε	3.6555	0.26612	13.736
Log(T)	0.0106	0.00124	8.5333
ϕ	1.4058	0.29487	4.7676
Adj. R^2 = 0.98; number of observations = 146; weighted least squares			

Table 2 captures the possibility of real wage differences between western and eastern Germany but it is still rather stringent to the extent that it assumes that within eastern

15 We also estimated the wage equation with ϕ using the housing stock H_s instead of our proxy for the housing price P_s and this also resulted in an insignificant ϕ coefficient.

16 *Ströhl* (1994) shows that the consumer price level in eastern German cities is 6% below the consumer price level in western German cities.

and western Germany real wage equalization holds. We thus still need equation (3) to estimate the wage equation. As our third estimation strategy we show how to estimate a wage equation that is based on a reduced form of equations (1) and (2) with its structural parameters without assuming real wage equalization beforehand. For this purpose it is necessary to simplify the price index defined in equation (2) by not considering all prices in all regions. Instead we consider only two prices: the price in region *r* of a manufactured good produced in region *r* and the *average* price outside region *r* of a manufactured good produced outside region *r*. For the determination of the simplified local price index for manufactures it also necessary to have a measure of *average* distance between region *r* and the regions outside. The distance from the economic center is an appropriate measure. This center is obtained by weighing the distances with relative Y.[17] The economic center of Germany turns out to be *Landkreis Giessen* (near Frankfurt), which is in the state of Hessen, West Germany. Equation (2) now becomes:

$$(4') \quad I_r = \left[\lambda_r W_r^{1-\varepsilon} + (1-\lambda_r)\left(\overline{W}_r T^{D_{r-center}}\right)^{1-\varepsilon}\right]^{1/1-\varepsilon},$$

where $\overline{W}_r$ is the average wage outside region *r*, $D_{r\text{-}center}$ is the distance from region *r* to the economic center, and weight λ_r is region *r*'s share of employment in manufacturing, which is proportional to the number of varieties of manufactures.

This simplified price index makes it possible to directly estimate wage equation. Since we apply the wage equation to Germany we also take into account that the marginal productivity of labor (MPL) in East Germany is lower than in West Germany. A uniform level of MPL in the West, θ_{west} and the East θ_{east}, is assumed but the MPL of the East is lower then the MPL in the West. Incorporating this difference means that the wage equation (1) and the simplified price index equation (2') change into:[18]

---

17 For each region *r* the weighted average distance to the other regions $\sum_s weight_s D_{rs}$ is calculated, using $weight_s = Y_s / \sum_j Y_j$. The region with the smallest average distance is the economic centre.

18 Employment in a typical Western firm in a typical Western region *r* for the production of manufacturing variety *i* is $\alpha + \beta x_{ir}$, where α is the fixed cost parameter and β is the marginal costs parameter. Employment in a typical Eastern firm in a typical Eastern region *r* is $\alpha + \beta x_{ir}(\theta_{west}/\theta_{east})$. We thus assume that marginal labor costs in East Germany are higher than in West Germany which is the same as assuming that $MPL_{west} > MPL_{east}$. Sales of a firm located in region *r* equals total demand for its product. Dropping subscript *i* for the individual firm:

$$\frac{(\varepsilon-1)\alpha}{\beta(\theta_{west}/\theta_r)} = \sum_{s=1}^{R}\left[\left(\frac{\varepsilon}{\varepsilon-1}\frac{\beta W_r T^{D_{rs}}(\theta_{west}/\theta_r)}{I_r}\right)^{-\varepsilon} T^{D_{rs}} \delta \frac{Y_r}{I_r}\right]$$

Which gives (1') above, where $\theta_{west}/\theta_r = 1$ if *r* is in West, and $\theta_{west}/\theta_r > 1$, if *r* is in East. Ideally, one would like to use district-data on productivity here, see for instance *Funke* and *Rahn* (2000).

$$(1') \quad W_r = \text{constant} \cdot \left(\frac{\theta_{west}}{\theta_r} \right)^{(1-\varepsilon)/\varepsilon} \left[\sum_{s=1}^{R} Y_s \left(T^{D_{rs}} \right)^{1-\varepsilon} I_s^{\varepsilon-1} \right]^{1/\varepsilon}$$

$$(2'') \quad I_r = \left[\lambda_r \left(W_r \frac{\theta_{west}}{\theta_r} \right)^{1-\varepsilon} + (1-\lambda_r) \left(\overline{W} T^{D_{r-center}} \right)^{1-\varepsilon} \right]^{1/(1-\varepsilon)}$$

Equation (2'') is finally substituted into (1'), which provides us with the reduced form of the equilibrium wage equation without having to invoke real wage equalization in order to approximate (2). The equation to be estimated is:

$$(6') \quad \log(W_r) = \kappa_0 + \varepsilon^{-1} \log \left[\sum_{s=1}^{R} Y_s \left(T^{D_{rs}} \right)^{1-\varepsilon} I_s^{\varepsilon-1} \right]$$

where $I_r^{1-\varepsilon} = \left[\lambda_r \left(W_r \left(1 + \kappa_1 D_{east} \right) \right)^{1-\varepsilon} + (1-\lambda_r) \left(\overline{W} T^{D_{r-center}} \right)^{1-\varepsilon} \right]$

and where D_{east} = dummy variable which equals 1 if r is an East German district.

Table 3 shows the regression results of estimating equation (6'). The parameter κ_1 is set equal to zero, as it turned out to be not significantly different from zero, implying that the productivity difference was not significant. An additional advantage of equation (6') compared to the basic wage equation (6) is that the share of income spent on manufactures δ (which we thus found to be rather large in our initial estimation in Table1a) does not need to be estimated now.

Table 3: Estimating Equation (6') for Germany

	Coefficient	Standard error	t-statistic
ε	4.3993	0.6311	6.9700
LogT	0.0073	0.00025	29.352
κ_0	1.7899	0.3011	5.9432
Adj. R^2 = 0.99; number of observations = 150; weighted least squares			

The results in Table 3 show that the distance parameter is significantly positive, and virtually identical to previous estimates, and the same holds for ε which indicates the robustness of the estimated parameters with respect to the estimated specification. Again, see equation (6'), the results support the notion that nominal wages in district *r* are higher if this region has a better access (in terms of distance) to larger markets.

All in all, the estimation of the basic wage equation (6) for Germany (see Table 1a) and the three alternative estimation strategies pursued in section 4 (see Tables 1b, 2 and 3) provide some support for the empirical relevance for the Helpman-Hanson model for Germany, that is to say, it turns out to come up with significant results for the key parameters. There are, however, two main "concerns". The first is that in the wage equation with the housing stock as independent variable, the share of income spent on manufactures is too large and this renders the housing sector as spreading force irrelevant. The second concern, and opposed to the first issue, is that when we use land prices instead of the housing stock the spatial wage structure seems only to depend on the fixed distribution of the housing stock because the no black hole condition is no longer met. The implication of the latter is that with a fixed spatial distribution of non-tradable goods (e.g. housing), changes in transportation costs will not lead to changes in existing core-periphery patterns in Germany. From the perspective of the convergence prospects following German re-unification this would not be very good news.

5 INCORPORATING GERMAN FEATURES: TRANSFERS AND RIGID LABOR MARKETS

Two features of the German economy might have special consequences for the spatial distribution of wages; interregional transfers and the functioning of the labor market. We will first turn to the issue of the transfers and then in section 5.2 to the labor market.

5.1 Distinguishing between GDP and Personal Income: The Role of Transfers

So far we took for the size of the market in a region, GDP (measured as value added) in that region, where region thus refers to one of the 114 city-districts or one of the 37 country-districts in our sample. The size of the market can also be approximated by taking personal income instead of GDP. In the absence of large intra-regional transfers, differences between the two measures will be small, but in the case of post-reunification Germany, one is less sure whether this is true. The main reason being the massive income transfers from western to eastern Germany. For eastern Germany as a region, income clearly is larger than GDP in 1995. To see whether this influences our results we replaced regional GDP by the regional local income tax base and with this alternative measure of Y_s we re-estimated the basic wage equation (6).[19] Table 4 gives the estimation results.

Table 4: Estimating Equation (6) with Income Instead of GDP[20]

	Coefficient	Standard error	t-statistic
ε	3.704	0.339	10.913
δ	1.486	0.095	15.640
log(T)	0.0067	0.0006	9.975
Adj. $R^2 = 0.95$; number of observations = 150; weighted least squares			

Comparing Table 4 with Table 1a makes clear that the results for the local income-regression are comparable to those for the local GDP-regression and that here also the share of income spent on manufactures exceeds 1.

Some regions receive more transfers than others. In order to control for these differences we constructed two new variables RY_{GDP} and RY_{inc} where RY_{GDP} = *(district GDP/German GDP)* and where RY_{inc} = *(district personal income tax base/German personal income tax base).* We are in particular interested in the ratio of these two variables (RY_{GDP}/RY_{inc}). A district for which this ratio is greater than 1 indicates that this region is a (net) donor of transfers (our measurement of transfers does not only include public transfers but the transfers of factor income as well). If this ratio is smaller than 1 this means this region's income and not so much its GDP exceeds the German averages and indicates that this region is a (net) recipient of transfers.

Given the massive transfers one might expect that (RY_{GDP}/RY_{inc}) is relatively low for East German districts. We checked for this (not shown here) and this is indeed the case. It is also true that (for instance due to commuting and subsidies to the agricultural sec-

19 To be able to compare results with the estimations with GDP (as shown by Table 2) we stick to equation (6). Data for 1995 on the income tax base (*Gesamtbetrag der Einkünfte)* at the district level were taken from the Statistik Regional database of the Federal Statistical Office of Germany.

20 Re-estimating this equation, using P_s instead of H_s, gives similar results.

tor) that this ratio is also relatively low for 37 country districts. One would like to know how this ratio affects regional wages. We therefore again estimated (6) but now with κ log(RY_{GDP}/RY_{inc}) added as an additional term and where κ is the coefficient to estimated. The results are summarized in Table 5.

Table 5: Estimating Relative Importance of Transfers on the Spatial Wage Structure

	Coefficient	Standard error	t-statistic
ε	4.077	0.3312	12.308
δ	1.3877	0.0608	22.823
log(T)	0.0076	0.00069	11.036
κ	0.3368	0.0650	5.1779
Adj. R^2 = 0.534; number of observations = 151; non-linear least squares			

The results for the three structural parameters are very similar to those reported in Tables 1a and 4. Our main interest here is with the κ-coefficient and here we find that regions with a relatively (RY_{GDP}/RY_{inc}) ratio have significantly higher wages which suggests that for local wages the economic size of that region in terms of its GDP matters more than its income. This also means that if a region receives a relatively large amount of transfers ($RY_{GDP}/RY_{inc}<1$) there is no upward effect on nominal wages. The main conclusion to be taken from Tables 4 and 5 is that the results for our central wage equation are not changed by much if we replace GDP by income.

5.2 A spatial (un)employment Structure and Land Price Structure?

The estimation results for Germany provide support for the Helpman-Hanson model in the sense that the key model parameters are found to be significant. Given the coefficients and in line with equilibrium wage equation (1), our central equation (6) illustrates that W_r is higher if district r is situated more closely to regions with a relatively high Y. We thus find confirmation for a spatial wage structure for Germany: regional wages become lower the further one moves away from manufacturing centers. To some extent this is a surprising result. Certainly compared to the case of the USA, the German labor market is considered to be rigid where the rigidity refers for instance to the idea agglomeration need not go along with interregional wage differences if, for whatever institutional reason, interregional wages are set at the same level. For a country like Germany one might thus very well expect that the spatial distribution of *Y* does *not* get reflected in spatial wage differences (see *Puga*, 2001 for this assertion for Germany).[21]

As we explained in section 3.1, the Helpman-Hanson model belongs to a class of new economic geography models in which a fall in transportation costs from a very high to a very low level typically results in spreading→(partial) agglomeration→ renewed spreading. If, however, agglomeration simply cannot lead to interregional wage differences the outcome will not only be, when trade costs fall from their intermediate to a very low level, that agglomeration continues to exist, but also that agglomeration "may get reflected instead into differences in unemployment rates" (*Puga*, 2001, pp. 18-19).

Is this last observation relevant for Germany? No definite answers are possible if only because we do not know if Germany 4 years into reunification was in 1995 anywhere near the regime of very low trade costs and we also only have cross-section data. In addition, *Puga* (1999, 2001) is also much more concerned with real instead of nominal wages. But the issue *Puga* (2001) raises is interesting in its own right: is there a spatial unemployment structure? Like most new economic geography models, the Helpman-Hanson model does not allow for unemployment so we can not test an unemployment version of equation (6). But, as a second best solution, we can use a market potential approach, which captures some important elements of the new geography approach, but in a less sophisticated way. The central idea is that unemployment of a specific region is a function of how easy this region has access to large surrounding regions. The more easy this access, the lower unemployment: if it is true that $W_r=W_s$ one would expect that "agglomerated" regions have a lower unemployment rate. Our market potential equation for unemployment has the following form[22]:

(7) $\log (U_r) = K_1 \log[\Sigma_s Y_r e^{-K2Djs}] + K_3 D_{east} + K_4 D_{country} + \text{constant}$

where U_r=unemployment rate in region *r,* unemployment data are for 1996.

---

21 Interregional wage differences are for instance not feasible if a union ensures centralised wage setting that is, irrespective of regional economic conditions, $W_r=W_s$ (see *Faini*, 1999). Centralised wage setting (at the industry level) is a tenet of the German labor market, see also Appendix 2.

22 The specification of equation (7) is similar to the "simple" wage equation used by Hanson, apart from the two dummies, the only difference is that U_r instead of W_r is the left-hand side variable. *Brakman*, *Garretsen*, *Schramm* (2000) test this market potential *wage* equation for the 114 German city-districts and find strong confirmation for the existence of a spatial wage structure or wage gradient. For the implications of introducing wage rigidity in a new economic geography model see *Peeters* and *Garretsen* (2000).

For a spatial *un*employment structure to exist it is crucial that coefficients K_1 and K_2 are significant. Table 6 shows, however, that this is not the case. The hypothesis of a spatial unemployment structure must be rejected and only the two dummy-variables are significant (with D_{east} capturing the idea that unemployment in Eastern Germany is indeed much higher than in Western Germany).[23]

Table 6: A Spatial Unemployment Structure?

	Coefficient	Standard error	t-statistic
K_1	0.0435	784.43	0.00055
K_2[F1]	-0.0001	0.2419	-0.000559
K_3	0.3519	0.0680	5.1357
K4	-0.1581	0.0399	-3.9534
Adj. R^2 = 0.368; number of observations = 151; non-linear least squares			

Because unemployment is to some extent a matter of definition we also turn to regional employment. With interregional nominal wage equalization (caused, for instance, by centralised wage setting) we test if we can observe a spatial *employment* structure under the restriction that $W=W_r=W_s$ due to centralised wage setting. In Appendix 2 this employment equation is derived and equation (8) below has been estimated (the scaling of employment is in line with Hanson (1998, 1999) who also estimates for a spatial employment structure. Hanson does, however, not derive the employment equation from the underlying model).

$$(8)\quad \log\left(L_r / area_r\right) = \text{constant} + \log\left(\sum_{s=1}^{R}\left[e^{-c_0 D_{rs}} Y_s\right]\right) + c_1 D_{east} + c_2 D_{country}$$

L_r = employment in district r measured in hours of employment in the manufacturing and mining sector scaled by the size of district r (in km^2), data for 1995.

23 This is not to deny that district unemployment is irrelevant. If we re-estimate equation (7) as a market potential function with wages W_r as the left hand side variable and with district unemployment as the additional explanatory variable U_s, it turns out that we still find a spatial wage structure but the unemployment-coefficient has a negative sign and is significant thereby suggesting, in line *with Blanchflower* and *Oswald* (1994), the existence of a wage-curve on the regional level where higher unemployment means lower local wages. The unemployment-coefficient is –0.024 (t-value 3.96). Since heteroscedasticity is not issue here, equation (7) is estimated with non-linear least squares.

The constant, c0, c1 and c2 are to be estimated. See Appendix 2 for the derivation that $c_0 = (\varepsilon\text{-}1)\log(T)$, and $c_1 D_{east} = (\varepsilon\text{-}1)\log(\vartheta_j)$, with ϑ_j being a measure of the productivity gap between East and West Germany.

Table 7 below shows that we can confirm the existence of a spatial employment structure because of the sign and significance of the c_0 coefficient which implies that employment in region *j* is higher if this regions is situated more closely to economic centers. Note that the c_1 and c_2 coefficients are also significant, indicating a lower employment in East German and country districts (given that heteroscedasticity was not found to be an issue, results in Table 7 are based on a NLS-estimation).[24]

Table 7: A Spatial Employment Structure?

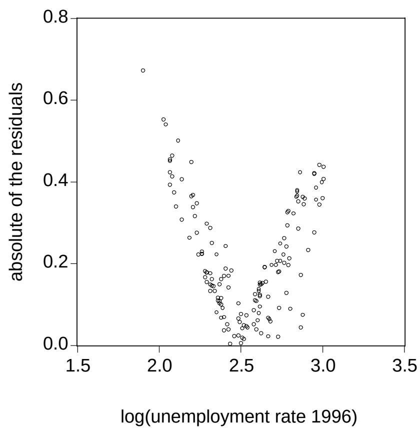
	Coefficient	Standard error	t-statistic
C_0	0.0029	0.0012	2.334
Const.[F2]	-0.8792	0.1402	-6.268
C_1	-9.3957	0.2515	-37.34
C_2	-2.1427	0.1077	-19.88
Adj. R^2 = 0.767; number of observations = 151; non-linear least squares			

The question arises how to reconcile a spatial employment structure with the absence of spatial *un*employment structure. After all, regional unemployment is the regional labor supply minus the regional employment. We can only speculate on some explanations, but empirical evidence indicates that unemployment is less responsive to agglomeration and spreading forces described in the model in an economy which relies relatively less on market forces and in which long-term unemployment leads to reduced employability thereby reducing effective labor supply (see also *Decressin* and *Fatas*, 1995). The fact that we find confirmation for the empirical relevance of both wage equation (6) and employment equation (8) is in line with the idea that Germany finds itself in middle position between the 2 extremes of full labor mobility and wage flexibility and complete labor immobility and wage rigidity. Finally, inspection of the residuals of the estimated unemployment equation (7), see Figure 1, reveals that with respect to regional unem-

24 To estimate equation (7) we used the following starting values for ε, log(T) and $\log(\vartheta)$ (based on our estimations in section 4): 4, 0.007 and 0.4.

ployment there are two distinct groups of districts. The first group (with almost every eastern German district) has a relatively high unemployment rate and the second group has a much lower unemployment rate. Estimating the unemployment equation (7) for the full-sample is therefore probably not appropriate to start with.

Figure 1: Residuals and Unemployment Rate (full sample estimation of eq. (7))



The reason to stick to the Helpman-Hanson model is not only that it seems to perform reasonably well for Germany but also that, as we have said before, it combines the best features of the demand linkages model due to *Krugman* (1991) with the input-output linkages model due to Venables (1996) and Krugman and Venables (1995). The inclusion of housing as a non-tradable consumption good lies at the very heart of *Helpman* (1998) and to illustrate (nothing more but certainly also nothing less) that housing prices may indeed act as a spreading force we have finally estimated equation (9). This estimation is also inspired by Maps 1-3 in section 2 where we showed that German states with relatively high wages and GDP also display higher land prices. As we explained before there are no district data on housing prices but we have German district data on land prices which serve as good 1st approximation. The estimation results, see Table 8, show that there is a "spatial land price" structure (see the coefficients K_1 and K_2) and this is precisely what the Helpman model predicts and this confirmation of such a structure also indicates that indeed the housing market can be looked upon a providing a spreading force. Also in line with other German evidence (see Sinn, 2000) is that land prices are significantly lower in country districts but notably also in East German districts. As with the estimation of the various wage equations throughout our paper, estimation of equation (9) indicated that heteroscedasticity is an issue so we used a weighted least squares estimation (WLS).

(9) $\log(LP_r) = K_1 \log[\Sigma_s Y_r e^{-K2Djs}] + K_3 D_{east} + K_4 D_{country} + \text{constant}$

where LP=land prices per m^2.

Table 8: A Spatial Land Price Structure

	Coefficient	Standard error	t-statistic
K_1	0.4077	0.0600	6.7874
K_2[F3]	0.0632	0.0165	3.8272
K_3	-0.5855	0.1425	-4.1079
K4	-1.3607	0.1299	-10.4685
constant	1.3881	0.6863	2.0224
Adj. R^2 = 0.726; number of observations = 151; weighted least squares			

6 CONCLUSIONS

The recent advances in the field of new economic geography have increased our understanding of spreading and agglomerating forces in an economy. Empirical testing, however, is difficult. Not only because the core models are characterized by multiple equilibria, but also because the lack of specific regional data makes approximations inevitable. Short-cuts cannot be avoided. Here we have tried to find evidence whether or not new economic geography models are in principle able to describe the spatial characteristics of an economy; here Germany. The answer basically is, yes. We found that the so-called Helpman-Hanson model, using data for Germany, confirms the idea of a spatial wage structure. The advantage of the Helpman-Hanson model is that it incorporates the fact that agglomeration of economic activity increases the prices of local (non-tradable) services. Thus providing a power full spreading force, and leading to less extreme outcomes than the core model of the new economic geography as described by *Krugman* (1991). The reason to choose Germany is that in the case of Germany and the fall of the Berlin Wall in 1989, there is a very obvious candidate for the kind of controlled or natural experiment that would address the problems with endogeneity that surround the estimation of new economic geography models. *Overman*, *Redding* and *Venables* (2001, p. 20) rightly point to *Hanson* (1997) as an example of such a natural experiment but we think the fall of Berlin Wall and hence the start of German re-unification is precisely the kind of exogenous shock one is looking for. However, in order to be able to say something conclusive about the relevance of the new econmic geography for the case of German re-unification we need to move away from the cross-section estimations upon which the present paper is based. In our ongoing research on Germany and the new economic geography we will therefore test the new economic geography approach before (1985), at about (1991) and after (1995) the start of German re-unification. The present paper only indicates that such an attempt is worthwhile because it is possible to test the key equations of the underlying model for Germany.

What are the next steps which have to be taken? The first thing that comes to mind is thus the question how the analysis can be made dynamic. In section 3.1 we indicated the importance of an Ω characteristic of these models, i.e. in the first phase of economic integration (relative high transportation costs) economic activity is spread across locations, when transportation costs starts to fall agglomeration starts, as demand and cost linkages make it advantageous to agglomerate, however, this drives up prices of local non-tradable services (here, housing) leading again to spreading of economic activity. Due to data limitations we have only cross-section estimations which gives no information about the position of the German economy on the Ω-curve. But for now we are satisfied that the recent theoretical advances in the field of new economic geography do find at least some support in the data.

Appendix 1: The Wage Equation with Incomplete Real Wage Equalization

Assume full real-wage equalization within East Germany and within West Germany, but not between East and West Germany (incomplete real-wage equalization).

We start with

$W_r = \left[\sum_s Y_s I_s^{\varepsilon-1} T^{D_{rs}(1-\varepsilon)}\right]^{1/\varepsilon}$ (wage equation (1)) and, instead of equation (3), $\frac{W_r}{P_r^{1-\delta} T_r^{\delta}} = \frac{W_s}{P_s^{1-\delta} T_s^{\delta}} . \varphi_{rs}$, where $\varphi_{rs}=\phi>1$, if r is West and s is East, and $\varphi_{rs}=1/\phi<1$, if r is East and s is West. ϕ Represents the real-wage gap between East and West Germany (incomplete real-wage equalization).

Substituting this last equation into wage equation (1) gives:

$$W_r = \left[\sum_s Y_s . \left(I_r . \left\{\frac{\varphi_{rs} W_s}{W_r}\right\}^{1/\delta} . \left\{\frac{P_r}{P_s}\right\}^{1-\delta/\delta} T^{-D_{rs}} \right)^{\varepsilon-1}\right]^{1/\varepsilon} \Rightarrow$$

$$W_r = \left[\sum_s Y_s . \left(I_r^{\varepsilon-1} . \left\{\frac{\varphi_{rs} W_s}{W_r}\right\}^{\varepsilon-1/\delta} . \left\{\frac{P_r}{P_s}\right\}^{(1-\delta)(\varepsilon-1)/\delta} T^{-D_{rs}.(\varepsilon-1)} \right)\right]^{1/\varepsilon} \Rightarrow$$

(A1) $$W_r = \omega_r^{1-\varepsilon/\delta\varepsilon} . \left[\sum_s Y_s . \left(\{\varphi_{rs} W_s\}^{\varepsilon-1/\delta} . \{P_s\}^{(1-\delta)(1-\varepsilon)/\delta} T^{-D_{rs}.(\varepsilon-1)} \right)\right]^{1/\varepsilon},$$

where ω is the real wage.

ω_r is ω_{WEST} for each district *r* in West Germany, ω_r is ω_{EAST} for each district r in East Germany; $\omega_{WEST}= \phi.\omega_{EAST}$. Note that in (A1) ω_r is a constant for region *r* because real wage equalization still holds within East and West Germany.

The logtransformation of (A1)

$$\log(W_r) = \left(1/\varepsilon\right)\log\left[\sum_s Y_s . \left(\{\varphi_{rs} W_s\}^{\varepsilon-1/\delta} . \{P_s\}^{(1-\delta)(1-\varepsilon)/\delta} T^{-D_{rs}.(\varepsilon-1)} \right)\right] + \left(1-\varepsilon/\delta\varepsilon\right)\log(\omega_r)$$

leads to the specification to be estimated:

$$(2)\ \log(W_r) = \left(1/\varepsilon\right)\log\left[\sum_s Y_s.\left(\{\varphi_{rs}W_s\}^{\varepsilon-1/\delta}.\{P_s\}^{(1-\delta)(1-\varepsilon)/\delta}T^{-D_{rs}.(\varepsilon-1)}\right)\right] + \alpha_0 + \alpha_1 Dummy_{EAST},$$

where $\alpha_1 = \dfrac{\varepsilon-1}{\delta\varepsilon}\log(\phi)$ and $\alpha_0 = \dfrac{1-\varepsilon}{\delta\varepsilon}\log(\omega_{WEST})$.

Appendix 2: Derivation of the Employment Equation

With centralised wage-setting:

$W_r = W_s = W$

Assume a productivity gap between East and West:

$L_{ir} = \alpha + \vartheta_{ir}\theta_{west}x_{ir}$

where L_{ir} is employment in firm *i* in region *r*, x is output, θ_{west} is the marginal productivity of labour in West Germany, and ϑ is 1 if region *r* is in West and ϑ is $\theta_{west}/\theta_{east} > 1$, if region *r* is in East.

Free entry and exit leads to the no-profit condition:

$$x_{ir} = \frac{\alpha.(\varepsilon-1)}{\vartheta_{ir}\theta_{west}}$$

Labour demand at the micro level in East and West is:

$L_{ir} = \alpha\varepsilon$

Output expressed in units of labour is:

$$x_{ir} = \left(\frac{\varepsilon}{\varepsilon-1}\right)\frac{L_{ir}}{\vartheta_{ir}\theta_{west}}$$

Using the Dixit-Stiglitz demand elasticities and dropping index i for the individual firm:

$$x_r = \sum_{s=1}^{R}\left[\left(\frac{\varepsilon}{\varepsilon-1}\frac{W_r T^{D_{rs}}(\theta_{west}\vartheta_r)}{I_s}\right)^{-\varepsilon} T^{D_{rs}}\delta\frac{Y_s}{I_s}\right]$$

where T is transport costs, and D_{rs} is the distance between regions *r* and *s*, I is the price index of manufactures.

The employment equation expressed in logarithms:

$$\log(L_r) = \text{constant} + \log(\vartheta_\text{r}) + \log(x_r) =>$$

$$\log(L_r) = \text{constant} + \log(\vartheta_\text{r}) + \sum_{s=1}^{R}\left[\left(\frac{\varepsilon}{\varepsilon - 1}\frac{W_r T^{D_{rs}}(\theta_{west}\vartheta_r)}{I_s}\right)^{-\varepsilon} T^{D_{rs}}\delta\frac{Y_s}{I_s}\right] =>$$

$$\log(L_r) = \text{constant} + \log(\vartheta_\text{r}) - \varepsilon\log(W_r) + \log\left(\sum_{s=1}^{R}\left[(\vartheta_r)^{-\varepsilon}\left(T^{D_{rs}}\right)^{1-\varepsilon} Y_s I_s^{\varepsilon-1}\right]\right)$$

because of the assumption of uniform nominal wages: $W_r = W$, and

$$\log(L_r) = \text{constant} + \log(\vartheta_\text{r}) + \log\left(\sum_{s=1}^{R}\left[(\vartheta_r)^{-\varepsilon}\left(T^{D_{rs}}\right)^{1-\varepsilon} Y_s I_s^{\varepsilon-1}\right]\right).$$

For an East German district *r* the employment equation is:

$$\log(L_r) = \text{constant} - (\varepsilon - 1)\log(\vartheta_\text{r}) + \log\left(\sum_{s=1}^{R}\left[\left(T^{D_{rs}}\right)^{1-\varepsilon} Y_s I_s^{\varepsilon-1}\right]\right)$$

For a West German district *r* the employment equation is:

$$\log(L_r) = \text{constant} + \log\left(\sum_{s=1}^{R}\left[\left(T^{D_{rs}}\right)^{1-\varepsilon} Y_s I_s^{\varepsilon-1}\right]\right)$$

To arrive at the specification to be estimated add a dummy variable that is 1 for East German districts and 0 for West German districts, the sign should be negative. Scale district employment by the variable area_r (=km^2 of a district) in order to account for the differences in district size in the sample. So the dependent variable becomes L_r/area_r. Using the long-run equilibrium in which real wages are equalized means that price indices of manufactures are equalized. So, the employment equation becomes:

$$\log(L_r) = \text{constant} + \log\left(\sum_{s=1}^{R}\left[\left(T^{D_{rs}}\right)^{1-\varepsilon} Y_s\right]\right) + c_1 D_{east}$$

As this equation shows, it is *not* possible to estimate the structural parameters ε and *T* separately. So the equation in the main text that has actually been estimated is (as with unemployment) closest tot he simple market-potential function (1'), with employment per km^2 as the dependent variable:

$$\log\left(L_r / area_r\right) = \text{constant} + \log\left(\sum_{s=1}^{R}\left[e^{-c_0 D_{rs}} Y_s\right]\right) + c_1 D_{east}$$

where the constant, c_0 and c_1 are to be estimated.

Note that $c_0 = (\varepsilon\text{-}1)\log(T)$, and $c_1 D_{east} = (\varepsilon\text{-}1)\log(\vartheta_r)$.

References

Anas, A.; R. Arnott; K.A. Small (1998):
Urban Spatial Structure, Journal of Economic Literature, Vol. XXXVI, pp.1426-1464

Blanchflower, D.G;. A.J. Oswald (1994):
The Wage Curve, MIT Press.

Brakman, S.; H. Garretsen (1993):
The Relevance of Initial Conditions for the German Unification, Kyklos, Vol. 46, pp. 163-181.

Brakman, S.; H. Garretsen; M. Schramm (2000):
The Empirical Relevance of the New Economic Geography: Testing for a Spatial Wage Structure in Germany, CESifo Working Paper, 395, Center for Economic Studies, Munich.

Brakman, S.; H. Garretsen; C. van Marrewijk (2001):
An introduction to Geographical Economics: Trade, Location and Growth, Cambridge University Press, Cambridge.

Decressin, J.; A. Fatas, (1995):
Regional Labour Market Dynamics in Europe, European Economic Review, 39, pp. 1627-1655.

Faini, R. (1999):
Trade Unions and Regional Development, European Economic Review, 43(2), pp. 457-474.

Funke, M.; J. Rahn (2000):
How Efficient is the East German Economy? An Exploration with Micro data, CESifo Working Paper, 397, Munich.

Fujita, M.; P. Krugman; A.J. Venables (1999):
The Spatial Economy, MIT Press.

Hanson, G.H. (1997):
Increasing Returns, Trade and the Regional Structure of Wages, Economic Journal, Vol. 107, pp. 113-133.

Hanson, G.H. (1998):
Market Potential, Increasing Returns, and Geographic Concentration, NBER Working Paper 6429, Cambridge Mass.

Hanson, G.H. (1999):
Market Potential, Increasing Returns, and Geographic Concentration, mimeo, University of Michigan (revised version of Hanson, 1998).

Hanson, G.H. (2000):
Scale Economies and the Geographic Concentration of Industry, NBER Working Paper 8013, Cambridge Mass (forthcoming in the Journal of Economic Geography).

Helpman, E. (1998):
The Size of Regions, in D. Pines, E. Sadka and I. Zilcha (eds.), Topics in Public Economics, Cambridge University Press.

Krugman, P. (1991):
Increasing Returns and Economic Geography, Journal of Political Economy, 3, pp. 483-499.
Krugman, P. (1998):
Space, the Final Frontier, Journal of Economic Perspectives, pp. 161-174.
Krugman, P.; A. Venables (1995):
Globalization and the inequality of nations, Quarterly Journal of Economics, Vol. 110, pp. 857-880.
Krugman, P.; A. Venables (1996):
Integration, Specialization, and Adjustment, European Economic Review, Vol. 40, pp. 959-967.

Ottaviano, G.I.P.; J-F Thisse (2001):
On Economic Geography in Economic Theory: Increasing Returns and Pecuniary Externalities, Journal of Economic Geography, 1(2), pp. 153-180.
Overman, H.; S. Redding; A.J. Venables (2001):
Trade and Geography: A Survey of the Empirics, mimeo, London School of Economics.

Peeters, J.; H. Garretsen (2000):
Globalisation, Wages and Unemployment: An Economic Geography Perspective, CESifo Working Paper, 256, Munich.
Puga, D. (1999):
The rise and fall of regional inequalities, European Economic Review, Vol. 43, pp. 303-334.
Puga, D. (2001):
European Regional Policies in Light of Recent Location Theories, CEPR Discussion Paper, 2767, London.

Neary, J.P. (2001):
Of hype and hyperbolas: introducing the New Economic Geography, Journal of Economic Literature, Vol. XXXIX, pp. 536-561.

Roos, M. (2001):
Wages and market potential in Germany, Jahrbuch für Regionalwissenschaft, 21, pp. 171-195.

Sinn, H-W. (2000):
Germany's Economic Unification, An Assessment after Ten Years, CESifo Working Paper, 247, Munich.
Ströhl, G. (1994):
Zwischenörtlichter Vergleich des Verbraucherpreisniveaus in 50 Städten, Wirtschaft und Statistik, 6, pp. 415-434.

Thomas, A. (1996):
Increasing Returns, Congestion Costs and the Geographic Concentration of Firms mimeo, International Monetary Fund, Washington.

Venables, A.J. (1996):
Equilibrium Locations of Vertically Linked Industries, International Economic Review, 37, pp. 341-359.