

**Tjalling C. Koopmans Research Institute**

*Tjalling C. Koopmans*



**Universiteit Utrecht**

**Utrecht School  
of Economics**

**Tjalling C.Koopmans Research Institute  
Utrecht School of Economics  
Utrecht University**

Vredenburg 138  
3511 BG Utrecht  
The Netherlands  
telephone (0031) 030 253 9800  
fax (0031) 030 253 7373  
website [www.koopmansinstitute.uu.nl](http://www.koopmansinstitute.uu.nl)

The Tjalling C. Koopmans Institute is the research institute and research school of the Utrecht School of Economics. It was founded in 2003, and named after Professor Tjalling C. Koopmans, the Dutch born Nobel Prize laureate in economics in 1975.

In the discussion papers series the Koopmans Institute publishes results of ongoing research for early dissemination of research results, and to enhance discussion with colleagues.

Please sent any remarks or questions on the Koopmans institute, or this series to [P.vanDriel@econ.uu.nl](mailto:P.vanDriel@econ.uu.nl)

ontwerp voorblad: WRIK Utrecht

**How to reach the authors**

*Please direct all correspondence to Harry Garretsen.*

Steven Brakman  
Department of Economics, University of Groningen,  
CDS (Centre for German Studies) and CESifo  
Email: [s.brakman@eco.rug.nl](mailto:s.brakman@eco.rug.nl)

Harry Garretsen  
Utrecht School of Economics  
University of Utrecht,  
CDS (Nijmegen) and CESifo  
p.a. Vredenburg 138  
3511 BG Utrecht  
The Netherlands  
tel: 00-31-30-2539810.  
Email: [h.garretsen@econ.uu.nl](mailto:h.garretsen@econ.uu.nl)

Marc Schramm  
Utrecht School of Economics  
and CDS  
Email: [m.schramm@econ.uu.nl](mailto:m.schramm@econ.uu.nl)

# The Spatial Distribution of Wages: Estimating the Helpman-Hanson model for Germany

Steven Brakman\*  
Harry Garretsen\*\*  
Marc Schramm\*\*\*

\*Department of Economics, University of Groningen

\*\*Utrecht School of Economics  
Utrecht University

\*\*\*Utrecht School of Economics  
Utrecht University

August 2003

## Abstract

Using German district data we estimate the structural parameters of a new economic geography model as developed by Helpman (1998) and Hanson (1998, 2001a). 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. Generalizing the Helpman-Hanson model, we also analyze the implications for the spatial distribution of wages once the assumption of real wage equalization is dropped. If we no longer assume real wage equalization we find support for a spatial wage structure as well as for the relevance of the structural parameters of the theoretical model.

**Keywords:** new economic geography, spatial wage structure

**JEL-code:** R10, R12, R23

## Acknowledgements:

We would like to thank the editors, three anonymous referees, Paul Elhorst, Michael Funke, Michael Pflüger, Michael Roos and seminar participants at the HWWA Hamburg, University of Groningen, University of Amsterdam, Utrecht University, CESifo Munich, Netherlands Bureau for Economic Policy Analysis (CPB) The Hague, and DIW Berlin for their comments. Finally we would like to thank Dirk Stelder for his help in drawing the maps.

## 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.<sup>1</sup> 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), this is no coincidence since there is still a lack of direct testing of the new economic geography models or its implications. In a review of Fujita, Krugman and Venables (1999), Neary (2001) concludes that empirical research is lagging behind and that much more work needs to be done. In order to make progress an empirical validation of the main theoretical insights is called for. The reason that the empirical research lags behind is that the new economic geography models are characterized by non-linearities and multiple equilibria and this makes empirical validation relatively difficult.

There is a substantial amount of empirical research that shows that location matters, but there are indeed still relatively few attempts to specifically estimate the structural parameters of new economic geography models (see the surveys by Overman, Redding and Venables (2001) and Head and Mayer (2003)). A notable exception is the work by Gordon Hanson (1998, 2001a). 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 but the underlying demand linkages are of limited geographical scope. In this paper we apply the Helpman-Hanson model to the case of Germany. The contribution of the paper is threefold.

First, we want to establish whether the Helpman-Hanson model can be verified, that is to say we want to know whether the key model parameters are significant and offer a

meaningful description of the basic spatial features of an economy, in our case the German economy. The main equation to be estimated will be a wage equation, central to this equation is the idea that wages will be higher in those regions that have easy access to economic centers because for those regions demand linkages are relatively strong. The main geographical unit of analysis is the German city-district (*kreisfreie Stadt*).

Second, we extend Hanson's analysis in two ways. We not only use the housing stock but also land prices. The latter variable is more in line with the Helpman-Hanson model, in particular with respect to the non-tradable good (housing) that plays a key role as a spreading force in the model. Our main extension is that we drop the assumption of real wage equalization as this may not be very realistic for most continental European countries and especially for re-unified Germany in the mid 1990s. This requires an alternative specification of the wage equation.

Third, we compare our estimation results for the wage equation with two alternatives by also estimating a simple market potential function for wages as well as a wage curve. The former encompasses a wide range of theoretical approaches that explicitly include a role for location or distance, and as such represents a summary of alternative explanations, see Harrigan (2001). The latter provides a spatial wage structure without using geography explicitly. Still, it should be emphasized that our main aim is to estimate (and not to test) the Helpman-Hanson model.

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 our new economic geography model is the "best" model to analyze Germany after the fall of the Berlin Wall in 1989. Similarly, it is also not the goal of this paper to come up with a model for the spatial distribution of wages that includes all potentially relevant variables that have been mentioned in the literature. We merely want to estimate a particular new economic geography model, here for the case of Germany.<sup>2</sup> The fall of the Berlin Wall creates a unique testing ground. A main problem in estimating new economic geography models is whether or not the economy is in a long-run

equilibrium and if so in which equilibrium. This problem is reduced for the German case. We can assume that a few years into the reunification process the spatial allocation of economic activity is not an equilibrium and we can assume *a priori* that versions that do not presuppose (real) spatial wage equalization are to be preferred over specifications that do. In our paper we take the approach recommended by Hanson (2001b) 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, 2001b, p.271).

The paper is organized as follows. In section 2 we visualize the idea that geography matters in Germany by showing a few maps. We also give a first indication of the importance of the housing prices, the most distinguishing variable of the Helpman-Hanson model. In section 3 we focus on the derivation of the empirical specification of the wage equation, discuss our data set and some estimation issues. Section 4 gives the main estimation results for the basic wage equation. In section 5 we address the implications of dropping the assumptions of real wage equalization. Section 6 briefly evaluates the value added of the new economic geography model by estimating two simple alternative models, the market potential function and the wage curve. Section 7 concludes the paper. We find the strongest support for a spatial wage structure and for the relevance of the structural parameters once we no longer assume real wage equalization. A comparison of our preferred model specification with the two alternative explanations of a spatial wage structure indicates that our model does not outperform these other specifications.

## **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 key economic variables of our model. Take, for

instance, a look at Map 1. This map gives GDP per km<sup>2</sup> (in millions of DM) for the 441 German districts (*Kreise*) that make up our data set. 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<sup>2</sup> is higher in urban districts and typically lower in eastern German districts.

### **MAP 1 HERE**

The key equation in the Helpman-Hanson 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 districts, but the map also shows 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<sup>2</sup>.

Agglomeration in new economic geography models is the result of the combination of agglomerating and spreading forces. An important agglomerating force is for instance the size of the market (see Map 1). Among the spreading forces is 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 somewhat less clear-cut than with respect to regional wages. Land prices per m<sup>2</sup> are particularly high in or close to economic centers. Land prices can be looked upon as a proxy for housing prices and therefore for prices of housing services, and as will become clear in Section 3 the prices of housing services 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, nominal wages, the size of the market (GDP) and prices of housing services

(here proxied by land prices). Taken together these maps indicate that there is no random distribution of economic activity across Germany and that high district wages go along with high GDP and high land prices. Also, there seems to be a clustering of the 3 district variables. Districts with relatively high wages, high GDP and high land prices are typically surrounded by similar districts, and the same holds for low wage districts. In our estimations we want to find out if indeed such a regional (wage) gradient exists, as predicted by the Helpman-Hanson model.

**MAP 2 HERE**

**MAP 3 HERE**

The inclusion of housing services as a non-tradable consumption good lies at the very heart of the Helpman (1998) model and to illustrate that prices of housing services may indeed display a spatial structure, as suggested by Map 3, we estimate a market potential function, equation (1).

$$(1) \quad \log(LP_r) = \kappa_1 \log[\sum_s Y_s e^{-\kappa_2 D_{rs}}] + \kappa_3 EG + \kappa_4 D_{country} + \text{constant}$$

where  $LP$ =land prices per m<sup>2</sup> in district  $r$  in 1995,  $Y$ =GDP in district  $r$  in 1994,  $D_{rs}$ =distance between districts  $r$  and  $s$ ,  $EG$  and  $D_{country}$  are two dummy variables for eastern German and country districts respectively, if district  $r$  is in eastern Germany  $EG = 1$  and if district  $r$  is not a city-district, then  $D_{country} = 1$ ;  $\kappa_1$ - $\kappa_4$  are coefficients to be estimated, for more details on the data see Section 3.2.

The estimation results, see Table 1, show that there is a “spatial land price” structure ( $\kappa_1 > 0$  and  $\kappa_2 > 0$ ) which means that land prices in district  $r$  are higher if this district is near to districts with a high GDP. This is precisely what drives the Helpman-Hanson model. 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.

**TABLE 1 HERE**

### **3.THE MODEL AND DATA**



### *The Helpman-Hanson Model*

The benchmark model of the new economic geography, developed by Krugman (1991), is in general not suited for empirical validation. In the long run it produces 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 because in reality we observe the co-existence of large and small locations. Furthermore, it lacks some of the spatial characteristics of agglomerations that 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).

Given the observation that complete agglomeration is not in accordance with the facts, there are two alternatives. First, there are the new economic geography models based on forward and backward linkages with no interregional labor mobility (Krugman and Venables, 1995, 1996, and Venables, 1996). However, direct testing of these models is rather cumbersome because it requires detailed information about input-output linkages between firms on a regional level (see Combes and Lafourcade (2001) for an empirical test of this class of models). Second, there is the Helpman-Hanson model which combines the “best of the two worlds”. It shares with Krugman (1991) the emphasis on demand linkages (which are more easy to test for than input-output linkages) and at the same time, through the inclusion of a non-tradable consumption good (i.e housing services), is capable of producing similar equilibria as the models based on input-output linkages and immobile factors of production. The price of housing services in the Helpman (1998) model which increases with agglomeration, serves as an analogous spreading force as the rising wages in for instance Puga (1999). In fact, it can be shown that in terms of equilibrium outcomes the Helpman model yields similar results as the model in Krugman and Venables (1996) where there is no interregional labor mobility and the possibility of agglomeration arises through intricate input-output linkages between firms.<sup>3</sup>

We briefly discuss the theoretical approach in Hanson (1998, 2001a) and focus on the equilibrium conditions because these are needed to arrive at the basic wage equation that will be estimated.<sup>4</sup> With a few exceptions (notably, the inclusion of a non-tradable good, housing services) the micro-foundation for the behavior of the individual consumers and producers is the same as in the seminal Krugman (1991) model. Since these basic features of the new economic geography are by now well-known 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 services which is a non-tradable 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 services (a non-tradable good) are 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 services) at the expense of a homogenous tradable good (agriculture), there are no fundamental differences in the model set-up between Krugman (1991) and Helpman (1998). In particular, in both models agglomeration is driven by the demand linkages and the interregional mobility of labor. Note, however, that the comparative statics of the 2 models differ, notably with respect to the implications of a change in transportation costs for the degree of agglomeration.<sup>5</sup> In section 4 we will address this issue in our discussion of the so called “no black hole” condition.<sup>6</sup>

This extension of core model with the non-tradable good, housing services, thus allows for a richer menu of equilibrium spatial distributions of economic activity than 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), see [Appendix A](#) for a derivation of equilibrium wage equation (2):

$$(2) \quad W_r = \left[ \sum_s Y_s I_s^{\varepsilon-1} T^{D_{rs}(1-\varepsilon)} \right]^{1/\varepsilon}$$

$$(3) \quad I_r = \left[ \sum_s \lambda_s (T^{D_{rs}})^{1-\varepsilon} W_s^{1-\varepsilon} \right]^{1/(1-\varepsilon)}$$

$$(4) \quad Y_r = \lambda_r L W_r$$

In which in equation (2)  $W_r$  is the region's  $r$  (nominal) wage rate,  $Y$  is income,  $I$  is the price index for manufactured goods,  $\varepsilon$  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  $\varepsilon$ , it can directly be seen from equation (2) 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  $I_s$ . As to the way in which this competition effect works, the price index  $I_s$  does not measure a competition effect in the sense in which this term is normally used (price mark-ups are fixed and there is no strategic interaction between firms). A low price index reflects that many varieties are produced in nearby regions and are therefore not subject to high transportation costs and this reduces the level of demand for local manufacturing varieties. Since firms' output level and price mark-up are fixed, this has to be off-set by lower wages. Hence, a low (high) price index  $I_s$  depresses (stimulates) regional wages  $W_r$ .

Equation (3) 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

(4) simply states income in region  $r$ ,  $Y_r$ , has to equal the labor income earned in that region, where  $\lambda_r$  is region  $r$ 's share of the total manufacturing labor force  $L$ .<sup>7</sup>

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 (2), exists for Germany. Equation (2) 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 (3)), the price index  $I$  is endogenous, and inter alia depends on each of the local wage rates, which makes a reduced form of equations (2) and (3) 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  $I$  in exogenous variables that can actually be observed for his sample of US counties.

First, he uses:

$$(5) \quad P_r H_r = (1 - \delta) Y_r$$

Equation (5) states that the market value of the housing services supplied equals the share of income spent on housing services, where  $P_r$  is the price of housing services in region  $r$ , and  $H_r$  is the *fixed* stock of housing in region  $r$ , which determines the total of housing services supplied, and  $(1-\delta)$  is the share of income spent on housing services and  $\delta$  is thus the share of income spent on manufactures.<sup>8</sup>

Second, real wage equalization between regions is assumed:

$$(6) \quad \frac{W_r}{P_r^{1-\delta} I_r^\delta} = \frac{W_s}{P_s^{1-\delta} I_s^\delta}$$

Equation (6) is quite important. In fact, 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).<sup>9</sup> 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 (2002, p. 389). We return to this issue in section 5 where we will drop the assumption of real wage equalization..

The importance of non-tradable housing services as a spreading force is implied by (6). A higher income  $Y_s$  implies, *ceteris paribus*, higher wages in region  $r$ , see equation (2), but it also, given the stock of housing, puts an upward pressure on the price of housing services  $P_r$ , equation (5). Combining (5) and (6) allows us to rewrite the price index in terms of the housing stock, income and nominal wages. First, we rewrite (6) as:

$$(6') \quad \frac{W_r}{P_r^{1-\delta} I_r^\delta} = \bar{\omega}$$

where  $\bar{\omega}$  is the real wage that is assumed to be constant and identical for all regions.

Second, the equilibrium condition for the market for housing services can be written as  $P_r = (1-\delta)Y_r/H_r$  and this expression can be substituted for  $P_r$  into equation (6') and this gives the price index  $I_r$  in terms of  $W_r$ ,  $Y_r$  and  $H_r$ . Substituting this in (1) results in a wage equation which can be estimated. This will also be the bench-mark wage equation in our empirical analysis:<sup>10</sup>

$$(7) \quad \log(W_r) = k_0 + \varepsilon^{-1} \log\left(\sum_s Y_s^{\varepsilon+(1-\varepsilon)/\delta} H_s^{(1-\delta)(\varepsilon-1)/\delta} W_s^{(\varepsilon-1)/\delta} T^{(1-\varepsilon)D_{rs}}\right) + err_r$$

Where  $k_0$  is a parameter and  $err_r$  is the error term. Equation (7) includes the three central structural parameters of the model, namely share of income spent on manufactures,  $\delta$ , the substitution elasticity,  $\varepsilon$  and the transport costs,  $T$ . Given the availability of data on wages, income, the housing stock, and a proxy for distance, equation (7) 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 (7), 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 (7) for our sample of German districts.

#### *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 and 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, our  $Y_s$  variable), 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 average hourly wage in the manufacturing and mining sector regional wage,  $W_r$ . Since we also want to analyze the cities' *Hinterland* we also included 37 aggregated (country) districts, constructed from a larger sample of 322 districts.<sup>11</sup> 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. 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 car travel time (in minutes) from district A to district B. The data are obtained from the Route Planner 2000 (Europe, And Publishers, Rotterdam).

For the data on the housing stock  $H_s$ , which are required to estimate equation (7), we use the number of rooms in residential dwellings per district. As a proxy for  $P_s$ , the average land price (*Baulandpreis*) is used. In our estimations we also include the following district variables to control for location-specific fixed effects: the employment structure (the respective shares for each district in agricultural, industrial, services and trade, and transport employment) and two variables indicating the skill level (proportion of workers that did not complete vocational training, proportion of workers that completed higher education/vocational training).

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 (7) in levels and we therefore restrict ourselves to cross-section estimations. The estimation of an equation like equation (7) 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 (7)). To check for this we have experimented with instrumental variables (IV) with respect to  $W_s$  in our estimations (not shown here, but available upon request). We used as instruments (inter alia) 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.

Secondly, if we were able to use multiple years of observation for each variable in estimating equation (7), the time-series element of the data would allow us, as in Hanson's work, to estimate in first differences. 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. As a next best strategy to deal with the possibility of time-invariant location specific fixed effects, we incorporate a number of district variables, in particular the skill level and the employment structure. These and a few other variables (see below) are used as controls for the presence of district or region specific determinants of wages  $W_r$ .

Thirdly, with respect to the geographical unit of analysis, the left and right hand side variables are both measured at the district level in our estimations. In Hanson (1998) or Roos (2001) the latter are measured at 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  $W_r$  has an impact on  $Y_s$  or  $W_s$ . On the other hand, a lower level of geographical aggregation of the data makes it less likely that location-specific shocks (via the error-term in equation (7)) have an impact on the independent variables.<sup>12</sup>

Finally, and related to this last observation, is the possibility that the variance of the error-term systematically varies across the various districts. To address the issue of heteroskedasticity we apply the Glejser-test and use weighted least squares (WLS)

estimations. Therefore, we estimate (via non-linear least squares, NLS) equation (7) or any other of our specifications and we then we regress the (absolute of the) resulting residuals on the right hand side variables. A significant impact of these variables on the residuals indicates heteroskedasticity and for every specification it turns out that this is indeed something that has to be taken into account. To deal with this we use 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

We now turn to the estimates of the structural parameters using the wage equation (7) for Germany. In doing so, we will not only be able to estimate the structural parameters  $\delta$ ,  $\varepsilon$  and  $T$  (and to establish the existence of a spatial wage structure), but also to verify the so-called no-black hole condition, which gives an indication for the convergence prospects in Germany. We first present the regression results of equation (7), which incorporates the assumption that the local housing stock  $H$  and GDP determine the local price of housing services. We then present the results of the wage equation (7'), in which the local land price  $LP$  is used as an explanatory variable, and where the housing stock  $H$  is omitted. Local land prices are considered as a proxy for the local prices of housing services  $P$  of the theoretical model, which makes equation (5), the equilibrium condition for the market of housing services, redundant (to arrive at equation (7') use equation (6) to express the price index  $I$  in terms of wages  $W$  and prices of housing services  $P$  and substitute this expression in wage equation (1)).

$$(7') \quad \log(W_r) = k'_0 + \varepsilon^{-1} \log\left(\sum_s Y_s LP_s^{(1-\delta)(1-\varepsilon)/\delta} W_s^{(\varepsilon-1)/\delta} T^{(1-\varepsilon)D_{rs}}\right) + err_r$$

where  $LP$  is land price.

Table 3 gives the basic estimation results for the estimation of equations (7) and (7'). We also included a dummy variable for eastern German districts and a dummy variable for country districts.<sup>13</sup> The dummy for eastern German districts is motivated by the fact that wages (and labor productivity) in eastern Germany are lower than in western Germany.<sup>14</sup> For example, in the east in 1995 average labor productivity in the mining sector was 68.3% of the level in the west and in the manufacturing sector it



just reached 55.1% (Federal Statistical Office, Wiesbaden). Furthermore, especially in East Germany, the central wage bargaining system in Germany prevents prices to perform their allocational role, that slows down the transition process towards a well functioning market economy. The inflexibility of the labor market has often been blamed for the present difficulties of the German economy. Besides the inclusion of the dummy variables for East German and country districts we also included 6 district-specific variables as control variables to control for location-specific effects on wages in all our estimations in the sections 4,5 and 6. Table 2 gives summary statistics on the district specific control variables (the standard deviation is a measure of the differences between districts).

### **TABLE 2 HERE**

In the estimations we make sure that the share variables do not sum to one, and are not perfectly collinear. With respect to the employment variable, public sector employment is excluded: with respect to vocation training we excluded all forms of vocational training except higher education.

From the remaining six variables the first two pertain to the skill level (proportion of workers that did not complete vocation training, proportion of workers that completed higher education/vocational training). The other 4 of these district variables pertain to the employment structure (the respective shares for each district in agricultural, industrial, services and trade, and transport employment). Note that, given the fact that we use WLS, the adjusted  $R^2$  is less informative. The corresponding NLS estimations (not shown) invariably resulted in an adjusted  $R^2$  in the range of 0.5-0.8.

### **TABLE 3 HERE**

In discussing our results the role of the EG dummy turns out to be quite important which explains the set-up of Table 3. First look at the columns I and II in Table 3. For the estimation of equation (7) without the EG dummy (see column I) all three structural parameters are found to be significant. They also have the correct sign thereby validating the Helpman-Hanson model. The substitution elasticity  $\varepsilon$  is significant and the coefficient implies a profit margin of 38% (given that  $\varepsilon/(\varepsilon-1)$  is the

mark-up), which is fairly reasonable<sup>15</sup>, although higher than found for the US by Hanson (1998, 2001a). Note that the value  $\varepsilon(1-\delta)$  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  $\varepsilon(1-\delta) < 1$  (see below)<sup>16</sup>. The coefficient for  $\delta$  is, however, implausibly large in the wage equation with the housing stock because it would indicate that Germans do not spend any part of their income on housing services (see equation (5)). Including a dummy for eastern German (EG) districts (column II) changes these results in an important way: the transportation cost coefficient now has the wrong sign, and  $\delta$  becomes statistically insignificant. So, estimating equation (7) for re-unified Germany yields two problems: (i) we find  $\delta$  to be (too) large, (ii) taking into account an EG dummy, because we expect for the mid 1990s wages in eastern Germany to be lower than in western Germany, does imply that we no longer find a spatial wage structure. One way to remedy these problems might be to dismiss the housing stock as the preferred variable to capture the spreading force of the non-tradable service (see below). Note that the high value found for  $\delta$  is, however, not at odds with the findings of Hanson, who also finds that  $\delta$  is large for the USA (above 0.9).<sup>17</sup> Finally, with respect to the control variables it turns out that not only the EG dummy is highly significant, but also the district specific share of workers with a higher-education contributes quite strongly to the explanation of spatial wages.

The first column of Table 3 also shows 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  $\delta$  is fixed, see however footnote 16). This implies that 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  $\varepsilon(1-\delta) > 1$  means that a region's share of manufacturing production is a function of its (fixed) relative housing stock only (Helpman, 1998, p. 40). The relevance of this conclusion with respect to the no black hole condition is, however, rather limited because specification (7) more or less breaks down when we take into account that the eastern German economy is rather different from its western German counterpart (column II).

One way to remedy the two problems in estimating equation (7) might be to dismiss the housing stock as the preferred variable to capture the spreading force of the non-tradable service, housing. Columns III and IV in Table 3 are therefore based on equation (7'), the wage equation with land prices (*Baulandpreise*) as an explanatory variable instead of the housing stock. From Table 1 we already know that land prices are higher in districts near to districts with a high GDP. *A priori*, estimation of the wage equation with these price data is preferred because it provides a more direct test of the Helpman-Hanson model: the influence of agglomeration on prices of local non-tradables is determining the strength of the spreading force in the Helpman-Hanson model. Also note, that by using a proxy for  $P_s$  variable we no longer need equation (5), the equilibrium condition for the housing market. However, we still need equation (6) that assumes real wage equalization.

Column III of Table 3 gives the estimation results with land prices for 146 districts, excluding the dummy for EG districts. The three structural parameters are significant with the correct sign. The most important difference with the previous estimations with the housing stock is that the  $\delta$ -coefficient is now significantly smaller than 1 which indicates that a significant part of income ( $1-0.684$ ) is indeed spent on housing services and that the housing sector can indeed act as a spreading force (the estimated share of income spent on manufactures in Germany of 0.684 is in line with the actual share of 0.68). This finding gives some support to the use of land prices instead of the housing stock. Another difference is that the “no black hole condition” is no longer met ( $\varepsilon(1-\delta)=1.30>1$ ). This implies 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 transport costs at all. More importantly though, is that the estimation results for equation (7') with land prices and the EG dummy again yields a wrong sign for the transportation cost parameter.

Estimating the Helpman-Hanson model for Germany by means of wage equations (7) and (7') thus leads to two conclusions. First, including land prices instead of the housing stock is a more direct test of the Helpman-Hanson model, and modestly improves the estimates, in the sense that the share of spending on housing is no longer

zero and is more in line with the actual share of spending on housing. The second and more important conclusion, however, is that the model should be amended, because incorporating an outstanding feature of the re-unified German Economy in 1994/1995, that is the inclusion of the EG dummy, gives the transport cost parameter the wrong sign. The most obvious amendment in the German case is to drop the assumption of (real) wage equalization. So, what happens when we, along with the housing equation (5), no longer use equation (6) in order to estimate our wage equation for Germany?

## 5. REAL WAGE DIFFERENCES

The assumption of real wage equalization (recall equation (6)) 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 implies that the economy is in a long-run equilibrium. Furthermore, and supported by our findings in the previous section, this assumption seems at odds with the stylized fact that (real) wages were different between eastern and western German regions at the start of the reunification process. Another problem with the application of wage equation (7) to a European economy like Germany is that the actual degree of real wage flexibility may be too low to bring about real wage equalization.

In this section we estimate a wage equation and the structural parameters *without* invoking real wage equalization and thereby extend Hanson's (2001a) analysis. We do so by deriving a wage equation that is based on a reduced form of the equilibrium wage equation (2) and the equilibrium price index equation (3). For this purpose we simplify the price index as defined in equation (3). For each district we focus on two prices: the price in district  $r$  of a manufactured good produced in district  $r$  and the *average* price outside district  $r$  of a manufactured good produced outside district  $r$ . The determination of the simplified local price index for manufactures requires a measure of *average* distance between region  $r$  and the regions outside. The distance from the economic center is an appropriate measure in our view. This center is obtained by weighing the distances with relative  $Y$ .<sup>18</sup> The economic center of

Germany turns out to be *Landkreis Giessen* (near Frankfurt), which is in the state of Hessen, West Germany. Equation (3) now becomes:

$$(3') \quad I_r = \left[ \lambda_r W_r^{1-\varepsilon} + (1 - \lambda_r) (\bar{W}_r T^{D_{r-center}})^{1-\varepsilon} \right]^{1/\varepsilon},$$

where  $\bar{W}_r$  is the average wage outside district  $r$ ,  $D_{r-center}$  is the distance from district  $r$  to the economic center, and weight  $\lambda_r$  is district  $r$ '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 the wage equation.

However, as in section 4, we also want to take into account that labor productivity in eastern Germany is lower than in western Germany. A relatively low marginal labor productivity (*MPL*) in eastern Germany implies *ceteris paribus* relatively high marginal costs in eastern Germany. This affects the price index equation (3').

Moreover, relatively high marginal costs imply adverse competitiveness for the eastern German firms, which needs to be corrected by relatively low nominal wages in eastern Germany. This affects the wage equation (2). By incorporating the *MPL*-gap, the wage equation (2) and the simplified price index equation (3') change into (see [Appendix B](#)):<sup>19</sup>

$$(2'') \quad W_r = \text{constant} \cdot \left( \frac{MPL_{west}}{MPL_r} \right)^{(1-\varepsilon)/\varepsilon} \left[ \sum_{s=1}^R Y_s (T^{D_{rs}})^{1-\varepsilon} I_s^{\varepsilon-1} \right]^{1/\varepsilon}$$

$$(3'') \quad I_r = \left[ \lambda_r \left( W_r \frac{MPL_{west}}{MPL_r} \right)^{1-\varepsilon} + (1 - \lambda_r) (\bar{W}_r T^{D_{r-center}})^{1-\varepsilon} \right]^{1/(1-\varepsilon)}$$

Equation (3'') is finally substituted into (2''), which gives the reduced form of the equilibrium wage equation without invoking real wage equalization in order to approximate (3). The equation to be estimated is:

$$(7'') \quad \log(W_r) = \kappa_0 + \varepsilon^{-1} \log \left[ \sum_{s=1}^R Y_s (T^{D_{rs}})^{1-\varepsilon} I_s^{\varepsilon-1} \right] + \kappa_1 EG$$

where  $I_r^{1-\varepsilon} = \left[ \lambda_r (W_r (1 + \gamma_r))^{1-\varepsilon} + (1 - \lambda_r) (\bar{W}_r T^{D_{r-center}})^{1-\varepsilon} \right]$ ,  $1 + \gamma_r \equiv e^{-\kappa_1 EG(r) \varepsilon / \varepsilon - 1}$ ,  $\gamma_r$

representing the productivity gap between western Germany and district  $r$ , which is  $(MPL_{west}/MPL_r) - 1$ , and  $EG(r)$  is a dummy variable which equals 1 if  $r$  is in eastern Germany.

An additional advantage of equation (7'') compared to the basic wage equation (7) is that the share of income spent on manufactures  $\delta$  (which we thus found to be rather large in our initial estimation in Table 3) does not need to be estimated now because the equilibrium condition for the market for housing services (equation (5)) is no longer needed to estimate the wage equation.

Table 4 shows the regression results of estimating equation (7'').<sup>20</sup> The results for equation (7'') in Table 4 show that the distance parameter is now significantly positive and the same holds for  $\varepsilon$ . In this case, see equation (7''), 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. That is to say, our alternative estimation strategy does yield a spatial wage structure for Germany.<sup>21</sup>

#### **TABLE 4 HERE**

The estimation results in Table 4 for Germany provide support for the Helpman-Hanson model in the sense that the model parameters are found to be significant. Our estimations of wage equation (7'') illustrate that  $W_r$  is higher if district  $r$  is situated more closely to regions with a relatively high  $Y$ . 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 which could imply in terms of our model that, for whatever institutional reason, interregional wages are set at the same level.<sup>22</sup> For a country like Germany one might thus very well expect that the spatial distribution of  $Y$  does *not* only get reflected in spatial wage differences but (also) in the spatial distribution of quantity variables like regional (un)employment (see also Puga, 2002, pp. 389-390, for this assertion for Germany).<sup>23</sup> The proper approach to deal with the implications of wage rigidity is to incorporate the implications of wage rigidity, most notably unemployment, into the model and then to (re-)estimate the structural parameters. This is left for future research. Here, we only refer to Brakman, Garretsen and Schramm (2002a) where using the, rather crude, wage rigidity assumption  $W_r=W_s$  (which we have shown in this section not to be the case, recall also Map 2!), an employment equation is derived. Estimation results confirm the existence of a spatial employment structure.<sup>24</sup>

Returning to the estimation results in Table 4 and following Hanson (1998, 2001a), we illustrate the strength of the interregional demand linkages that give rise to the spatial wage structure, through the following experiment. Based on the estimated coefficients in Table 4 we derive the impact on regional wages from a 10% GDP increase in the city-district of Munich. The localization of demand linkages turns out to be quite strong. The positive GDP shock leads to an increase of wages in Munich itself by 0.8% and the impact on wages in other regions strongly decays with distance. In Berlin, for instance, the impact is a mere 0.08%. This is in line with the findings of Hanson (2001a) for the USA and Roos (2001) for Germany. Figure 1 shows the results of our “Munich-experiment” for each of the German districts. It clearly shows that the impact of the GDP-shock on wages rapidly declines the further one moves away from Munich.

**FIGURE 1 HERE**

## **6. TWO SIMPLE ALTERNATIVE MODELS FOR A SPATIAL WAGE STRUCTURE**

So far, the estimation results provide some support for the Helpman-Hanson model once we no longer have to assume real wage equalization. Even though our main aim is to estimate and not test this model, our results in Table 4 raise the question how well this model specification performs against alternative models that also include the possibility of a spatial wage structure. In this section we therefore briefly compare the estimated wage equation (7'') with two simple alternative models, a market potential function and a wage curve. The market potential function encompasses a wide range of theoretical approaches and as such represents a summary of alternative explanations, see Harrigan, 2001. Equation (1) is an example of a market potential function and here we use the same specification but now with wages as the dependent variable, see equation (1').

$$(1') \quad \log(W_r) = \kappa_1 \log \left[ \sum_s Y_s e^{-\kappa_2 D_{rs}} \right] + \kappa_3 EG + \kappa_4 D_{country} + \text{constant}$$

where  $W_r$  = wages in district  $r$ ,  $Y_r$  = income in district  $r$ ,  $D_{rs}$  = distance between districts  $r$  and  $s$ ,  $EG$  and  $D_{country}$  are two dummies for eastern-German and country districts respectively.

The wage curve, see equation (8), states that regional wage differences reflect regional differences in GDP or unemployment but it does not include geography in the sense that distance between regions is not used as an explanatory variable (Blanchflower and Oswald, 1994).

$$(8) \quad \log (W_r) = \beta_1 \log (y_r) + \beta_2 \log (U_r) + \text{constant}$$

where  $W_r$  =wages in district  $r$ ,  $y_r$  =GDP per capita in district  $r$ ,  $U_r$  = unemployment rate in district  $r$ .

We proceed as follows. We first estimate the market potential function (1') and the wage curve (8) with the same variables controlling for location-specific effects as used in specification (7''), and we then compare the models. Table 5 gives the estimation results for (1') and (8) and also gives the Akaike information criterion (AIC) which is used for model selection.

#### **TABLE 5 HERE**

Both the market potential function and the wage curve give the expected results and thus also give rise to a spatial wage structure. The parameters are significant and have the right signs. A low value for the AIC test statistic is preferred over a higher one. According to this criterion it seems that the market potential function (1') must be slightly preferred over the wage curve (8) and our wage equation (7'').<sup>25</sup> We, however, prefer the structural model over the market potential function or a wage curve, because of its direct link to theory, and because it allows us to estimate the structural variables. Moreover, the findings seem plausible for the German case. The market potential approach is in fact a reduced form with no clear link to theory, whereas the the wage curve does not include a spatial structure. More research is clearly needed here, but at present we must conclude that from an empirical point of view the choice is more difficult.

## **7. 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 shortcuts inevitable. In this paper 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. Using data for German districts we use the so-called Helpman-Hanson model to investigate the existence of a spatial wage structure and to estimate the key model parameters.

We modify and extend the work by Hanson (1998, 2001a) in two ways. First, in order to deal with housing market we not only use data on the housing stock but also on land prices. Subsequently, and more importantly, we drop the assumption of real wage equalization. It is only when we no longer have to assume that real wages are not equalized across re-unified Germany that we find clear cut evidence for both a spatial wage structure and the relevance of structural model parameters. A first brief comparison of our preferred model specification with two alternative explanations suggests that the new economic geography approach is a serious alternative to other explanations of the spatial distribution of wages. A straightforward next step is to analyze the development of a spatial wage structure over time using a new economic geography approach and to look at other European countries besides Germany. In doing so, a major challenge will be to take into account that many European countries are characterized by various forms of wage rigidity.

## Appendix A Derivation of the Spatial Wage Equation (2)

First look at the demand side of the Helpman (1998) model. Assume an economy with two sectors, Housing (H) and Manufacturing (M). Every consumer in the economy shares the same, Cobb-Douglas, preferences for both types of commodities:

$$U = M^\delta H^{(1-\delta)}$$

The parameter  $\delta$  is the share of income spent on manufactured goods. Where M is a CES sub-utility function of many varieties.

$$M = \left( \sum_{i=1}^n c_i^\rho \right)^{1/\rho}$$

Maximizing the sub-utility subject to the income constraint gives the demand for each variety, j:

$$c_j = p_j^{-\varepsilon} I^{\varepsilon-1} \delta Y, \text{ where } I = \left[ \sum_i (p_i)^{(1-\varepsilon)} \right]^{1/(1-\varepsilon)}$$

is the price index for manufactures,  $\varepsilon = \frac{1}{1-\rho}$  the elasticity of substitution, and Y=income.

Next, turn to the supply side. Each variety, i, is produced according to:

$$l_i = \alpha + \beta x_i$$

where,  $l_i$  is the amount of labor necessary to produce  $x_i$  of variety  $i$ . The coefficients  $\alpha$  and  $\beta$  describe, respectively, the fixed and marginal labor requirement. Maximizing profits gives the familiar mark-up pricing rule:

$$p \left( 1 - \frac{1}{\varepsilon} \right) = w \beta,$$

Using the zero profit condition and the mark-up pricing rule above, gives the break-even supply of a variety i (each variety is produced by a single firm):

$$x_i = \frac{\alpha(\varepsilon-1)}{\beta}$$

Furthermore, transportation of manufactures is costly. Transportation costs are so-called iceberg transportation costs,  $T = T^{D_{12}} > 1$ , where  $D_{12}$  is the distance between region 1 and 2. Also assume, for illustration purposes, that there only two regions, 1, and 2. Total demand for a product from, for example region 1, now comes from two regions, 1 and 2. The consumers in region 2 have to pay transportation costs on their imports. This leads to the following total demand for a variety produced in region 1:

$$x_1 = \delta(Y_1 p_1^{-\varepsilon} I_1^{\varepsilon-1} + Y_2 p_1^{-\varepsilon} (T^{D_{12}})^{-\varepsilon} I_2^{\varepsilon-1})$$

We already know that the break-even supply equals  $x_1 = \frac{\alpha(\varepsilon-1)}{\beta}$ , equating this to

total demand gives (note that the demand from region 2 is multiplied by  $T^{D_{12}}$  in order to compensate for the part that melts away during transportation):

$$\frac{\alpha(\varepsilon-1)}{\beta} = \delta(Y_1 p_1^{-\varepsilon} I_1^{\varepsilon-1} + Y_2 p_1^{-\varepsilon} (T^{D_{12}})^{1-\varepsilon} I_2^{\varepsilon-1})$$

Inserting the mark-up pricing rule in this last equation and solving for the wage rate gives the two-region version of the equilibrium wage equation (2) in the main text. Also, inserting the mark-up pricing rule into the price index above, gives this index in terms of the wage rate, as in equation (3) in the main text.

## Appendix B Derivation of the Reduced-Form Wage Equation (7'') with a Marginal Labor Productivity Gap between eastern and western Germany

Assume a productivity gap between eastern and western Germany:

$$L_{ir} = \alpha + (1 + \gamma_{ir})\beta x_{ir}$$

where  $L_{ir}$  is employment in firm  $i$  in region  $r$ ,  $x$  is output,  $\gamma_{ir}$  measures the marginal labor productivity gap between western Germany and a firm  $i$  in region  $r$  and is equal to  $(MPL_{west}/MPL_{ir}) - 1$ .

Assume a uniform level of MPL in western Germany, and a uniform level of MPL in eastern Germany: for any firm  $i$  in region  $r$  in western Germany  $MPL_{ir} = MPL_{west}$ , and for any firm  $i$  in region  $r$  in eastern Germany  $MPL_{ir} = MPL_{east} = \frac{MPL_{west}}{1 + \gamma}$ ; so  $\gamma_{ir} = 0$  if region  $r$  is in western Germany, and  $\gamma_{ir} = \gamma > 0$ , if region  $r$  is in eastern Germany.

Free entry and exit and using the zero-profit condition leads to the equilibrium output for firm  $i$  in region  $r$  (see Brakman, Garretsen and van Marrewijk, 2001, pp 78-79):

$$x_{ir} = \frac{\alpha(\varepsilon - 1)}{(1 + \gamma_{ir})\beta}$$

The equilibrium demand facing each firm  $i$  is

$$x_d = \frac{p^{-\varepsilon}}{I^{(1-\varepsilon)}} \delta Y$$

Summing over all firms, using the mark-up pricing rule  $p = \frac{\varepsilon}{\varepsilon - 1}(1 + \gamma_r)\beta W$ , and

taking iceberg transport costs into account gives:

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

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

This expression is equal to the break-even supply of each firm:

$$\frac{\alpha(\varepsilon - 1)}{(1 + \gamma_r)\beta} = \delta \left( \frac{\varepsilon}{\varepsilon - 1} W_r (\beta(1 + \gamma_r)) \right)^{-\varepsilon} \sum_{s=1}^R \left[ \left( \frac{T^{D_{rs}}}{I_s} \right)^{1-\varepsilon} Y_s \right]$$

The wage rate in region  $r$  determined by solving this equation for the wage rate, giving:

$$W_r = (1 + \gamma_r)^{1-\varepsilon/\varepsilon} \left( \frac{\varepsilon - 1}{\varepsilon\beta} \right) \left( \frac{\beta\delta}{(\varepsilon - 1)\alpha} \right)^{\frac{1}{\varepsilon}} \sum_{s=1}^R \left[ \left( \frac{T^{D_{rs}}}{I_s} \right)^{1-\varepsilon} Y_s \right]$$

The log transformation of this expression results in the log transformation of wage equation (2'):

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

where  $\frac{1-\varepsilon}{\varepsilon} < 0$ .

The productivity gap between western and eastern Germany also affects the price index equation (3'), because marginal costs changes into:

$$MC_{ir} = W_{ir}\beta(1+\gamma_{ir}),$$

and so the simplified price index equation (3') becomes –dropping subscript  $i$ :

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

## References

- Anas, A., R. Arnott, and K.A. Small. 1998. Urban Spatial Structure, *Journal of Economic Literature*, Vol. XXXVI, 1426-1464.
- Anselin, L. and S. Hudak. 1992. Spatial Econometrics in Practice. *Regional Science and Urban Economics*, Vol. 22, 509-536.
- Blanchflower, D.G. and A.J. Oswald. 1994. *The Wage Curve*, MIT Press, Cambridge Mass.
- Brakman, S. and H. Garretsen. 1993. The Relevance of Initial Conditions for the German Unification, *Kyklos*, vol. 46, 163-181.
- Brakman, S., H. Garretsen and 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, and Ch. van Marrewijk. 2001. *An introduction to Geographical Economics: Trade, Location and Growth*, Cambridge University Press, Cambridge.
- Brakman, S., H. Garretsen, and M. Schramm. 2002a. New Economic Geography in Germany: Testing the Helpman-Hanson Model, *HWWA Discussion Paper 172*, Hamburg (see <http://www.hwwa.de/hwwa.html>).
- Brakman, S., H. Garretsen, and M. Schramm. 2002b. The Final Frontier? Border Effects and German Regional Wages, *HWWA Discussion Paper 197*, Hamburg (see <http://www.hwwa.de/hwwa.html>).
- Combes, P.-H., M. Lafourcade .2001. Transport Costs Decline and Regional Inequalities: Evidence from France, *CERAS Working Paper*, 01-01, Paris.
- Faini, R. 1999. Trade Unions and Regional Development, *European Economic Review*, 43(2), 457-474.
- Fujita, M., P. Krugman and A.J. Venables. 1999. *The Spatial Economy*, MIT Press.
- Hall, R.E. 1988. The Relation between Price and Marginal Cost in U.S. Industry, *Journal of Political Economy*, no. 5, 921-947.
- Hanson, G.H. 1997. Increasing Returns, Trade and the Regional Structure of Wages, *Economic Journal*, vol. 107, 113-133.
- Hanson, G.H. 1998. Market Potential, Increasing Returns, and Geographic Concentration, *NBER Working Paper 6429*, Cambridge Mass.
- Hanson, G.H. 2001a. Market Potential, Increasing Returns, and Geographic Concentration, *mimeo*, Graduate School of International Relations and Political Studies, UC San Diego (revised version of Hanson. 1998). (<http://www-irps.ucsd.edu/faculty/gohanson/Uscenty.pdf>)
- Hanson, G.H. 2001b. Scale Economies and the Geographic Concentration of Industry, *Journal of Economic Geography*, Vol. 1(3), 255-276.
- Harrigan, J. 2001. Specialization and the Volume of Trade: Do the data Obey the Laws?, *NBER Working Paper 8675*, Cambridge Mass.
- Head, K and T. Mayer. 2003. The Empirics of Agglomeration and Trade, in V. Henderson and J-F. Thisse (eds.) *The Handbook of Regional and Urban Economics, volume IV*, North Holland, forthcoming.
- Haskel, J., C. Martin, I. Small. 1995. Price, Marginal Cost and the Business Cycle, *Oxford Bulletin of Economics and Statistics*, no. 1, 25-41.
- 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*, nr. 3, 483-499.
- Krugman, P. 1998. Space, the Final Frontier, *Journal of Economic Perspectives*, 161-174.
- Krugman, P., and A.Venables. 1995. Globalization and the inequality of nations, *Quarterly Journal of Economics*, Vol. 110, 857-880.
- Krugman, P, and A. Venables. 1996. Integration, Specialization, and Adjustment, *European Economic Review*, vol. 40, 959-967.
- Norrbin, Stefan C. 1993. The Relation between Price and Marginal Cost in U.S. Industry: A Contradiction, *Journal of Political Economy*, no. 6, 1149-1164.
- Ottaviano, G.I.P. and J-F Thisse. 2001. On Economic Geography in Economic Theory: Increasing Returns and Pecuniary Externalities, *Journal of Economic Geography*, 1(2), 153-180.
- Overman, H., S. Redding and A.J. Venables. 2001. Trade and Geography: A Survey of the Empirics, *mimeo*, London School of Economics, forthcoming in the *Handbook of International Trade*, Basil Blackwell.
- Peeters, J. and H. Garretsen. 2000. Globalisation, Wages and Unemployment: An Economic Geography Perspective, *CESifo Working Paper*, no. 256, Munich, forthcoming in S. Brakman and B. Heijdra (eds.). 2003. *The Monopolistic Competition Revolution in Retrospect*, Cambridge University Press, Cambridge.
- Puga, D. 1999. The rise and fall of regional inequalities, *European Economic Review*, Vol. 43, 303-334.
- Puga, D. 2002. European Regional Policies in Light of Recent Location Theories, *Journal of Economic Geography*, Vol. 2, 373-406.
- Neary, J.P. 2001. Of hype and hyperbolas: introducing the New Economic Geography, *Journal of Economic Literature*, vol. XXXIX, 536-561.
- Roos, M. 2001. Wages and market potential in Germany, *Jahrbuch für Regionalwissenschaft*, 21, 171-195.
- Sinn, H-W. 2000. Germany's Economic Unification, An Assessment after Ten Years, *CESifo Working Paper*, no. 247, Munich.
- 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, 341-359.

<i>Table 1 Spatial land price structure</i>	
$\kappa_1$	0.408 (6.8)
$\kappa_2$	0.063 (3.8)
$\kappa_3$	-0.586 (-4.1)
$\kappa_4$	-1.361 (-10.5)
Adjusted R <sup>2</sup>	0.726

# obs.=146, estimation method: weighted least squares (WLS), result for constant not reported, t-statistic between brackets.

*Table 2: District-specific control variables (cross-section of 151 German districts) measured in shares of total employment*

	Vocational training (1998)			Sectoral employment (1995)				
	no	intermediate	high	agriculture	industry	services (TTC)*	other services	Public Sector
West German average	.203 (.029)	.636 (.039)	.075 (.035)	.028 (.025)	.347 (.091)	.198 (.037)	.225 (.051)	.202 (.071)
East German average	.105 (.014)	.695 (.051)	.105 (.040)	.035 (.026)	.350 (.076)	.175 (.021)	.204 (.040)	.236 (.063)

Unweighted standard deviation between brackets;

\* Trade, transport & communication.



	(I), Eq.(7) with housing stock and no EG dummy, (Number of obs. 151)	(II), Eq(7) with housing stock and EG dummy (Number of obs. 151)	(III), Eq. (7'), with land prices and no EG dummy, (Number of obs. 146)	(IV), Eq. (7') with land prices and EG dummy, (Number of obs. 146)
$\delta$	1.333 (3.0)	12.801 (0.1)	0.684 (28.4)	0.536 (12.0)
$\varepsilon$	3.623 (9.2)	3.105 (1.9)	4.213 (15.9)	4.987 (6.6)
Log(T)	0.007 (8.2)	-0.001 (-2.0)	0.010 (10.2)	-0.001 (-4.6)
District Specific Control Variables (dummies)				
EG	--	-.781 (-12.2)	--	-.811 (-12.1)
D <sub>country</sub>	-.056 (-1.4)	-.029 (-.9)	-.128 (-2.9)	-.024 (-.6)
Industry	.031 (.1)	1.235 (5.6)	-.494 (-1.7)	1.434 (6.3)
Other services	1.165 (2.4)	1.738 (4.6)	.409 (.8)	1.873 (4.8)
Low-skilled	-.386 (-.8)	-.686 (-1.3)	.228 (.4)	-1.286 (-2.1)
High-skilled	2.840 (4.2)	6.087 (10.8)	2.316 (3.3)	6.238 (11.0)
Adjusted R <sup>2</sup>	0.99	0.99	0.99	0.99
Implied values:				
$\varepsilon/(\varepsilon-1)$	1.38	1.48	1.31	1.25
$\varepsilon(1-\delta)$	-1.21 (-1.3)*	-36.64 (-0.1)*	1.33 (2.5)*	2.31 (4.5)*

Estimation method: weighted least squares (WLS); results for the constants, the dummy for the city of Erlangen are not reported, as well as the result for the dummy for Hamburg in (7'); t-statistic between brackets; \* H<sub>0</sub>:  $\varepsilon(1-\delta) = 1$ . The District specific variables, Agriculture, and Trade, Transport and Communication were insignificant, and are not reported.

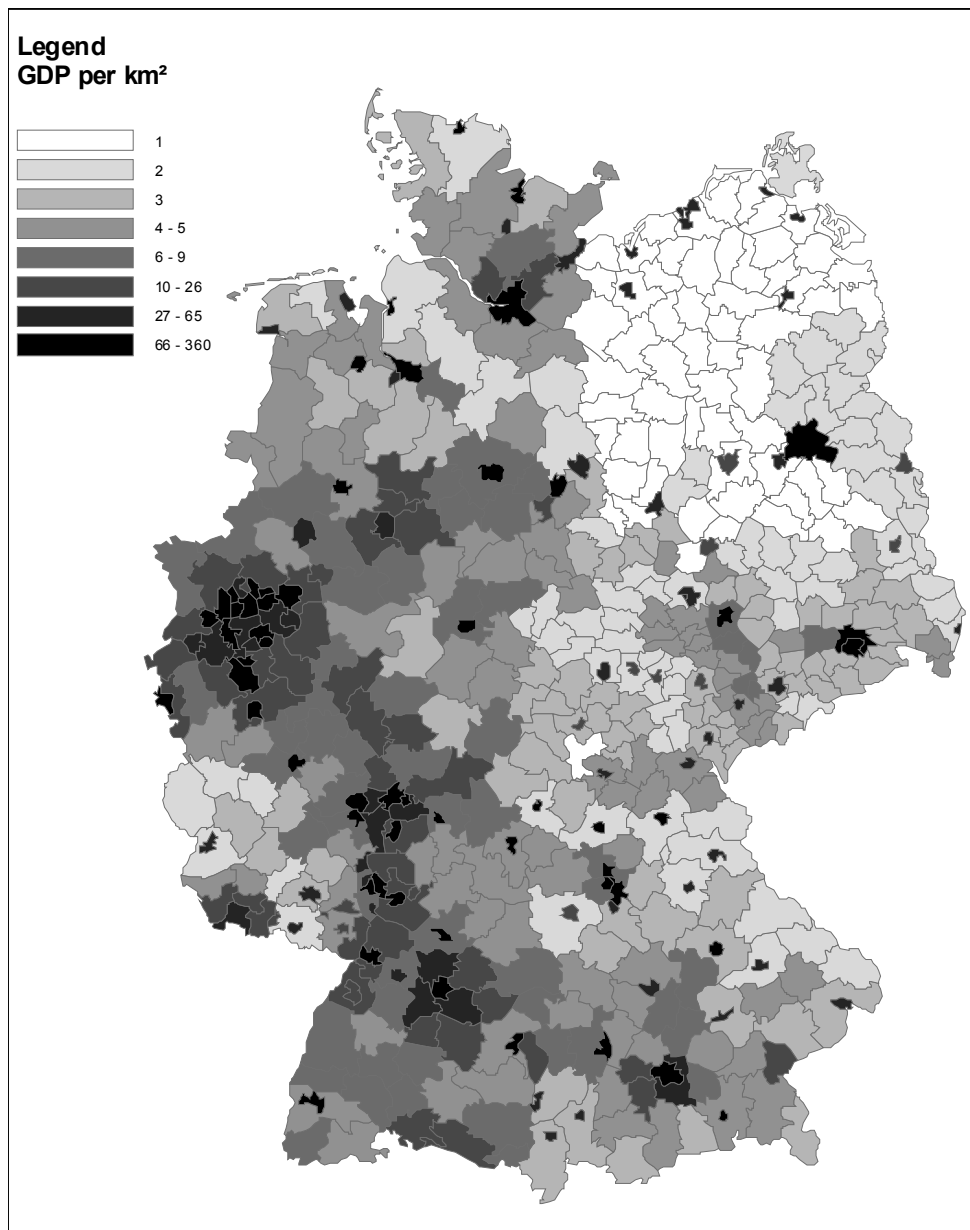
<i>Table 4 No Real Wage Equalization</i>	
	Equation (7''), with EG dummy
$\varepsilon$	3.652 (23.4)
Log(T)	0.003 (13.7)
District Specific Control Variables (dummies)	
EG	-0.633 (-16.2)
$D_{\text{country}}$	-0.056 (-1.4)
Industry	1.052 (5.8)
Other services	1.983 (5.4)
High-skilled	5.456 (11.7)
Adjusted R <sup>2</sup>	0.99

Estimation method for equation (7''): WLS; number of observations 151; district-specific control variables that are not statistically significant are omitted; result for the dummy for the city of Erlangen is not reported, t-statistic between brackets.

<i>Table 5 Alternative models</i>		
	<i>Market potential equation (1')</i>	<i>Wage curve (8)</i>
$\kappa_1$	0.049 (3.5)	--
$\kappa_2$	0.092 (2.0)	--
$\kappa_3$	-0.655 (-8.778)	--
$\kappa_4$	-0.126 (-3.1)	--
$\beta_1$	--	0.104 (2.4)
$\beta_2$	--	-0.002 (-0.4)
District Specific Control Variables (dummies)		
EG	-.654 (-9.8)	-.653 (-8.7)
D <sub>country</sub>	-.126 (-3.1)	--
Industry	.787 (3.8)	.905 (4.7)
Other services	1.150 (3.0)	1.641 (4.4)
Low-skilled	-.924 (-2.0)	-1.025 (-2.0)
High-skilled	4.823 (9.2)	4.979 (9.0)
Adj. R <sup>2</sup>	0.99	0.99
Akaike information criterion (AIC)	-0.989	-0.947
AIC for eq. (7'), with simplified price index (reduced form)	-0.931	

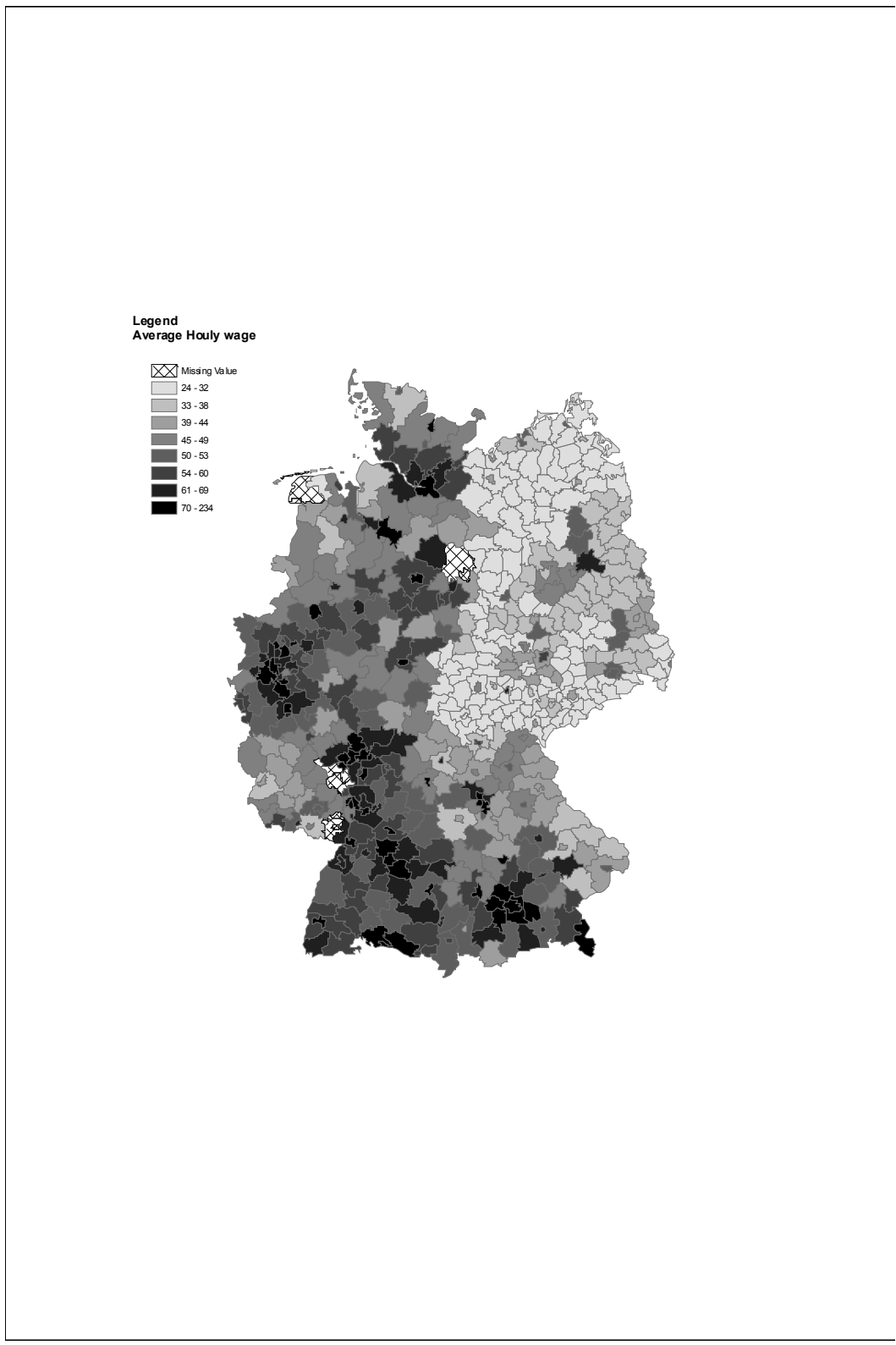
# obs.=151, estimation method: weighted least squares (WLS), result for constant not reported, result for dummy for the city of Erlangen (8') is not reported, t-statistic between brackets. District-specific control variables that are not statistically significant are omitted. Results for Akaike info criterion are based on non linear least squares (NLS) estimations, this criterion requires that the dependent variables are the same which is not the case with WLS.

**Map 1. GDP per km<sup>2</sup> (in millions of DM, 1994)**



**Source: Federal Statistical Office, Wiesbaden**

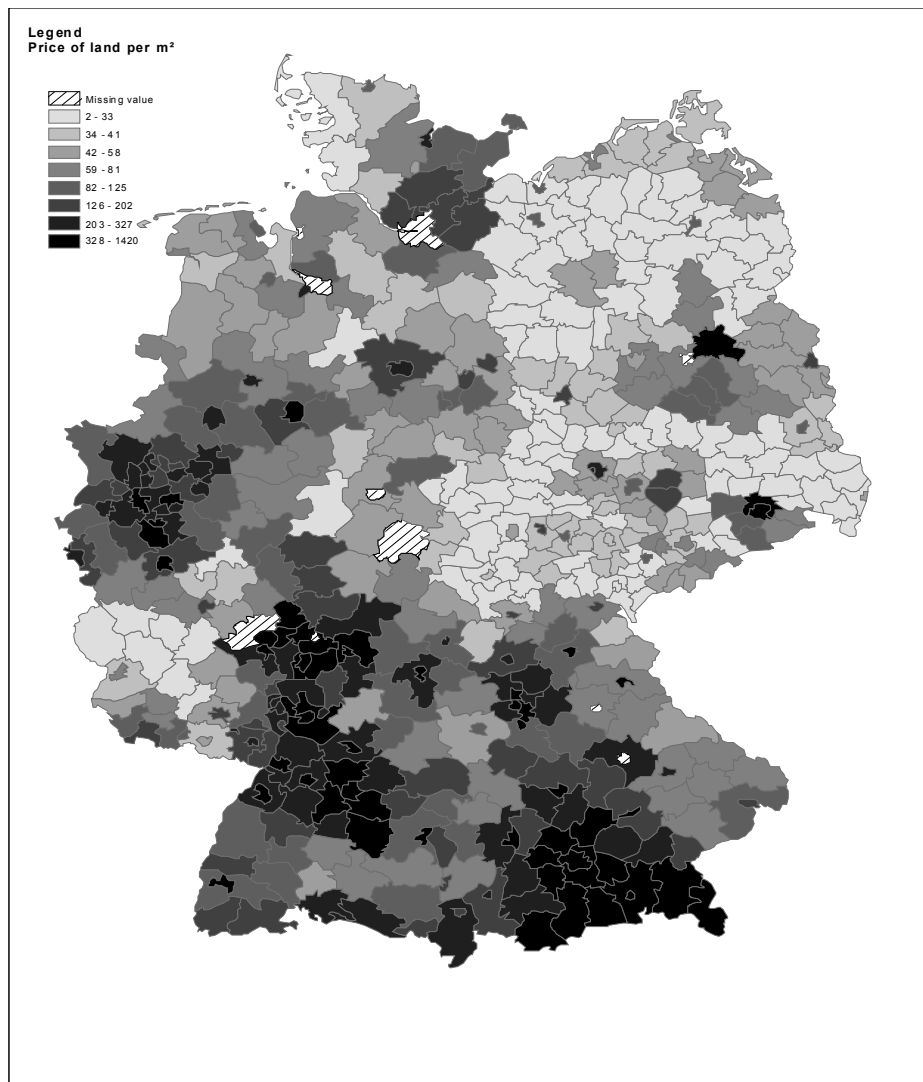
**Map 2. Average Hourly Manufacturing Wage (DM, 1995)**



#obs.=432. Data on wages in the districts Alzey-Worms, Aurich, Emden, Gifhorn, Landau in der Pfalz, Mainz-Bingen, Neustadt an der Weinstraße, Suedliche Weinstraße, Wolfsburg are missing (see the shaded areas on the map)

**Source: Federal Statistical Office, Wiesbaden.**

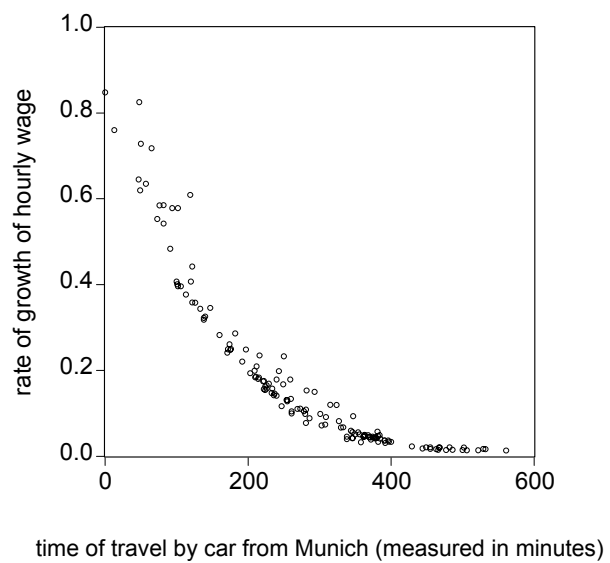
**Map 3. Land prices per m<sup>2</sup> (in DM, 1995)**



**Source: Federal Statistical Office, Wiesbaden**

#obs= 433. Data on land prices in districts Amberg (city district), Bremen, Bremerhaven, Hamburg, Hersfeld-Rotenburg, Kassel (city district), Offenbach, Potsdam, Regensburg (city district) and Rheingau-Taunus-kreis and are missing (these are shaded in the map).

**Figure 1. Wage growth (%) and distance, following a 10% increase of GDP in Munich**



Source: own calculations.

---

<sup>1</sup> 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.

<sup>2</sup> For a similar attempt see Roos (2001).

<sup>3</sup> For a discussion on 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 analytically similar to the input-output models, see Puga (1999, p.324), Puga (2002, p. 388), Ottaviano and Thisse (2001, p. 175).

<sup>4</sup> For an in-depth analysis of the Krugman (1991) model see Fujita, Krugman and Venables (1999, chapters 4 and 5) or Brakman, Garretsen and van Marrewijk (2001, chapters 3 and 4). For the full specification of the Helpman-Hanson model, see Hanson (1998, pp. 9-12).

<sup>5</sup> In Krugman (1991) low (high) transportation costs lead to agglomeration (dispersion), whereas in Helpman (1998) it is the other way around.

<sup>6</sup> This condition determines whether or not transport costs matter. If the condition is not met, agglomeration always occurs, irrespective of transport costs.

<sup>7</sup> Note that the housing stock is owned by absentee landlords. No income is gained from owning houses. Dropping this assumption (by endowing each worker with a share of the housing stock) has no impact on the empirical specifications below (equation (7)), compare Hanson (1998) with Hanson (2001a).

<sup>8</sup> Note that direct observation of a price index of housing services would imply that one only has to use equation (6) and that the use of equation (5) is no longer necessary. We will return to this in section 4.

<sup>9</sup> Overman, Redding and Venables (2001, p. 17) and Head and Mayer (2003) discuss how the model used by Hanson can be seen as a specific version of a more general new economic geography model.

<sup>10</sup> Note that this equation resembles the spatial lag model of Anselin and Hudak (1992), although that model is linear whereas equation (7) is non-linear.

<sup>11</sup> From a total of 441 districts there are 322 districts that do not correspond with a city. Each of these 322 districts comprise of several villages and towns. In order to simplify the distance matrix considerably we aggregated these 322 districts at the higher level of *Bezirke*. This leads to 37 *country districts*.

<sup>12</sup> In Hanson (2001a, pp. 15-16) a different aggregation scheme is used. For each US county (the most disaggregated level of geographical analysis), surrounding counties are grouped within concentric distance bands (up to a distance of a 1000 km. these bands have a width of 100 km.) and then the independent variables  $Y_s$  and  $W_s$  are aggregated across the counties within each band.

<sup>13</sup> Inspection of the wage data revealed that there is one very large outlier, the city of Erlangen (Bavaria), which is the city referred to as the German capital of the pharmaceutical industry and the city in which the Siemens company has its headquarters. Erlangen has by far the highest wage, so we included a dummy for this district as well.

<sup>14</sup> 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 the outcomes but this was not the case. We did not control for fixed regional endowments like climate. Hanson (2001a) does control for these endowments in his study for the USA but for a relatively small country like Germany these kind of differences are assumed not to be relevant.

<sup>15</sup> Estimates of the margin of price over marginal costs in industry are ranging from 14% in the US (Norrbinn, 1993), 100% in the UK (Haskel, Martin & Small, 1995), to over 100% in the US (Hall, 1988).

<sup>16</sup> In Krugman (1991) the no black hole condition is met if  $\varepsilon(1-\delta)>1$ . Helpman (1998) shows that 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 services which are in fixed supply).

<sup>17</sup> Restricting  $\delta$  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 variance in wages, as compared to 48% in the unrestricted specification. A likelihood ratio test



indicates that the restricted model has to be rejected as being inferior compared to the unrestricted model.

<sup>18</sup> For each district  $r$  the weighted average distance to the other districts  $\sum_s weight_s D_{rs}$  is calculated, using  $weight_s = Y_s / \sum_j Y_j$ . The district with the smallest average distance is the economic center.

<sup>19</sup> Under the assumptions that there is a uniform level of  $MPL$  in western Germany ( $MPL_{west}$ ), which is higher than the uniform level of  $MPL$  in eastern Germany; see Appendix B for the derivation of (2') and (3'').

<sup>20</sup> We also captured the possibility of no real wage equalization in a more simple way by sticking to wage equation (7) and by allowing only for a real wage differential between on the one hand western

and on the other hand eastern German regions by changing equation (6) into  $\frac{W_r}{P_r^{1-\delta} I_r^\delta} = \frac{W_s}{P_s^{1-\delta} I_s^\delta} \phi_{rs}$ ,

where  $\phi_{rs} = \phi > 1$ , if  $r$  is West and  $s$  is East, and  $\phi_{rs} = 1/\phi < 1$ , if  $r$  is East and  $s$  is West.  $\phi$  represents the real-wage gap between East and West Germany (incomplete real-wage equalization). It turned out that  $\phi = 1.406$  (t-value 4.768) thereby validating that real wages are higher in western German regions.

<sup>21</sup> As a verification on the validity of the results in Table 3 we calculated Moran's correlation coefficient to check for spatial dependencies (and to test for possible mis-specification). It turns out that, using the residuals from the estimation of equation (7''), that there is no spatial dependency.

<sup>22</sup> One might expect that the massive transfers between western and eastern Germany are also an important institutional feature to take into account. We checked for this by using personal income (which includes transfers) instead of GDP as a measure for the variable  $Y$ . Also, we included the ratio of each district's GDP to personal income as an additional explanatory variable in equation (7).

However, this had virtually no impact on the estimation results for the key model parameters.

<sup>23</sup> 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.

<sup>24</sup> Based on a simple market potential function like equation (1), Brakman, Garretsen and Schramm (2002a) test for a spatial employment as well as and an unemployment structure and they find confirmation for the former but, also due to data limitations, not for the latter. Peeters and Garretsen (2000) show how unemployment can be incorporated into a basic new economic geography model.

<sup>25</sup> Using a similar test, the Schwarz-criterion, Hanson (2001a) finds for the case of the USA that wage equation (7), estimated in first differences, has always to be preferred compared to the market potential function.