HWWA HAMBURG

DISCUSSION PAPER

Market Access and Regional Disparities

New Economic Geography in Europe

Annekatrin Niebuhr

HWWA DISCUSSION PAPER

269

Hamburgisches Welt-Wirtschafts-Archiv (HWWA)
Hamburg Institute of International Economics
2004
ISSN 1616-4814

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

The HWWA is a member of:

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

HWWA Discussion Paper

Market Access and Regional Disparities

New Economic Geography in Europe

Annekatrin Niebuhr

HWWA Discussion Paper 269
http://www.hwwa.de

Hamburg Institute of International Economics (HWWA)
Neuer Jungfernstieg 21 - 20347 Hamburg, Germany
e-mail: hwwa@hwwa.de

I would like to thank seminar participants at the ERSA Congress 2003, University of Nijmegen, HWWA Hamburg, TU Dresden and the meeting of the Ausschuss für Regionaltheorie und –politik des Vereins für Socialpolitik at the University of Eichstätt-Ingolstadt for helpful comments on an earlier version of this paper. Financial support from the Fritz Thyssen Stiftung is gratefully acknowledged as part of the project "Border Regions, Border Impediments and European Integration".

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

Edited by the Department European Integration
Head: Dr. Konrad Lammers

HWWA DISCUSSION PAPER 269
February 2004

Market Access and Regional Disparities
New Economic Geography in Europe

ABSTRACT

New Economic Geography (NEG) has reached a theoretical consolidation while related empirical tests are still scarce. The present paper aims at providing some evidence on the validity of forces emphasised by NEG. The analysis starts from the nominal wage equation derived from the Krugman "core-periphery model" and focuses on one of the main propositions of NEG that access advantages raise factor prices. The paper investigates the significance of market access for regional wages and the geographic extent of demand linkages for a cross section of European regions, also taking into account the effects of national borders. The regression analysis covers the period between 1985 and 2000. The results are consistent with the implication of NEG that demand linkages affect the geographic distribution of economic activities, confirming the basic findings of previous analyses. However, regarding the spatial extent of demand linkages, our results differ significantly from previous findings that point to highly localised effects.

Keywords: New economic geography, market access, Europe
JEL-Classification: C21, F12, R12

Annekatrin Niebuhr
Hamburg Institute of International Economics
Department of European Integration
Neuer Jungfernstieg 21
D 20347 Hamburg - Germany
Phone: +49 40 42834-410
Fax: +49 40 42834-451
E-mail: niebuhr@hwwa.de

1 Introduction

Initiated by the seminal contribution by Krugman (1991), New Economic Geography (NEG) developed as a theory that explains the emergence of a heterogeneous economic space. In NEG models, agglomeration of economic activities and population bases on increasing returns to scale and transport costs. Various models investigate the interaction of centrifugal and centripetal forces which shape the economic landscape. In a recent review, Neary (2001) states that "New economic geography has come of age". Indeed, Ottaviano (2002) concludes that NEG has reached a theoretical consolidation. Both authors emphasise two directions for future research: policy and empirics. As Head and Mayer (2003) note in a survey, empirical research on NEG is lagging behind. Moreover, results of existing studies are difficult to compare, since different types of methodology are in use.

The aim of the present paper is to add to the empirical evidence on NEG. The analysis focuses on one of the main propositions of NEG that access advantages raise factor prices. We investigate whether there exists a spatial wage structure in the European Union, i.e. whether there is a positive correlation between regional wage level and access to purchasing power. The study departs from the nominal wage equation derived from the Krugman "core-periphery model". However, following the advice by Leamer and Levinsohn (1995), we "estimate and don't test". What this means is, that instead of testing a specific NEG model against alternative explanations of agglomeration, a specific model as well as the so called market potential concept by Harris (1954), consistent with different NEG models, is applied. By investigating the validity of forces emphasised by NEG in a more general manner, we are able to compare our findings with results of related studies. In this context, we pay special attention to the geographic extent of demand linkages identified in different analyses.

There is a long tradition of studies dealing with the significance of market access for regional disparities in Europe. Early studies by Clark et al. (1969) and Keeble et al. (1982) investigate the impact of European integration on the spatial structure of economic activity in the EU, based on market access considerations. However, the definite relevance of market access for regional development remains vague in these analyses, since a firm theoretical fundament is missing. NEG can supply this theoretical base, because it establishes the missing economic link between market access and regional development. Current analyses use this theoretical framework in order to investigate mechanisms of regional development. Previous studies have tended to focus on evidence for single countries: Tests of spatial wage structures have been conducted for US and German regions by Hanson (2000), Roos (2001) and Brakman et al. (2002) and for Italy by Mion (2003). In contrast, this paper aims at analysing the spatial wage structure for a cross section of European regions. The regression analysis covers the period between 1985 and 2000.

The rest of the paper is organised as follows. Section 2 comprises a brief description of the theoretical model that constitutes the theoretical framework of the empirical analysis. In section 3, the regression models are derived and the empirical implementation is illustrated. In section 4 previous empirical evidence on the market potential is summarised. Data and regional system are described in section 5. The results of the regression analysis are presented in section 6. Section 7 concludes.

2 The Theoretical Model

In NEG models, the interaction of transport costs and increasing returns to scale generates demand linkages and serves as an explanation for agglomeration. The location of firms and consumers is endogenous. There is no need for exogenous location advantages or disadvantages in these models to explain location choices. Agglomeration is caused by a circular relationship in which the spatial concentration of manufacturing both creates and follows market access. Firms locate in regions with large markets and in turn market size in these regions increases due to the growth of manufacturing. Preferred locations are those with good access to the market and to the supply of manufacturing goods (see Krugman 1992).

Recent empirical investigations on market access and interregional wage differences depart from Krugman's core-periphery model (Krugman 1991) and the version by Helpman (1998) respectively. The general structure of both models is similar. The main difference concerns the immobile factor introduced in the model, farmers in the Krugman model and housing in the Helpman model. The present analysis is based on the Krugman model, since data on housing, necessary for a test of the Helpman model, is not available for a cross section of European regions. The structure of the Krugman model is as follows. The economy consists of a manufacturing and an agricultural sector. All consumers have identical Cobb-Douglas tastes for two types of goods:

(1) $U = C_M^{\mu} C_A^{1-\mu}$

where μ is the share of expenditure on manufactures, C_A is the quantity of the agricultural good consumed, and C_M is a composite of symmetric product varieties given by:

(2) $C_M = \left[\sum_{k=1}^{K} c_k^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$

σ is the constant elasticity of substitution between any pair of varieties and K is the number of varieties. Consumers have a love for variety. With increasing σ, the substitutability among varieties rises, thus the desire to spread consumption over manufactured goods declines.

There are two factors of production: mobile manufacturing workers L_M and immobile farmers L_A. Labour is distributed across a number of regions $j = 1,...,J$. Agricultural production takes place under constant returns to scale. By choice of units, farm labour in region j can be set equal to agricultural production in region j:

(3) $L_{Aj} = Q_{Aj}$

In contrast, there are increasing returns in the production of each individual variety of manufactured goods:

(4) $L_{Mjk} = a + bQ_{Mjk}$

Because of increasing returns, each variety is only produced in one region. Thus regions do not produce the same set of products, but differentiated bundles of manufactured goods. The number of produced varieties is proportional to the region's manufacturing labour force. If labour supply increases due to immigration, the number of supplied goods will increase.

Manufactured goods are traded among regions incurring iceberg transport costs, i.e. a fraction of any good shipped melts away and only the part υ_{ij} arrives at its destination:

(5) $\upsilon_{ij} = e^{-\tau d_{ij}}$

where τ is transport costs per distance unit and d_{ij} is the distance between the regions i and j.

Manufacturing workers move from regions with below-average real wages towards locations with above-average real wages. In equilibrium, workers have no incentive to relocate. Immobility of workers will be ensured, if real wages are equal in all regions:

$$(6)\ \frac{w_i}{T_i^{\mu}} = \frac{w_j}{T_j^{\mu}}$$

The nominal wage is given by w_j, and T_j is the price index for manufactures in region j.[1] Real wages are influenced by the geographic distribution of industry. Backward and forward linkages might cause a spatial concentration of workers and firms. The concentration of firms raises real wages in the corresponding region via a decline of the price index of manufacturing goods and thereby increases the attractiveness of the location for mobile workers (forward linkage). Large markets, i.e. locations with many workers, in turn are attractive production sites for the industrial sector and allow firms to reward their workers with higher wages (backward linkage). Thus there is a mechanism of cumulative causation which might result in spatial concentration of manufacturing. The distribution of firms and workers across space depends on the relative strength of centripetal and centrifugal forces. The centrifugal force in this model bases on the exogenous location of agricultural workers and the desire of manufacturing producers to get away from competitors. The attractiveness of agglomeration for firms and workers constitutes the centripetal force (see Krugman 1992).

The price index for manufactures in region j is given by:

$$(7)\ T_j = \left[\sum_{i=1}^{J} \lambda_i (w_i e^{\tau d_{ij}})^{1-\sigma} \right]^{\frac{1}{1-\sigma}}$$

The price index accounts for transport costs. It is higher in regions where a large fraction of manufactured goods has to be imported from distant locations. λ_i is the share of

1 All prices and wages are measured in terms of the agricultural good.

firms (and hence manufactured goods produced) in region *i*. Finally, it can be shown that the equilibrium wage rate is given by:[2]

$$(8)\quad w_j = \left[\sum_{i=1}^{J} Y_i e^{-\tau(\sigma-1)d_{ij}} T_i^{\sigma-1} \right]^{1/\sigma}$$

with Y_i as income in region *i*. This is the so-called nominal wage equation which is point of departure of most studies that investigate the existence of a spatial wage structure for different cross sections. According to equation (8), the nominal wage level in region *j* depends on a weighted sum of purchasing power in all accessible regions *i*, whereby the weighting scheme is a function declining with increasing distance between locations *i* and *j*. As Hanson (2000) notes, equation (8) can be thought of as a spatial labour demand function in an economy with perfect labour mobility. Labour demand and wages are relatively high in locations close to high consumer demand. Regional wages increase with income of neighbouring regions and decline with rising transport costs to these locations.

The nominal wage equation represents one of the main propositions emerging from NEG models mentioned by Head and Mayer (2003): access advantages raise local factor prices. More precisely, production sites with good access to major markets because of relatively low trade costs tend to reward their production factors with higher wages and land rentals. Moreover, the equation resembles the market potential concept introduced by Harris (1954). The market potential concept states that the attractiveness of a region as a production site depends on its access to markets.

3 Regression Model

The nominal wage equation (8) cannot be estimated directly since data for regional price indices T_i are not available. In principle, there are two strategies to eliminate T_i in order to arrive at an estimable specification. Firstly, the equilibrium condition (6) can be used to substitute $T_{i.}$ The corresponding regression model is given by:

2 For a detailed description of the derivation of equilibrium equation (8) see Krugman (1992).

$$(9)\quad \log(w_j) = \theta + \frac{1}{\sigma}\log\left(\sum_{i=1}^{J} Y_i w_i^{\frac{\sigma-1}{\mu}} e^{-\tau(\sigma-1)d_{ij}}\right) + \varepsilon_j = \alpha_0 + \alpha_1 \log\left(\sum_{i=1}^{J} Y_i w_i^{\alpha_3} e^{-\alpha_2 d_{ij}}\right) + \varepsilon_j$$

where θ is a function of fixed parameters and ε_j is an error term.

The second method to avoid the requirement of regional price data is to assume that the price index is equal in all regions ($T_i = T$):

$$(10)\qquad \log(w_j) = \alpha_0 + \alpha_1 \log\left(\sum_{i=1}^{J} Y_i e^{-\alpha_2 d_{ij}}\right) + \varepsilon_j$$

This so-called market potential equation simply states that wages in a given region are determined by the proximity to consumer markets. The equation can be regarded as reduced form of several NEG models. Therefore, the market potential approach is a fairly fundamental test of NEG (see Roos 2001). In this context, we have to consider that assuming $T_i = T$ implies that the effect of competition from other manufacturing production sites is missing in the regression model given by equation (10).

However, the market potential function as well as the model given by equation (9) are fairly restricted explanations of regional wage differences. There are probably a number of additional factors which determine the spatial distribution of economic activities and the regional wage level such as local amenities, sectoral composition of the regional economy or qualification of work force.[3] In order to deal with these issues and to check whether the mechanisms emphasised by NEG prove to be robust, control variables are included in the regression models:

$$(11)\qquad \log(w_j) = \alpha_0 + \alpha_1 \log\left(\sum_{i=1}^{J} Y_i e^{-(\alpha_2 d_{ij})}\right) + \sum_{n=1}^{N} \gamma_n X_{jn} + \varepsilon_j,$$

where X_{jn} is a control variable and γ_n is the corresponding coefficient. So as to control for effects of local agglomeration, we include the population density as an additional explanatory variable. Furthermore, control variables comprise indicators for sectoral composition of regional economies, the presence of local amenities, human capital and dummies for countries and outlying regions if necessary.

3 Head and Mayer (2003) discuss some alternative explanations of agglomeration such as natural advantages, human capital externalities and knowledge spillovers.

The analysis investigates the existence of a spatial wage structure for a cross section of more than 150 European regions. Hence we have to consider trade of manufacturing goods across national borders, i.e. foreign markets. The results of several studies point to significant border impediments even between highly integrated EU member states (see Bröcker 1998, Nitsch 2000). The regression models should therefore take into account border effects that reduce the accessibility of foreign markets:

$$(12) \quad \log(w_j) = \alpha_0 + \alpha_1 \log\left(\sum_{i=1}^{J} Y_i e^{-(\alpha_2 d_{ij} + \alpha_4 B_{ij})}\right) + \sum_{n=1}^{N} \gamma_n X_{jn} + \varepsilon_j$$

with B_{ij} as border variable. In the regression analysis, border impediments are approximated by estimated border effects. $B_{ij} = 0$, if i and j are located in the same country. B_{ij} will be approximated by estimated impediments, if the regions i and j are located in two different EU member states.[4] A significant and positive coefficient would indicate that the impact of purchasing power on the wage in a given region is significantly reduced by crossing a national border.

Another estimation issue concerns unobserved and time-invariant characteristics of regions. This problem can be solved by controlling for fixed location effects via estimating a specification in time differences (see Hanson 2000). Equation (10) becomes then:

$$(13) \quad \Delta \log(w_{jt}) = \alpha_1 \left[\log\left(\sum_{i=1}^{J} Y_{it} e^{-\alpha_2 d_{ij}}\right) - \log\left(\sum_{i=1}^{J} Y_{it-k} e^{-\alpha_2 d_{ij}}\right)\right] + \Delta\varepsilon_{jt},$$

with Δ as difference operator and t as time index. The errors $\Delta\varepsilon_{jt}$ are assumed to be *i.i.d.*, uncorrelated with the regressors as well as across regions. Changes in the market potential of regions should reflect long term structural changes of the spatial distribution of economic activity. Thus they should be uncorrelated with time- and region-specific shocks.

Finally, we have to consider the endogeneity problem, i.e. right hand side variables, such as regional income are not exogenous, possibly causing inconsistent estimates. Therefore, apart from nonlinear least squares (NLS) method, we apply nonlinear instrumental variables estimation to address this issue. Historical data on regional Gross Value Added (GVA) and population, lagged by 10 years, are used as instruments for

4 For a detailed description of data on border impediments and data sources see section 5.

contemporary income. We also test for spatial error autocorrelation and estimate spatial econometric models if necessary, i.e. in case of a misspecification of the nominal wage equation.

In the analysis, all Western European countries are included as sources of demand for products of a specific location. In other words, main trading partners of the countries and regions investigated in the analysis are taken into account. However, we cannot consider demand from Central and Eastern European countries because of data restrictions with respect to the period under consideration.

4 Previous Empirical Evidence on Market Access

In principle, there are two different groups of empirical studies dealing with the significance of market access for regional development. The first group comprises some early studies that apply market potential arguments in order to analyse spatial integration effects. Clark et al. (1969) and Keeble et al. (1982) investigate effects of European integration by analysing changes in regional accessibility and market potential induced by a reduction of tariff barriers. They use the market potential concept as proposed by Harris (1954). The analyses assume that accessibility is important for investment decisions and, therefore, regional growth. A high market potential is rated as a locational advantage. Thus, the most densely populated areas and central locations in Europe should realise the highest integration benefits.

According to the results of Keeble et al. (1982), the most inaccessible regions, marked by extremely low market potentials, are located in the geographical periphery. In contrast, high accessibilities and market potentials are estimated for regions in the north-east of Europe, covering large parts of the Netherlands, Belgium and West Germany. Moreover, the results point to a widening of regional disparities in accessibility and market potential: Enlargement as well as faster growth of more accessible regions tended to favour the central areas in Europe in the 1960s and 1970s. As Keeble et al. (1982) point out, the basic pattern of the market potential reflects historic processes, e.g. industrialisation and urbanisation. The effects of integration induce only slight changes in the market potential of European regions. However, the positive effect ascribed to the change of the market potential is not based on a well defined theoretical approach in the above mentioned studies. The significance of the market potential for regional development remains an unclear matter – from a theoretical as well as from an empirical point of view.

With the development of NEG, theoretical deficiencies concerning market access have been remedied. The second, more recent group of studies that investigate the empirical significance of the market potential focuses on tests of corresponding theoretical models. Hanson (1998, 2000) conducted the leadoff analysis of this kind for US counties. The findings of the regression analyses point to strong, but highly localised demand linkages between regions. A higher effective consumer demand, determined by income and transport costs, tends to raise wages in a given location. Regional wages decline with increasing distance to consumer markets. According to the results, purchasing power in regions more than 1,000 kilometres away from a location does not affect demand for goods of that location and has, therefore, no impact on local wages. Brakman et al. (2000) estimate the market potential function for German districts with data for the year 1995. They find strong confirmation of the significance of a spatial wage structure in Germany. Regional wages are affected by economic activity and demand in neighbouring regions. Again, the effects of demand are highly localised, i.e. distance matters a lot for interregional demand linkages. Adding market access to Germany's main trading partners has no significant effect on the regression results. The authors, therefore, conclude that economic activity beyond national borders seems not to affect the spatial wage structure in Germany.

The analysis of Roos (2001) affirms the empirical evidence provided by Hanson (2000) and Brakman et al. (2000). A positive relationship between regional wage and purchasing power in neighbouring locations marks the analysed cross section of West German NUTS 3 regions. However, the findings of Roos imply that empirical relevance of the applied Helpman model for Germany is restricted compared with market potential approach. Moreover, market access only seems to affect the wages of skilled workers, whereas there is no significant effect on the wages of unskilled workers. Finally, Mion (2003) estimates the Helpman model for Italian NUTS 3 regions applying dynamic panel techniques in order to deal with the problem characteristic for NEG models that almost all variables are endogenous. The regression results are consistent with the NEG hypothesis that demand linkages affect the spatial structure of economic activities.

To sum up, the findings of current empirical investigations are in line with specific implications of NEG models. However, recent studies on the relevance of market access focus on spatial wage structures in single countries. In contrast, the following regression analysis considers a cross section of European regions.

5 Data and Regional System

5.1 Data

The dependent variable in the regression analysis of European regions is the log compensation per employee and the log change in compensation per employee respectively. In the level specifications of the market potential function, data on regional compensation per employee in 1985, 1990, 1995 and 2000 are included. For the time-difference specifications, data for the period 1995-2000 are used. The dependent variable is given for 158 European regions. Regional income, i.e. purchasing power, is approximated by GVA in 205 European regions. Indicators for the sectoral composition of regional economies base on GVA data by NACE-CLIO R6 classification (agricultural, forestry and fishery products, manufactured products, building and construction, market services, non-market services). The corresponding GVA shares, i.e. the percentages of regional GVA in agriculture, manufacturing et cetera, are used as control variables. The data were taken from Cambridge Econometrics' European regional databank. Information on local amenities (e.g. length of the seashore, mean annual sunshine radiation, concentration of cultural sites) were taken from the databank generated in the course of the Study Programme on European Spatial Planning (SPESP).[5] In order to control for effects resulting from regional differences with respect to the qualification of the work force, we also considered the share of human resources in science and technology in total population. Human resources in science and technology (S&T) are defined as people who successfully completed education at tertiary-level in an S&T field of study or people without the formal qualification who are employed in S&T occupations where the qualifications are normally required. The corresponding data were taken from the Eurostat Regio databank. In the Krugman model the wage level on the right hand side of the equation was approximated by GVA per capita, since regional data on compensation per employee is not available for Sweden, Norway and Switzerland.

5 See BBR (Eds.): Study Programme on European Spatial Planning. Final Report. Forschungen 103.2.- Bonn 2001.

5.2 Distances and Border Impediments

Distance is measured by travel time in minutes between the centres of regions.[6] With respect to internal distances of the European regions, we use an approach frequently applied in the corresponding literature, i.e. internal distance is modelled as proportional to the square root of the region's area.[7] The area of each region is approximated with a disk in which all production activity is concentrated in the centre and consumers are distributed evenly across the area. Under these assumptions, the average distance between consumers and producers in the region can be estimated as a function of the square root of the area. Following Bröcker (1999), we determine the internal distance of region *i* in minutes of travel time as:

$$(14) \quad d_{ii} = 0.75 \cdot \sqrt{A_i} \, ,$$

where A_i denotes the area of region *i*. Crozet (2000) notes that this kind of determination might not be as precise as other specifications applied for interregional distances. In particular, two biases are relevant in this context. Firstly, both consumers and firms tend to be located in or around cities. Thus actual distances between consumers and producers should be smaller than those implied by the disk approximation. Secondly, the approximation is simultaneously affected by downward bias, since internal distance is measured ”as the crow flies“. The biases work in opposite directions, and the effects might just level out each other (see Crozet 2000).

In order to generate data on bilateral border impediments for the EU, different sources are used. Bröcker (1998) estimates bilateral trade impediments for several European countries for the year 1994. The factors, by which international trade is reduced compared with intranational trade, range between 7 and 117. On average, trade is reduced by a factor of 20 due to crossing a border. These results are in line with the empirical evidence provided by Mc Callum (1995). However, they are rather high compared to the estimates of Nitsch (2000) or Wei (1996) for the EU. By applying a gravity model to EU-trade, Nitsch (2000) estimates border effects between 6 and 16. The findings of Wei (1996) suggest border impediments around a factor of 10. Moreover, Wei's results

6 I would like to thank Johannes Bröcker for the provision of interregional distances.
7 See Head and Mayer (2000), Nitsch (2000) and Crozet (2000).

imply that border effects in the EU declined by 50% between 1982 and 1994. Estimates of Head and Mayer (2000) range between 12 and 20.

The analysis by Nitsch (2000) provides more information on the development of border impediments in the EU between 1979 and 1990. According to the results, there has been a pronounced decline of border effects between 1979 and 1982 and since then the impediments gradually declined from factor 12 to factor 10. Head and Mayer (2000) analyse the change of border impediments in the EU as well. Their empirical evidence suggests that the border effects have decreased from a factor of 20 in the late 1970s to roughly 13 in 1993/1995 after the completion of the Single European Act. Head and Mayer report a rapid decline until 1985 and only small change thereafter. The estimates on bilateral trade impediments in 1994 and on the development of average border impediments in the EU since the end of the 1970s are combined in order to generate data on bilateral border effects for the period under consideration.

5.3 Regional System

Two cross sections have to be distinguished in the present analysis. One cross section concerns the dependent variable and comprises 158 EU regions. The second cross section consists of all regions the income of which is included in the market potential, in total 205 European regions. The regional system largely corresponds with the NUTS 2 level. Exceptions concern in particular Denmark (3 former NUTS regions) Belgium, Germany (NUTS 1 level) and Sweden (NUTS 3 level). The following regions are not considered because of data restrictions: Berlin and all NUTS 2 regions in East Germany, Départements d'outre-Mer (France), Açores, Madeira (Portugal), Ceuta y Melilla, Canarias (Spain). Norway (19 Fylke) and Switzerland (7 Grossregionen) are included in the larger cross section for estimation of the market potential. With respect to the left hand side of the regression model, Sweden, Norway and Switzerland could not be considered because of data restrictions. A more detailed description of the cross section is given in the appendix.

6 Regression Results

A number of different specifications, as described in section 3, are estimated. The regression analysis comprises the equations (9) to (13) applying NLS, nonlinear instrumental variables estimator and spatial regression methods. The time difference specification is estimated in order to deal with the problem of time-invariant, unobserved determinants of regional wages, as in Hanson (2000). In the following, generally only significant control variables are considered in the displayed regression results or congruent groups of control variables in order to allow comparisons of different models.

Table (1) shows regression results for the nominal wage equation derived from the Krugman model (equation 9) for 1985 and 2000. In column (1) estimates for the nominal wage equation in 1985 without control variables are given. The coefficients are significant and the signs correspond with theoretical expectations. However, as the results in column (2) indicate, these findings are not robust. The inclusion of control variables drives the coefficients α_1 and α_3 to insignificance. Only the coefficient of distance remains significant at the 0.01 level. Moreover, the implied value for the share of expenditure on manufactures μ is not in line with theory. Regression results for other years are even worse, as the estimates in the columns (3) and (4) for the spatial wage structure in 2000 exemplify. The relevant variables exert no important influence on regional wages. Furthermore, implied values for μ and σ are implausibly large. Altogether, the estimates provide no persuasive evidence for the relevance of this specific model.

Table 2 summarises estimates of the market potential models given by equations (10) and (11) for the log compensation per employee in 2000. GVA, applied as income measure, population density and sectoral composition refer to 2000, human capital to 1999. In column (1) the estimates for the basic market potential equation are presented. The coefficients are highly significant with signs in accordance with theoretical fundamentals. 50% of the variation in regional wages is already explained by market access. The coefficient of the market potential α_1 declines if control variables are included, but the effect remains significant (column 2). However, the models given in columns (1) and (2) are marked by outlying observations and spatial autocorrelation in the regression residuals.[8] The outlying regions do not correspond with the spatial wage structure determined by the majority of observations. Outliers will seriously affect the coefficient estimates, if they are influential leverage points, i.e. outlying observations with regard to the market potential. In order to control for effects of outlying observations, dummy variables for the outliers are introduced. The most significant outlier is the region Brussels. Moreover, most Portuguese and Austrian regions are outliers as well. To control for the latter observations, country dummies were included.

8 We identified outliers as those regions the standardised residuals of which exceed the critical value $|2.5|$.

Table 1: Regression results for the nominal wage equation - Krugman model

	Dependent variable: Compensation per employee NLS			
	(1) $\log(w_{j,\,1985})$	(2) $\log(w_{j,\,1985})$	(3) $\log(w_{j,\,2000})$	(4) $\log(w_{j,\,2000})$ a)
α_0	-0.15 (0.20)	-1.80* (2.40)	4.21** (13.40)	5.44** (8.21)
α_1	0.12** (5.14)	0.04 (1.83)	0.01 (0.64)	0.004 (0.29)
α_2	0.0095** (5.44)	0.0119** (2.77)	0.1013 (0.64)	0.0681 (0.29)
α_3	7.24** (3.93)	10.25 (1.72)	64.96 (0.65)	63.29 (0.28)
σ	8.29	22.55	110.35	235.74
τ	0.0013	0.0006	0.0009	0.0003
μ	1.01	2.10	1.68	3.71
Agriculture$_t$		-0.06** (2.79)		-0.06** (2.69)
Manufacturing$_t$		0.57** (8.09)		0.16* (2.45)
Market Services$_t$		1.03** (6.91)		0.46** (3.57)
Non-Market Services$_t$		0.51** (6.79)		0.08 (1.25)
Seashore		0.002** (3.53)		0.0004 (0.97)
Sunshine		-0.15** (5.04)		-0.15** (4.50)
Emission		-0.09* (2.26)		-0.05 (1.92)
Density Cultural Sites		-0.001* (2.47)		-0.0004 (0.86)
Dummy Brussels		-1.12** (12.81)		-1.05** (10.78)
Dummy Portugal		-0.61** (6.15)		
Dummy Austria		0.47** (9.14)		0.28** (6.80)
Adj. R^2	0.77	0.93	0.68	0.81

Notes: *t*-statistics are based upon White's heteroscedasticity-adjusted standard errors.
** significant at the 0.01 level, *significant at the 0.05 level.
a) The dummy variable for Portugal had be excluded to achieve convergence.

Table 2: Regression results for market potential function 2000

	Dependent variable: Compensation per employee 2000			
	NLS			Spatial Error Model (ML)
	(1)	(2)	(3)	(4)
α_0	7.67** (23.50)	8.10** (20.20)	8.43** (25.71)	3.69** (27.26)
α_1	0.18** (8.53)	0.11** (6.00)	0.08** (5.16)	0.09** (7.31)
α_2	0.0035** (3.20)	0.0037* (2.75)	0.0036* (2.17)	-
σ	5.65	9.41	12.01	10.75
τ	0.00075	0.00044	0.00033	–
λ				0.49** (6.27)
Population density$_{2000}$		-0.04 (1.73)	-0.03 (1.39)	-0.02 (1.38)
Agriculture$_{2000}$		-0.07* (2.05)	-0.10** (4.31)	-0.11** (5.63)
Market Services$_{2000}$		0.20* (2.57)	0.18** (3.01)	0.09 (1.47)
Human Capital$_{1999}$		0.54 (1.94)	0.54* (2.10)	0.06* (2.02)
Hazard		-0.09** (5.84)	-0.09** (6.99)	-0.02** (4.33)
Protected Areas		0.03** (3.09)	0.03** (3.17)	0.01** (3.21)
Cultural Sites		0.004* (2.38)	0.003* (2.51)	0.001** (2.75)
Dummy Brussels			-1.01** (15.25)	-0.53** (9.06)
Dummy Portugal			-0.45** (5.04)	-0.23** (6.01)
Dummy Austria			0.31** (10.52)	0.12** (6.27)
Moran	8.52**	4.43**	3.60**	-
LM_{ERR}	69.00**	18.21**	11.90**	-
LM_{LAG}	6.95**	0.13	0.08	1.28
Adj. R^2	0.50	0.70	0.85	0.80

Notes: *t*-statistics are based upon White's heteroscedasticity-adjusted standard errors.
** significant at the 0.01 level, *significant at the 0.05 level.

Including dummies for outlying observations does not alter the estimates considerably (see column 3). In particular, the coefficients of the market potential and distance remain fairly unchanged, though α_1 is further reduced. However, as the tests for spatial autocorrelation indicate, the model is still misspecified.[9] The Moran test as well as the Lagrange Multiplier test for spatial error autocorrelation (LM_{ERR}) are highly significant and point to ignored spatial effects. In contrast, the test for spatial lag dependence (LM_{LAG}) does not hint at a misspecification. This suggests that the problem is not caused by ignored spatial interaction. In fact, the spatial autocorrelation seems to pertain to the error term and might be caused by measurement problems, such as a poor match between the spatial pattern of the analysed phenomenon and the units of observation. The spatial correlation of residuals implies that LS estimators remain unbiased, but are no longer efficient. In order to check the consequences with respect to the importance of market access, we estimate a linearised version of the wage equation, taking as given the distance decay determined in column (3). The results in column (4) base on a Maximum Likelihood estimation of the nominal wage equation with a spatially dependent error term. Taking into account the spatial autocorrelation does not change the implications regarding the relevance of the market potential.

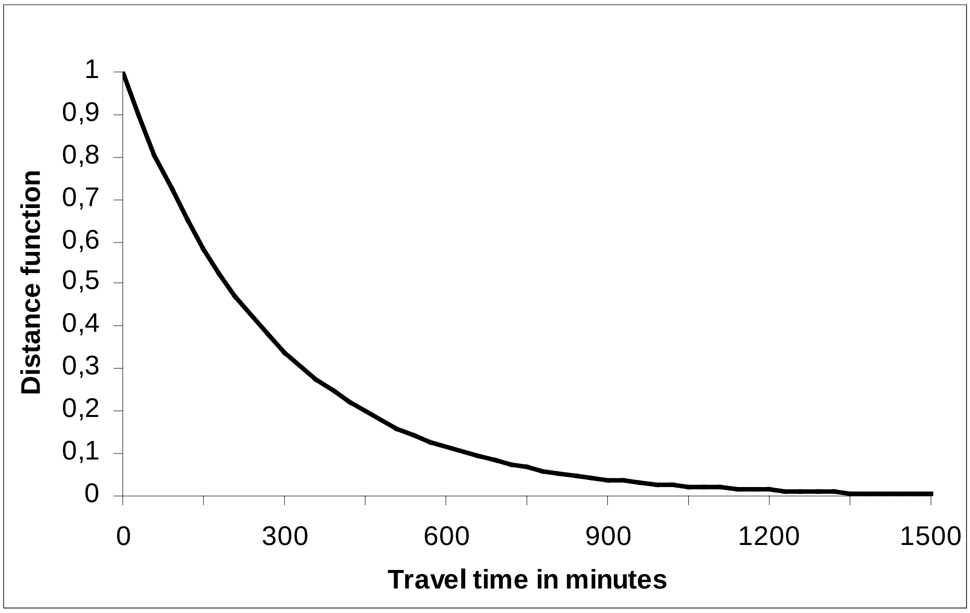
According to the results, market access has a positive effect on the wage level. Control variables and dummy variables improve the fit of the regression considerably. The model presented in column 3 explains more than 80% of the variation in regional wages. Apart from market access, the wage level is also influenced by population density and sectoral composition of the regional economy. Qualification of labour force and local amenities matter as well. The results suggest that, controlling for market access, sectoral composition, qualification and some basic characteristics of the regions, highly agglomerated regions are not characterised by an above average wage level. Thus, high wages in densely populated regions of the European Union can be traced back to their favourable market access, their specialisation, skills of labour force and local amenities. Settlement structure affects the wage level, but regional wages tend to decline with increasing agglomeration. A high percentage of agricultural production has a depressant impact on the wage level, as one would expect. In contrast, a comparatively high share of GVA in market services tends to exert a positive influence on regional wages. Moreover, human capital as well as the availability of protected areas and cultural sites tend

9 A binary contiguity matrix was applied as spatial weights matrix.

to increase the regional wage level, whereas a high risk of natural disasters has a negative effect on wages.[10]

According to the estimates for α_1, the elasticity of substitution σ is 12.0, i.e. consistent with theory larger than 1.[11] The coefficient α_1 points to increasing returns to scale since $\sigma/(\sigma\text{-}1) = 1.09$. These results basically confirm the evidence provided by Roos (2001), Brakman et al. (2000) and Hanson (2000). The coefficient α_2 can be interpreted as a spatial discount factor that determines changes in the weight of purchasing power with increasing distance. The estimated distance decay coefficient implies that the intensity of demand linkages declines by 50% over a range of roughly 190 minutes of travel time. Figure 1 displays the corresponding distance function. Assuming an average

Figure 1: Estimated distance decay function



Source: Estimates based on data from Cambridge Econometrics' European regional databank, own calculations.

10 We may question whether the signs of all control variables are consistent with the theoretical framework. For instance, in case amenities such as cultural sites increase utility of workers, then their influence on wages should be negative.

11 Estimates of σ range between 4.9 and 13.0 for the market potential equation which is roughly in accordance with recent estimates in the empirical literature. See also Head and Mayer (2003) as well as Hanson (2000) for a comparison.

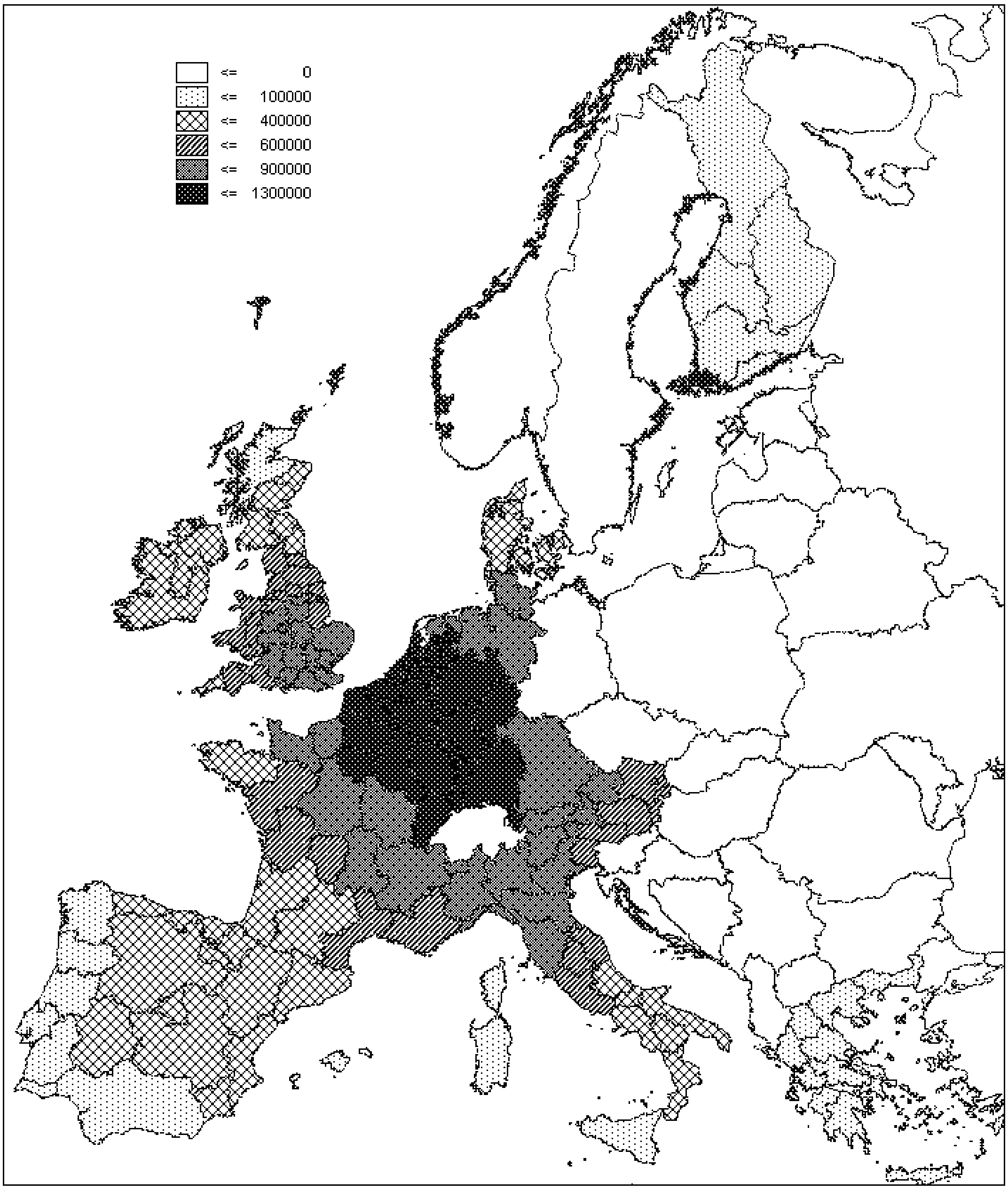
speed of 80 km/h gives a half-life distance of approximately 260 km. Compared with previous findings, this geographical scope of demand linkages is rather large. The regression results in Roos (2001) imply that the intensity of spatial effects halves after 5 to 30 minutes of travel time, depending on the specification. The half-life distance derived from the estimates in Brakman et al. (2002) ranges between 2 km and 8 km.[12] The results derived by Hanson (2000) imply that transport of goods over 2 kilometers increases prices by a factor of more than 50 (see Head and Mayer 2003). The estimated price increases based on the results in Roos (2001) and Brakman et al. (2000, 2002) are relatively moderate varying between a factor of almost 3 and 1.04. In the present analysis, the factor ranges from 1.007 to 1.01, however, referring to an additional travel time of 2 minutes.

Our findings with respect to the distance decay and the corresponding significance of trade costs are in line with findings generated in quite different empirical approaches. Bröcker (2003) determines a similar distance function based on estimates of a gravity model of international trade. Analyses of Fürst et al. (1999) suggest that a distance decay parameter of 0.007 for car traffic and 0.003 for trucks is appropriate for the calculation of European accessibility measures. Finally, literature on logistic costs points at a share of transport costs in sales volume that varies between 1.8% and 5.2% in Europe (see Weber 2002). These independent estimates are not consistent with large distance decay parameters identified in some recent studies on spatial wage structures.

Head and Mayer (2003) suggest that the "implausibly large estimate" of the distance coefficient in Hanson (2000) could be caused by the form of the distance function, i.e. applying a negative exponential function instead of a power function. However, our regression results, based on an exponential function as well, indicate that the functional form is probably not responsible for the high distance decay. A comparison of cross sections analysed in different studies provides an alternative explanation. The pronounced differences in the geographic extent of demand linkages could be a consequence of the regional system under consideration, in particular of the average size of the observational units. Whereas Roos (2001), Brakman et al. (2002) and Mion (2003) investigate fairly small NUTS 3 regions in a single European country, our analysis refers to larger NUTS 1 and NUTS 2 regions in several EU member states. The regional system analysed by Hanson (2000) also consists of small US counties. So the size of the esti-

12 In the present analysis, half-life distance ranges between 190 km and 270 km for the market potential function (150 to 200 minutes of travel time).

Figure 2: Estimated Market Potential 2000



Source: Estimates based on data from Cambridge Econometrics' European regional databank, own calculations.

Table 3: Sensitivity of results – estimates for the market potential function 1985, 1990, 1995 and 2000

	Dependent variable: Compensation per employee			
	(1) log($w_{j,\,1985}$)	(2) log($w_{j,\,1990}$)	(3) log($w_{j,\,1995}$)	(4) log($w_{j,\,2000}$)
α_0	9.93** (14.28)	7.32** (14.56)	7.95** (24.63)	8.52** (27.47)
α_1	0.20** (8.84)	0.17** (7.85)	0.10** (6.54)	0.08** (5.26)
α_2	0.0038** (3.51)	0.0040** (2.68)	0.0047* (2.38)	0.0042* (1.96)
σ	4.89	5.88	9.54	13.00
τ	0.00095	0.00082	0.00055	0.00035
Population density$_t$	-0.13** (3.78)	-0.16** (4.28)	-0.07** (2.70)	-0.03 (1.49)
Agriculture$_t$	-0.19** (5.75)	-0.21** (6.63)	-0.12** (5.10)	-0.10** (4.33)
Market Services$_t$	0.30** (4.13)	0.38** (5.36)	0.31** (5.49)	0.20** (3.40)
Hazard	-0.13** (6.61)	-0.10** (4.82)	-0.10** (8.08)	-0.09** (7.45)
Protected Areas	0.02 (1.35)	0.01 (0.89)	0.03** (2.80)	0.03** (3.63)
Cultural Sites	0.006** (3.34)	0.006** (3.41)	0.005** (4.11)	0.003* (2.56)
Dummy Brussels	-1.29** (11.64)	-1.36** (13.10)	-1.01** (11.65)	1.02** (14.55)
Dummy Portugal	-0.59** (5.67)	-0.53** (5.67)	-0.40** (4.82)	-0.43** (4.65)
Dummy Austria	0.48** (9.80)	0.41** (7.46)	0.44** (13.18)	0.32** (11.13)
Moran	8.97**	10.17**	6.23**	3.45**
LM_{ERR}	76.59**	98.78**	36.61**	10.93**
LM_{LAG}	0.39	0.04	0.00	0.03
Adj. R^2	0.87	0.84	0.85	0.84

Notes: *t*-statistics are based upon White's heteroscedasticity-adjusted standard errors.
** significant at the 0.01 level, *significant at the 0.05 level.

mated distance decay seems to increase with the declining size of regions. This might imply that the nature of spatial effects detected in previous studies and in the present analysis is not the same.

The estimate of the distance decay parameter combined with regional data on GVA allows to calculate the market potential. Figure 2 shows the results for the year 2000 and a distance decay α_2 = 0.0036. The spatial pattern of the market potential resembles accessibility measures and peripherality indices calculated by Keeble et al. (1982) or Schürmann and Talaat (2000).[13] Regions marked by low market potentials are located in the geographical periphery, comprising in particular Finland, Greece, Portugal, the south of Spain and Italy. In contrast, high accessibility and market potentials are estimated for regions in the north-east of Europe, covering large parts of the Netherlands, Belgium, Germany and the north of France.

Table 3 summarises results of a sensitivity analysis with respect to changes over time. Altogether, the specification including the market potential, control variables and dummies generates fairly robust estimates for different years. Estimates are given for the log compensation per employee in 1985, 1990, 1995 and 2000. The regression results indicate that the market potential has become less important for the regional wage level. Moreover, the implied values for the unit transport costs τ point to declining transport costs over time – a plausible result.[14]

In Table 4 the findings with respect to border impediments and specifications in time differences are summarised. The corresponding results are rather disappointing. The inclusion of border effects in the basic market potential function presented in column 1 does not yield significant border impediments that hamper demand linkages between domestic and foreign regions. Moreover, the coefficient α_4 is negative, which contradicts empirical evidence on trade reducing effects of national borders. The negative coefficient implies that crossing a national border increases the weight of demand, i.e. the importance of purchasing power in foreign regions is ceteris paribus higher than the importance of domestic demand. Including control variables brings about a correct sign

13 However, the study by Schürmann and Talaat (2000) also comprises East-European countries.

14 In contrast, the results of Hanson (2000) suggest, counterintuitively, that transport costs have risen slightly over time in the US. Test on spatial autocorrelation in Table 3 point to ignored spatial effects. However, unreported results for linear spatial error models indicate that controlling for autocorrelation does not affect the findings with respect to the significance of the market potential.

Table 4: Sensitivity of results – border effects and first differences

	Dependent variable: Compensation per employee			
	(1) $\log(w_{j,\,2000})$	(2) $\log(w_{j,\,2000})$	(3) $\Delta\log(w_{j,\,1995\text{-}2000})$	(4) $\Delta\log(w_{j,\,1995\text{-}2000})$
α_0	7.85^{**} (24.78)	8.40^{**} (24.19)	0.19^{**} (15.66)	0.23^{**} (10.39)
α_1	0.16^{**} (7.66)	0.08^{**} (4.58)	-0.07 (1.59)	-0.11^{*} (2.19)
α_2	0.0039^{**} (3.27)	0.0031 (1.94)	-0.22 (0.31)	-0.19 (0.38)
α_4	-0.0199 (1.05)	0.0183 (0.44)		
σ	6.06	11.82	-13.71	8.73
τ	0.00077	0.00029	-0.01526	-0.11456
Population density$_t$		-0.03 (1.39)		-0.07 (0.15)
Agriculture$_t$		-0.09^{**} (4.18)		0.05 (1.43)
Construction$_t$				0.30^{**} (3.80)
Manufacturing$_t$				0.24 (1.38)
Market Services$_t$		0.19^{**} (3.11)		0.21 (0.68)
Non-Market Services$_t$				0.37 (2.84)
Human Capital$_t$		0.55^{*} (2.11)		
Hazard		-0.09^{**} (6.98)		
Protected Areas		0.03^{**} (3.11)		
Cultural Sites		0.003^{*} (2.53)		
Dummy Brussels		-1.00^{**} (14.23)		
Dummy Portugal		-0.42^{**} (4.13)		
Dummy Austria		0.34^{**} (6.02)		
Adj. R^2	0.49	0.85	0.01	0.22

Notes: *t*-statistics are based upon White's heteroscedasticity-adjusted standard errors.
** significant at the 0.01 level, *significant at the 0.05 level.

for α_4. However, the effect remains insignificant. Furthermore, the simple difference specification without control variables, shown in column 3, achieves only a poor fit. The relevant coefficients α_1 and α_2 are neither consistent with theory nor significant. This applies to the extended model in column 4 as well.

Finally, some remarks on regression methods are due. Unreported estimates based on nonlinear instrumental variables regression will resemble the results generated by NLS, if we choose starting values close to the NLS coefficients. However, significant variations of the starting values might result in a collapse of the regression procedure or a failure to converge. Altogether, the nonlinear least squares method has proven to partly generate fairly unstable results. Some estimates are highly sensitive to changes of the specification and starting values for the coefficients.[15] This suggests that in future research the robustness of the findings with respect to different regression methods should be investigated when necessary data are available.

7 Conclusions

The results of the present analysis confirm in parts evidence on the relationship between regional wages and market potential provided by previous studies. Regional wages in Europe tend to rise with increasing market potential of the location. However, with respect to the Krugman model the evidence is fairly weak. The market potential function generates more persuasive results in this respect. Consistent with the conclusions of Roos (2001) the market potential approach yields more favourable results than an equation directly linked to a specific NEG model. The results suggest that the Krugman model is not suitable for an explanation of agglomeration in Europe. Yet the findings confirm the relevance of mechanisms emphasised by NEG. Admittedly, this confirmation is restricted to backward linkages.

Furthermore, our regression results indicate that the significance of market access for the wage level seems to decline over time. This is in contrast to the findings of Hanson (2000) who detects growing demand linkages over time. Moreover, compared with the evidence provided by other studies of the nominal wage equation, the estimated geographical scope of demand linkages is rather large in the present analysis. A comparison of the relevant analyses suggests that estimates of the distance decay might be in-

15 The additional regression results are available from the author upon request.

fluenced by the average size of the analysed regions. We think that, regarding the scope of demand linkages, our results are plausible for a highly integrated area such as the European Union. Finally, there is no indication for important border impediments affecting the spatial wage structure in Europe.

Significant differences in the detected geographic extent of demand linkages raise the issue what is meant by core or centre in NEG models: single cities or agglomerations at the European scale as the Ruhr area comprising several cities? Does the centre-periphery structure in NEG model refers to a relationship between urban and rural areas and is the relevant market a regional market? This interpretation of NEG is in line with the highly localised effects detected in some previous studies. Based on our estimates for the geographic extent of demand linkages, we prefer to think of the centre-periphery structure on a national or international scale with the relevant market being the EU. Relevant market areas are probably smaller for most services, but NEG deals with the spatial distribution of manufacturing that should be on average characterised by a much larger market size. Finally, keeping this in mind implies that future research on the importance of the market potential has to consider the enlarged EU.

References

Brakman, S., Garretsen, H.; Schramm, M. (2002):
New Economic Geography in Germany: Testing the Helpman-Hanson Model, HWWA Discussion Paper No. 172, Hamburg.

Brakman, S., Garretsen, H.; Schramm, M. (2000):
The empirical relevance of the new economic geography: Testing for a spatial wage structure in Germany, CESifo Working Paper No. 395, Munich.

Bröcker, J. (2003):
Regionale Wirkungen der Europäischen Währungsunion. Jahrbuch für Regionalwissenschaft 23, 3-22.

Bröcker, J. (1999):
Wie wirken sich neue Verkehrstechnologien und die Entwicklung transeuropäischer Netze auf die Europäische Standortverteilung aus? Forschungsbericht, TU Dresden.

Bröcker, J. (1998):
How would an EU-membership of the Visegrád-countries affect Europe's economic geography? Annals of Regional Science 32, 91-114.

Bundesamt für Bauwesen und Raumordnung (2001):
Study Programme on European Spatial Planning. Final Report. Forschungen 103.2, Bonn.

Clark, C., Wilson, F.; Bradley, J. (1969):
Industrial Location and Economic Potential in Western Europe, Regional Studies 3, 197-212.

Crozet, M. (2000):
Do migrants believe in market potential, TEAM – Paris 1 Panthéon-Sorbonne University, CNRS.

Fürst, F.; Hackl, R.; Kramar, A.; Schürmann, C.; Spiekermann, K.; Wegener, M. (1999):
The SASI Model: Model Implementation. Berichte aus dem Institut für Raumplanung No. 49, Universität Dortmund:
(http://irpud.raumplanung.uni-dortmund.de/irpud/pro/sasi/ber49.pdf).

Hanson, G.H. (1998):
Market Potential, Increasing Returns, and Geographic Concentration, NBER Working Paper No. 6429, Cambridge Massachusetts.

Hanson, G.H. (2000):
Market Potential, Increasing Returns, and Geographic Concentration, University of Michigan and NBER.

Harris, C. (1954):
The Market as a Factor in the Localization of Industry in the United States, Annals of the Association of American Geographers 44, 315-348.

Head, K.; Mayer, T. (2000):
Non-Europe: The Magnitude and Causes of Market Fragmentation in the EU, Weltwirtschaftliches Archiv 136, 284-314.

Head K.; Mayer T. (2003):
The Empirics of Agglomeration and Trade, in: Henderson, V.; Thisse, J.F. (eds.), Handbook of Regional and Urban Economics Vol. 4, forthcoming.

Helpman, E. (1998):
The Size of Regions, in: Pines, D.; Sadka, E.; Zilcha I. (eds.), Topics in Public Economics, Cambridge University Press, pp. 33-54.

Keeble, D.; Owens, P.L.; Thompson, C. (1982):
Regional Accessibility and Economic Potential in the European Community, Regional Studies 16, 419-432.

Krugman, P. (1991):
Increasing Returns and Economic Geography, Journal of Political Economy 99, 483-499.

Krugman, P. (1992):
A Dynamic Spatial Model, NBER Working Paper No. 4219, Cambridge, Massachusetts.

Leamer, E.E.; Levinsohn, J.A. (1995):
International trade theory: the evidence, in: Jones, R.W.; Grossman, G.M. (eds.), Handbook of International Economics, Vol. 3, Amsterdam, North-Holland, pp. 1339-1394.

Mc Callum, J. (1995):
National Borders Matter: Canada – U.S. Regional Trade Patterns, American Economic Review, 85, 615-623.

Mion, G. (2003):
Spatial Externalities and Empirical Analysis: The Case of Italy, ZEW Discussion Paper No. 03-38, Mannheim.

Neary, P. (2001):
Of Hype and Hyperbolas: Introducing the New Economic Geography, Journal of Economic Literature 39, 536-561.

Nitsch, V. (2000):
National borders and international trade: evidence from the European Union, Canadian Journal of Economics 22, 1091-1105.

Ottaviano, G. (2002):
Regional policy in the global economy: Insights from New Economic Geography, HWWA Discussion Paper No. 211, Hamburg.

Roos, M. (2001):
Wages and Market Potential in Germany, Jahrbuch für Regionalwissenschaft 21, 171-195.

Schürmann, C.; Talaat, A. (2000):
Towards a European Peripherality Index: Final Report, Berichte aus dem Institut für Raumplanung, No. 53, Dortmund.

Weber, J. (2002):
Logistikkostenrechnung, Springer, Berlin, Heidelberg, New York.

Wei, S.-J. (1996):
Intra-National versus International Trade: How Stubborn are Nations in Global Integration? NBER Working Paper No. 5531, Cambridge Massachusetts.

Appendix

Cross section

Two cross sections have to be distinguished: a smaller cross section concerning the dependent variable and a larger one that consists of all regions the income of which is included in the market potential (in total 205 European regions). The regional system largely corresponds with the NUTS 2 level. Exceptions concern in particular Denmark (3 former NUTS regions) Belgium, Germany (NUTS 1 level) and Sweden (NUTS 3 level). The following regions are not considered because of data restrictions: Berlin and all NUTS 2 in East Germany, Départements d'outre-Mer (France), Açores, Madeira (Portugal), Ceuta y Melilla, Canarias (Spain). Moreover, Norway (19 Fylke) and the Switzerland (7 Grossregionen) are included in the larger cross section.

In the cross section for the dependent variable 158 EU regions are included. Sweden, Norway and the Switzerland are not considered in this cross section because of data restrictions. The 158 regions used in the sample are:

Belgium (3):	Bruxelles, Vlaams Gewest, Région Wallonne
Denmark (3):	Hovedstadsregionen, Ost for Storebaelt, ex.Hovedst, Vest for Storebaelt
Finland (6):	Uusimaa, Etelä-Suomi, Åland, Itä-Suomi, Väli-Suomi, Pohjois-Suomi
Germany (10):	Schleswig-Holstein, Hamburg, Bremen, Niedersachsen, Nordrhein-Westfalen, Hessen, Rheinland-Pfalz, Saarland, Baden-Württemberg, Bayern
Greece (13):	Sterea Ellada, Peloponnisos, Ionia Nisia, Thessalia, Dytiki Makedonia, Kentriki Makedonia, Anatoliki Makedonia, Thraki, Ipeiros, Kriti, Voreio Aigaio, Notio Aigaio, Attiki, Dytiki Ellada
Spain (16):	Galicia, Principado de Asturias, Cantabria, Pais Vasco, La Rioja, Comunidad Foral de Navarra, Castilla y León, Comunidad de Madrid, Castilla-la Mancha, Extremadura, Aragón, Cataluña, Islas Baleares, Comunidad Valenciana, Región de Murcia, Andalucia
France (22):	Rhône-Alpes, Picardie, Auvergne, Provence-Alpes-Côte d'Azur, Champagne-Ardenne, Midi-Pyrénées, Languedoc-Roussillon, Basse-Normandie, Poitou-Charentes, Centre, Limousin, Bourgogne, Bretagne, Aquitaine, Franche-Comté, Haute-Normandie, Pays de la Loire, Lorraine, Nord - Pas-de-Calais, Alsace, Île de France, Corse
Ireland (2):	Border, Midland and Western, Southern and Eastern
Italy (20):	Valle d'Aosta, Piemonte, Liguria, Lombardia, Emilia-Romagna, Trentino-Alto Adige, Veneto, Friuli-Venezia Giulia, Toscana, Marche, Umbria, Lazio, Abruzzo, Molise, Puglia, Campania, Basilicata, Calabria, Sicilia, Sardegna
Luxembourg (1)	
Netherlands (12):	Groningen, Friesland, Drenthe, Overijssel, Gelderland, Flevoland, Utrecht, Noord-Holland, Zuid-Holland, Zeeland, Noord-Brabant, Limburg (NL)

Portugal (5): Norte, Centro, Lisboa e Vale do Tejo, Alentejo, Algarve

Austria (9): Burgenland, Niederösterreich, Wien, Kärnten, Steiermark, Oberösterreich, Salzburg, Tirol, Vorarlberg

United Kingdom (36): Tees Valley and Durham, Cumbria, Northumberland and Tyne and Wear, East Riding and North Lincolnshire, North Yorkshire, South Yorkshire, West Yorkshire, Derbyshire and Nottinghamshire, Leicestershire, Rutland and Northamptonshire, Lincolnshire, East Anglia, Bedfordshire and Herefordshire, Berkshire, Buckinghamshire and Oxfordshire, Surrey, East and West Sussex, Essex, London, Hampshire and Isle of Wight, Kent, Gloucestershire, Wiltshire and North Somerset, Cornwall and Isles of Scilly, Devon, Dorset and Somerset, Herefordshire, Worcestershire and Warwickshire, Shropshire and Staffordshire, West Midlands, Cheshire, Greater Manchester, Lancashire, Merseyside, East Wales, West Wales and The Valleys, Eastern Scotland, South Western Scotland, North Eastern Scotland, Highlands and Islands, Northern Ireland

Data

Cambridge Econometrics regional data bank

Regional wage level: approximated by compensation per employee

Regional income: gross value added (GVA),

Sectoral composition: shares of sectors in total GVA of region (NACE-CLIO R6 classification: agriculture, manufacturing, building and construction, market services, non-market services)

Eurostat Regio Data

Human capital: Share of human resources in science and technology (S&T) in total population; human resources are people who successfully completed a third level education in an S&T field of study (according to the International Standard Classification of Education – ISCED 1997)

Data from the Study programme for European Spatial Planning (SPESP)

Seashore: Length of seashore in percentage of region's perimeter,

Sunshine: Mean annual sunshine radiation in kWh/m^2,

Emission: Emissions of acidifying gases – 3 classes,

Hazard: Natural hazards – 7 risk classes (earthquakes, volcanic activity, tidal waves, snow avalanches, slope instability),

Protected areas: Designated or protected areas – 5 classes,

Cultural sites: Number of registered monuments/cultural sites,

Density of cultural sites: Number of cultural sites by total area.

Missing regional data for Denmark and Norway was completed by data from the corresponding national statistical offices.