Relative investment specialisation inside EU:

An econometric analysis for EU-Regions

Claudia Stirboeck

Centre for European Economic Research (ZEW), Mannheim, Germany

March 2002

Paper presented at the 42nd European Regional Science Association Congress,
Dortmund 2002

Abstract

This paper analyses the relative distribution of gross fixed capital formation across industrial sectors in EU-regions, i.e. the level of relative regional investment specialisation, for the period between 1985 and 1994. Controlling for heteroscedasticity, potential endogeneity as well as spatial dependence, we get consistent econometric results. Larger market and regional sizes diminish relative investment specialisation while a higher unemployment rate, population density, the fact of being a central region, the distance to the economic centre, and economic liberalisation increase its level. The variation of the specialisation level of one region over time, however, cannot be explained econometrically, it thus might underlie random disturbances.

JEL classification: C30, F15, F2, R12

Keywords: Economic Geography, Regional Economics, Capital Allocation, Spatial Econometrics

Claudia Stirboeck
ZEW
PO Box 10 34 43
D-68034 Mannheim
Tel.: +49-621-1235-218
Fax: +49-621-1235-223
E-Mail: stirboeck@zew.de

Non-technical summary

Since Krugman has assumed increasing core-periphery tendencies due to stronger industrial agglomeration in an economically integrated Europe, controversial discussions about this prediction have risen. However, the seminal study of Krugman comparing EU and US regional specialisation levels bases on mobile labour. But this high mobility of employees across US states contrasts sharply with the low mobility of labour across the EU. Results on specialisation tendencies for US states which base on employment data can therefore not simply be extended to EU countries or regions. Neglecting capital, though, might lead to disturbed results – especially inside the EU where we face an increasing mobility of capital which is enforced by the EMU and financial market integration. This study therefore aims at identifying the economic determinants of the (un)even relative regional allocation of capital across industries. Its focus is on regional patterns as a profound analysis of the homogeneity of regional, not national, economic structures is still missing in recent research.

We consider regional data of the Eurostat REGIO database for the period between 1985 and 1994. Our focus is on NUTS 1- as well as on NUTS 2-regions – the latter being further disaggregated than the former. 56 NUTS 2- and 33-NUTS 1-regions from seven EU countries are included: France, Italy, Belgium, UK, Ireland, Denmark, and Luxembourg. Ireland, Denmark, and Luxembourg, however, are mono-regional countries at NUTS 1- and NUTS 2-level, i.e. they are not disaggregated any further. In order to capture the degree to which EU-regions differ in their investment structures, indices of regional investment specialisation are calculated for each region and each year. In order to capture relative regional specialisation levels, investment shares are analysed in relation to the average EU sectoral shares.

One important, but unsurprising feature that can be detected is the higher level of relative investment specialisation when regarding the more disaggregated NUTS 2-regions in comparison to the NUTS 1-regions. In a descriptive analysis of the most and the least specialised regions within each country, we find higher specialised regions to perform worse in economic terms than lower specialised regions with respect to unemployment rate, number of patents, total regional GDP and total regional GFCF. The fact that the distribution of relative investment shares of peripheral regions is more uneven than in the core regions while these regions are of poorer economic performance, already stresses the importance for the EU and its member countries not to neglect their focus on economic development of peripheral regions. Economic centres, especially the region of Bruxelles and the Île de France, are highly specialised as well. However, they demonstrate a large potential of high economic performance. As no causal relationship can be derived from this purely descriptive analysis, econometric analyses are conducted to test for the significance of potential determinants of the even or uneven relative allocation of investments across sectors within a region.

The theoretical basis for the empirical investigation of the determinants of regional specialisation and the concentration of economic activity is vast. According to traditional trade theories, regional specialisation takes place in accordance with comparative advantages. The regional economics' polarisation theory stresses the potential cumulative causation of factor agglomeration in the centre and backwash effects for peripheral regions. Gravity models focus on centripetal forces such as market size and centrifugal forces such as transport costs and imperfect economic integration or liberalisation. Finally, the New Economic Geography finds transaction costs and economies of scale to explain among others the concentration of sectors in space. The core thus specialises in sectors with high economies of scale, the periphery in sectors with low or constant economies of scale.

We conduct GLS-estimates accounting for potential heteroscedasticity, instrumental-variable estimates accounting for possible endogeneity between the dissimilarity indicator and some of its determinants as well as a dynamic specification capturing possible first-order serial correlation effects. In addition, a logistic transformation was used to take account of the endogenous variable's restriction to the range between zero and one. Since we deal with regional developments, we cannot exclude spatial interdependencies as well as measurement errors leading to spatial autocorrelation effects. We therefore additionally run spatial econometric estimates to control for the potential spatial dependence and to check for the robustness of our results obtained by classical econometric methods. All estimations lead to very similar results and provide evidence of a high importance of regional size,

market size, the unemployment rate as well as the location in the centre or the periphery, population density, economic openness as well as capital market integration.

The bigger the size of a region is, the higher is the similarity of relative investments. Market size reflects the economic and demand potential of a region: The higher it is, the lower the level of relative specialisation in terms of investments tends to be. This is in contrast to the results of recent empirical studies on sectoral agglomeration which found market size to have an increasing influence on the concentration across space. While firms tend to locate close to large markets and high demand (thus increasing sectoral concentration), regions with a large market seem to attract capital of all types of sectors with a rather even relative allocation (thus decreasing relative regional dissimilarities). This effect is counteracted by an apparently strong tendency towards high specialisation of central, economically most important regions who demonstrate to have a significantly higher level of relative investment specialisation. Equally, population density increases the specialisation level. The unemployment rate, finally, reflects negative economic performance of a region (not accounting for migration effects etc.). The higher it is, the stronger the level of relative regional specialisation turns out to be. The higher the distance of a region to the economic centre is, the less similar are its investment shares to EU average. Peripheral regions are thus stronger specialised in terms of relative investments than regions closer to the centre.

In addition, the impacts of economic openness and the influence of capital market integration are tested in separate estimates. Both indicators, however, are not available at the regional level, but only at the country level. They therefore might pick up country-specific effects in cross-sectional analyses. However, reliable results on the impact of liberalisation tendencies are gained in the pooled regressions when we are able to exploit the indicators' variation over time while efficiently controlling for country-specific effects. Both, the extent of capital market integration as well as economic openness consistently seem to have a significant increasing impact on relative specialisation levels of gross fixed capital formation.

The spatial econometric analyses, controlling for spatial interdependencies, provide strong evidence for the robustness of the above described economic phenomena. In addition, the regressions display negative spatial dependence either due to negative spillovers or simply due to data measurement errors.

Comparing results for NUTS 2- and the more aggregated NUTS 1-level, we find them to be similar. Only the regional size does not show a significant influence on the similarity of relative investment shares in NUTS 1-regions. As NUTS 1-regions are higher aggregated and usually consist of a number of NUTS 2-regions, their size differs much less than at NUTS 2-level. In addition, bigger regions are logically more diversified in relative investments than smaller regions. Thus, this result is not surprising.

By means of fixed effects estimates, there are only few significant impacts to be detected. The region-specific constants all turn out to be highly significant in the fixed effects estimates. Consistently, we find only low explanatory power of further exogenous variables when controlling for region-specific effects. Regional characteristics like regional gross domestic product, the regional unemployment rate, regional size, the distance to the economic centre as well as the fact of being a central region thus seem to determine the respective level of specialisation to a large extent. The variation over time, though, is not explained by our regressions. In addition, the extent of market integration and thus of European integration seems to strengthen relative specialisation tendencies.

I Introduction

Since Krugman (1991) has assumed increasing core-periphery tendencies due to stronger industrial agglomeration in an economically integrated Europe, controversial discussions about this prediction have risen. On the one hand, regional specialisation leads to a number of advantages. Firms or whole industries are able to benefit from economies of scale and intra-industrial linkages to a higher extent. On the other hand, the strong regional specialisation increases the potential risk of asymmetric shocks. Specialisation tendencies thus lead to the need of improved or new shock absorbing mechanisms on the national and even regional level, whereas inside EMU, monetary policy is centralised at the European level.

The seminal study of Krugman which confronts a lower level of specialisation in the EU with a higher one in the US is based on the analysis of mobile labour. But the high mobility of employees across US states contrasts sharply with that across the EU. Results on specialisation tendencies for US states which rely on employment data can therefore not simply be extended to EU countries or regions. In addition, inside the EU, an increasing mobility of capital enforced by European Monetary Union (EMU) and increased financial market integration can be observed[1]. Due to the possible substitution of capital and labour as factors of production, the allocation of capital might reflect specialisation tendencies inside the EU which cannot be detected when restricting investigations to labour. However, up to now theoretical and descriptive analyses on the distribution of economic activity focus on the production factor labour and are only rarely extended to production or trade data. The allocation of capital (investments and/or direct investments), however, has been neglected so far, but is subject of the analysis in Stirboeck (2001) and in European Commission (1999). In addition, a profound analysis of the homogeneity of regional, not national, economic structures is still neglected in recent research. Exceptions are the studies of Kalemli-Ozcan, Sorensen and Yosha (1999), Tirado, Paluzie and Pons (2000) as well as Stirboeck (2001).

The purpose of this study is to investigate the regional distribution of investment across industrial sectors, i.e. the level of regional specialisation, inside the EU with the emphasis on the identification of important determinants of relative investment dissimilarities in EU-regions and the impact of EU-integration on these regional developments. Determinants of regional specialisation tendencies are given by theoretical models of the international trade theory, regional economics as well as the new economic geography. In section II, these are shortly summarised and the results of recent econometric analyses on the allocation of production and their empirical determinants are given.

Acknowledgements: Financial support of the Volkswagen-Stiftung within the project No. II/76547 is gratefully acknowledged. I would like to thank Laura Manthey, Paula Monotya, Jennifer Säwe, Elena Todorova, and especially Simone Giloth and Felix Morck for their enduring help in setting up the extensive regional data sets. I am also indebted to Herbert S. Buscher, Jürgen Kaehler and Michael Schröder for helpful comments. All remaining errors, however, are my own.

[1] All restrictions on long- and even short-term capital movements within EU are formally prohibited since the passing of the EU-Council directive 88/361/EEC in June 1988. However, a large number of exceptions to this directive have existed for very long. Even according to the IMF Annual Report on Exchange Restrictions 2000, some of them were still persisting in 1999 – the year of the instauration of EMU. In addition to these national administrative barriers, the segmentation of European capital markets also impeded the perfect liberalisation of capital movements. However, the European Common Market as well as the preparation and installation of EMU have exerted a great pressure on the implementation of perfect and unhindered capital mobility and have already improved market integration. For details, see e.g. the recent studies of Santillán, Bayle and Thygesen (2000), European Commission (2000) as well as Danthine, Giavazzi and von Thaden (2000).

Our econometric investigations on the determinants of the similar or dissimilar relative allocation of sectoral gross fixed capital formation (GFCF) within regions are then presented in section III. Data refer to the period 1985 to 1994 and to two different aggregation levels of the Eurostat nomenclature of territorial units (NUTS – Nomenclature des unités territoriales) of seven EU countries: NUTS 1- and NUTS 2-regions. NUTS 2-regions, the less aggregated geographical entities, are of particular importance as EU regional policy is implemented at NUTS 2-level. We run, for the 10 separate years, cross-sectional as well as pooled regressions by use of GLS- and IV-estimates as well as pooled regressions in logit terms. In addition, fixed effects (within) estimates and between estimates are conducted to control for different effects in the process of the potentially dissimilar regional specialisation of EU-regions. Since we deal with regional developments, we cannot exclude spatial interdependencies as well as measurement errors leading to spatial error autocorrelation. Spatial econometric estimates are therefore additionally run to control for the potential spatial dependence and to check for the robustness of our results obtained by classical econometric methods.

In these econometric analyses, we can detect that a higher market size as well as the size of a region decrease the regional level of relative investment specialisation while a higher rate of unemployment, number of patents as well as population density, the fact of being a central region, and the distance to the economic centre increase it. A higher level of economic openness or of capital market integration also leads to stronger structural dissimilarity. These results are confirmed in the spatial econometric estimates where we are confronted with an additional negative spatial dependence between nearby regions.

II Explaining regional specialisation tendencies

In traditional trade theory, the level of regional specialisation and thus the concentration of economic activity is assumed to be in accordance with comparative advantages. Regions specialise to a larger extent in case of economic openness and market integration. Agglomeration tendencies such as a high density of population, capital or (sectoral) economic activity in only one regional area and a disequilibrium in economic developments are not to be expected. Gravity models in international economics (Tinbergen, 1962; Linnemann, 1966) explain economic flows between regions through gravitational and resistance forces such as market size or market potential, distance, barriers to international activity etc. The spatial concentration of e.g. investments can thus be the result of gravitational forces. These forces will gain importance when resistance forces, such as transport costs or imperfect integration, go down.

The new economic geography has sharply increased the importance of regional economic theory in the 1990s. It has induced a new wave of attention to concentration and specialisation patterns. However, already before the 1990s, polarisation theories in the framework of regional economics have provided explanations of circular and cumulative agglomeration tendencies due to “forward and backward linkages” (Hirschman, 1958) or “backwash-effects” (Myrdal, 1957) which are unfortunate for peripheral regions. Since Krugman (1991), the new economic geography has gained a special focus of attention as according to these models, specialisation need not – like in the neo-classical world – to develop according to the comparative advantage of regions, but can be the result of historical conditions and macroeconomic (partly random) processes. Thus even similar regions can develop differently and the resulting patterns of specialisation are ex ante unpredictable. Due to the existence of economies of scale at the level of plants (further increased by economies of localisation at the industrial level), firms do not produce at each single place of local demand. Instead, the production of each differentiated good is locally concentrated and close to large markets. The core thus specialises in economic activity underlying increasing economies of scale, the periphery in agriculture or industries with constant or low economies of scale. In case of high transport costs, the allocation of production over space is rather persistent. Transport costs (also proxying transaction costs) are a centrifugal force working against the spatial concentration of production. Decreasing transport costs, though, strengthen the centripetal force of economies of scale and might thus be a trigger for the concentration of production.

However, the concentration of production in the core can decrease again as soon as transport costs reach a critical low level (Venables, 1996; Krugman and Venables 1995). The reason for this is the dispersing influence of low wages in the periphery which is strong enough when transport costs are sufficiently low. Production in the periphery will then be expanded as immobile labour does not migrate to the core regions although these are marked by higher wages. This theoretical approach therefore predicts an inverse U-shaped development of concentration.

In spite of the growing number of theoretical and empirical studies on the location of industries, econometric studies on regional economic structures are still rare. Recent econometric studies investigating concentration tendencies of production across EU countries are Amiti (1999) and Haaland et al. (1999) as well as Tirado, Paluzie and Pons (2000) focusing on industrial shares in Spanish regions in the 19th century. The level of the industrial specialisation of regions of – among others – three EU countries, finally, is focussed on by Kalemli-Ozcan, Sorensen and Yosha (1999).

Amiti (1999) gets evidence for significant, positive effects of economies of scale and intermediate good intensities on the concentration of industries across five European countries in addition to mostly significant positive fixed industry and time effects which are not explained by the model. Fixed time effects have been increasing over time which according to the author might pick up trade liberalisation effects. In Haaland et al. (1999), the most important determinant for the relative sectoral concentration turns out to be market size, i.e. industries tend to locate close to larger markets. A smaller, but significant effect is also found for labour intensity, i.e. skill-intensive industries seem to concentrate in countries offering highly skilled labour. By use of spatial econometrics, Tirado, Paluzie and Pons (2000) empirically confirm that the industrial intensity (i.e. the share of industrial production in total production) of Spanish provinces strongly increased in the 19th century. It was influenced negatively by the province's share of population, and positively by human capital endowment, relative size of the province's market, and the extent of large scale production (approximated by the province's average tax payment by taxpayer). As the authors compare two points in time before and after the Spanish market integration, they conclude that market integration can be regarded as a trigger of the sharply increased agglomeration of economic activities. Kalemli-Ozcan, Sorensen and Yosha (1999) find higher population density, lower per capita gross domestic product, lower population as well as a higher degree of risk sharing (supposed to represent financial market integration or development) to have a significant increasing impact on the level of industrial specialisation of regions.

The presented theoretical frameworks point to a number of centripetal forces such as market potential, economies of scale or local demand which seem to be important determinants in the explanation of the concentration of sectors across space. Transport or transaction costs play an essential role as a centrifugal force according to the theory as well. In case of mobile capital, transaction costs are largely determined by capital market integration and liberalisation. Both, transaction costs and integration variables will capture effects of EU-integration on regional development. The econometric investigations of concentration tendencies in production mostly find some support for the traditional trade theory as well as the new economic geography approaches. Several important exogenous variables are evident. The concentration of industries across space appears to be determined by market size, human capital or labour intensity, scale intensities, and intermediate goods intensities (or market linkages). In addition, integration seems to have an increasing effect on sectoral concentration.

In this study, however, we focus on regional developments and the similarity of sectoral structures of EU-regions. Hereby, we analyse the allocation of different industrial sectors within a region and hence, to what extent a country or region is specialised sectorally. By this, we do not explain why a region is especially strong or specialised in a particular sector (what is in the focus of traditional trade theory). Instead, our attention is on the determinants of an uneven allocation of relative investment shares within a region, i.e. level of relative regional specialisation. As long as regional specialisation arises along with the concentration of sectors across space, its level might equally be influenced by market size or the gravitational force of the centre. In addition, the distance to the centre and integration or liberalisation impacts (extending potential markets and enforcing gravitational forces) seem to be relevant determinants. Finally, a very important aspect to focus on when analysing regional investments is the potential of economic growth within a region, e.g. number of patents and the regional economic situation.

III Empirical evidence: What determines the level of relative investment specialisation in EU-regions?

As a measure of relative investment similarity, Gini-coefficients[2] are constructed like in the studies of Krugman (1991), Brülhart (1998), Klüver and Rübel (1998), and Amiti (1999)[3]. In order to abstract from size and classification effects (i.e. the differing importance of sectors and the possibly inadequate disaggregation of economic activity in subsectors), we do not calculate the Gini-coefficient over absolute investment shares, but over relative ones, i.e. investment shares in relation to an economy of reference. This is important as the absolute allocation of production across sectors does not tell anything about a particularly high level of sectoral engagement of that region while this is what we focus on: relative allocation and

[2] The Gini-coefficient is well known from the analysis of problems of distribution and is expressed as the ratio of twice the area between the Lorenz curve and the 45°-line. The Lorenz curve is constructed by plotting the - in ascending order - cumulated relative sectoral shares. The Gini-coefficient can be used to focus either on relative or absolute similarity depending on the precise construction of industrial shares whose distribution is analysed (see below). It gives strong weight to the middle parts of the distribution of relative sectoral shares. As a consequence, changes in industrial sectors similar to the median structure have a larger effect on the value of the Gini-coefficient than changes in industrial sectors at the outer sides of the distribution (Cowell, 1995). However, the coefficient's range between 0 (low concentration) and 1 (high concentration) usually reflects well differences in the level of concentration. Therefore, the Gini-coefficient is the most widely used inequality measure in the analysis of the spatial allocation of sectors or sectoral allocation of regional economic activity.
In addition to the calculation of the Gini-coefficient as a measure of the level of relative regional specialisation, we also calculated the Finger-Kreinin-index as well as the coefficient of variation of the adapted Balassa-indices. The latter stresses changes at the outer sides of the distribution of relative sectoral shares which is in contrast to the weighting of the Gini-coefficient. A graphical comparison of all three indicators shows a similar development over time, the main difference is a generally lower value of the coefficient of variation which, however, does not influence its course. We therefore get similar results when using these alternative measures.

[3] However, Sapir (1996) analysing absolute country specialisation with export data made use of the Herfindahl index instead, Greenaway and Hine (1991) of the Finger-Kreinin index. Kalemli-Ozcan, Sorensen and Yosha (1999) used an adaptation of the Finger-Kreinin index which is based on variance-measuring and consistent with their focus on risk-sharing. For a discussion of these measure and their more detailed presentation see Stirboeck (2001).

hence, the degree of relative specialisation. It is the unequal size of regions or sectors that generally causes the difference between absolute and relative investment patterns[4].

The relative regional distribution of capital within EU-regions can, of course, be investigated applying two different perspectives. First, it is possible to analyse the regional investment structure in relation to the national one which would be a national perspective. Second, the regional investment structure can be compared to the average EU structure, thus adopting a European perspective. Both perspectives lead to slightly different patterns of specialisation. We now focus on the European perspective as we aim at capturing potential impacts of EU integration on regional specialisation[5]. Relative investment indices have therefore been constructed to measure the sectoral investment share of the respective region in relation to the average investment share of the sector in EU as a whole[6]: $B_{sr} = (I_{sr} / \sum_s I_{sr}) / (\sum_r I_{sr} / \sum_r \sum_s I_{sr})$ with I as investment and s (r) as the sectoral (regional) index. As a result, this adapted „Balassa-index“ reflects the relative investment performance of a region in a sector. If the region's investment in one sector is relatively strong (low) compared to the other regions, the index is higher (smaller) than 1.

Measuring the level of relative investment specialisation, the Gini-coefficient is calculated in the following way by first ranking the B_{sr} in ascending order:

$$G_r = \frac{\sum_{s=1}^{N} (2s-1) B_{sr}}{N} - 1 \ .$$

It thus captures the degree of homogeneity of these relative investment performance indices for the respective region[7]. In case of very similar relative investment shares for the different sectors (i.e. similar relations of the regional investment shares for all sectors to their respective average share in the reference economy), we get a Gini-coefficient close to zero representing a low level of relative investment specialisation. This Gini-coefficient ranges between 0 and (N-1)/N. A standardised Gini-coefficient G*N/(N-1), referred to as the Lorenz-Münzner-coefficient, is used in the estimates.

[4] While measures of absolute concentration are influenced by regional size and sectoral classification, measures of relative concentration are influenced by the sectoral patterns of either the economy of reference or the average pattern of the group of countries included. In case of a very special pattern of the reference economy, the relative specialisation pattern of the economic entities analysed can be biased. Further details on the construction of different relative and absolute concentration indices can be found in Stirboeck (2001) as well as Krieger-Boden (1999).

[5] Using the relative investment indices in the national perspective does not affect the regression outputs.

[6] Up to 17 differentiated sectors – consistent to the industrial classification of Nace Rev. 1 - Nomenclature des activités économiques dans les Communautés Européennes - are available in the REGIO database.

[7] An alternative way of calculating the similarity of relative patterns by use of the Gini-coefficient was applied by e.g. Krugman (1991) and Amiti (1999). According to their measure, the cumulative sums of sectoral shares of the given regions are to be plotted against those of the reference economy ranked according to their relative shares (i.e. the adapted Balassa-indices). Both sectoral structures (and not their relation) are thus directly compared. In case of very equal sectoral structures, we also get a Gini-coefficient near zero. However, both sectoral shares, the one of the respective region as well as the one of the reference economy, influence the value of the Gini-coefficient. If a large sectoral share in e.g. the reference economy is confronted with an even larger sectoral share in the region in focus, the value of the Gini-coefficient is largely determined by this economically important sector. This effect influences the value of the level of relative specialisation for some regions. Regression results for this kind of indicator, however, do only slightly change.

III. 1 Descriptive features of the similarity of relative gross fixed capital formation in EU-regions

The maximum number of regions included at the NUTS 2-level is 56 while it is 33 at the NUTS 1-level[8]. These regions belong to Belgium, Denmark, France, Luxembourg, Ireland, Italy as well as Great Britain (only NUTS 1). Details are given in the appendix. For all other countries, the sectoral data availability is not sufficient for our kind of analysis. The disaggregation of EU countries into NUTS-regions is primarily based on political or administrative entities. Such "normative" regions are regarded for practical reasons of data availability in the REGIO database but also in accordance with the implementation of regional policies[9]. These regions are not grouped together on the basis of economic criteria. This is often criticised by economists as this might not give us the actual degree of specialisation of economic entities. However, the definition of economic regions might differ for each variable or even sector regarded, i.e. a general specification of regional disaggregation is inappropriate. In addition, the analysis of normative regions disaggregated according to NUTS allows us to focus on the degree of specialisation of a territorial community which is authorised to implement regional policies or is in the focus of regional structural programmes. As the debate about how specialised EU's regions are originates in the question about their regional shock absorbing potential and the necessity of improving regional policies, the analysis of *administrative* regional entities is one relevant empirical aspect.

Table A4 presents average relative Gini-coefficients for all regions for the time period 1985 to 1994 as far as data has been available. One important feature that can be detected when analysing French, Italian, and Belgian regions (the only countries with available investment data for both NUTS 1- and NUTS 2-levels) is the higher level of specialisation when regarding the more disaggregated NUTS 2-regions in comparison to the NUTS 1-regions. Though this is not surprising as relative sectoral dissimilarities are very likely to be aggregated away in bigger economic entities.

Further insights into the process of regional specialisation can be gained from a descriptive comparison of the most "extreme" regions. Table A5 focuses on the two least and the two most specialised NUTS 2-regions. With respect to GDP per capita, GFCF in percent of GDP, and net migration[10], there are no systematic differences between the regions analysed. But those regions with a more uniform relative allocation of investments across industrial sectors are also marked by a higher number of patents, higher absolute GFCF as well as consequently by higher absolute GDP. Higher specialised regions, however, seem to perform worse in economic terms than lower specialised regions with respect to the unemployment rate, the share of regional to total employment, the number of patents, and total regional GDP as well as total regional GFCF. Exceptions though are a number of regions which are – in economic terms – among the most important and which usually are located in the centre of the respective countries[11]. As no causal relationship can be derived from this purely descriptive analysis, econometric analyses are conducted in the following to test for significance and importance of potential determinants.

---

[8] As Ireland, Denmark and Luxembourg are monoregional countries at the NUTS 1- as well as the NUTS 2-level, their relative concentration can only be calculated in relation to EU average sectoral shares. Otherwise, the specialisation indices would all be 1 per definition, and the concentration indices thus 0.

[9] Since the 1961 Brussels Conference on Regional Economies, regional policies are generally applicated in NUTS 2-regions (Eurostat, 1999).

[10] Here, only data since 1997 is available.

[11] This effect becomes even more obvious when regarding Bruxelles and the Île de France at the NUTS 1-level.

III. 2 Econometric evidence on the level of relative investment specialisation

As presented above, a number of determinants from different theoretical approaches are supposed to explain the level of regional specialisation of gross fixed capital formation. However, explanatory variables added in this analysis are to some extent limited by the data availability. Including the core variables mentioned above, we test the following specification:

$$GCCFEU_i = \beta_0 + \beta_1 MAR_i + \beta_2 AREA_i + \beta_3 PAT_i + \beta_4 PODEN_i + \beta_5 UEWP_i + \beta_6 CENTR_i + \beta_7 INT_i + \beta_8 ZENTRREG_i$$

The market size (MAR) of region i is approximated by gross domestic product (GDP)[12]. Additional important exogenous variables are the size of a region (AREA), the number of patents (PAT), population density (PODEN), unemployment rate in percent of working population (UEWP) which are all taken from the REGIO database. The distance of the region to the economic centre (CENTR) capturing effects of peripheral location[13], variables representing European integration (INT), i.e. economic openness and capital account liberalisation indices, as well as an indicator variable for central, economically most important regions (ZENTRREG) are added. Details on all these variables are given in the data appendix. Unfortunately, data on patent applications has not been available for Corse and Northern Ireland. As a consequence, these two regions are dropped in the estimations when using patents as explanatory variables. In addition, data on patents are only available since 1989. Separate estimations for the shorter time period are therefore presented when patents turn out to be an important variable in panel estimates.

The inclusion of the presented potential explanatory variables in the final estimation models of the cross-sectional analysis has been determined by the result of a likelihood ratio (LR) test at the 10%-level of significance[14]. The decision between two different (non-nested) models for the same period has been made in accordance with the Akaike information criterion (AIC). We estimated variance-corrected standard errors by generalised least squares (GLS) to prevent that potential heteroscedasticity influences the coefficients' significance.

However, we cannot exclude from pure theory a potential problem of reverse causation between the level of investment specialisation and regional gross domestic product or the regional rate of unemployment. In order to control for potential endogeneity problems, instrumental variable regressions have been conducted additionally for both GCCF and GCCFEU.

[12] Further variables reflecting market size are the regional level of gross fixed capital formation, value added at factor costs, total population, total employment as well as (aggregated) compensation of employees. Unsurprisingly, these all turn out to be highly correlated with gross domestic product (see Table A6a and Table A6b in the appendix). Due to this fact, only one of these variables can be included in the explanation of the strength of specialisation of gross fixed capital formation. Preference has hereby been accorded to the GDP variable (which is a good proxy for market size). In most cases, the substitution of GDP by any of the other variables does change the results only negligibly, i.e. the estimated coefficients and models are very robust. In some cases, the use of another variable instead of GDP, e.g. population, would have improved the overall goodness of the estimation. However, GDP was used in the estimations what did not lead to differing results.

[13] In the analysis of sectoral agglomeration or concentration, distances are usually supposed to capture trade costs. Up to a certain level, decreasing transport costs might strengthen agglomeration tendencies. But once reaching this level, theory tells that dispersion factors (such as factor costs in the periphery) can be stronger. The variable's sign is thus not expected ex ante by theory. However, as we focus on regional aspects in this analysis, the simplest and most consistent interpretation of the variable "distance to the economic centre" is that it captures the effect of being far away from the economically most important regions, i.e. the impacts of the peripheral location of a given region.

[14] However, most coefficients as well as the LR tests on the inclusion of further variables are significant at the 5%- or even the 1%-level of significance.

Following one common approach in econometric analysis, lagged values of the unemployment rate as well as of gross domestic product are included as instruments. As a consequence, results are very similar, and most coefficients are even nearly identical[15].

III. 2.1 Cross-sectional analyses

Cross-sectional analyses are conducted at the NUTS 2- as well as the more aggregated NUTS 1-level. Results for the NUTS 2-level are displayed in Table 1. The higher gross domestic product and the bigger the size of the region, the more similar relative sectoral shares of gross fixed capital formation turn out to be. In contrast, an increase in the unemployment rate as well as the fact of being a central (economically important) region increase the level of investment specialisation. In some years, population density seems to capture the effect usually picked up by the indicator variable for the central region. Both variables therefore appear to reflect a form of centrality effect. In some years, the number of patents shows a significant increasing effect on the level of specialisation as well.

Only the interpretation of the empirical influence of the liberalisation indicator taken from Quinn (1997, 2000)[16] is ambiguous as it shows first negative and, since 1990, positive signs. Since these liberalisation indicators are not available at the regional level, but only at the country level, they might pick up country-specific effects in cross-sectional analyses. Controlling for this potential effect, country-specific dummies are added to the presented estimations. These country dummies are relative to Italy as the Italian data is available for all sectors and years. As a result, the liberalisation indicators indeed loose their significance in most cases or are even dropped due to problems of high collinearity with the country dummies. However, only for 1990, the equation including country-specific indicators is statistically better according to the AIC and the likelihood ratio test result[17]. In addition, the country-specific dummies are – except for the Belgian dummy - rarely significant. This indicates that on the one hand the effects captured by Quinn's liberalisation indicator cannot be picked up equally by country-specific dummies for the different years. On the other hand, we cannot exclude that in these cross-section regressions the liberalisation indicator does not actually measure additional effects or even other impacts. More reliable results will be gained in the pooled regressions as the indicators' variation over time is exploited.

Besides these inconclusive impacts of the liberalisation indicator, we find plausible and consistent signs for all other coefficients for which significant results are demonstrable. In addition, these coefficients are very robust with respect to the estimation method (i.e. OLS, GLS, IV estimates) as well as to changes in the estimated model or the substitution of gross domestic product by one of the other variables proxying market size such as population. Only for the years 1993 and 1994, the econometric results are not as reliable as in the other years. This is due to the lower number of regions with available data - e.g. for all French regions data is missing.

---

[15] As data is only available for 1985 to 1994, instrumental variable cross-section analyses for 1985 had to be omitted. For the same reason, only 30 out of 53 (56) observations can be included in the regressions for 1986 so that the IV estimates cannot capture the same effects as the simple GLS regression. However, for all other years, the same effects can be demonstrated as explained above.

[16] This indicator differentiates between varying levels of liberalisation over time. Its construction is explained in the appendix.

[17] Results for the regressions including country dummies are available from the author upon request.

Table 1: Cross-sectional analysis: Determinants of GCCFEU, NUTS 2-regions

year	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994
constant	2.2736	0.4643	0.4230	0.4238	0.4563	-0.0480	-1.2992	-0.9289	0.4454	0.4353
	3.06	10.34	10.29	11.95	15.58	-0.32	-3.27	-1.73	13.9	11.23
gdp	-0.0025	-0.0009	-0.0008	-0.0007	-0.0009	-0.0009	-0.0008	-0.0007	-0.0014	-0.0011
	-4.44	-2.04	-2.02	-1.8	-2.81	-2.89	-2.91	-2.31	-3.78	-2.76
zentrreg	0.1891	0.1031	0.1210	0.0841	0.1281	0.1097	0.1861			
	3.7	2.22	2.8	1.97	3.23	2.58	3.6			
uewp	0.0078	0.0085	0.0111	0.0093	0.0065	0.0057	0.0078	0.0105	0.0086	0.0081
	2.21	2.26	3.23	3.28	2.46	1.97	2.96	2.71	3.23	2.82
poden								0.0489	0.0487	0.0517
								2.73	3.65	3.35
area		-0.0039	-0.0039	-0.0031	-0.0025					
		-3.92	-4.2	-3.38	-2.91					
quinn_openn	-0.1798					0.0375	0.1269	0.0989		
	-2.46					3.41	4.39	2.54		
centr	0.1550									
	2.35									
no. obs.	30	56	56	56	56	56	53	56	34	34
Prob Chi²	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Note: Z-values of GLS-estimates are given in the line below the coefficients. The probability of the Chi²-test gives the overall fit of the model. For abbreviations see Table A2 and A3 in the appendix.

Table 2: Cross-sectional analysis: Determinants of GCCFEU, NUTS 1-regions

year	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994
constant	1.2197	0.2591	0.3824	0.4627	0.4411	0.4402	0.4031	0.2310	0.4548	0.2947
	4.55	5.99	12.67	17.99	20.33	21.86	10.38	6.43	12.34	6.36
gdp	-0.0030		-0.0009	-0.0008	-0.0007	-0.0006	-0.0008		-0.0012	
	-5.38		-2.97	-2.52	-2.75	-2.85	-3.36		-2.83	
zentrreg	0.2154	0.0784	0.1297	0.0678	0.0788	0.1290	0.2184	0.0904		
	4.24	2.05	3.22	1.84	2.36	4.42	5.78	2.61		
uewp		0.0119					0.0050	0.0163		0.0073
		3.34					1.7	4.82		2.73
poden		0.0310	0.0361		0.0345			0.0589	0.0625	0.0757
		2.15	2.82		2.7			4.38	4.34	7.53
area									0.0019	0.0026
									2.07	3.19
quinn_openn	-0.0694									
	-2.87									
pat										-0.0002
										-2.43
centr	0.1722		0.1400							
	3.11		2.6							
no. obs.	22	33	33	33	33	33	30	33	25	17
Prob Chi²	0.0000	0.0000	0.0000	0.0124	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Note: see Table 1.

Changing our focus from NUTS 2- to the more aggregated NUTS 1-level, results – displayed in Table 2 – are very similar[18]. Gross domestic product and the unemployment rate prove to have a significant impact on the level of GCCFEU in many years. Central or core regions still show a higher level of relative investment specialisation at NUTS 2-level while population

[18] Again, instrumental variables estimations have been conducted by use of the lagged unemployment rate and the lagged gross domestic product variable as instruments. Like for the NUTS 2-level, results have been very similar to the simple GLS regressions at the NUTS 1-level.

density now demonstrates a joint increasing effect on specialisation in many years. However, if both variables are jointly significant, the size of both coefficients is consistently lower. Only the size of the region does not show a significant and consistent influence at this aggregation level. As NUTS 1-regions are higher aggregated and usually consist of a number of NUTS 2-regions, their size differs much less. In addition, bigger regions are naturally more diversified, so this result is not surprising.

Only in 1986 and 1992, the estimations significantly improve according to the AIC and the LR test results when adding country-specific variables. Ireland, Denmark, and Luxembourg prove to have significant country dummies in some years as well. However, their significance is rather inconsistent over time. The inclusion of country dummies therefore does not really improve estimates at NUTS 1-level. Analogously, the Quinn liberalisation indicator is also not significant any more – besides of the year 1985. This might be the result of the decreased number of regions included at NUTS 1-level and therefore even lower variation of the openness indicator. When adding country-specific dummies, the Quinn-indicator of openness of 1985 still remains significant while the country dummies do not improve the estimation.

III. 2.2 Pooled estimates and panel data analyses

In the following, a number of results for pooled data and panel estimations are presented. Pooled data estimates, given in Table 3, display results very similar to those of the cross-section analyses. Since regressions including patents as explanatory variables only start in 1989 (due to the data availability), two regression models are displayed for the two different time periods when relevant. However, all regressions are robust as the coefficients for the other exogenous variables do hardly change. And – like in the cross-sectional estimates – the instrumental variable estimations – displayed in Table A8[19] - lead to extremely similar results.

Again, higher gross domestic product and a bigger size of a region diminish the level of specialisation of this region while a higher unemployment rate, population density[20] and the fact of being a central region increase its level. The influence of the distance to this central region on the level of specialisation is positive as well. The more peripheral a region is, the less similar is its sectoral structure to the average structure of the EU.

This time, the inclusion of country-specific effects, again relative to Italy, generally improves the estimates according to the AIC and LR tests. All country-specific indicator variables are generally significant. Besides the Belgian regions, which are marked by a significantly higher level of specialisation than the Italian regions, all other regions have a significantly lower level of specialisation. But we cannot be sure that the significant positive fixed country effect for Belgium is due to country-specific characteristics. However, this dummy might also capture the poor quality of the Belgian data. In contrast to fixed country effects, fixed time effects estimated relative to 1994 never improve the estimates.

We are now, finally, aware of a consistent positive impact of liberalisation on the level of specialisation. The Quinn openness indicator is not only significant in most of the pooled regressions, it also remains significant in the improved estimates including country-specific

[19] In this table, we display the results for the first lag of GDP and UEWP used as instruments. We also tested for robustness by using GDP and UEWP lagged two periods. However, the results did hardly change.

[20] The coefficient for population density, however, changes its sign in the estimates at NUTS 2-level when including country-specific dummies. Running separate estimates for either population density or the central region dummy, we get positive signs again.

dummies. Again, we get no differing results when conducting instrumental-variable regressions to exclude potential problems of endogeneity (see Table A8).

Table 3: Panel estimates, pooled regressions: Determinants of GCCFEU

	NUTS 2				NUTS 1			NUTS 2		NUTS 1
	1985-94		1989-94		1985-94			1985-94	1989-94	1985-94
constant	0.3453	0.3112	0.2734	0.3687	0.2593	0.2778	constant	0.3696	0.0604	0.3790
	9.57	8.6	4.5	6.25	6.35	6.75		3.51	0.47	3.95
gdp	-0.0009	-0.0010	-0.0015	-0.0020	-0.0004	-0.0005	gdp	-0.0009	-0.0016	-0.0004
	-7.14	-7.38	-5.01	-6.04	-4.16	-3.8		-7.16	-5.28	-4.25
zentrreg	0.1216	0.2342	0.0849	0.1856	0.1013	0.1558	zentrreg	0.1217	0.0790	0.1020
	6.58	8.5	4.08	5.81	7.02	6.83		6.57	3.78	7.08
uewp	0.0087	0.0075	0.0085	0.0075	0.0072	0.0067	uewp	0.0088	0.0085	0.0074
	8.54	7.66	6.58	6.37	6.73	5.67		8.47	6.57	6.87
poden	0.0136	-0.0134	0.0267	0.0017	0.0359	0.0321	poden	0.0135	0.0277	0.0355
	2.3	-2	3.58	0.21	8.06	5.23		2.28	3.74	8.00
area	-0.0032	-0.0021	-0.0017	0.0002	-0.0003	0.0002	area	-0.0032	-0.0015	-0.0003
	-8.76	-3.94	-3.66	0.29	-1.85	1.06		-8.63	-3.12	-2.00
quinn_openn	0.0072	0.0068	0.0130	0.0058	0.0063	0.0062	quinn_openn	0.0051	0.0274	-0.0029
	2.6	2.63	2.88	1.36	1.97	2		0.69	3.05	-0.44
pat			0.0001	0.0002			pat		0.0001	
			2.77	3.8					3.09	
centr	0.0513	0.1237			0.0606	0.0356	centr	0.0517		0.0615
	3.00	6.66			3.03	1.68		3.02		3.09
dum_fra		-0.0208		-0.0681		-0.0506	year_1985	0.0156		-0.0313
		-1.99		-4.25		-3.57		0.44		-1.00
dum_bel		0.0920		0.0539		-0.0430	year_1986	-0.0013		-0.0287
		6.6		3.49		-2.35		-0.04		-0.97
dum_ire		-0.0504		-0.1136		-0.0967	year_1987	-0.0086		-0.0291
		-0.97		-1.75		-2.78		-0.27		-0.98
dum_lux		-0.1970		-0.2613		-0.0927	year_1988	-0.0142		-0.0167
		-4.82		-5.32		-2.56		-0.48		-0.60
dum_den		-0.1249		-0.1654		-0.1485	year_1989	-0.0004	0.0612	-0.0081
		-3.02		-3.31		-4.77		-0.01	1.90	-0.29
dum_ukd		--		--		-0.0011	year_1990	0.0031	0.0239	0.0001
						-0.10		0.14	1.09	0.00
							year_1991	-0.0083	0.0022	-0.0014
								-0.41	0.11	-0.06
							year_1992	0.0190	0.0269	0.0090
								0.94	1.36	0.42
							year_1993	0.0068	0.0119	0.0134
								0.31	0.47	0.60
no. obs.	487	487	282	282	292	292	no. obs.	487	282	292
SSR	3.9500	3.3305	2.1533	1.6841	1.4399	1.2745	SSR	3.9062	2.1079	1.4157
Log Likeli	481.32	522.86	287.22	321.87	361.25	379.06	Log Likeli	484.03	290.23	363.72
Prob Chi²	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	Prob Chi²	0.0000	0.0000	0.0000
AIC	-1.944	-2.094	-1.98	-2.191	-2.42	-2.5	AIC	-1.918	-1.966	-2.375

Note: see Table 1.

In contrast to the cross-section estimates presented above, time correlation effects might be of importance in panel data. This means that autocorrelation of the residuals cannot be excluded a priori. In order to take account of these potential effects, the lagged endogenous variable is included in the estimates to capture possible first-order serial correlation effects. The use of other, more sophisticated specifications of dynamic adjustments is possible[21]. However, the methodological discussion about the optimal dynamic specification in the econometric analysis of panel data is still ongoing. We therefore only include the endogenous variable lagged one period when checking the robustness of the above presented results. In the estimates with country fixed effects presented in Table A8, almost all coefficients remain significant and

[21] See Baltagi (1995) for an account of this issue.

their sign unchanged[22] – only the population density as well as the regional size variable loose their explanatory power. However, the empirical impact of the variables identified so far is mostly confirmed.

Throughout all regressions, the predicted values were in the range between 0 and 1, so that our restricted endogenous variable did not impose an econometric modelling problem. The use of a logistic transformation of the regression specification – limiting the range of y to lie in the interval 0,1 – is a more sophisticated procedure, but does not change our results as none of the predicted values lies at the outer sides of the given range. The results for additional pooled regressions in logit terms are presented in the appendix Table A8.

Table 4: Panel estimations: Determinants of investment specialisation

	Nuts 2	Nuts 1	Nuts 2	Nuts 1
	WE	WE	BE	BE
constant	0.6376	0.6551	0.4289	0.3203
	8.64	7.20	13.1	8.43
gdp		0.0004	-0.0008	-0.0004
		1.800	-2.46	-1.72
zentrreg			0.0108	0.0842
			3.58	2.9
uewp				0.0095
				3.52
poden	-0.5514	-0.6475	0.1200	0.0375
	-2.29	-2.87	3.13	3.38
area			-0.0033	
			-3.95	
pat				
no. obs.	487	292	487	292
no. groups	56	33	56	33
R^2 within	0.0121	0.0397	0.0000	0.0122
R^2 between	0.1026	0.3518	0.4371	0.6543
R^2 overall	0.0886	0.2924	0.3492	0.5048
Prob F	0.0225	0.0055	0.0000	0.0000

Note: T-values are given in the line below the coefficients. For abbreviations, see Table A2 and A3 in the appendix. WE stands for fixed effects (within) estimates and BE for between estimates.

In order to further exploit the time and cross-sectional structure of the data and to check for robustness of our results, fixed effects (within) estimates and between estimates are conducted additionally[23]. Fixed effects estimates control for region-specific differences by adding a specific constant for each region[24]. Within estimates ("WE") explain the deviation of the regional observations from their respective group (regional) means, thereby excluding all region-specific determinants from these fixed effects estimates. Within estimates thus provide us with further insights on the variation within the groups (in our case within the regions). Be-

[22] The size of the exogenous variables' coefficients naturally changes due to the dynamic relation specification of the relationship. We have excluded the British regions from the dynamic specification for NUTS 1-regions. The quality of the British regional data is rather low since many sectors are missing. This does not disturb our results for pooled estimates, but for the dynamic specification.

[23] First-difference estimates – explaining the change in the level of specialisation – are conducted additionally, but are not convincing neither at the NUTS 2- nor at the NUTS 1-level. This might largely be due to the fact that many of the explanatory variables are simply eliminated by the construction of the first difference. This is the case for the indicator variable for central regions, but also for the only slightly varying variables such as the size of a region, the degree of liberalisation, population density as well as the distance to the economic centre of the respective country.

[24] Using this approach, we have to omit the overall constant.

tween effects ("BE"), in contrast, explain the variation between the different groups only focusing on the groups' means and on region-specific characteristics[25].

By means of fixed effects within estimates, only few significant impacts can be detected. For the panel of all NUTS 2- (NUTS 1-)regions, only population density consistently proves to be a relevant exogenous variable in explaining the variation of the level of regional specialisation over time. Since the variation of population density over time within a region is of minor extent, this effect is not really convincing. All in all, the within-estimates do not contribute to the explanation of regional specialisation patterns within the groups.

In additional fixed effects estimates with explicit region-specific constants, we find that the region-specific constants are generally significant at the 1%-level and turn out to be the most important determinants in the explanation of the specialisation levels. The region-specific constants' levels differ largely in these estimations. The one of Bruxelles-capital is the highest, at a large distance to even the second largest regional constant. We find all regional constants to be significant for the specification at the NUTS 2-level. However, the fixed effects estimates at NUTS 1-level strongly improve when excluding the population density variable from the specification. And again, we find the fixed regional effects to be the only dominant impact on specialisation in these estimates. Though we have a – sometimes not negligible – variation of the specialisation level of one region over time, this variation can only be explained by the presented estimates to a minor extent. The change in the level of specialisation of one region over time might therefore not underlie systematic changes but rather be the result of random disturbances[26].

In contrast to the rather low explanatory power of these within estimates are the results of between effects estimates ("BE") which explain the variation of the level of specialisation between regions. The results are very similar to those of the cross-sectional and the pooled regressions presented above. Gross domestic product, regional size, the unemployment rate, population density, the indicator variable for the central region as well as the distance to the economic centre of the respective country again have significant explanatory power. In addition, the coefficients are of the same sign and about the same size as in the other regressions.

All explanatory variables are consistently significant in pooled and between estimates and are of strong importance in the explanation of systematic differences in the specialisation levels. The variation over time, though, is not explained by our regressions. It is now evident that regional characteristics determine the respective level of specialisation while the variation within a region over time cannot be found to be of systematic nature.

III.2.3 Spatial econometric estimates

Regional data, due to its spatial nature, potentially underlies spatial dependence or interaction. Standard regressions do not account for spatial dependence or autocorrelation thus leading to inefficient or even biased estimates in case of significant spatial processes. In addition to simple OLS estimates presented so far, we therefore refer to models of spatial econometrics in this section, which explicitly take account of spatial interaction (see e.g. Anselin, 1988). The

[25] For further details on panel data analyses, see e.g. Baltagi (1995).

[26] It is also possible that measurement errors as well as the changing number of available sectors have some influence on the variation of the level of regional concentration over time which then naturally cannot be explained by economic determinants.

structure of spatial interconnectedness is usually given or rather imposed by so-called spatial weights matrices (W).

A number of different spatial econometric models – as well as combinations of those[27] – can be formulated. In a spatial autoregressive error model, λ[28] captures the spatial autoregression of the error term ε while u is an independently and normally distributed error term with constant variance:

$$Y = X\beta + \varepsilon , \qquad \varepsilon = \lambda\, W_{\varepsilon}\, \varepsilon + u ,$$

$$u \sim N\,(0, \sigma^2_u\, I) .$$

In a spatial lag model, ρ is the spatial autoregressive parameter which measures the reaction of Y to surrounding economic developments, i.e. spatial spillovers or the influence exerted by the neighbouring regions on the level of specialisation in region i:

$$Y = \rho W_y Y + X\beta + u , \qquad u \sim N\,(0, \sigma^2_u\, I) .$$

In addition to spatially interdependent endogenous variables, neighbouring regions might be affected by spatially interdependent exogenous variables, thus giving rise to a regressive spatial model with spatially lagged explanatory variables:

$$Y = X_1\beta_1 + W_x X_2 \beta_2 + u , \qquad u \sim N\,(0, \sigma^2_u\, I) .$$

A model specification with spatially lagged explanatory variables only can be estimated by simple OLS estimates[29]. However, to prevent inefficient or even biased estimates in case of spatial lag or error dependence, we have to refer to different estimation methods. A standard technique to deal with spatial lag dependence or spatial autoregressive error terms is to conduct Maximum-Likelihood (ML)-estimates[30].

The standard software (SpaceStat1.90) for this kind of analysis does not capture time-space-models[31], while the simple dynamic relationship we estimated provided evidence for significant serial first-order correlation. However, the dynamic specification also confirmed our results for the main determinants of the level of regional specialisation – with the exception of population density and regional size. Due to this, we now limit our estimations to pure space

[27] In addition to spatial error models with a spatial autoregressive error term, the disturbance term can also follow a spatial moving-average process. For a discussion of different first- or higher order spatial processes combining spatial autoregressive dependent variables or error terms, spatial moving-average error terms as well as spatially lagged external variables in such called "SARMA"- or "SARMAX"-models, see e.g. Anselin/Bera (1998: 251f).

[28] In spatial processes, the spatial autoregressive parameters are not restricted to the usual interval –1, +1. The parameter space is instead restricted by $1/\omega_{min}$ and $1/\omega_{max}$ with $1/\omega_{min}$ and $1/\omega_{max}$ as the smallest and largest eigenvalues of the spatial weights matrix implemented in the regression (Anselin, 2001: 321). Thus, the spatial autoregressive parameter can be smaller than –1.

[29] See e.g. Haining (1990: 344-50). A problem of multicollinearity (between X and WX), however, arises in case of spatially autocorrelated external variables. In this case, estimated parameters have to be interpreted carefully.

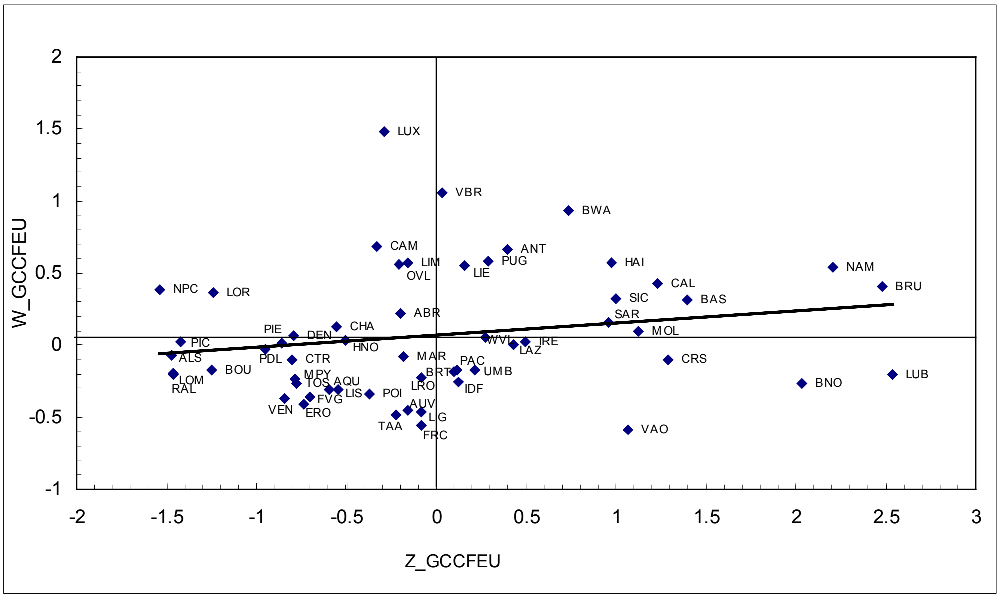
[30] While OLS provides biased estimates in case of spatial lag dependence, it leads to unbiased, but inefficient estimates in case of spatial autocorrelation of the error terms. Since the autocorrelation parameter λ is unknown, we cannot simply conduct weighted least squares estimates, however, and have to refer to maximum-likelihood estimates as well. For further details on this topic, see e.g. Anselin (1988, 1999a).

[31] The implementation of simultaneous time-space effects would have high computational costs and be rather complicated. In the past, different solutions to this problem have been suggested. Schulze (1982) e.g. uses a four-step-Aitken procedure with, first, the elimination of serial correlation and, second, the modelling of spatial dependence.

models. Since the tests on spatial dependence require a normal distribution in the errors of the models estimated, we use restricted datasets so that the non-normality of errors can be rejected for each model presented in the following. We therefore eliminated outlying observations to achieve a normal distribution of the error terms[32].

In order to prevent that our findings are due to the formulation of spatial dependence imposed by the spatial weights matrix, we test the sensitivity of our results with respect to different weights matrices[33]. First, we include two inverse distance matrices and, second, a neighbourhood contiguity matrix. The distance matrices are based on Euclidean distances between administrative centres of the regarded regions as well as between regional centres as provided by the ArcView software. We use the squared inverse of both distance matrices which thus reflects a decreasing intensity of influence of nearby regions with increasing distance. Assuming such a decreasing influence is economically more plausible than a constant strength of interaction[34] as two neighbouring regions can be expected to have stronger interactions than two regions at a high distance from each other. In addition, we use a neighbourhood contiguity matrix, with the element $w_{ij} = 1$ in case of a common border of the regions i and j, and 0 otherwise, while the diagonal is set to 0.

Graph 1: Moran Scatterplot for NUTS2-regions, average level of relative specialisation 1985-94, squared inverse distance between regional capitals



Note: Standardised z_scores of GCCFEU assure the interregional comparability.

The Moran scatterplot given in Graph 1 displays the spatial association between the 56 regions (Anselin, 1996) with respect to their average level of specialisation and the weighted average of the neighbouring values. The levels of specialisation are taken as deviations from their means, the scatterplot is thus centred around 0,0. In the upper right and the lower left quadrant, those regions are displayed which are surrounded by similarly specialised regions

[32] Tests on normality of errors are based on skewness tests as well as Kiefer-Salmon tests – both are applied in such a way that the non-normality of errors cannot even be assumed at the 10% level of significance. By this, we can conclude on a normal distribution of the residuals.
[33] However, a common procedure is also to test a variety of slightly differing distance matrices in order to find the spatial weights matrix that best fits to the underlying process of spatial dependence like e.g. in Molho (1995) or Niebuhr (2001).
[34] We additionally tested for the potential influence of inverse distance matrices with constant influence. In many cases, we got similar results, though the estimates' fit was generally not as good as the one for the preferred weights matrices capturing a decreasing influence of neighbourhood economic activity.

and are thus marked by positive spatial association. Regions with dissembling neighbours are located in the upper left (regions with low specialisation surrounded by highly specialised regions) and the lower right quadrants (vice versa). Using the weights matrix of the squared inverse distances of regional capitals we find four outlying regions[35]: Brussels, Namur, Luxembourg (Belgium) and Basse-Normandie all with a strongly uneven allocation of relative investment shares. The two former, however, are surrounded by similarly specialised regions while the two latter are surrounded by dissimilar regions.

The degree of linear association between the vectors y and Wy is formally indicated by the Moran I statistic[36]. The Moran I coefficient is centred around its theoretical expected mean which is -1/(N-1). Values larger than its expected mean (-0.002 in our case of 469 observations[37]) display positive spatial autocorrelation. The local Moran coefficient points to a significant positive spatial autocorrelation of the level of specialisation, i.e. regions with similar levels of specialisation are more spatially clustered than in the case of random patterns. In other words, regions with a high (low) level of specialisation are more likely to be surrounded by highly (low) specialised regions. For the two squared inverse distance matrices, the significant local Moran I-value is 0.1077 (distance between regional capitals) and 0.1158 (distance between regional centres) and it is even higher – 0.2000 – for the neighbourhood-contiguity matrix. However, from this kind of analysis, we only get information about spatial associations, i.e. the spatial clustering of similar or dissimilar regions. Evidence on spatial dependencies or even causal interactions have to be derived from spatial regression analyses.

Table 5: Diagnostics on Spatial Dependence, OLS-estimates

	ID10K_2		ID10Z_2		NGH_NVDU	
Moran's I (error)	-5.2009	***	-2.7259	***	0.2113	
Lagrange Multiplier (error)	14.6786	***	5.7395	***	0.1976	
Robust LM (error)	2.9375	*	1.0681		0.0047	
Lagrange Multiplier (lag)	25.5496	***	13.3197	***	3.8984	**
Robust LM (lag)	13.8085	***	8.6483	***	3.7055	*

Note: Spatial weights matrices are defined as follows: ID10K_2: squared inverse of distance between regional capitals; ID10Z_2: squared inverse of distance between regional centres; NGH_NVDU: neighbourhood contiguity matrix with the element $w_{ij} = 1$ in case of a common border of the regions i and j, and 0 otherwise.

Table 5 displays the diagnostics on the potential spatial structure to be found in the disturbance terms of simple OLS-estimates. Evidence on the spatial dependence of the residual of simple OLS estimates is provided by a number of tests. It is stronger for the squared inverse distance matrices than for the neighbourhood matrix what is not confirmed in the spatial estimates presented below. The Moran I test (an extended version of the local Moran test presented above[38]) provides evidence for a negative spatial autocorrelation of the residuals as the two significant Moran I values are smaller than its expected value of –0.002.

The significant spatial structure in the residuals, indeed provides evidence that the GLS-estimates presented above suffer from a misspecification. Our aim is now to investigate these effects in more detail in order to check for robustness of the determinants and efficiency of

[35] Outliers are defined as being "extreme with respect to the central tendency reflected by the regression slope", i.e., "they do not follow the same process of spatial dependence as the bulk of the other observations" (Anselin, 1996: 117). The SpaceStat software identifies them as those values larger (smaller) than the third (first) quartile plus (minus) 1.5 times the interquartile range. In the standardised Moran scatterplot, these are the values further than two units away from the origin (Anselin, 1995: 45).
[36] For further details see Anselin (1996: 115ff) and Anselin (1992: 132f).
[37] We now refer to the 469 observations of the restricted dataset which lead to normally distributed residuals in the OLS-estimates of the given specification.
[38] For further details, see Anselin and Bera (1998: 265ff).

the coefficient tests identified in the classical econometric estimates. In all cases, the Lagrange Multiplier (LM)-error test suggested by Burridge (1980) has a lower probability value than the LM-lag-test suggested by Anselin (1988), thus pointing to a model of spatial lag dependence rather than of spatial error dependence. The robust LM-lag and LM-error tests[39] are supposed to be more "suitable for the identification of the source of dependence" (Anselin et al. 1996: 77). In the analysis of our Gini-coefficient, they clearly point to a spatial lag model.

With respect to the spatial lag specification, we get a significant spatial lag differing between –0.07 and –0.60 for the three weights matrices. Thus, we find a significant impact of the level of specialisation in nearby regions. The high level of specialisation in the neighbouring regions significantly reduces specialisation in the region in focus. Therefore, we are confronted with a significant negative spatial interaction of the level of regional specialisation. Regarding the other explanatory variables, the model's results coincide with the above found results. While the LR-tests confirm the spatial lag dependence, the test diagnostics for further spatial error dependence do not provide any further spatial structure.

Table 6: Maximum-Likelihood Estimates of Spatial Lag Models (469 observations)

VARIABLE \ weights matrix	ID10K_2	z-value	ID10Z_2	z-value	NGH_ND	z-value
W_GCCFEU	-0.5964	-4.46	-0.3097	-2.80	-0.0652	-1.79
CONSTANT	0.5157	8.06	0.4017	7.10	0.3076	8.36
GDP	-0.0008	-7.43	-0.0008	-7.89	-0.0008	-8.10
ZENTRREG	0.2117	9.61	0.2213	10.03	0.2198	9.69
UEWP	0.0107	11.91	0.0092	10.86	0.0078	9.82
PODENNEU	-0.0091	-1.73	-0.0107	-1.98	-0.0076	-1.40
AREA	-0.0018	-4.41	-0.0018	-4.20	-0.0017	-4.10
QUINN_OP	0.0087	4.23	0.0086	4.12	0.0083	3.91
CENTR	0.1200	8.15	0.1251	8.35	0.1152	6.15
DUM_FRA	-0.0381	-4.61	-0.0372	-4.42	-0.0331	-3.93
DUM_BEL	0.1279	9.14	0.1123	8.16	0.0897	7.76
DUM_IRE	-0.0562	-1.38	-0.0607	-1.48	-0.0835	-1.93
DUM_LUX	-0.0775	-2.21	-0.1368	-4.04	-0.1685	-5.18
DUM_DEN	-0.1081	-3.27	-0.1208	-3.65	-0.1558	-4.34
Breusch-Pagan test	98.1411	***	99.1209	***	98.80191	***
LR-Test on spatial lag dependence	23.7133	***	10.2056	***	3.5954	*
LM-Test on spatial error dependence	0.0065		1.4692		1.278618	
LIK	621.39		614.64		611.34	
AIC	-1214.79		-1201.28		-1194.67	
SC	-1156.68		-1143.17		-1136.56	

Note: For spatial weights matrices see Table 5.

ML-estimates of the spatial error specification are presented in Table 7. Again, we find a negative spatial correlation which is displayed in the negative autocorrelation coefficient of the error terms[40]. Neglecting the population density variable, all our determinants are robust like in the spatial lag specification. Except for the estimates referring to the neighbourhood contiguity matrix, the LR-tests on spatial error correlation confirm the spatial autoregressive error dependence and the tests on further spatial lag dependence are not significant[41]. The

[39] The robust LM-tests are modifications controlling for a joint significance of both, spatial lag and error dependence, developed by Bera and Yoon (1993) and are discussed in detail in Anselin et al. (1996).

[40] As explained above, the spatial error correlation coefficient is restricted to the range between $1/\omega_{min}$ and $1/\omega_{max}$. Using id10k_2, the minimal spatial error correlation coefficient is thus –1.654, in case of id10z_2 it is 1.552 and of the neighbourhood contiguity matrix it is –1.405. For our three estimates, this condition is fulfilled.

[41] However, the two tests on common factor hypothesis point to an "inherent inconsistency" of the spatial error model. This might be caused by a further influence of spatially lagged explanatory variables (see Anselin, 1992: 212) which is worth being analysed in future research.

results for the neighbourhood contiguity matrix, however, are in line with the spatial dependence diagnostics obtained for the classical regressions which are given in Table 5.

Table 7: Maximum-Likelihood Estimates of Spatial Error Models (469 observations)

VARIABLE \ weights matrix	ID10K_2	z-value	ID10Z_2	z-value	ngh_nvdu	z-value
CONSTANT	0.2469	9.32	0.2563	9.24	0.2542	8.77
GDP	-0.0008	-7.90	-0.0009	-8.71	-0.0009	-8.87
ZENTRREG	0.2280	10.42	0.2377	10.48	0.2413	10.61
UEWP	0.0098	18.76	0.0091	15.47	0.0087	12.27
PODEN	-0.0015	-0.26	-0.0133	-2.39	-0.0113	-2.06
AREA	-0.0007	-1.75	-0.0011	-2.61	-0.0013	-3.02
QUINN_OP	0.0078	3.92	0.0088	4.22	0.0088	4.16
CENTR	0.1162	9.82	0.1157	9.01	0.1298	9.18
DUM_FRA	-0.0372	-6.11	-0.0382	-5.59	-0.0323	-4.39
DUM_BEL	0.0824	9.42	0.0872	9.45	0.0915	8.92
DUM_IRE	-0.1302	-3.39	-0.1174	-2.90	-0.1008	-2.45
DUM_LUX	-0.0488	-1.67	-0.1330	-4.03	-0.1653	-5.05
DUM_DEN	-0.1701	-5.45	-0.1665	-5.10	-0.1496	-4.54
LAMBDA	-1.4844	-13.60	-1.0209	-4.57	-0.3032	-1.58
Breusch-Pagan test	107.5154	***	110.2553	***	105.3065	***
LR-Test on spatial error dependence	43.2261	***	11.4204	***	0.9424	
LM-Test on spatial lag dependence	0.1409		1.6321		0.6174	
LIK	631.15		615.25		610.01	
AIC	-1236.30		-1204.49		-1194.02	
SC	-1182.34		-1150.54		-1140.06	

Note: For spatial weights matrices see Table 5.

Both spatial models (spatial lag as well as spatial autoregressive error dependence) generate the same empirical relationships between the level of specialisation and the explanatory variables we found in the GLS-estimation results without taking account of spatial dependence[42]. In addition, both models appear to be a better specification than the GLS-model without spatial dependence with an AIC of –1193.07 and a log likelihood of 609.54. However, both do not solve the problem of heteroscedasticity. Comparing the two specifications, we consistently get better (i.e. lower) information criteria for the spatial error specification. In contrast to this stand the spatial dependence diagnostics of the OLS-estimates which all pointed to a better specification of the spatial lag model.

IV Economic perspectives and effects of EU integration

The regression results we find in cross-sectional and pooled regression analyses as well as between effects and spatial econometric estimates consistently point to a high importance of market size, regional size, the location in the centre, the distance to the centre, and the population density of a region as well as the unemployment rate and economic and capital market integration in the explanation of relative regional investment specialisation. Our results on the impacts of market size, the population density of a region and of capital market or economic liberalisation on specialisation levels are in line with Kalemli-Ozcan, Sorensen and Yosha (1999) who found lower population, higher population density and higher risk-sharing (as a proxy for capital market integration) to increase regional specialisation of production. In ad-

[42] Population density is insignificant in the GLS-estimates with the restricted dataset of 469 observations as well as in some of the spatial dependence models. In case of significance, it displays a negative sign, what is not consistent with most the different GLS-estimates.

dition, we controlled for the effect of further economic variables, in particular the unemployment rate, the number of patents of a region as well as its distance from the centre.

Table 8: Impact of economic variables on the level of relative specialisation

Economic variable	Sign of impact on GCCFEU
Gross domestic product	-
Fact of being a central region	+
Unemployment rate	+
Population density	+
Size if a region	-
Economic openness	+
Distance to economic centre	+
Number of patents	+

Market size reflects the economic as well as demand potential of a region. The better it is, the more similar relative investment shares tend to be. However, this significant impact of market size is in contrast to the increasing impact of market size on sectoral agglomeration consistently found by Haaland et al. (1999) and Tirado, Paluzie and Pons (2000). While – according to their results – firms tend to locate close to large markets, our empirical results show that regions with a larger market seem to attract capital of all types of sectors with a more even relative allocation (hence showing a lower level of relative regional specialisation) than smaller markets. Economic activity in regions with a lower gross domestic product seems to be specialised to a higher degree. However, the determinants of the location of particular sectors across EU-regions is not subject of this study.

In addition to regions with a larger market size, regions with a bigger size tend to have a more similar relative distribution of investments. But the significance of the regional size variable might simply be due to the fact that bigger regions are logically more diversified in their production structure than smaller regions. The fact that NUTS 1-regions are less varying in their size as well as more evenly diversified than NUTS 2-regions leads to an insignificant size variable at the NUTS 1-level. This demonstrates that it is very important to analyse equally big regions. Thus preference has to be accorded to regions as small as possible to avoid that aggregation cancels out potential specialisation patterns. Controlling for regional size effects in the estimates is therefore essential.

In contrast to the fact that larger markets do not show a strong investment specialisation in only few sectors is the result we find for a number of economically very important regions: Central regions (and equally regions with a high population density) demonstrate a significantly higher level of relative investment specialisation. This increasing effect counteracts the decreasing impact of the large market size. But, the high importance of the centrality indicator variable which captures an outstanding strong fixed effect for Brussels, Lazio, and the Île de France is in line with polarisation theory and new economic geography which predict cumulative causation and self-reinforcing agglomeration. Regions having once gained a particularly high potential of market or factor access attract further firms. Supposedly, sectors with firms underlying positive economies of scale or economies of localisation expand in the core.

In addition, the impact of the distance of a given region to its (economic) centre on the level of specialisation is positive. This means that peripheral regions are more different from the average EU structure than regions closer to the centre. Both, the high specialisation of core as well as of peripheral regions is in line with new economic geography predictions. In contrast to the specialisation of core regions, we expect the specialisation of peripheral regions to occur in sectors with low economies of scale and possibly high labour-intensity. However, the sectoral patterns of regional specialisation remain the subject of further research.

The impact of European integration on specialisation patterns is captured by the influence of market liberalisation and openness. In this context, capital market integration which improves the potential mobility of capital is especially important when regarding dissimilarities in the allocation of relative investments. In addition to the direct measuring of the impacts of economic openness presented above we therefore also tested for the influence of capital market integration in the pooled estimates. This is displayed in Table A8. All explanatory coefficients are in general very similar to the coefficients in the models with the openness indicator. The correlation between both indicators is about 0.77 (0.68) for the regions at NUTS 2-level (NUTS 1-level). The indicator of liberalisation of capital accounts ranges between 0.5 and 4 while the economic openness indicator goes up to 14 when perfect openness is reached. As a consequence the coefficient of the capital account liberalisation indicator is higher than the other coefficient in our estimates. However, we cannot conclude on a difference in the strength of the indicators' influence from this fact.

The results of our regressions clearly demonstrate an increasing impact of liberalisation on investment specialisation, so that the perfect liberalisation and capital market integration within EMU seem to further augment specialisation patterns. Instead of a stronger diversification, European regions might end up with an increasingly different relative investment structure in the process of market integration. If this effect continues, further liberalisation would lead to a higher risk of asymmetric shocks. However, specialisation need not always be negative even though production structures become less diversified. The specialisation in an industrial sector providing a high growth potential might be an asset and improve the regional competitiveness in spite of a highly asymmetric industrial structure.

The unemployment rate reflects negative economic performance of a region (not accounting for migration effects etc.). The worse the economic performance is, the stronger the structural dissimilarity to the average patterns of EU turns out to be in our empirical analyses. The number of patents of a region is significant in many estimations as well, though not in all. Its coefficient is mostly positive, i.e. we have a first indication that a higher number of patents increases relative regional specialisation as well. Possibly patents only attract investments of very particular sectors as they play a strong role in many important, but not in all sectors.

The spatial econometric analyses largely confirm the described impacts of economic variables on the regional level of specialisation while controlling for spatial interdependencies. Thus, the determinants identified by classical econometric methods remain robust. In addition, the regressions provide evidence of negative spatial interactions. While the OLS test diagnostics on spatial dependence point to a spatial lag dependence, the information criteria indicate a better performance of the spatial error model. In the spatial lag specification, we find a significant negative spatial spillover between the regional levels of specialisation. Economically more plausible, however, is the interpretation of the negatively spatially correlated error terms as a nuisance term due to measurement errors. Data inconsistencies as well as incompatibilities due to shortcomings of the regional databases or a poor fit of the units of observation with actual economic regions can be the initial reason for such spatial nuisances in the data.

Further improvements of this study are possible. Regarding the econometric analysis, it is possible to further elaborate the weights matrices and to ameliorate the model specification by e.g. investigating spatially lagged external variables. In addition, our analysis only focuses on the determinants of regional investment specialisation levels. We find some evidence for the stronger specialisation of regions in the extreme localisations core and periphery. As mentioned above, the analysis of sectoral patterns of relative specialisation is an important aspect when focusing on regional specialisation to detect possible regional imbalances by e.g. core-periphery patterns of the localisation of capital-intensive or growth-oriented sectors.

Bibliography

AMITI, M. (1999): Specialization Patterns in Europe, Weltwirtschaftliches Archiv 135 (4), 573-593.

ANSELIN, L. (1988): Spatial econometrics: methods and models, Dordrecht, Kluwer Academic Publishers.

ANSELIN, L. (1992): SpaceStat Tutorial, A Workbook for Using SpaceStat in the Analysis Spatial Data, available: http://www.spacestat.com/docs/spacestat_tutorial.pdf.

ANSELIN, L. (1995): SpaceStat Version 1.80 User's Guide, available: http://www.spacestat.com/docs/V180man.pdf.

ANSELIN, L. (1996): The Moran scatterplot as an ESDA tool to assess local instability in spatial association, in: Fisher, M. and D. Unwin (eds.): Spatial analytical perspectives on GIS, Taylor and Francis, London.

ANSELIN, L. (1999a): Spatial econometrics, mimeo, available: http://geog55.gis.uiuc.edu/~luc/papers/baltchap.pdf.

ANSELIN, L. (1999b): Spacestat Version 1.90, Ann Arbor, BioMedware.

ANSELIN, L. (2001): Spatial Econometrics, in: Baltagi, B (Hrsg.): A Companion to Theoretical Econometrics, Oxford, Basil Blackwell, 310-330.

ANSELIN, L. und A. K. BERA (1998): Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics, in: Ullah, A. und D.E.A. Giles (Hrsg.): Handbook of Applied Economic Statistics, New York, 237-289.

ANSELIN, L., BERA, A.K., FLORAX, R. AND M.J. YOON (1996): Simple diagnostic tests for spatial dependence, Regional Science and Urban Economics 26 (1996), 77-104.

BALTAGI, B.H. (1995): Econometric Analysis of Panel Data, Wiley, Chichester.

BERA, A.K. AND M.J. YOON (1993): Specification Testing with Misspecified Alternatives, Econometric Theory, 9, 649-658.

BRÜHLHART, M. (1998): Trading Places: Industrial Specialisation in the European Union, Journal of Common Market Studies 36 (3) 1998, 319-346.

BURRIDGE, P. (1980): On the Cliff-Ord Test for Spatial Autocorrelation, Journal of the Royal Statistical Society, 42, 107-108.

COWELL, F. A. (1995): Measuring Inequality, London, Prentice Hall.

DANTHINE, J.-P., F. GIAVAZZI AND E-L. VON THADEN (2000): European Financial Markets after EMU: A First Assessment, CEPR Diskussion Paper, No. 2413.

ESRI (2001): ArcView 3.2, http://www.esri.com/software/arcview/index.html.

EUROPEAN COMMISSION (1999): Specialisation and (Geographic) Concentration of European Manufacturing – Background Paper for "The Competitiveness of European Industry: 1999 Report", Enterprise DG, Working Paper N° 1.

EUROPEAN COMMISSION (2000): Report from the Commission - Economic Reform: Report on the functioning of Community product and capital markets, COM(2000) 881 final, Brussels.

EUROSTAT (1999): Statistik kurzgefaßt, *Allgemeine Statistik* 1-3/1999, Luxembourg.

GREENAWAY, D., AND R.C. HINE (1991): Intra-industry Specialization, Trade Expansion and Adjustment in the European Economic Space, Journal of Common Market Studies 29(6), 389-403.

HAALAND, J.I. , KIND, H.J., MIDELFART-KNARVIK, K.H. AND J. TORSTENSSON (1999): What determines the economic geography of Europe?, Centre for Economic Policy Research, Discussion Paper No. 2027, London.

HAINING, R. (1990): Spatial data analysis in the social and environmental sciences, Cambridge University Press, Cambridge.

HIRSCHMAN, A. (1958): The Strategy of Economic Development, Yale University Press, New Haven.

IMF (several years): Annual Report on Exchange Arrangements and Exchange Restrictions, Washington.

KALEMLI-OZCAN S., B.E. SORENSEN AND O. YOSHA (1999): Risk Sharing and Industrial Specialization: Regional and International Evidence, mimeo.

KLÜVER, A. AND G. RÜBEL (1998): Räumliche Industriekonzentration und die komparativen Vorteile von Ländern – eine empirische Studie der Europäischen Union, Passauer Diskussionspapiere, Diskussionsbeitrag No. 5-98, Passau.

KRIEGER-BODEN, C. (1999): Nationale und regionale Spezialisierungsmuster im europäischen Integrationsprozeß, Die Weltwirtschaft (2), 234-254.

KRUGMAN, P. (1991): Geography and Trade, Cambridge/Mass., MIT Press.

KRUGMAN P. AND A.J. VENABLES (1995): Globalization and the Inequality of Nations, Quarterly Journal of Economics 110(4), 857-880.

LINNEMANN, H. (1966): An Econometric Study of International Trade Flows, North-Holland, Amsterdam.

MOLHO, I. (1995): Spatial Autocorrelation in British Unemployment, Journal of Regional Science, Vol. 53, 4, 641-658.

MYRDAL, G. (1957): Economic Theory and Under-Developed Regions, London.

NIEBUHR, A. (2001): Convergence and the Effects of Spatial Interaction, Jahrbuch für Regionalwissenschaft, 21, 113-133.

QUINN, D. (1997): The Correlates of Change in International Financial Regulation, American Political Science Review, 91(3), 531-551.

QUINN, D. (2000): Democracy and International Financial Liberalization, mimeo.

SANTILLAN, J., M. BAYLE AND C. THYGESEN (2000): The Impact of the Euro on Money and Bond Markets, Occasional Paper No. 1, European Central Bank

SAPIR, A. (1996): The Effects of Europe's Internal Market Program on Production and Trade: A First Assessment, Weltwirtschaftliches Archiv 132 (3), 457-475.

SCHULZE, P.M. (1982): Zur Schätzung regionaler Konsumfunktionen, Jahrbuch für Regionalwissenschaft, 116-128.

STIRBÖCK, C. (2001): Agglomeration Tendencies in EU Regions: Where Does Capital Go?, ZEW Discussion Paper No.01-28.

TINBERGEN, J. (1962): Shaping the World Economy, Suggestions for an International Economic Policy, Twentieth Century Fund, New York.

TIRADO A., E. PALUZIE AND J. PONS (2000): Economic integration and industrial location: The case of Spain before WWI, paper presented at the EEA-meeting in Bolzano in September 2000.

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

Appendix

Data description

The regional disaggregation of the data is given according to the Nomenclature of Territorial Units for Statistics (NUTS - Nomenclature des unités territoriales statistiques). The REGIO database disaggregates data for the three aggregation levels NUTS 1, 2 and 3. However, data for GFCF is not available further disaggregated than the NUTS 2-level. In addition, it is not complete (with regard to the regional and/or the sectoral disaggregation – the latter needed for the calculation of the specialisation indices). Data availability is sufficient for the seven countries given below. Here, the UK does not provide data disaggregated further than NUTS 1-level. Luxembourg, Denmark as well as Ireland are only regarded as one single region at the NUTS 1- as well as at the NUTS 2-level (=monoregional countries). The maximum number of regions available is therefore 33 at the NUTS 1-level and 56 at the NUTS 2-level.

Table A1: Regional data for GFCF from the REGIO database

Country	NUTS level	Respective national disaggregation level	Number of regions NUTS 1	Number of regions NUTS 2
UK	1	Groups of Counties or local authority regions	11 (with 3 n.a.)	n.a.
Belgium	2	Provinces	3	11
France	2	Régions	8	22
Italy	2	Regioni	11	20
Denmark	1&2	-	1	1
Ireland	1&2	-	1	1
Luxembourg	1&2	-	1	1
Total number of regions			33 (+ 3 n.a.)	56

Note: Version of *NUTS 1995*. French oversea departments (DOM – départements outre-mer) are not counted in total sums for France as well as for the EU.

Data is taken from the Eurostat REGIO Database (yearbooks up to 2000) which – for gross fixed capital formation - comprises data for the years 1985 to 1994. All data included in the analysis is based on ESA79.

Table A2: List of explanatory variables, REGIO Database

abbreviation	variable	unit
gfcf	Gross Fixed Capital Formation	Currency: Billions of ECU
totem	Total Employment	in 1000 persons
coe	Compensation of employees	Currency: Billions of ECU
vafp	Gross value added at factor costs	Currency: Billions of ECU
gdp	Gross domestic product	Currency: Billions of ECU
pat	European R&D patent applications	total number
uewp	Total Unemployment rates	in % OF WORKING POPULATION
pop	Total annual average population	in Mio. PERSONS
poden	Population density	in 1000 INHABITANTS/KM2

In addition to the available national account data, a number of further variables has been used in the econometric analysis. The distance to the centre (**centr**) captures peripheral effects. It is measured by the optimal route distance between the regional capital and the centre of the respective country. Centres are Paris, Rome, London and Brussels. The distance is defined to be 1 for Denmark, Luxembourg as well as Ireland, and it is equally 1 for the regions containing the capital of the respective country. Central and economically important regions (**zentrreg**) in the analysis are Île de France (France), Brussels (Belgium), and Lazio (Italy).

Table A3: List of further explanatory variables

abbreviation	variable	unit
centr	distance to centre, proxy for transport costs	metres
zentrreg	regional dummy set for central region	0 or 1
quinn_openn	indicator of openness per country	0-14 (variation by 0.5)
quinn_ca	indicator of capital account liberalisation per country	0-4 (variation by 0.5)

Available indicators of liberalisation arising from official sources are mostly indicator variables being either 0 or 1. However, such indicator variables do not allow to differentiate the varying levels of control or to capture a decreasing level of control over time. Measuring a level of integration for each year is therefore a better solution from an econometric point of view. Quinn (1997, 2000) has constructed such a yearly index of openness on the basis of those restrictions published by the IMF since the 1950s. This index is scaled from 0 (highest degree of restrictions) up to 14 (highest degree of liberalisation) and aggregates the different indicators of liberalisation progress in seven specified fields (capital in – and outflows, im– and exports of goods and of services as well as international conventions of liberalisation) with a respective degree of liberalisation between 0.5 and 2.

Quinn weighs quantitative restrictions of imports for example the highest (i.e. he attributes the lowest partial liberalisation index of 0 in case of full and 0.5 in case of partly quantitative restrictions), existence of laws requiring the approval of international transactions are scored 1, taxes 1.5 and finally free trade 2. With regard to capital account liberalisation, Quinn attributes 0 in case of required approval for capital transactions which are rarely granted, 0.5 (1) in case of occasional (frequent) approval and finally 1.5 in case of taxing measurements (without the need of an official approval). A subindex of the overall liberalisation index is a financial liberalisation indicator ranging on a score between 0 and 4 which is aggregated from restrictions of capital inward and outward flows in the way explained above. All named potential indicators, however, are only available at country, not regional, level, which has to be taken into account in econometric analysis. Detailed restrictions for Luxembourg are not available as Luxembourg and Belgium are part of a common monetary union since the 1950s. In our analysis the „Quinn-indicator“ for Luxembourg is therefore naturally set equal to the one of Belgium.

Table A4: Investment specialisation levels in EU-regions in 1985 to 1994 in relation to EU as a whole (Standardised Gini-coefficients/ 17 sectors)

France		Italy		Belgium		United Kingdom	
Nuts 2, 1986 to 1992		**Nuts 2**		**Nuts 2**			
Basse-Normandie	0.672	Basilicata	0.607	Luxembourg (B)	0.725	**Note:** Sectoral availability is strongly varying in Belgium from 4 to 11 sectors. However, mostly 11 sectors are included. Due to a change in sectoral availability of British data (9 sectors prior to 1987, 4 since 1988), results for the time since 1988 are rather primarily and not directly comparable to prior years. Data has not been available for three of the eleven NUTS 1-regions.	
Corse	0.596	Calabria	0.590	Bruxelles-capitale	0.719		
Ile de France	0.475	Molise	0.579	Namur	0.691		
Bretagne	0.474	Valle d'Aosta	0.572	Hainaut	0.563		
Provence-Alpes-Côte d'Azur	0.473	Sicilia	0.566	Brabant Wallon	0.538		
Languedoc-Rousillon	0.454	Sardegna	0.562	Antwerpen	0.503		
Franche-Comté	0.454	Lazio	0.507	West-Vlaanderen	0.491		
Auvergne	0.446	Puglia	0.492	Liège	0.479		
Poitou-Charentes	0.424	Umbria	0.485	Vlaams Brabant	0.466		
Haute-Normandie	0.410	Liguria	0.454	Limburg (B)	0.446		
Limousin	0.406	Marche	0.443	Oost-Vlaanderen	0.441		
Champagne-Ardenne	0.405	Abruzzo	0.442				
Aquitaine	0.400	Trentino-Alto Adige	0.439				
Midi-Pyrénées	0.381	Campania	0.428			**Monoregional countries**	
Centre (F)	0.379	Friuli-Venezia Giulia	0.390			Denmark	0.380
Pays de la Loire	0.364	Emilia-Romagna	0.386			Ireland	0.513
Lorraine	0.334	Toscana	0.382			Luxembourg	0.432
Bourgogne	0.333	Veneto	0.375				
Picardie	0.315	Piemonte	0.373				
Rhône-Alpes	0.311	Lombardia	0.311				
Alsace	0.310						
Nord - Pas-de-Calais	0.303						
Nuts 1, 1986 to 1992		**Nuts 1**		**Nuts 1**		**Nuts 1, 1985 to 1987**	
Ile de France	0.475	Sicilia	0.566	Bruxelles-capitale	0.719	Yorkshire and the Humber	0.471
Méditerranée	0.445	Sardegna	0.562	Région Wallonne	0.439	Wales	0.449
Ouest	0.405	Lazio	0.507	Vlaams Gewest	0.271	West Midlands	0.441
Sud-Ouest	0.361	Sud	0.500			East Midlands	0.435
Bassin Parisien	0.327	Abruzzo-Molise	0.446			East Anglia	0.424
Centre-Est	0.316	Campania	0.428			Scotland	0.399
Nord - Pas-de-Calais	0.303	Emilia-Romagna	0.386			South West	0.390
Est (F)	0.289	Centro (I)	0.375			Northern Ireland	0.378
		Nord Est	0.359			North	n.a.
		Nord Ovest	0.344			North West	n.a.
		Lombardia	0.311			South East	n.a.

Table A5: Characteristics of most/least specialised NUTS 2 regions – specialisation relative to EU structure (average 1985-94 unless indicated in brackets)

Region	Index of investment specialisation (GCCFEU)	Index of employment specialisation	Regional per capita GDP in Mio ECU per 1000	GFCF in billion ECU	GFCF in % of GDP	Number of patents [89-94]	Employment share (in%)	Unemployment rate (in % of working population)	Net migration rate 1997 in ‰ of population
NUTS 2, Most specialised regions									
Bruxelles-capitale	0.719	0.515 [85-92]	23.17	n.a.	n.a.	71	18.0	10.82	0.1
Luxembourg (B)	0.725	0.542 [85-92]	12.57	n.a.	n.a.	6	2.1	7.52	2.8
Basse-Normandie	0.672 [86-92]	0.410 [85-89]	13.15	4.247 [86-92]	24.22	59	2.5	9.87	0.4
Corsica	0.596 [86-92]	0.523 [85-89]	12.40	0.583 [86-92]	19.26	n.a.	0.4	11.18	3.1
Basilicata	0.607	0.648	8.62	1.462	28.0	1	0.9	16.67	3.1
Calabria	0.590	0.697	7.90	4.222	25.56	3	2.9	19.31	-3.2
NUTS 2, Least specialised regions									
Limburg (B)	0.446	0.302 [85-92]	14.05	n.a.	n.a.	21	6.7	10.34	0
Oost-Vlaandern	0.441	0.397 [85-92]	14.06	n.a.	n.a.	71	12.0	6.67	1.5
Alsace	0.310 [86-92]	0.247 [85-90]	16.2	6.058 [86-92]	23.45	215	2.9	6.28	1.1
Nord-Pas-de-Calais	0.303 [86-92]	0.265 [85-89]	13.02	9.877 [86-92]	19.67	125	5.9	13.12	-3.9
Piemonte	0.373	0.367	15.48	12.865	19.12	306	8.4	7.08	2.8
Lombardia	0.311	0.373	17.55	27.849	17.90	806	17.2	4.97	3.9
Region including national capital									
Ile de France	0.475 [86-92]	0.277 [85-89]	24.30	52.066 [86-92]	20.61	2232	22.6	8.29	-4.9
Lazio	0.507	0.359	15.18	16.626	21.24	147	9.3	9.75	5.2
South-East	n.a.	n.a.	15.08	n.a.	n.a.	779	33.8	7.86	25.0

Table A6a: Correlation matrix NUTS 1-level: multicollinear variables

	gdp	gfcf	vafp	pop	totem	coe
gdp	1					
gfcf	0.9456	1				
vafp	0.9984	0.9555	1			
pop	0.8535	0.8267	0.8574	1		
totem	0.9383	0.877	0.936	0.9496	1	
coe	0.9826	0.9044	0.973	0.8456	0.9349	1

Table A6b: Correlation matrix NUTS 2-level: multicollinear variables

	gdp	gfcf	vafp	pop	totem	coe
gdp	1					
gfcf	0.9039	1				
vafp	0.9994	0.9027	1			
pop	0.9118	0.8473	0.9138	1		
totem	0.9667	0.8755	0.967	0.9726	1	
coe	0.9936	0.9217	0.9906	0.9129	0.9652	1

Table A7a: Correlation matrix NUTS 1-level: explanatory variables

	gccfeu	gdp	zentrreg	quinn_openn	uewp	poden	centr
gccfeu	1						
gdp	-0.3161	1					
zentrreg	0.4739	0.2648	1				
quinn_openn	0.0361	0.1657	-0.0209	1			
uewp	0.4081	-0.4353	-0.0743	-0.0676	1		
poden	0.5095	-0.1072	0.61	-0.0669	-0.0076	1	
centr	-0.0517	-0.049	-0.4859	0.0721	0.2763	-0.3271	1

Table A7b: Correlation matrix NUTS 2-level: explanatory variables

	gccfeu	gdp	zentrreg	quinn_openn	uewp	poden	centr
gccfeu	1						
gdp	-0.3152	1					
zentrreg	0.2018	0.4548	1				
quinn_openn	0.0147	0.1289	-0.0093	1			
uewp	0.2367	-0.1423	-0.0344	-0.0671	1		
poden	0.2547	0.0684	0.618	-0.0448	-0.0345	1	
centr	-0.0631	0.0062	-0.3231	0.0886	0.4613	-0.2537	1

Table A8: Robustness estimates, 1985-94, pooled regressions

	NUTS 2	NUTS 1		NUTS 2	NUTS 1		NUTS 2	NUTS 1		NUTS 2	NUTS 1
	IV 2SLS			dynamic model			logit terms estimation		capital market integration		
constant	0.2943	0.2835	constant	0.0387	0.0087	constant	-0.7781	-0.9509	constant	0.3254	0.2809
	7.31	6.17		1.36	0.33		-4.94	-5.4		8.71	6.95
			gccfeu AR(1)	0.7623	0.7769						
				23.24	17.08						
gdp (IV)	-0.0009	-0.0005	gdp	-0.0002	-0.0001	gdp	-0.0041	-0.0021	gdp	-0.0009	-0.0005
	-6.85	-3.71		-2.58	-1.84		-7.25	-3.84		-7.28	-3.76
zentrreg	0.2371	0.1563	zentrreg	0.0602	0.0551	zentrreg	0.9874	0.6541	zentrreg	0.2328	0.1552
	8.00	6.36		2.89	3.38		8.24	6.71		8.44	6.81
uewp (IV)	0.0075	0.0061	uewp	0.0014	0.0016	uewp	0.0307	0.0277	uewp	0.0073	0.0066
	6.74	4.59		1.95	2.00		7.22	5.5		7.5	5.62
poden	-0.0141	0.0315	poden	-0.0019	0.0049	poden	-0.0511	0.1461	poden	-0.0132	0.0323
	-1.95	4.78		-0.40	1.38		-1.75	5.56		-1.96	5.26
area	-0.0020	0.0002	area	-0.0004	0.0001	area	-0.0082	0.0010	area	-0.0021	0.0002
	-3.46	0.86		-1.00	0.99		-3.62	1.22		-3.94	1.04
quinn_openn	0.0078	0.0064	quinn_openn	0.0045	0.0045	quinn_openn	0.0282	0.0282	quinn_ca	0.0200	0.0207
	2.75	1.89		2.41	2.57		2.5	2.12		2.11	1.96
centr	0.1269	0.0353	centr	0.0342	0.0346	centr	0.5121	0.1440	centr	0.1238	0.0359
	6.31	1.53		2.48	2.45		6.35	1.59		6.65	1.69
dum_fra	-0.0196	-0.0485	dum_fra	-0.0021	-0.0091	dum_fra	-0.0887	-0.2172	dum_fra	-0.0185	-0.0480
	-1.72	-3.14		-0.29	-1.07		-1.96	-3.58		-1.72	-3.3
dum_bel	0.0963	-0.0426	dum_bel	0.0213	0.0047	dum_bel	0.3992	-0.1964	dum_bel	0.0909	-0.0437
	6.41	-2.17		2.06	0.45		6.58	-2.51		6.51	-2.39
dum_ire	-0.0353	-0.0753	dum_ire	0.0089	-0.0256	dum_den	-0.5318	-0.6208	dum_den	-0.1237	-0.1480
	-0.63	-2.01		0.24	-1.22		-2.96	-4.66		-2.99	-4.75
dum_lux	-0.2111	-0.1156	dum_lux	-0.0822	-0.0237	dum_ire	-0.2406	-0.4164	dum_ire	-0.0455	-0.0925
	-4.77	-2.92		-2.79	-1.14		-1.07	-2.8		-0.88	-2.67
dum_den	-0.1249	-0.1457	dum_den	-0.0316	-0.0341	dum_lux	-0.8493	-0.4059	dum_lux	-0.1978	-0.0928
	-2.80	-4.34		-1.07	-1.76		-4.78	-2.62		-4.83	-2.56
dum_ukd	---	0.0069	dum_ukd	---	---	dum_ukd	---	-0.0040	dum_ukd	---	-0.0071
		0.59						-0.09			-0.64
no. obs.	431	259	no. obs.	431	192	no. obs.	487	292	no. obs.	487	292
Prob F	0.0000	0.0000	Prob Chi^2	0.0000	0.0000	Prob Chi^2	0.0000	0.0000	Prob Chi^2	0.0000	0.0000

Note: GDP and UEWP have been instrumented by their first lag. T-values are given in the IV-estimates, z-values for the dynamic model.