



ELSEVIER

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Regional Science and Urban Economics 35 (2005) 597–624

regional
SCIENCE
& urban
ECONOMICS

www.elsevier.com/locate/econbase

An account of geographic concentration patterns in Europe

Marius Brühlhart^{a,*}, Rolf Traeger^{b,1}

^a*University of Lausanne, Switzerland, and CEPR*

^b*United Nations Economic Commission for Europe, Geneva, Switzerland*

Received 12 July 2004

Available online 30 November 2004

Abstract

We use entropy indices to describe sectoral location patterns across Western European regions over the 1975–2000 period. Entropy measures are decomposable, and they lend themselves to statistical inference via associated bootstrap tests. We find that the geographic concentration of aggregate employment, as well as of most market services, has not changed statistically significantly over our sample period. Manufacturing, however, has become significantly more concentrated relative to the distribution of aggregate employment (increased “relative concentration”), while becoming significantly less concentrated relative to physical space (decreased “topographic concentration”). The contribution of manufacturing to the topographic concentration of aggregate employment has fallen from 26% to 13% over our sample period.

© 2004 Elsevier B.V. All rights reserved.

JEL classification: R12; R14; F15

Keywords: Geographic concentration; EU regions; Entropy indices; Bootstrap inference

* Corresponding author. Département d'économétrie et économie politique, Ecole des HEC, Université de Lausanne, CH-1015 Lausanne, Switzerland. Tel.: +41 21 692 3471.

E-mail address: Marius.Brulhart@hec.unil.ch (M. Brühlhart).

¹ Any opinions expressed in this paper are those of the authors and do not necessarily reflect those of UNECE or its member countries.

1. Introduction

Empirical research in spatial economics is flourishing. The recent theoretical advances of the “new economic geography” and the ongoing erosion of distance- and border-related transaction costs give rise to a demand for both stylised facts and rigorous hypothesis tests on the location of economic activity. This demand is particularly strong in Western Europe, where the spatial concentration forces that characterise the recent location models are perceived by some as a looming threat. Numerous researchers have therefore examined the data in a quest for robust evidence on geographic concentration patterns in Europe.²

It has proven difficult to distil strong stylised facts from this research. Sectoral relocation in Europe is a slow and multifaceted process that does not leap out from the data. Overman et al. (2003) summarise the available evidence as follows: “In contrast to the US, EU countries are becoming increasingly specialised (. . .), although the changes are not particularly large.” This diagnosis of a slowly more concentrated European industrial geography is supported by the majority of analyses, but there are numerous exceptions.³ Furthermore, most of the available evidence relates to the distribution across *countries* of *manufacturing* sectors. Yet, rather little is still known about geographic concentration of sectors at sub-national level and across the full range of economic activities. In addition, existing studies use a number of different concentration measures that are chosen largely for their intuitive simplicity; they are based on data with varying coverage and disaggregation; and they do not attempt to gauge the statistical significance of observed patterns.

The aim of this paper is therefore to provide a comprehensive and methodologically rigorous account of sectoral concentration patterns across Western European regions, in a quest for empirically well-founded stylised facts. Our study distinguishes itself from the existing literature in four principal respects.

First, we apply entropy indices to measure geographic concentration. These indices have distinct advantages over the conventional measures in this literature. One advantage lies in their suitability to inequality decomposition analysis. This allows us to compare within-country concentration to between-country concentration in conceptually rigorous fashion. In addition, we can quantify how much each sector contributes to the geographic concentration of aggregate activity, by decomposing aggregate concentration into component “sector contributions”.

Second, we employ bootstrap inference to test the statistical significance of changes in observed concentration measures. These tests have been shown to be particularly accurate when used in conjunction with entropy measures.

² For studies of geographic concentration patterns in Europe using sectoral output or employment data, see Aiginger and Davies (2001), Aiginger and Leitner (2002), Aiginger and Pfaffermayr (2004), Amiti (1999), Barrios and Strobl (2004), Brühlhart (2001a,b), Clark and van Wincoop (2001), Haaland et al. (1999), Hallet (2000), Helg et al. (1995), Imbs and Wacziarg (2003), Kalemli-Ozcan et al. (2003), Krugman (1991), Midelfart Knarvik et al. (2002), Paci and Usai (2000), Peri (1998), and Storper et al. (2002). Combes and Overman (2004) provide a comprehensive survey.

³ Decreasing trends in sectoral specialisation of countries and/or geographic concentration of sectors have also been found by Aiginger and Davies (2001), Aiginger and Leitner (2002), Aiginger and Pfaffermayr (2004), Hallet (2000), Midelfart Knarvik et al. (2002), Paci and Usai (2000), and Peri (1998).

Third, we address aggregation biases that arise in regional data and are often overlooked. Consideration of this issue leads us to compute separate indices for “relative concentration”, where we measure the degree to which sectors are concentrated relative to the geographic distribution of aggregate activity, and for “topographic concentration”, where we measure the degree to which sectors are concentrated relative to physical space. Our results show that this conceptual distinction has substantial empirical relevance.

Fourth, our study is based on comprehensive regionally and sectorally disaggregated data. Our main data set provides us with a balanced panel of employment in eight economic sectors in 236 NUTS-2 and NUTS-3 regions belonging to 17 Western European countries over the 1975–2000 period.⁴ The eight sectors of this data set cover the full range of economic activities, including agriculture and services. Through the use of employment as the size measure we can avoid problems of currency conversion inherent in value data. As a complement to the main data set, we use a second data set that disaggregates manufacturing value added into nine industries for 116 EU-15 NUTS-1 and NUTS-2 regions over the 1980–1995 period.

Some existing studies use similar methodologies to ours, but none covers all four elements. Indeed, while entropy indices are common in the income distribution literature (see e.g. Cowell, 2000), their use in spatial contexts has remained relatively rare. A number of researchers have used entropy measures and their decompositions to describe the spatial inequality of *aggregate income* in Europe (e.g. De la Fuente and Vives, 1995; Duro and Esteban, 1998; Duro, 2001; Combes and Overman, 2004). The application of entropy measures to *sectoral* data for Europe has been pioneered by Aiginger and Davies (2001) and Aiginger and Pfaffermayr (2004). These studies are based on *country*-level data, they do not exploit the indices’ decomposability, and they perform no statistical inference. Finally, the paper by Mori et al. (2004) resembles ours in some key respects: topographic concentration patterns are described using an entropy index and exploiting its decomposability. Mori et al. (2004) do not, however, explicitly address regional aggregation biases, and they use a method for statistical inference that requires strong assumptions.⁵

Our main results are as follows. We find that the topographic concentration of *aggregate* employment has not changed significantly over our sample period. The concentration of European *manufacturing*, however, has indeed changed statistically significantly: manufacturing has become more geographically concentrated relative to the spatial spread of total employment (increased relative concentration), but it has become less geographically concentrated relative to physical space (decreased topographic concentration). This likely explains the differences in diagnoses of European concentration trends cited above. In addition, the topographic spread of

⁴ Nomenclature of territorial units for statistics (NUTS) is Eurostat’s classification of sub-national spatial units, where NUTS-0 corresponds to the country level and increasing numbers indicate increasing levels of sub-national disaggregation.

⁵ Furthermore, Mori et al. (2004) base their study on data for *Japanese* regions. We discuss their proposed methodology in Section 2.3.

manufacturing in part explains our observation that the contribution of this sector to the topographic concentration of aggregate employment has fallen from 26% to 13% over our sample period. As to *services*, we detect a significant decrease in concentration, both in relative and topographic terms, for the transport and telecommunications sector. The geographic concentration of the remaining market service sectors (financial services, distribution, and other services), however, has not changed significantly over our sample period.

Our paper is organised as follows: Section 2 provides a detailed presentation of the entropy measures we use, their associated bootstrap tests, and our data resources. In Sections 3 and 4, we describe Western European geographic concentration patterns using the entropy measures and their decompositions. Specifically, Section 3 describes relative concentration while Section 4 describes topographic concentration. Section 5 concludes with a brief summary and some discussion of the results.

2. Measurement, inference and data

Following [Krugman \(1991\)](#), “locational Gini indices” have become the measure of choice for studies of geographic specialisation patterns.⁶ The Gini index has strong intuitive appeal, but it is not ideally suited to our analysis. One feature that we seek in a measure of geographic concentration is decomposability into its within-country and between-country components. The Gini index is only decomposable if the range of the values taken by the variable of interest does not overlap across subgroups of individual observations ([Cowell, 1980](#)). This is evidently not the case in our context: regions in different countries may well have similar degrees of specialisation in a particular sector. Another desirable characteristic of any retained measure would be its suitability for statistical inference.

It turns out that measures belonging to the single-parameter generalised entropy class perform particularly well on both those counts: they are additively decomposable both by population subgroup and by sectors, and they lend themselves particularly well to bootstrap-based statistical inference.⁷

⁶ Other popular concentration measures include the Herfindahl index, x -region concentration ratios and [Krugman's \(1991\)](#) bilateral similarity index. In terms of our discussion below, they share the limitations of the Gini index.

⁷ [Combes and Overman \(2004\)](#), based on [Duranton and Overman \(2002\)](#), list seven criteria for a “good” measure of geographic concentration. The measures we employ offer a methodological improvement over standard measures in terms of their second criterion (spatial decomposability) and of their fourth criterion (amenability to significance testing). To meet their remaining five criteria, measures would need to be based on plant-level data. Note that, where applicable, we have computed Gini indices as well as entropy measures. The choice of index did not affect our qualitative findings, and we therefore report only the entropy-based results. All results are available on request.

2.1. Entropy indices

Consider a population of spatial *basic units* $i \in \{1, 2, \dots, N\}$, where each basic unit is associated with a unique value of the measured variable y_{si} , representing economic activity (measured in terms e.g. of employment or value added) in a particular sector $s \in \{1, 2, \dots, S\}$. A basic unit is defined as a square kilometre of land area.⁸ Country or region boundaries partition this population exhaustively into non-overlapping *subgroups* of basic units $k \in \{1, 2, \dots, K\}$.

Members of the generalised entropy (GE) class of inequality indices are defined by the following expression:

$$GE(\alpha)_s = \frac{1}{\alpha^2 - \alpha} \left[\frac{1}{N} \sum_{i=1}^N \left(\frac{y_{si}}{\bar{y}_s} \right)^\alpha - 1 \right] \tag{1}$$

where

$$\bar{y}_s = \frac{1}{N} \sum_{i=1}^N y_{si} = \frac{Y_s}{N},$$

Y_s is activity in sector s summed across all N basic units, and α is a sensitivity parameter. α measures the weight given to distances among values taken by y_{si} at different parts of the distribution over i . It can in principle be set to any real number. The neutral parameter value is 1. If $\alpha < 1$, then a bigger weight is attributed to the dispersion of y_{si} in the lower tail of the distribution of y_{si} over i , and if $\alpha > 1$, then a bigger weight is attributed to the dispersion in the upper tail. Like the Gini, these indices increase in the degree of inequality.

Following standard practice, we confine our analysis to the cases where $\alpha=1$ and $\alpha=2$. Using L'Hopital's rule on Eq. (1), the first case yields the *Theil index* of inequality:

$$GE(1)_s = \frac{1}{N} \sum_{i=1}^N \frac{y_{si}}{\bar{y}_s} \log \frac{y_{si}}{\bar{y}_s}, \tag{2}$$

where

$$0 \leq GE(1)_s \leq \log N.$$

The second case yields *half the squared coefficient of variation*, CV:

$$GE(2)_s = \frac{1}{2} CV_s^2, \tag{3}$$

⁸ For our subsequent analysis to be strictly valid (i.e. unbiased), we would need to define basic units as infinitesimally small areas. We take a square kilometre as a heuristic approximation of an infinitesimally small area. Note also that we assume throughout that economic activity is distributed continuously in space, i.e. we abstract from the geographic “lumpiness” of production on discrete sites. Given the relatively large spatial scale and low level of sectoral disaggregation we work on, this assumption is unlikely to affect our results.

where

$$CV_s = \frac{1}{\bar{y}_s} \left[\frac{1}{N} \sum_{i=1}^N (y_{si} - \bar{y}_s)^2 \right]^{\frac{1}{2}},$$

and

$$0 \leq GE(2)_s \leq \frac{1}{2}(N - 1).$$

A simple illustration of the behaviour of these measures is given in Appendix A.

2.2. Decompositions

Entropy indices describe the concentration of the distribution of y_{si} over i through a single number. It can be interesting to decompose such summary concentration measures. In the geographic context, the most obvious decompositions of interest are (a) to separate within-country from between-country concentration, and (b) to identify the contributions of individual sectors to the geographic concentration of aggregate activity. Entropy indices are ideally suited for such exercises.

These indices are *decomposable by population subgroups* in particularly appealing fashion. Each GE index can be decomposed additively as:

$$GE(\alpha)_s = GEw(\alpha)_s + GEb(\alpha)_s, \quad (4)$$

where *GEw* and *GEb* stand for within-subgroups and between-subgroups general entropy, respectively. Subgroups can stand for countries, regions, or groupings thereof, but in this paper we think of them in terms of countries. Between-country inequality, *GEb*, is computed by applying Eq. (1) to the K country means \bar{y}_{sk} instead of the N observations on y_{si} . The contribution of within-country inequality is computed as follows:

$$GEw(\alpha)_s = \sum_{k=1}^K \left(\frac{n_k}{N} \right)^{1-\alpha} \left(\frac{Y_{sk}}{Y_s} \right)^\alpha GE(\alpha)_{sk}, \quad (5)$$

where $GE(\alpha)_{sk}$ is the GE index as defined by Eq. (1) but confined to observations belonging to country k (so that N becomes n_k). Country GE indices are therefore calculated as if each country were a separate population. It is evident from Eq. (5) that the *GE(1)* index weights individual country indices by countries' y shares. The *GE(2)* index decomposition implies weights that are based on the n_k shares as well as the Y_{sk} shares. For decompositions by population subgroups, *GE(1)* is generally preferred to *GE(2)*, because for *GE(2)* the weights used to compute *GEw* are not independent from *GEb*.⁹

⁹ Bourguignon (1979) and Shorrocks (1980) have proven that *GE(0)* and *GE(1)* are the only additively decomposable scale invariant inequality measures for which the weights of the within-subgroup inequalities sum to a constant (i.e. 1) and are independent of *GEb*. Shorrocks (1984) showed that even if one relaxes the *additively decomposable* constraint by allowing weaker aggregation properties, the admissible set of indices expands only to monotonic transformations of the *GE(α)* family.

For a *decomposition of overall concentration by sectors*, we seek a rule according to which we can express a measure of the geographic concentration of aggregate activity Y_i , which we denote C , as the sum of the contributions from all sectors, so that sector s provides a disequalising contribution if $\Phi_s > 0$, and an equalising contribution if $\Phi_s < 0$:

$$C = \sum_{s=1}^S \Phi_s(C).$$

Functions that generate suitable values of sector contributions Φ_s are referred to as “decomposition rules”. The adoption of such a rule is necessary to apportion concentration contributions exhaustively and uniquely to individual sectors when the locational patterns of sectors are correlated. In general, there is an infinite possible number of such rules, these rules have different properties depending on the precise index C chosen, and the choice is arbitrary. However, [Shorrocks \(1982\)](#) has proven that under some weak and plausible assumptions one arrives at the following unique decomposition rule for proportional sectoral contributions $\phi_s(C)$:

$$\phi_s(C) = \frac{\Phi_s(C)}{C} = \rho_s \frac{\sigma(\mathbf{y}_s)}{\sigma(\mathbf{y})} = \frac{\text{cov}(\mathbf{y}_s, \mathbf{y})}{\sigma^2(\mathbf{y})},$$

where $\mathbf{y}=(y_1, \dots, y_N)$ is the vector of aggregate activity across basic units, $\mathbf{y}_s=(y_{s1}, \dots, y_{sN})$ is the vector of sector s activity across basic units, σ denotes the standard deviation, and ρ_s is the correlation between \mathbf{y}_s and \mathbf{y} .¹⁰ This decomposition rule is especially appealing, since, as shown by [Shorrocks \(1982\)](#), it yields the same set of proportional sector contributions ϕ_s irrespective of the concentration index C that is chosen. In terms of the proportional sector contributions, the choice of concentration measure therefore becomes irrelevant. However, it is standard practice to resort in this context to the GE(2) index, for which the Shorrocks decomposition rule happens to be the “natural rule”, since:

$$\phi_s = \rho_s \frac{\bar{y}_s}{\bar{y}} \sqrt{\frac{\text{GE}(2)_s}{\text{GE}(2)}}. \tag{6}$$

Hence, a certain sector s ’s proportional contribution to the geographic concentration of aggregate activity is the product of (a) the correlation of \mathbf{y}_s with \mathbf{y} , (b) s ’s share in aggregate activity Y , and (c) the inequality in that sector relative to total inequality, measured using GE(2).¹¹

2.3. Spatial aggregation: topographic versus relative concentration

Our most disaggregated *observed* spatial units are NUTS-2 and NUTS-3 sub-national regions (see Appendix B). These regions should not be interpreted as the basic

¹⁰ ϕ_{s_s} , of course, corresponds to the slope coefficient from a regression of \mathbf{y}_s on \mathbf{y} .

¹¹ Φ_s can be interpreted in two different ways; (a) as the inequality that would be observed if sector s were the only source of geographic concentration, Φ_s^A , and (b) as the inverse of the amount by which aggregate geographic concentration would change if the spatial concentration of sector s were reduced to zero, Φ_s^B . [Shorrocks \(1982\)](#) has shown that, for the GE(2) index, $\Phi_s = 1/2(\Phi_s^A + \Phi_s^B)$, whereas for most other inequality indices there exists no such obvious connection between Φ_s and $\{\Phi_s^A, \Phi_s^B\}$.

units, because they differ significantly in terms of both geographic and economic size, and it is well known that spatial inequality measures are sensitive to the definition of regions. This is commonly referred to as the “modifiable areal unit problem” (MAUP), according to which the results of statistical analysis of data for spatial zones can be varied at will by changing the zonal boundaries (Arbia, 1989). The problem has two components; a problem of scale, involving the aggregation of smaller units into larger ones, and a problem of alternative allocations of component spatial units to zones (gerrymandering).¹²

To acknowledge that regions should not be treated as basic units in their own right still leaves open a number of possible alternatives. The main issue concerns how we weight economic activity in each basic unit. The choice of weights may seem innocuous, but in fact it implies fundamentally different underlying meanings of “geographic concentration”. Our results show that empirical results are highly sensitive to this choice.

When we measure economic activity per square kilometre *without weighting*, the no-concentration benchmark obtains where an activity is spread perfectly evenly over physical space. Conversely, any departure from such an even spatial spread will register as geographic concentration, irrespective of the spatial distribution of endowments or of other economic sectors. We refer to this conception of geographic concentration as *topographic concentration*.

Alternatively, we weight sectoral activity per square kilometre by the amount of aggregate economic activity on that square kilometre. In other words, we condition physical space by the distribution of aggregate activity. If, for example, we measure activity as employment, the no-concentration benchmark implies that the employed persons on that square kilometre allocate their working time across sectors exactly according to the proportions corresponding to those sectors’ use of employed labour across all locations. This is the concept of concentration that has been used (often implicitly) in most previous studies and that seems economically most relevant. We shall refer to this definition as *relative concentration*.

In a nutshell, given the spatially uneven distribution of aggregate employment, a sector that happens to be perfectly evenly spread in physical space would have zero topographic concentration but positive relative concentration. Conversely, a sector that is spread exactly proportionally to total employment would have zero relative concentration but positive topographic concentration. We use the expression “geographic concentration” as the general term that encompasses both the “topographic” and the “relative” definition.¹³

To formalise this issue, note that our observed regions $r \in \{1, 2, \dots, R\}$ are sets of basic units i . The size of each region is defined in terms of the number of basic units it contains,

¹² Measures that equate observed units with basic units have come to be labelled “absolute” concentration indices (Aiginger and Leitner, 2002; Aiginger and Pfaffermayr, 2004; Haaland et al., 1999). As pointed out by Combes and Overman (2004), the no-concentration benchmark implied by “absolute” concentration is that an industry has identical employment/output in all regions irrespective of those regions’ size, which is difficult to reconcile with any market-based location model.

¹³ Our *relative* versus *topographic* definition is equivalent to the distinction in spatial statistics between *heterogeneous* and *homogeneous* space (see e.g. Marcon and Puech, 2003).

$n_r \geq 1$, such that $\sum_r n_r = N$.¹⁴ We denote the observed region-sector specific activity variable as \tilde{Y}_{sr} . Depending on the type of geographic concentration we seek to measure, this observed variable corresponds to *unweighted* region totals of unobserved basic-unit realisations, $\tilde{Y}_{sr} = \sum_i y_{sir}$ (topographic concentration), or to *weighted* totals of those unobserved realisations, $\tilde{Y}_{sr} = \sum_i \frac{y_{sir}}{\sum_s y_{sir}}$ (relative concentration).

In this setting, the expressions for the two basic entropy indices become:

$$GE(1)_s = \sum_{r=1}^R \frac{n_r}{N} \frac{\bar{y}_{sr}}{\bar{y}_s} \log \frac{\bar{y}_{sr}}{\bar{y}_s}, \tag{7}$$

and

$$CV_s = \frac{1}{\bar{y}_s} \left[\sum_{r=1}^R \frac{n_r}{N} (\bar{y}_{sr} - \bar{y}_s)^2 \right]^{\frac{1}{2}}, \tag{8}$$

where

$$\bar{y}_{sr} = \frac{\tilde{Y}_{sr}}{n_r}, \text{ and } \bar{y}_s = \frac{\tilde{Y}_s}{N},$$

and where $\tilde{Y}_s = \sum_r \sum_i y_{sir}$ for topographic concentration and $\tilde{Y}_s = \sum_r \sum_i \frac{y_{sir}}{\sum_s y_{sir}}$ for relative concentration.¹⁵ Simple illustrations of the behaviour of these measures for both changes in \tilde{Y}_{sr} and changes in n_r are given in Appendix A.

A number of *potential biases* warrant discussion. First, these measures are true representations of actual geographic concentration only if geographic concentration among basic units inside regions is zero. If intra-regional concentration exists, which of course applies in reality, the weighted measures will underestimate total concentration. This downward bias in measured geographic concentration rises with the level of spatial aggregation. It is a manifestation of the scale-related MAUP. By size-weighting the GE indices in expressions (7) and (8), we minimise the downward bias given the data at hand, but we cannot eliminate it.¹⁶

For the second component of the MAUP, the arbitrariness inherent in administrative region borders, given a certain distribution of region sizes, there is no methodological

¹⁴ With countries as our subgroups k , we can write that $N > R > K$.

¹⁵ Note that, for topographic concentration, \bar{y}_{sr} corresponds to region r 's activity in sector s divided by that region's *area*; while, for relative concentration, \bar{y}_{sr} corresponds to region r 's activity in sector s divided by that region's *total activity* summed across all sectors. Equivalently, for topographic concentration, \bar{y}_s corresponds to the sum of all sample regions' activity in sector s divided by the sum of all those regions' areas; while for relative concentration, \bar{y}_s corresponds to the sum of all sample regions' activity in sector s divided by those regions' total activity summed across all sectors.

¹⁶ One approach used in the income inequality literature to deal with grouped data is to estimate a certain distribution function parametrically using maximum likelihood, and to calculate inequality indices over the estimated distribution. We do not follow this route for two reasons. First, we have no priors as to the functional form of such a distribution. Second, there is no clear case based on empirical work for favouring either our non-parametric approach or the parametric method (Slottje, 1990).

palliative. In addition, broad statistical definitions of *sectors* may also obscure economically relevant concentration patterns, if offsetting concentration structures of sub-sectors are blurred by the aggregation of those sub-sectors. Absolute levels of the indices, and decompositions thereof, must therefore be interpreted with caution. However, the focus of this study is on changes in geographic concentration patterns over time, and if biases due to the MAUP and to sectoral aggregation are stable intertemporally, their absolute magnitude will not distort our inference or our conclusions.¹⁷

Finally, Mori et al. (2004) highlight a potential bias in the second moment of relative concentration indices if the scaling variable n_r is defined as the sum of sectoral levels of activity. Specifically, the larger the share of a sector in aggregate activity, the lower is the possible upper bound of measured relative concentration of that sector. This is an additional reason for treating intersectoral comparisons with caution.

2.4. A bootstrap test for the significance of changes in geographic concentration

Any concentration index describes the dispersion of a distribution through a scalar, and it therefore has its own sampling distribution. Traditionally, inference on entropy measures has been based on asymptotic results obtained through the delta method. For a test of the equality of two distributions on the same units at different times, however, this method requires cumbersome covariance calculations to take account of the intertemporal dependencies in the data. Furthermore, the finite-sample properties of such tests are unknown.

One solution is to ignore data dependencies and assume instead that the spatial distributions to be compared are independent. Mori et al. (2004) have adopted this assumption in the geographic concentration context and developed a simple formula to compute confidence intervals around the Theil index. They are aware, however, of the limiting nature of the independence hypothesis, calling it “a convenient fiction”.¹⁸

In the income inequality literature, Biewen (2001) and Mills and Zandvakili (1997) have argued in favour of using bootstrap inference. With this approach, the sampling distribution of an inequality index is estimated by multiple random resampling with replacement from the data set at hand. Through the bootstrap one can account for dependencies in the data without having to estimate covariance matrices explicitly. Biewen (2001) proved that the bootstrap test for inequality changes over time is consistent for any inequality statistic that can be expressed in terms of population

¹⁷ Evidence on the co-location of firms at the micro-geographic level points to the importance of narrowly confined clusters. According to Duranton and Overman (2002), the relevant distance for geographical clusters of British manufacturing firms is mostly smaller than 50 km. In comparison, the radius of a circle with a surface corresponding to the average area of regions in our data set 1 (15,000 km²) is 69 km. The UK is among the more densely populated European countries, and the Duranton–Overman result may therefore provide a lower-bound estimate. Nonetheless, a rigorous study of the accuracy with which regional data reflect patterns and changes in these fundamental distributions would be very useful.

¹⁸ In making direct intersectoral comparisons, they also implicitly assumed away MAUP-related biases. Given their spatially very disaggregated data (Japanese counties, with an average economically relevant area of 37 km²), such an assumption may be defensible; but the same could not be said for our European regional data (the average area of our sample regions being 15,000 km²).

moments—which includes the GE class of indices but not the Gini index. This result is shown by Biewen to be valid also for grouped data (i.e. for observed units that are aggregates of basic units). Using Monte Carlo simulations, he demonstrated that this approach achieves a finite-sample coverage accuracy that is equivalent to that obtained through analytically derived (but asymptotic) tests. Mills and Zandvakili (1997) found that the bootstrap estimated standard errors were closer to the corresponding asymptotic estimates for the Theil index than for the Gini index, and they too therefore preferred the entropy measure.

The standard use of the bootstrap is as a method for making probabilistic statements about population parameters based on a data sample drawn randomly from that population. One interpretation of this test in our context would therefore be to consider our yearly sets of regional observations as random draws from the universe of (industrialised) world regions.

Alternatively, and more plausibly, one can consider the set of Western European regions as the population, and search for specifically Western European parameters. In this setting, bootstrap inference remains useful, considering that the data are measured with error, and that the measurement error is distributed stochastically across observations (assuming that measurement errors are distributed independently from y). The principal attraction of the bootstrap in this case is that it absolves us from making assumptions on the form of the measurement error distribution across observations.¹⁹

Finally, one might posit as a null hypothesis that the spatial configuration that would result from the profit-maximising choices of well informed firms is constant over time, but allow for informational imperfections and motivational idiosyncrasies among firms, which add a stochastic element. The bootstrap test then pits this null against the alternative hypothesis that the profit-maximising equilibrium spatial configuration is changing over time.

By treating all observations equally in the resampling process, the standard bootstrap method implies that the disturbances attached with each observation are *iid* draws from the population distribution of disturbances. This assumption is difficult to justify in the context of our study, as we have strong reason to believe that measurement errors are to a large extent country-specific (i.e. spatially autocorrelated). We therefore apply block-wise resampling, defining countries as blocks. For each replication, a sample is drawn randomly among K blocks of regions, where each block has sample size R_k . Since we have no priors on the distribution of disturbances across countries, we attach equal probability weights to those K sets of observations in the resampling procedure.²⁰ All bootstrap results are based on 10,000 replications.

¹⁹ An alternative strategy for inference on concentration indices in exhaustive samples of grouped data with measurement error is to assume certain distributions of those measurement errors and to simulate corresponding distributions for the concentration indices (Bourguignon and Morrisson, 2002). That approach requires strong assumptions on the distributional forms of measurement errors.

²⁰ We ran all tests also with region-level resampling. As expected, this yielded generally tighter confidence intervals, but the higher moments of the distributions underlying those intervals were not affected significantly. Note that our country-level resampling procedure is more appropriate if measurement error (rather than random firm-level idiosyncrasies) is the main source of stochastic variation in the data.

2.5. Data

We draw on two complementary data sets, both of which are described in detail in Appendix B. Data set 1, compiled by Cambridge Econometrics, provides a balanced panel of sectoral employment for 17 West European countries, the 15 EU member states (pre-2004) plus Norway and Switzerland (collectively referred to as WE17). Except for Luxembourg, all country data are disaggregated into NUTS-2 or NUTS-3 regions, giving a total of 236 region-level observations per sector and year. The number of regions within countries ranges from 2 (Ireland) to 37 (UK). Employment is reported annually for eight sectors, covering the full range of economic activities, over the period 1975–2000.

Fig. 1 illustrates the evolution over our sample period of the relative sizes of the eight sectors in data set 1. It emerges clearly that the WE17 economies have been marked in the last quarter century by pronounced growth in the relative sizes of the tertiary sector, at the expense of the primary and the secondary sectors. This fact alone provides strong motivation for studying geographic specialisation patterns not just for manufacturing industries, but across the full spectrum of economic activities. Since our principal aim is to provide a comprehensive description of sectoral concentration patterns, and not to test market-based location models, we include non-market services in our data sample, even though locational determinants in this sector are largely of a political nature.

Data set 2, compiled by Hallet (2000), reports gross value-added (GVA) of nine manufacturing sectors across the 15 EU member states (referred to as EU15). For eight countries, the data are disaggregated into either NUTS-1 or NUTS-2 regions, giving a total of 109 regions. The remaining seven countries appear in the data as single regions. Among the countries that are subdivided, the number of regions ranges from 5 (Portugal) to 23 (UK). The period covered is 1980–1995.

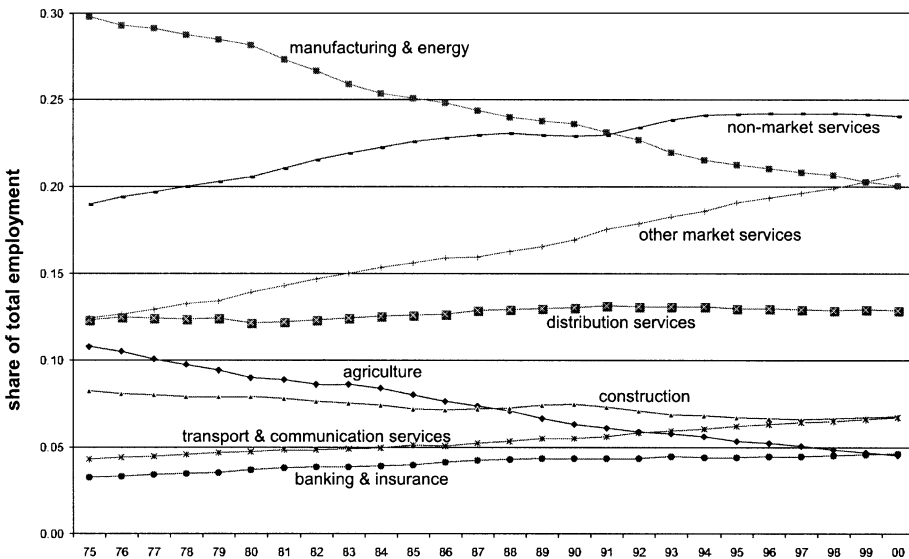


Fig. 1. Sector shares in total employment, 1975–2000.

The two data sets differ in terms of geographic and sectoral disaggregation, but they are complementary. The time span of the second is encompassed by that of the first. Moreover, data set 1 offers a broader base for comparison of agglomeration between and within countries, because it is more regionally disaggregated. We consider employment data as preferable to data based on production values, because the former are not subject to the problems associated with price conversions across countries and years. The comparative attraction of data set 2 is the detail it provides on manufacturing sectors, which facilitates comparisons with previous research findings by bringing us closer to the data sets that have been used in most existing studies.

3. Relative concentration

3.1. Relative concentration across all regions

3.1.1. All sectors

Sectoral Theil indices of relative concentration across the full spectrum of activities in WE17 regions (i.e. using data set 1) are reported in [Table 1](#) and [Fig. 2](#). These indices are computed according to Eq. (7), using total regional employment as the weighting variable n_r .

On average over our sample period, agriculture turns out to be by far the most concentrated sector (note the log scale of [Fig. 2](#)), and manufacturing is second-most concentrated, while construction is the most dispersed.

These results seem plausible. In view of the regional and sectoral aggregation problems, however, our analysis focuses not on levels but on changes over time. In [Table 1](#), we report changes in relative concentration (i) over our entire sample period 1975–2000, (ii) over the subperiod 1975–1987 and (iii) over the subperiod 1987–2000. The sample period is divided in this way since 1987 coincides with the entry into force of the Single European Act and thus the launch of the EU's Single Market programme. Hence, one might interpret the second subperiod as a time of particularly strong policy-

Table 1
Relative concentration of sectors 1975–2000^a (employment, 236 regions)

Sector	Avg GE(1) ^b	Δ GE(1) _{75–00}	Δ GE(1) _{75–87}	Δ GE(1) _{87–00}	Share ^c
Agriculture	0.474	0.029	0.008	0.021	0.07
Manufacturing, energy	0.055	0.020**	0.004	0.016**	0.24
Banking, insurance	0.053	0.004	–0.012	0.016	0.04
Non-market services	0.041	–0.023	–0.022	–0.001	0.22
Transport, communication	0.036	–0.043**	–0.036**	–0.007*	0.05
Distribution	0.031	0.007	0.002	0.004	0.13
Other market services	0.030	–0.005	–0.008	0.003	0.16
Construction	0.025	0.019	–0.012*	0.031**	0.07

^a **/* denotes rejection of H0 that Δ GE(1)=0, based on bootstrap 95%/90% confidence intervals (10,000 replications).

^b Average annual GE(1) index (employment weighted) over 1975–2000 period.

^c Sector share in total employment over the full sample period.

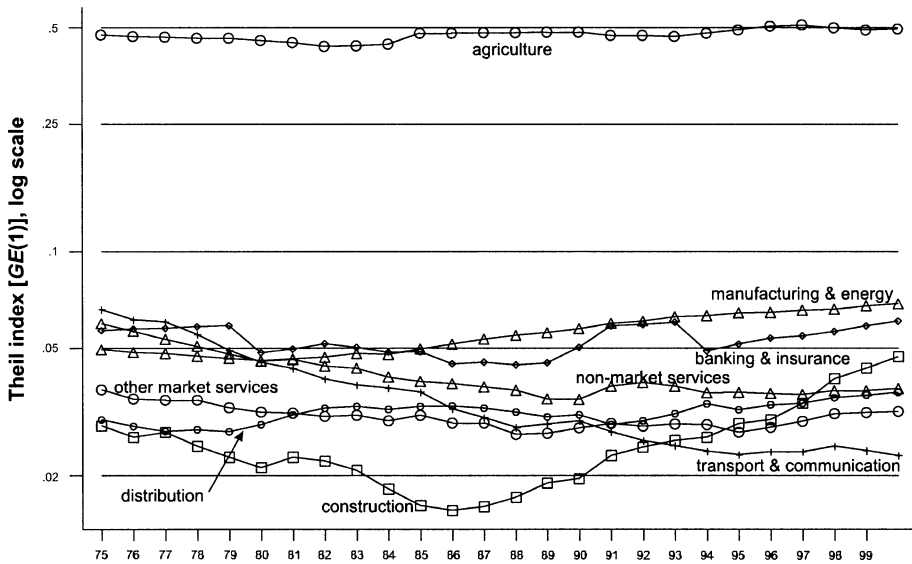


Fig. 2. Relative concentration of sectors (Theil index, employment), 1975–2000.

led integration. Table 1 also reports statistical significance levels according to the bootstrap test described above.

We find that manufacturing is the only sector that has seen a monotonic and statistically significant increase in relative concentration. This increase was more pronounced in the post-Single Market subperiod than in the earlier subperiod. Our analysis therefore confirms that European manufacturing is becoming more geographically concentrated.

Our results of Table 1 furthermore show that, with the exception of the “transport and communications” sector, which has become significantly more dispersed, no service sector exhibits a statistically significant change in relative concentration over the full sample period. On the whole, therefore, the evidence does not support the view of strong sectoral reallocation trends across the spectrum of economic activities. Looking at the subperiods, however, we find that the tendency to concentrate (disperse) geographically is stronger (weaker) in the second subperiod than in the first subperiod for all eight sectors. This finding is consistent with the view that the deepening of European integration through the Single Market programme has favoured geographic concentration forces.

3.1.2. Manufacturing

In Table 2, we report indices of relative concentration for disaggregated manufacturing sectors across EU15 regions, calculated from our data set 2. As noted above, these findings are not strictly comparable with those based on data set 1, due to differences of measurement units (value added instead of employment) and to narrower regional and time coverage.

Table 2

Relative concentration of manufacturing sectors, 1980–1995^a (gross value added, 116 regions)

Sector	Avg GE(1) ^b	Δ GE(1) _{80–95}	Δ GE(1) _{80–87}	Δ GE(1) _{87–95}	Share ^c
Ores, metals	0.389	–0.0555	–0.0551*	–0.0004	0.04
Textiles, clothing, footwear	0.379	0.1649**	0.0534**	0.1115**	0.08
Transport equipment	0.163	0.0196	0.0216	–0.0020	0.10
Chemicals	0.152	0.0003	0.0085	–0.0082	0.10
Non-metallic minerals	0.142	0.0171	0.0016	0.0156	0.06
Misc. manufactures	0.111	–0.0044	–0.0064	0.0020	0.09
Machinery, electronics	0.109	–0.0057	0.0180**	–0.0238**	0.31
Paper products	0.104	0.0098	–0.0022	0.0120	0.08
Food, drink, tobacco	0.082	0.0114	0.0026	0.0088	0.13
Total manuf.	0.043	0.0023	0.0071**	–0.0048	1.00

^a **/* denotes rejection of H0 that Δ GE(1)=0, based on bootstrap 95%/90% confidence intervals (10,000 replications).

^b Average annual GE(1) index (GVA weighted) over 1980–1995 period.

^c Sector share in total employment over the full sample period.

The results of the two data sets are consistent in so far as they both show a trend towards stronger relative concentration of total manufacturing for the first subperiod (although not for the second one).

The strongest increase in relative concentration is found for the textiles, clothing and footwear sector—a tendency which is particularly pronounced in the post-1987 subperiod but statistically significant throughout. This is in line with earlier findings whereby the strongest relocation tendencies in European manufacturing are in relatively low-tech and labour-intensive sectors (Brühlhart, 2001a).

We do not find a statistically significant change in the concentration index over the full 1980–1995 period for any other manufacturing sector. Six of the nine sectors display stronger concentration trends post-1987 than pre-1987. Here too, we can therefore retain as a stylised fact that EU industries exhibit weak overall concentration pressures, with some evidence of a strengthening subsequent to 1987.²¹

3.2. Relative concentration: between-country and within-country components

Exploiting the decomposability of entropy indices according to Eq. (4), we can track the evolution of the within-country and between-country components of geographic concentration.²²

²¹ According to the last row of Table 2, total manufacturing seems to have become more concentrated pre-1987 and more dispersed thereafter, which is not consistent with the concentration time profile found in the employment data. However, this turns out to be driven largely by “machinery, electrical and electronics”, the largest of our nine manufacturing industries, for which we find a significant initial increase and a significant subsequent decrease in concentration. Inspection of the data suggests the post-1987 decrease is primarily driven by a drop in reported value added of this sector in the West German regions. Given the estimated nature of the statistics for Germany in our data set 2, this result might be influenced by measurement problems (see Hallet, 2000).

²² In the context of relative concentration, a “factor decomposition” of total concentration is meaningless, since the concentration of total employment across regions weighted by total employment is zero.

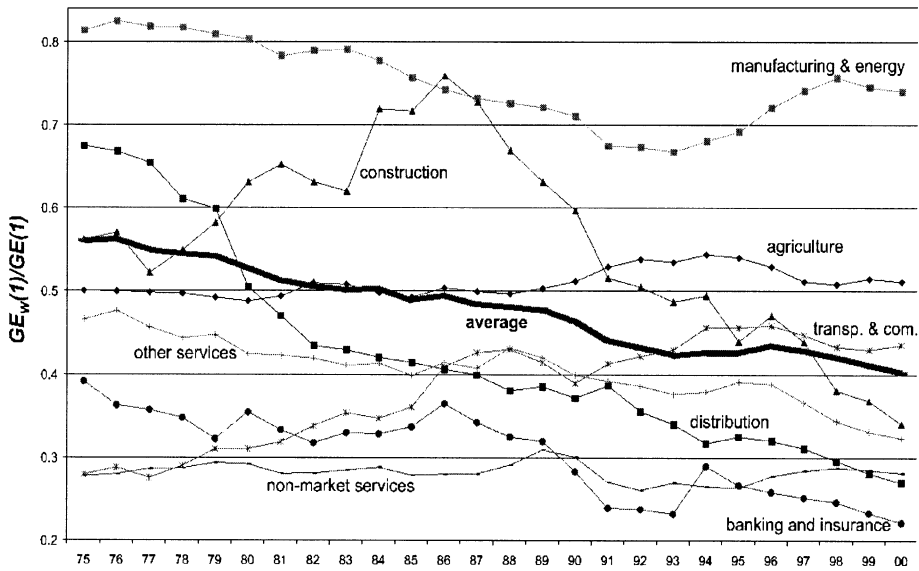


Fig. 3. Within-country share in overall relative concentration (employment), 1975–2000.

3.2.1. All sectors

Using data set 1, we have computed within-country shares of relative concentration ($GE_w(1)/GE(1)$) across all sectors. The results are reported in Fig. 3.

On average, most of the concentration of service sectors is between countries rather than within countries. The opposite applies to manufacturing: within-country concentration largely dominates between-country concentration. Whilst it would be tempting to interpret this finding (e.g. that manufacturing is still overly dispersed across national borders, as each country protects its industrial champions), the aggregation biases discussed in Section 2.2 probably loom especially large here and caution against such conjectures.

In terms of changes over time, we observe that the within-country share of relative concentration has fallen over our sample period for a majority of sectors. Hence, between-country concentration forces seem to have been relatively stronger than within-country concentration forces. Given that countries' internal markets were already liberalised in 1975, whereas our sample period was marked by strong between-country liberalisation, this result is in line with the view that European integration opens scope for between-country specialisation, which hitherto had existed only at the within-country level.²³

Unlike any other sector, relative concentration of manufacturing exhibits a trend break in the early 1990s towards a re-increase in the within-country share. It thus appears that, after a period of more pronounced inter-country concentration processes, intra-country

²³ The most evident feature of Fig. 3 is a strong non-monotonic evolution for the construction sector. It is impossible to explain this pattern with the available data, but we can point out that, given the highly dispersed nature of that sector, small variations in its distribution can produce seemingly large swings in the share measure reported here.

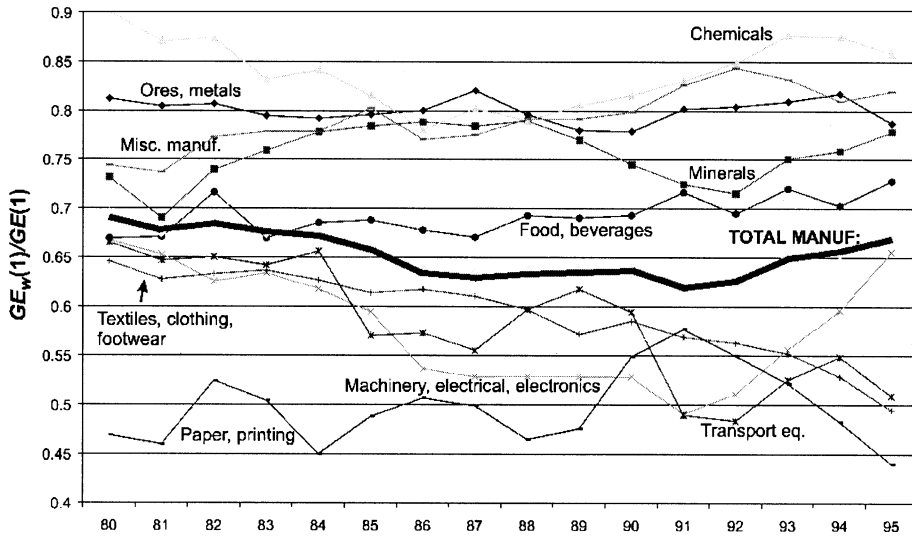


Fig. 4. Within-country share in overall relative concentration of manufacturing sectors (GVA), 1980–1995.

concentration forces have come to dominate relocation of manufacturing employment in the 1990s.

3.2.2. Manufacturing

Within-country shares of relative concentration for the manufacturing sectors, based on data set 2, are given in Fig. 4. In this data set too, the within-share of relative concentration of total manufacturing shows a u-shaped time profile—declining in the 1980s but increasing since the early 1990s.

Most sub-industries do not exhibit pronounced time patterns. The sector that emerges with the clearest trend is textiles, clothing and footwear, which exhibits a steady decline in the within-country share of geographic concentration, thus suggesting that between-country relocation has been particularly pronounced in this sector. This detected pattern complements that found in Table 2: the textile sector not only exhibited the most pronounced increase in concentration, but this concentration trend was effective mostly between, rather than within, countries.

4. Topographic concentration

4.1. Topographic concentration across all regions

As discussed in Section 2.2, the choice of spatial weights, which might appear at first an arcane technicality, turns out to be empirically important. Table 3 and Fig. 5 report indices of topographic concentration, computed for data set 1. The difference compared to the relative concentration indices is most evident for agriculture. Of our eight sample sectors, agriculture exhibits the highest average level of relative concentration but the

Table 3
Topographic concentration of sectors, 1975–2000 (employment, 236 regions)

Sector	Avg GE(1) ^a	$\Delta GE(1)_{75-00}^b$
Other market services	1.039	-0.016
Transport, communication	1.028	-0.148**
Banking, insurance	1.008	-0.024
Distribution	0.938	-0.052
Non-market services	0.890	-0.140*
Manufacturing, energy	0.868	-0.161**
Construction	0.738	0.008
Agriculture	0.490	0.104**
Total employment	0.810	-0.002

^a Average annual GE(1) index (area weighted), 1975–2000.

^b ***/* denotes rejection of H0 that $\Delta GE(1)=0$, based on bootstrap 95%/90% confidence intervals (10,000 replications).

lowest level of topographic concentration. In both cases the gap separating agriculture from the most similarly concentrated sector is large. These results are of course entirely consistent. While agriculture is spread out more than the other sectors in line with total land area, it is typically concentrated in regions with low employment densities, and hence it is concentrated strongly when we condition the spatial distribution of agricultural employment by the distribution of total employment. Another difference between topographic and relative concentration is that service sectors are by far the most concentrated ones in the former case, whereas in terms of relative concentration they are less concentrated than manufacturing as well as agriculture. The obvious interpretation,

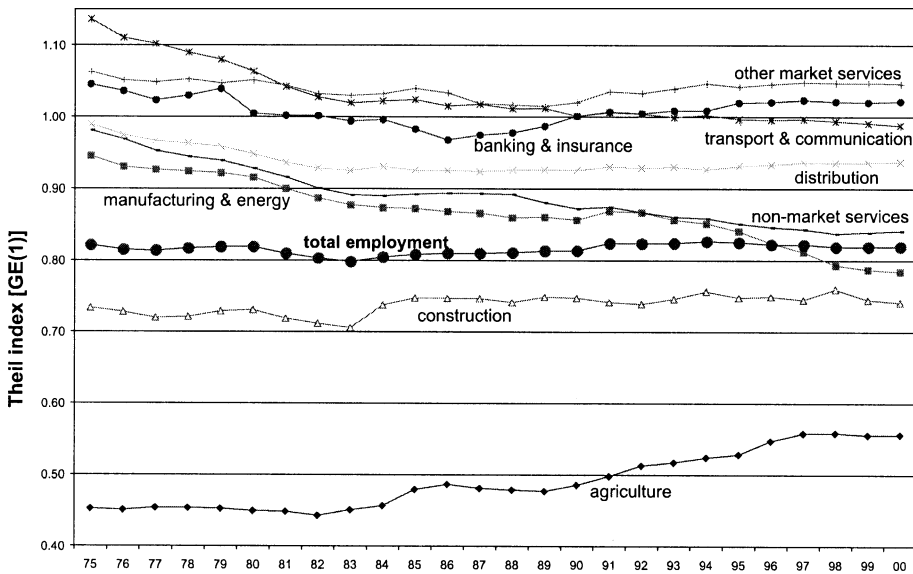


Fig. 5. Topographic concentration of sectors (Theil index, employment), 1975–2000.

aggregation biases notwithstanding, is that service jobs are concentrated in high-density (urban) areas, agriculture is concentrated in low-density (rural) areas, and manufacturing is located in-between.

Turning to the time profiles of our topographic concentration measures, Fig. 5 suggests that the topographic concentration of aggregate employment has remained stable over the sample period, and the bootstrap test does not reject the null hypothesis of identical concentration indices in the base and end periods. We do not, therefore, detect a systematic tendency for aggregate employment to concentrate or disperse spatially in Western Europe.

The evident stability in the topographic distribution of total employment, however, masks offsetting changes in the topographic concentration of individual sectors. The most pronounced trends are an increase in topographic concentration of agriculture and a simultaneous decrease in the concentration of manufacturing. These changes are statistically significant.

The decrease in topographic concentration of manufacturing, together with the detected increase in relative concentration, suggests that manufacturing jobs have moved from regions with high employment density towards regions with low employment density. This result may also provide the explanation for the apparent inconsistency in the literature that we mention in the Introduction. In fact, Overman et al. (2003) diagnose a rise in manufacturing concentration based on a survey of studies that predominantly use *relative* measures. The reverse result of Aiginger and Pfaffermayr (2004), in turn, is based on measures of *absolute* concentration (that is, they define countries as basic units, without weighting). Absolute concentration measures resemble topographic concentration measures in so far as, compared to relative concentration measures, both will attach bigger weights to countries/regions with comparatively low economic density and smaller weights to countries/regions with comparatively high economic density. Hence, some apparently contradictory results in the literature may simply be due to the different spatial weights that are used (implicitly, in many cases) to compute the concentration indices.

4.2. Topographic concentration: decompositions

4.2.1. Between-country and within-country components

The decomposition of aggregate topographic concentration into its within-country and between-country components for each of the eight sectors is reported in Fig. 6. On average, service sectors have the highest share of within-country concentration, again as opposed to the patterns observed for relative concentration. Nevertheless, the two types of measures share a trend: as in the case of relative concentration, we detect a falling tendency of the within-country share for a majority of sectors. The 1990s, however, are characterised by an apparent reversal in this tendency, that is by an increase in the within-country share of topographic concentration. That reversal is again most manifestly evident for the manufacturing sector. This suggests that, after a period of dominant inter-country reallocations of manufacturing employment, the 1990s have been dominated by intra-country geographic shifts in manufacturing employment.

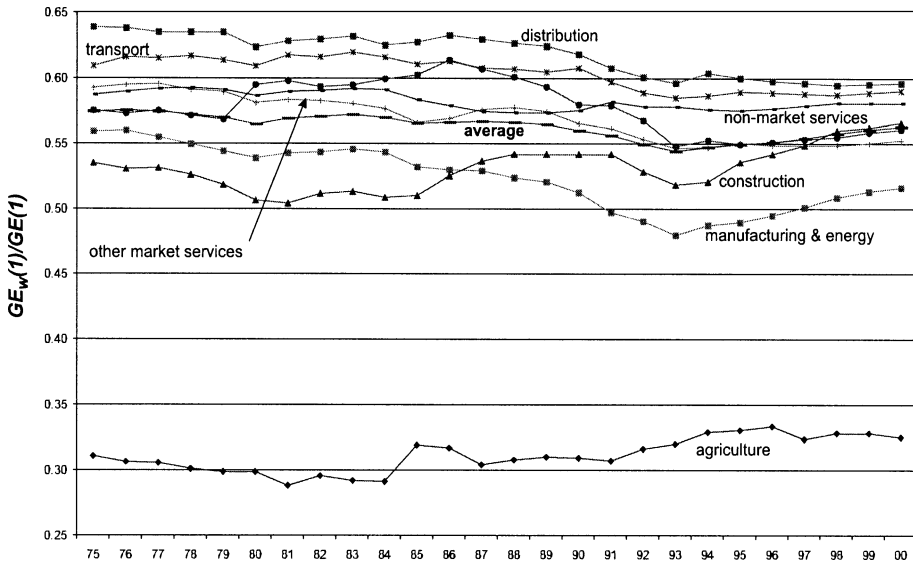


Fig. 6. Within-country share in overall topographic concentration (employment), 1975–2000.

4.2.2. Sector decomposition

In Fig. 7, we report proportional sector contributions (ϕ_s) based on a decomposition of the topographic concentration of total employment using the GE(2) index (Eq. (8)) and the decomposition rule of Eq. (6).

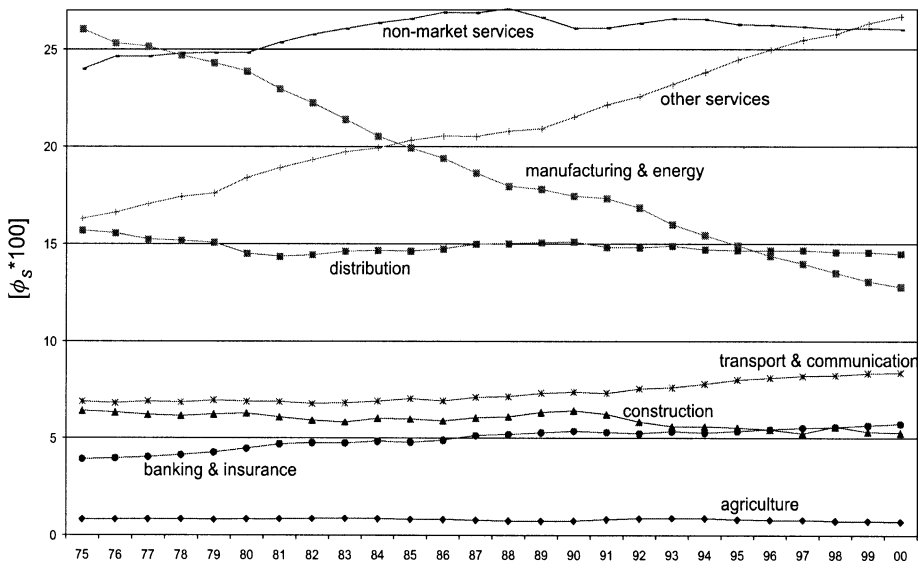


Fig. 7. Sectoral “factor contributions” to topographic concentration (employment, GE(2) index), 1975–2000.

These decompositions make it evident that total topographic concentration is determined mainly by the concentration of tertiary activities. This is primarily a consequence of the growing share of services in aggregate employment (Fig. 1). The sector-decomposition analysis also shows that, of our eight individual sectors, non-market services on average account for the largest share of total topographic concentration. Hence, public-sector employment appears as the biggest contributor to the uneven geographical spread of economic activity.

In contrast, the manufacturing sector accounted for a continuously decreasing contribution to the topographic concentration of total employment: while concentration of the manufacturing sector accounted of over 26% of aggregate topographic concentration in 1975, its share had shrunk to 13% by the year 2000. This result is consistent with the declining share of manufacturing in total employment (Fig. 1) and its decreasing topographic concentration (Fig. 5)—two factors that correspond to the second and third term, respectively, in the “natural” decomposition rule expressed by Eq. (6).

Finally, agriculture accounts for the lowest and largely constant share of total topographic concentration. Agriculture’s declining share of total employment (Fig. 1) was largely offset by an increase in its level of topographic concentration (Fig. 5).

5. Conclusions and conjectures

We provide an account of geographic concentration patterns in a broad range of sectors across Western European regions and countries from 1975 to 2000. Geographic concentration is quantified using entropy indices. These indices present two major advantages: they are decomposable, and they lend themselves to statistical inference through bootstrap tests. We distinguish between “relative” concentration, where location patterns are expressed relative to the spatial distribution of aggregate economic activity, and “topographic” concentration, where location patterns are expressed relative to physical space.

We find that the topographic concentration of aggregate employment has not changed significantly over our sample period. This stability of the geographic concentration of overall activity masks some distinct evolutions at the sectoral level.

Our study describes a European *manufacturing* sector that is slowly becoming more geographically concentrated relative to the spatial spread of total employment. Relative to physical space, however, manufacturing concentration has been decreasing. We find that both these processes are statistically significant. Due to the decrease in the topographic concentration of manufacturing and to the reduction in the share of manufacturing jobs in total employment, the contribution of the manufacturing sector to the topographic concentration of aggregate employment has fallen from 26% to 13% over our sample period. Among manufacturing sub-industries, the most pronounced increase in relative concentration is observed in the textiles, clothing and footwear sector. Finally, the evolution of the within-country share in the topographic concentration of total manufacturing is non-monotonic, with a decrease in the 1970s and 1980s and an increase in the 1990s.

Service sectors generally appear more concentrated than manufacturing and agriculture in topographic terms. A significant decrease in concentration, both in relative and topographic

terms, is observed for the transport and telecommunications sector. The geographic concentration of the remaining market service sectors (financial services, distribution, and other services), however, has not changed significantly over our sample period.

The main aims of this paper were to propose versatile measures for the description of geographic concentration patterns, and to provide a characterisation of locational trends in Western Europe. We believe that a rigorous and detailed description of changing concentration patterns is of interest in itself. Yet, conjectures of relevance to related studies are possible. For example, [Ciccone \(2002\)](#), drawing on a methodology developed in [Ciccone and Hall \(1996\)](#), has estimated the extent to which the topographic concentration of total employment (i.e. regional employment density) increases regional labour productivity, which he called “agglomeration effects”. While the estimation approach cannot identify the nature of spatial externalities underlying these estimated effects, it is based on careful instrumenting of topographic concentration so as to establish causality that runs from concentration to productivity. Using regional cross-section data sets for the five largest West European countries, [Ciccone \(2002\)](#) found that productivity rises in topographic concentration, with a remarkably robust and precisely estimated elasticity of around 4.5%. Our analysis shows no significant change in the topographic concentration of total employment over time, but statistically significant changes in the topographic concentration of individual sectors ([Table 3](#)). Hence, it would be revealing to exploit the complementarity between approaches by extending Ciccone’s study to changes in *sectoral* agglomeration effects over time, and to consider relative as well as topographic concentration measures.

Additional extensions to our work are not difficult to conceive. For example, it would be interesting to describe evolutions of the full distribution of sectoral location patterns including transitions over time of region-sector observations inside those distributions, and to compute measures of spatial separation so as to assess the contiguity of sectoral clusters. The biggest constraint on the quality of research on location patterns in Europe, however, is the quality of available sub-national data. Our analysis cannot entirely escape the spatial and sectoral aggregation biases inherent in conventional regional statistics, even though we do our best to minimise their distorting impact. If it were possible to merge plant-level micro-geographic data sets that have been collected in several European countries, ideally encompassing services as well as manufacturing establishments, the description of the European economic geography could take a quantum leap in terms of accuracy, comparability and potential for theory-based inference.

Acknowledgements

We thank Philip Lane, two anonymous referees, and seminar participants at ETSG and ERSA annual conferences, at the CEPR “Topics in Economic Geography” workshop in London, at the Hamburg Institute of International Economics (HWWA), at the Vienna Institute for International Economic Studies (WIIW), and at the Universities of Dijon, Dublin (Trinity College), Madrid (Complutense), Milan (Bocconi), Nottingham and Utrecht for helpful comments. Martin Hallet has generously allowed us to use his data set. Financial support from the Swiss National Science Foundation (grant 612-65970) and

from the UK Economic and Social Research Council (ESRC grant L138251002) is gratefully acknowledged.

Appendix A. Illustrations of geographic concentration indices

We provide some examples of the changes in our indices for two simple scenarios of changing geographic concentration patterns. In both scenarios, we assume a universe of two observed units (i.e. regions), and we do not consider concentration patterns inside of those observed units. In scenario I, we assume that the two regions are identical in every respect bar their shares of Y (i.e. activity in the sector of interest). One can therefore abstract in this example from weighting issues, and treat the observed units as if they were basic units. We track the values of our measures as activity in the sector of interest changes from being fully concentrated in one region to being perfectly dispersed across the two regions. In scenario II, we assume that activity in the sector of interest remains equally split between the two regions, and we vary the underlying sizes of those regions instead. We can thus no longer treat regions, the observed units, as if they were the basic units. We track the values of our measures as the sizes of two regions move from being very unequal to being perfectly identical.

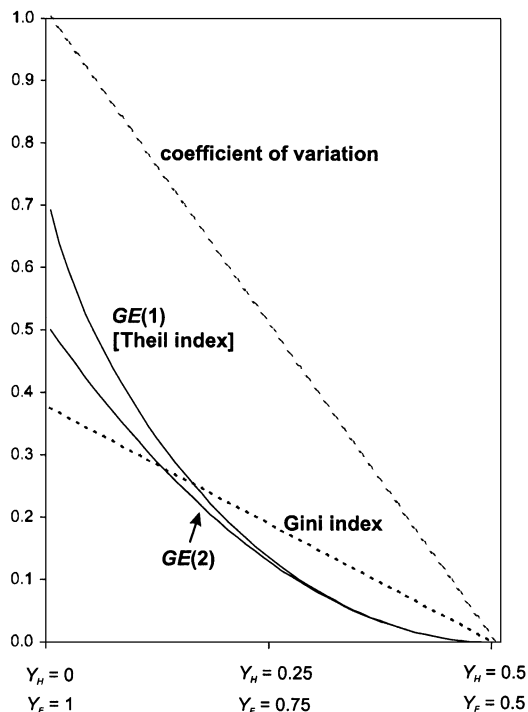


Fig. A1. Sectoral relocation between two regions (Scenario 1).

The two scenarios thus illustrate the two possibilities of changing geographic concentration of an individual sector: relocation of the sector of interest, or changes in the region sizes with unchanged location of the particular sector. As pointed out by Mori et al. (2004), the two types of changes are not necessarily independent. If we compute measures of relative concentration of a large sector, then the geographic concentration of that sector will affect relative region sizes. We abstract from this issue here and assume the two components to be independent (which strictly applies to the case of topographic concentration and of relative concentration of infinitesimally small sectors).

Both our scenarios simulate a reduction in geographic concentration. The graphs show that our indices always fall in geographic dispersion, and that all indices are monotonic but nonlinear transformations of each other.

Scenario 1: Suppose two identical regions, H and F . The world size of the sector, Y , is assumed constant and equal to 1, but its distribution across H and F is allowed to change. Moving from left to right in Fig. A1, we start from a situation where all of that sector’s activity is concentrated in region F , so that $Y_H=0$, and then gradually move activity out of region F and into region H , until $Y_H=Y_F=0.5$.

Scenario 2: Suppose the two regions can have different sizes, n_H and n_F , but that $Y_H=Y_F=0.5$ throughout. The size of the world is set to 100 ($N=n_H+n_F=100$). Moving from left to right in Fig. A2, we start from a situation with very unequally sized regions, where $n_H=1$ and $n_F=99$, and then gradually equalise region sizes, until $n_H=n_F=50$.

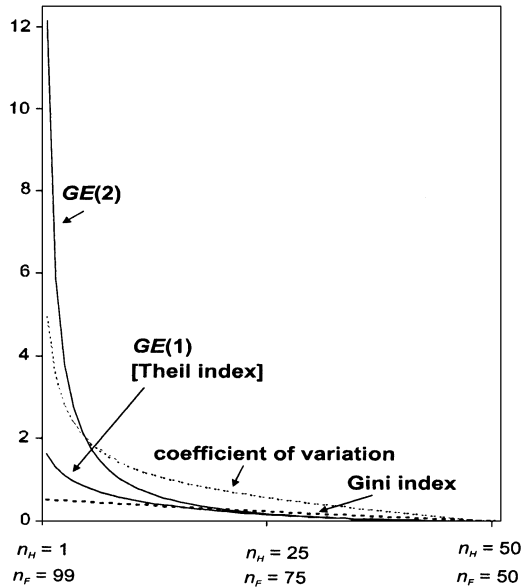


Fig. A2. Unchanged sectoral location in two changing regions (Scenario 2).

Appendix B. Data

B.1. Data set 1

- Source: Cambridge Econometrics Regional Database (based on Eurostat's REGIO and national sources).
- Variable: employment
- Time dimension: annual averages, 1975–2000.
- Sectors: agriculture; manufacturing and energy; construction; distribution; transport and communications; banking and insurance; other market services; non-market services (eight sectors, based on NACE-CLIO classification).
- Regional breakdown: 236 regions, see [Table A1](#).
- Number of observations: 49,088.

Table A1
Regional breakdown of data sets 1 and 2

Country	Number of regions for which data are available	Administrative units	Classification level	Observations
<i>Data set 1</i>				
Belgium	10	Provinces	NUTS 2	Vlaams Brabant and Brabant Wallon clustered as one region
Denmark	3	Regions	TL 2	
Germany	31	Regierungsbezirke	NUTS 2	Neue Länder excluded
Greece	13	Development regions	NUTS 2	
Spain	18	Comunidades autónomas+ Ceuta y Melilla	NUTS 2	
France	22	Régions	NUTS 2	DOMs excluded
Ireland	2	Regions	NUTS 2	
Italy	20	Regioni	NUTS 2	
Luxembourg	1			
Netherlands	12	Provincies	NUTS 2	
Austria	9	Bundesländer	NUTS 2	
Portugal	5	Comissões de coordenação regional	NUTS 2	Regiões autónomas excluded
Finland	6	Suuralueet	NUTS 2	
Sweden	21	Län	NUTS 3	
United Kingdom	37	Counties or groups of unitary authorities	NUTS 2	
Norway	19	Fylker	TL 3	
Switzerland	7	Grandes régions	TL 2	
Total EU15	210			
Total WE17	236			
<i>Data set 2</i>				
Belgium	11	Provinces	NUTS 2	
Denmark	1			

(continued on next page)

Table A1 (continued)

Country	Number of regions for which data are available	Administrative units	Classification level	Observations
<i>Data set 2</i>				
Germany	10	Länder	NUTS 1	Berlin and neue Länder excluded
Greece	1			
Spain	18	Comunidades autónomas+ Ceuta y Melilla	NUTS 2	
France	22	Régions	NUTS 2	DOMs excluded
Ireland	1			
Italy	20	Regioni	NUTS 2	
Luxembourg	1			
Netherlands	12	Provincies	NUTS 2	
Austria	1			
Portugal	5	Comissões de coordenação regional	NUTS 2	Regiões autónomas excluded
Finland	1			
Sweden	1			
United Kingdom	11	Government office regions	NUTS 1	According to NUTS 95 classification
Norway				
Switzerland				
Total EU15	116			

B.2. Data set 2

- Source: Hallet (2000) (based on Eurostat's REGIO and national sources).
- Variable: gross value added.
- Time dimension: annual averages, 1980–1995.
- Sectors retained: ores and metals; non-metallic minerals; chemicals; metal goods, machinery and electrical goods; transport equipment; food products; textiles, clothing and footwear; paper and printing products; misc. manufactured goods (nine industrial sectors, based on NACE-CLIO classification).
- Regional breakdown: 116 regions, see Table A1 (French “Départements d'outre-mer” as well as Madeira and Açores were dropped from Hallet's original data set, in order to enhance comparability with data set 1).
- Number of observations in full data set: 32,368.

References

- Aiginger, Karl, Davies, Steve W., 2001. Industrial specialisation and geographic concentration: two sides of the same coin? Working Paper #23, University of Linz.
- Aiginger, Karl, Leitner, Wolfgang, 2002. Regional concentration in the United States and Europe: who follows whom? *Review of World Economics* (Weltwirtschaftliches Archiv) 138 (4), 652–679.
- Aiginger, Karl, Pfaffermayr, Michael, 2004. The single market and geographic concentration in Europe. *Review of International Economics* 12 (1), 1–11.

- Amiti, Mary, 1999. Specialisation patterns in Europe. *Review of World Economics* (Weltwirtschaftliches Archiv) 135 (4), 573–593.
- Arbia, Giuseppe, 1989. *Spatial Data Configuration in Statistical Analysis of Regional Economic and Related Problems*. Kluwer.
- Barrios, Salvador, Strobl, Eric, 2004. Industry mobility and geographic concentration in the European Union. *Economics Letters* 82 (1), 71–75.
- Biewen, Martin, 2001. Bootstrap inference for inequality, mobility and poverty measurement. *Journal of Econometrics* 108 (2), 317–342.
- Bourguignon, François, 1979. Decomposable income inequality measures. *Econometrica* 47 (4), 901–920.
- Bourguignon, François, Morrisson, Christian, 2002. Inequality among world citizens: 1820–1992. *American Economic Review* 92 (4), 727–744.
- Brühlhart, Marius, 2001a. Evolving geographic concentration of European manufacturing industries. *Review of World Economics* (Weltwirtschaftliches Archiv) 137 (2), 215–243.
- Brühlhart, Marius, 2001b. Growing alike or growing apart? Industrial specialisation of EU countries. In: Wyplosz, C. (Ed.), *The Impact of EMU on Europe and the Developing Countries*. Oxford Univ. Press.
- Ciccone, Antonio, 2002. Agglomeration effects in Europe. *European Economic Review* 46 (2), 213–227.
- Ciccone, Antonio, Hall, Robert E., 1996. Production and the density of economic activity. *American Economic Review* 86 (1), 54–70.
- Clark, Todd, van Wincoop, Eric, 2001. Borders and business cycles. *Journal of International Economics* 55 (1), 59–85.
- Combes, Pierre-Philippe, Overman, Henry G., 2004. The spatial distribution of economic activities in the European Union. In: Henderson, V., Thisse, J.-F. (Eds.), *Handbook of Regional and Urban Economics*, vol. 4. Elsevier-North Holland.
- Cowell, Frank A., 1980. On the structure of additive inequality measures. *Review of Economic Studies* 47 (3), 521–531.
- Cowell, Frank A., 2000. Measurement of inequality. In: Atkinson, A.B., Bourguignon, F. (Eds.), *Handbook of Income Distribution*, vol. 1. Elsevier-North Holland.
- De la Fuente, Angel, Vives, Xavier, 1995. Infrastructure and education as instruments of regional policy: evidence from Spain. *Economic Policy* 20, 11–40.
- Duranton, Gilles, Overman, Henry G., 2002. Testing for localisation using micro-geographic data. CEPR Discussion Papers 3379.
- Duro, Juan A., 2001. Regional income inequalities in Europe: an updated measurement and some decomposition results. Mimeo, Institute for Economic Analysis, Universitat Autònoma de Barcelona.
- Duro, Juan A., Esteban, Joan, 1998. Factor decomposition of cross-country income inequality, 1960–1990. *Economics Letters* 60 (3), 269–275.
- Haaland, Jan I., Kind, Hans J., Midelfart Knarvik, Karen H., Torstensson, Johan, 1999. What determines the economic geography of Europe? CEPR Discussion Papers 2072.
- Hallet, Martin, 2000. Regional specialisation and concentration in the EU. European Commission Economic Papers 141.
- Helg, Rodolfo, Manasse, Paolo, Monocelli, Tommaso, Rovelli, Riccardo, 1995. How much (a)symmetry in Europe? Evidence from industrial sectors. *European Economic Review* 39 (5), 1017–1041.
- Imbs, Jean, Wacziarg, Romain, 2003. Stages of diversification. *American Economic Review* 93 (1), 63–86.
- Kalemli-Ozcan, Sebnem, Sorensen, Bent E., Yosha, Oved, 2003. Risk sharing and industrial specialization: regional and international evidence. *American Economic Review* 93 (3), 903–918.
- Krugman, Paul, 1991. *Geography and Trade*. MIT Press.
- Marcon, Eric, Puech, Florence, 2003. Measures of the geographic concentration of industries: improving distance-based methods. Mimeo, TEAM-University of Paris 1.
- Midelfart Knarvik, Karen H., Overman, Henry G., Redding, Stephen J., Venables, Anthony J., 2002. The location of European industry. *European Economy, Special Report* vol. 2/2002. European Commission, Brussels.
- Mills, Jeffrey A., Zandvakili, Sorushe, 1997. Statistical inference via bootstrapping for measures of inequality. *Journal of Applied Econometrics* 12 (2), 133–150.
- Mori, Tomoya, Nishikimi, Koji, Smith, Tony E., (2004) Relative entropy index of industrial localization: a statistic for large sample analysis. Mimeo, Institute of Economic Research, Kyoto University.

- Overman, Henry G., Redding, Stephen, Venables, Anthony J., 2003. The economic geography of trade, production and income: a survey of empirics. In: Kwan Choi, E., Harrigan, J. (Eds.), *Handbook of International Trade*. Blackwell.
- Paci, Raffaele, Usai, Stefano, 2000. Technological enclaves and industrial districts: an analysis of the regional distribution of innovative activity in Europe. *Regional Studies* 34 (2), 97–114.
- Peri, Giovanni, 1998. Technological growth and economic geography. IGIER Working Paper #139, Bocconi University.
- Shorrocks, Anthony F., 1980. The class of additively decomposable inequality measures. *Econometrica* 48 (3), 613–625.
- Shorrocks, Anthony F., 1982. Inequality decomposition by factor components. *Econometrica* 50 (1), 193–211.
- Shorrocks, Anthony F., 1984. Inequality decomposition by population subgroups. *Econometrica* 52 (6), 1369–1385.
- Slottje, Daniel J., 1990. Using grouped data for constructing inequality indices: parametric vs. nonparametric methods. *Economics Letters* 32 (2), 193–197.
- Storper, Michael, Chen, Yun-chung, De Paolis, Fernando, 2002. Trade and the location of industries in the OECD and European Union. *Journal of Economic Geography* 2 (1), 73–107.