Regional Science and Urban Economics xxx (2008) xxx–xxx

Contents lists available at ScienceDirect

Regional Science and Urban Economics

journal homepage: www.elsevier.com/locate/regec

Using entropy measures to disentangle regional from national localization patterns

Eleonora Cutrini *

Dipartimento di Studi sullo Sviluppo Economico, University of Macerata, Piazza Oberdan, 3, 62100 Macerata, Italy

ARTICLE INFO

Article history:
Received 8 February 2007
Received in revised form 29 March 2008
Accepted 20 August 2008
Available online xxxx

JEL classification:
C43
L16
O18
R12

Keywords:
Localization
Concentration
Specialization
Dissimilarity entropy indices
Within–between country decomposition

ABSTRACT

Localization of economic activities is a manifestation of two closely-related economic phenomena: the specialization of geographical units and the spatial concentration of industries. Nonetheless, the direction of changes in concentration and specialization, across national boundaries, may differ from those occurring within countries. Combining a regional approach with an international perspective, the paper introduces an entropy index of overall localization that allows specialization to be conceptualized as the mirror image of concentration, and also focuses attention on the possible divergence in agglomeration patterns at the different spatial scales.

1. Introduction

Economists and geographers have always shared an interest in the analysis of interregional and international spatial distribution of economic activities.

While international patterns of localization have been related to the existence of comparative advantages, localization at the regional level is mainly considered to be the outcome of increasing returns derived by firms belonging to the same industry, as a result of their agglomeration in a specific geographical area. Since Marshall's (1964) pioneering theoretical insights, knowledge spillovers, the presence of a thick labour market, and the easy reach of firms producing tailored inputs have been considered the main determinants of industry-specific external economies.

From a methodological point of view, ever since the early studies on international economics and location theory, the concept of localization has been related to divergence in the spatial spreading of an economic activity, with respect to a "theoretical case" in which simple location characteristics are at work (Ohlin, 1933). Hoover (1936) explicitly defined localization as "the degree of dissimilarity between the geographical distribution of an industry with reference to area, to other industries, to population, to raw-material supply, and so on."

Disproportionality in the spatial distribution of industries across regions, relative to the overall economic activity, is a clear sign of the specialization of regions, with respect to a reference area. The equivalence between the two phenomena entails recognition that any analysis of the spatial concentration of economic activities should explicitly or implicitly involve its mirror image: the specialization of geographical entities. However, empirical studies have disregarded the connection between the two phenomena; thus, a comprehensive analytical system with which to conceptualize the two dimensions of localization as unquestionably intertwined economic phenomena has hitherto been lacking in the literature.

Moreover, with few exceptions, economists assess localization patterns at mainly a single geographical level, even though the coexistence of different institutional levels calls for rigorous methodologies, in order to analyse structural changes at different spatial scales simultaneously.

The aim of this paper is to provide a framework for the analysis of localization, with regional specialization and geographical concentration considered as observable processes unquestionably tied together. Instead of developing a purely *ad hoc* tool, the paper draws heavily on insights from a wider and long-standing literature regarding the concepts and techniques of Information Theory, and their applicability in several disciplines. From a two-fold geographical perspective, the measure proposed makes it possible to disentangle spatial patterns of localization *between* and *within* countries.

* Tel.: +39 0733 2583931; fax: +39 0733 2583970.
E-mail address: eleonora.cutrini@unimc.it.

0166-0462/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.regsciurbeco.2008.08.005

The paper is organised as follows. Section 2 reviews the main, recent, methodological advances in measuring geographical concentration. Section 3 is devoted to the notation and the basic definitions of the dissimilarity indices of specialization and concentration used throughout the paper. Section 4 introduces the index of overall localization, whose main properties also clarify the economic interpretation of the entire hierarchical system introduced. Section 5 presents the results for European location patterns. Section 6 concludes the paper.

2. Recent advances in measuring geographical concentration

Over the past two decades, the resurgence of interest in the localization of economic activities has not only generated a large body of theoretical literature – which has come to be called New Economic Geography (Krugman, 1991) – it has also led to a surge of new statistical tools for empirical research, in an attempt to highlight stylised facts on agglomeration and dispersion tendencies.

Traditionally, the most frequently-used absolute indices[1] in regional studies have been the Gini coefficient, the Herfindhal index (Sapir, 1996; Aiginger and Pfaffermayr, 2004, among others) and the absolute entropy index (Aiginger and Pfaffermayr, 2004; Aiginger and Davies, 2004, among others), while relative concentration and relative specialization have commonly been measured through the Gini location quotient, based on the Hoover–Balassa Index (Kim, 1995; Amiti, 1999; Haaland et al., 1999; Brülhart, 2001; Midelfart et al., 2004, among others) or the K-concentration (or K-specialization) index (see, for example, Krugman, 1991; Hallet, 2000; Midelfart et al., 2004; Mulligan and Schmidt, 2005).

There have been further, recent, methodological developments in measuring geographical concentration. Ellison and Glaeser (1997) introduced a relative measure of concentration, which controls for the industrial structure and the overall agglomeration of manufacturing. The "dartboard approach" was adopted in successive empirical studies, to investigate the spatial distribution of manufacturing industries in France and the United Kingdom (Maurel and Sédillot, 1999; Devereux et al., 2004). The Ellison–Glaeser index (1997) allows a localization resulting from a highly-concentrated industrial structure to be distinguished from a localization arising from the spatial clustering of small-sized firms in the same geographical area. Moreover, on the basis of a specific location model, this analytical framework makes it possible to ascertain whether the observed concentration differs significantly from randomness.

However, the methodology introduced by Ellison and Glaeser (1997) can be used to evaluate the geographical concentration of each industry at a single geographical scale only.

This limitation has been overcome by the latest improvements in measuring geographical concentration over continuous space. Distance-based methods have been implemented to evaluate the spatial distribution of establishments in France and the United Kingdom, respectively (Marcon and Puech, 2003; Duranton and Overman, 2005). The spatial concentration of an industry is assessed with reference to a theoretical random distribution (homogeneous space), or to the spatial concentration of overall manufacturing (heterogeneous space). In the latter case, the geographical concentration of each manufacturing industry is determined by comparing the extent to which the specific industry is spatially concentrated (at some distance) against the degree of concentration of the aggregate manufacturing activity. By means of this procedure, manufacturing industries are detected as "dispersed at some distance" simply because their degree of concentration at that distance is lower than that of the reference distribution. Distance-based approaches provide the leading-edge methodology for understanding the geographical scope of a single industry's localization (dispersion). However, they require data on the location of individual establishments within each industry — data which are frequently not available.

Moreover, for policy purposes, it may be necessary to rely on decomposition properties, which make it possible to ascertain the extent to which localization occurs and changes over time, at the different spatial levels at which policies are designed and implemented. To this purpose, economists have begun to consider adoption of generalised entropy measures, such as those that have been widely used in several disciplines,[2] because they are endowed with the desirable decomposition properties. Brülhart and Traeger (2005) adopted generalised entropy measures, clearly distinguishing between *topographic* and *relative concentration*. While *topographic concentration* measures the extent to which an industry spreads, disproportionally to physical space,[3] *relative concentration* is more closely connected to the notions of urbanization, and localization economies.[4]

Relative concentration indices (e.g., the raw G-index of Ellison and Glaeser (1997), the relative concentration measure proposed by Brülhart and Traeger (2005), the Gini location quotient, based on the Hoover–Balassa index) are corrected for the overall distribution of the aggregate activity selected as a benchmark. They therefore gauge the extent to which the distribution across observed units of a specific sector departs from the regional distribution of the aggregate manufacturing activity.

Brülhart and Traeger (2005) highlighted the importance of defining appropriate regional weights, since empirical results are greatly affected by this choice. Recently, Bickenbach and Bode (2006)

[1] Absolute indices compare a distribution to the uniform distribution, while relative measures encompass the comparison between two non-uniform distributions. In the case of specialization, relative measures compare the sectoral distribution of employment in a given region to that of a supra-regional reference. In the case of concentration, the employment distribution across observed units (e.g., regions) is contrasted with the geographical distribution of a wider sector.

[2] Although a comprehensive account of the huge body of literature on Information Theory would be beyond the scope of this paper, mention should be made of some of the most important contributions, with examples of the existing applications in several disciplines. The concepts and techniques of Information Theory, introduced by Shannon's seminal work (Shannon, 1948; Shannon and Weaver, 1949), have been under development for a long time, in parallel with their implementation in many disciplines, among them, physics, biology, sociology and economics. Since the Information Theory approach adopted by Jaynes in statistical mechanics (Jaynes, 1957), *Tsallis entropies* have been introduced for the analysis of complex systems, i.e., systems involving interactions that occur over long distances, or which have long persistency of perturbations (Tsallis, 1988). In biology, the concepts and techniques of entropy have mainly been adopted to account for species diversity (see Jost, 2006, for clarification of the difference between entropy and diversity in ecology). Hill (1973) generalized different-order entropies, after Renyi (1961), while Taneja (1989) addressed the decomposition properties of the Shannon entropy, and relative entropy measures. In sociology, Shannon entropy has been regarded as the most satisfactory index for the analysis of racial segregation from a conceptual and mathematical point of view (Reardon and Firebaugh, 2002; Reardon and O'Sullivan, 2004). Theil's (1967) book provided the first, and most promising, applications of Information Theory in the economic field. Thereafter, entropy measures were adopted for economic forecasting (Zani, 1974; Mattioli, 1978, among others), for the evaluation of divergence from the risk-neutral probability distribution in the finance literature (see Stutzer, 1994, 1995), for macroeconomic modelling (Aoki, 1996), and for assessment of income inequality (Sala-i-Martin, 2002).

[3] Mori et al. (2005) also measured *topographic concentration* for Japan, using the Kullback–Leibler distance (Kullback and Leibler, 1951), and compared the distribution of establishments over the distribution of economic land area.

[4] In the regional science literature, distinction between localization and urbanization economies dates back to Marshall (1964) and Jacobs (1969), respectively. Localization economies are a mixture of advantages accruing to firms from being located close to other establishments in the same industry (industry-specific knowledge spillovers, labour market pooling, sharing of specialized suppliers). Instead, urbanization economies are common to all firms, and arise as the benefits of diversity, and of inter-industry spillovers among firms. Since the city is regarded as the cradle of diversity and innovation, it is diversity, rather than specialization, which fosters growth in cities, as stressed by Jacobs (1969). When localization economies predominate, space tends to be structured in specialized poles (and high concentration mirrors high specialization); when urbanization economies predominate, space is characterised by highly-diversified regions (low concentration mirrors low specialization).

Please cite this article as: Cutrini, E., Using entropy measures to disentangle regional from national localization patterns, Regional Science and Urban Economics (2008), doi:10.1016/j.regsciurbeco.2008.08.005

have stressed the distinction between reference distribution and weighting scheme,[5] and defined possible combinations of disproportionality measures: weighted absolute, weighted relative, unweighted absolute and unweighted relative.[6]

This review of the literature reveals that recent advances in measuring localization have focused on only one side of the question, namely the geographical concentration of industries, without any explicit relationship to regional specialization, although this is a dimension of analysis which is inherently involved in any assessment of the location of economic activities. Moreover, while several measures have been developed in a relatively *ad hoc* manner, this paper introduces a new measure of *overall localization* which is embedded in a system of entropy measures drawn from a large and long-standing body of literature. Decomposability underpins the hierarchical system of the dissimilarity measures adopted, and makes it possible to ascertain whether, and to what extent, localization occurs within and across countries. Bearing in mind the need for careful theoretical grounding, the comprehensive approach to the measurement of localization, introduced here, allows explicit recognition to be made of the equivalence between specialization and concentration.

3. The statistical toolbox: dissimilarity entropy measures of concentration and specialization

3.1. Notation and basic definition

Throughout the paper, the following notations are used: x denotes the variable of main interest (employment hereafter); subscripts i, j, and k index the country, region and industry, respectively; and superscripts ○ and c identify the supranational economy and the country, respectively (the geographical benchmarks for specialization measures). Thus:

x_{ijk}	number of workers in industry $k(k=1,\ldots,n)$, in region $j(j=1,\ldots,r_i)$, belonging to country i $(i=1,\ldots,m)$
x_{ij}	total employment in region ij
x_{ik}	total employment in industry k, in country i
x_i	total employment in country i
x_k	total employment in industry k, in the supranational economy
x	total employment in the supranational economy

Hence:

$LQ^*_{ijk} : = \frac{\frac{x_{ijk}}{x_{ij}}}{\frac{x_k}{x}}$ is the region-industry location quotient, relative to the supranational economy

$LQ_{ijk} : = \frac{\frac{x_{ijk}}{x_{ij}}}{\frac{x_{ik}}{x_i}}$ is the region-industry location quotient, relative to the country

$LQ_{ik} : = \frac{\frac{x_{ik}}{x_i}}{\frac{x_k}{x}}$ is the country-industry location quotient, relative to the supranational economy.

In what follows, the concept of overall localization refers to the pattern of an aggregate economic activity (manufacturing hereafter) which is composed of N industries and spans across R regions (observed units). Perfect regularity arises when all the industries composing the aggregate economic activity are spatially distributed, proportionally to total employment; and, accordingly, each region of the entire area has the same economic structure as the area as a whole (the supranational geographical benchmark). The absence of localization is a theoretical case. Ellison and Glaeser (1997) introduced the idea that observed industrial distribution will not be perfectly regular, even if firms choose their locations independently of each other. There are two reasons for this: stochasticity and industrial concentration (the non-independence of employment distribution). Therefore, Ellison and Glaeser also suggest an analytical way to compute confidence intervals for the null hypothesis of random location.

Localization of an aggregate economic activity encompasses two dimensions: relative concentration of industries, and relative specialization of regions. Localization is therefore a higher-order concept than specialization and concentration.

The agglomeration of an industry, k, is defined here as the divergence in the spatial distribution of that industry, controlling for the spread of overall economic activity (the reference distribution, in the case of concentration, e.g., manufacturing). The specialization of an observed unit (e.g., region j of country i) is taken to be the dissimilarity[7] between the regional economic structure (i.e., the allocation of the main interest variable across all the industries that constitute the overall economic activity in the region) and the economic structure of the geographical unit selected as the benchmark (i.e., the allocation of the main interest variable across all the industries that constitute the overall economic activity in the whole area).

3.2. Measuring relative concentration and relative specialization

The dissimilarity Theil index (Theil, 1967; see also Maasoumi, 1993, for its statistical properties) is particularly useful for studying the spread of economic activities across space, and the structural differences between geographical units. All raw measures of geographic concentration and specialization can be derived as dissimilarity Theil indices.

The basic dissimilarity Theil index used to measure the geographic concentration of one industry, k, corresponds to the entropy index adopted by Brülhart and Traeger (2005) to measure *relative concentration*:

$$T_k = \sum_{i=1}^{m} \sum_{j=1}^{r_i} \frac{x_{ijk}}{x_k} \ln\left(LQ^*_{ijk}\right). \quad (1)$$

The upper bound of the Theil index of relative concentration is given by log(Z), where Z is the number of "basic units" (i.e., number of employed workers, in the context of this paper). The index is unbounded only where the index is defined over a continuum, i.e., where basic units are infinitesimally small, which is not the case here. If $T_k=0$, it must be that $LQ^*_{ijk}=1$, for each region in the area. In this specific case, the localization of manufacturing industry k overlaps with the distribution of overall manufacturing activity in such a way that, in industry k, no region shows a specialization, with respect to the supranational economy; $T_k=0$ occurs when industry k is distributed across the regions in the same way as total manufacturing spans across the same regions of the entire area. Over time, an increasing relative concentration index denotes a process of regional specialization in that industry, somewhere in the entire considered economy.

The degree of geographic concentration of each industry (T_k) can be conceived as a measure of the strength of localization economies and/or the importance of industry-specific natural advantages. In fact, in the case of perfect regularity ($T_k=0$), the location of the industry is mainly due to the advantages of being located in those regions with the highest density of aggregate economic activity. If all industries

[5] The reference distribution is the distribution to which regional employment shares is compared, while the weighting scheme defines the weights applied to regions.

[6] From this perspective, all the raw entropy indices of specialization and concentration in the present paper are *unweighted relative measures*, because manufacturing employees are used as basic units (see Brülhart and Traeger, 2005 for a proper definition of the basic unit concept), while the reference is a non-uniform distribution. The indices used to assess overall localization are instead defined as relative weighted measures.

[7] In the present work, dissimilarity and disproportionality are synonyms, while dissimilarity and regularity are antonyms.

Please cite this article as: Cutrini, E., Using entropy measures to disentangle regional from national localization patterns, Regional Science and Urban Economics (2008), doi:10.1016/j.regsciurbeco.2008.08.005

follow the regular case (employment is allocated across regions in the same way as total employment), this indicates that industries are affected neither by localization economies (e.g., intra-industry spillovers, labour market pooling) nor by industry-specific natural advantages (see Ellison and Glaeser, 1997).

The two geographical components of the concentration index for each industry k can be easily derived by factor decomposition (see Appendix A for details on the formal decomposition of the geographical concentration index defined in Eq. (1)). Hence:

$$T_k^w = \sum_{i=1}^{m} \sum_{j=1}^{r_i} \frac{x_{ijk}}{x_k} \ln(LQ_{ijk}) \tag{2}$$

evaluates the geographical concentration of industry k *within countries*, while:

$$T_k^b = \sum_{i=1}^{m} \frac{x_{ik}}{x_k} \ln(LQ_{ik}) \tag{3}$$

assesses the geographical concentration of industry k *between countries*.

- $T_k^w = 0$ defines a benchmark of *perfect regularity within countries*, which implies that $LQ_{ijk} = 1$, for each region. If the *within* component of the relative concentration index is 0, then industry k is proportionally distributed to total manufacturing employment, in the internal regions of each country. The higher the domestic component, the more the inner allocation of regional comparative advantages of each country is uneven. An increasing value of the *within* component is related to a process of growing dissimilarity in the spatial distribution of the industry, internal to the countries, and therefore to the increasing importance of regional localization economies in industry k.
- $T_k^b = 0$ defines a benchmark of *perfect regularity between countries*, since when $T_k^b = 0$, it must be that $LQ_{ik} = 1$, for each country i. This case is associated with an international distribution of industry k, which perfectly overlaps with the allocation of manufacturing as a whole, across countries. In other words, countries reveal neither a comparative advantage nor a comparative disadvantage in the specific industry k analysed, with respect to the overall area. Accordingly, the higher the between-country component, the more the allocation of national comparative advantages in industry k is unbalanced. An increase in the between-country component of relative concentration indicates an increasingly unequal allocation of comparative advantages, associated with the process of country specialization.

Turning to the specialization side, it is possible to define several measures, by varying the observed region and/or the supra-regional references. Clearly, the choice of the observed unit, and the reference, depends upon the purpose of the analysis:

$$T_{ij}^{\circ} = \sum_{k=1}^{n} \frac{x_{ijk}}{x_{ij}} \ln\left(LQ_{ijk}^{*}\right). \tag{4}$$

The index T_{ij}° evaluates the dissimilarity between the economic structure of one region (composed of N industries), compared to that of a supranational geographical unit (Europe, USA, world).

To measure the relative specialization of region ij, with respect to the country i, or with respect to the supranational area, the following indices are derived:

$$T_{ij}^{c} = \sum_{k=1}^{n} \frac{x_{ijk}}{x_{ij}} \ln(LQ_{ijk}) \tag{5}$$

assesses the divergence between the regional manufacturing allocation among industries, and the country structure; and,

$$T_{i}^{\circ} = \sum_{k=1}^{n} \frac{x_{ik}}{x_{i}} \ln(LQ_{ik}) \tag{6}$$

is a measure of the "distance" (see Maasoumi, 1993 for a definition of the Theil index as a quasi-distance index) between the industrial composition of the country, and the industrial composition of the wider region (relative specialization of region i, with respect to the supranational area).

If T_{ij}° takes the value 0, then region j, located in country i, has a manufacturing structure which is proportional to the manufacturing structure of the overall economy. An increasing value of relative specialization indicates that the regional manufacturing structure is becoming increasingly different from the supranational economic structure. Similarly, if $T_{ij}^{c} = 0$, then the manufacturing allocation of employment across industries of region j, located in country i, mirrors the manufacturing structure of country i, while an increasing value of the relative specialization index denotes a divergent structural change, with respect to the country's economic transformation. Finally, if $T_i^{\circ} = 0$, then the manufacturing structure of country i is exactly proportional to the supranational distribution of employment, across industries.

When dissimilarity logic[8] is adopted, country specialization, relative to a supranational unit (T_i°), can be envisioned as a residual of the averaged regional specialization, relative to the same benchmark — once the divergence of regional manufacturing structures related to that country has been accounted for.

Adopting a regional standpoint, country specialization is best defined as the averaged regional specialization indices, relative to the supranational unit (aRS_i°); and, it consists of two elements: a within-country component (aRS_i^{c}), which accounts for the average specialization of internal regions (with respect to the country), and the national component, associated with national characteristics — in other words, the country specialization, relative to the entire area (T_i°). The following relation holds:

$$aRS_i^{\circ} = aRS_i^{c} + T_i^{\circ}, \tag{7}$$

where:

$$aRS_i^{\circ} = \sum_{j=1}^{r_i} T_{ij}^{\circ} \frac{x_{ij}}{x_i}, \tag{8}$$

and

$$aRS_i^{c} = \sum_{j=1}^{r_i} T_{ij}^{c} \frac{x_{ij}}{x_i}. \tag{9}$$

In this setting, each country specialization, relative to the supranational unit (T_i°), is simply the difference between the two country-based averaged regional specialization measures:

$$T_i^{\circ} = \sum_{j=1}^{r_i} \left(T_{ij}^{\circ} - T_{ij}^{c}\right) \frac{x_{ij}}{x_i}. \tag{10}$$

To conclude, the weighted average of each country's regional specialization indices, relative to the supranational unit (aRS_i°), is decomposable into a within-country component of regional specialization (aRS_i^{c}), and a between-country component (T_i°).

4. Measuring overall localization: the equivalence between concentration and specialization

4.1. The entropy index of overall localization

This paper conceives overall localization as the aggregate outcome of two closely-related economic phenomena: on the one hand, the relative specialization of basic units (regions), the divergence between

[8] "Dissimilarity logic" is the approach that underpins this study: precisely, it is the construction of indices based on a comparison between two overlapping distributions, and aims at measuring their "distance" or "divergence".

Please cite this article as: Cutrini, E., Using entropy measures to disentangle regional from national localization patterns, Regional Science and Urban Economics (2008), doi:10.1016/j.regsciurbeco.2008.08.005

regional economic structures, and the economic structure of the entire area; on the other, the relative concentration of each industry composing the aggregate economic activity.

Since the relative concentration of industrial activities is reflected in the relative specialization of geographical entities, if individual industries tend to be concentrated in a rather small number of regions, it should be the case that some regions are highly specialized in these industries. By contrast, if all industries are spreading across space, proportionally to the aggregate economic activity, then regions will exhibit a low level of specialization. Therefore, specialization and concentration are strictly connected conceptually[9] and can be condensed into the concept of overall localization. From a purely statistical viewpoint, measuring overall localization involves evaluating the entire distribution of economic activities — across regions and across industries.

The entropy index proposed here, to measure overall localization (*L-index*), is a weighted sum of the logarithms of location quotients, where the weights are the industry-region shares of aggregate manufacturing ($\frac{x_{ijk}}{x}$):

$$L=\sum_{k=1}^{n}\sum_{i=1}^{m}\sum_{j=1}^{r_i}\frac{x_{ijk}}{x}\ln\left(LQ^*_{ijk}\right). \tag{11}$$

Since $\frac{x_{ijk}}{x}=\frac{x_{ij}}{x}\cdot\frac{x_{ijk}}{x_{ij}}=\frac{x_k}{x}\cdot\frac{x_{ijk}}{x_k}$, it is possible to rewrite the entropy index of overall localization (L) as follows:

$$L=\sum_{j=1}^{r}\sum_{k=1}^{n}\frac{x_{ij}}{x}\cdot\frac{x_{ijk}}{x_{ij}}\ln\left(LQ^*_{ijk}\right)=\sum_{k=1}^{n}\sum_{j=1}^{r}\frac{x_k}{x}\cdot\frac{x_{ijk}}{x_k}\ln\left(LQ^*_{ijk}\right). \tag{12}$$

Substituting Eqs. (4) and (1) in the left and right sides of Eq. (12), respectively, a first intuition of the two-fold interpretation of the entropy index of overall localization can be derived. The *L-index* is a summary statistic of both relative specialization indices and relative concentration ones:

$$L=\sum_{j=1}^{r}\frac{x_{ij}}{x}T^{\circ}_{ij}=\sum_{k=1}^{n}\frac{x_k}{x}T_k. \tag{13}$$

From the specialization point of view, the aggregation gives an idea of the average dissimilarity between the regional employment proportions across industries and the manufacturing structure of the supranational economy selected as benchmark. Similarly, from a concentration standpoint, the composite measure of localization informs us about the average dissimilarity between the distribution of industries across geographical units and the spread of overall manufacturing across geographical units. Since the size of the regional units affects the degree of relative specialization, all of the raw specialization indices are weighted by the region's share in total manufacturing employment ($\frac{x_{ij}}{x}$). Similarly, the difference in industry size is accounted for by weighting each relative concentration index by the industry's share in total manufacturing employment ($\frac{x_k}{x}$).

Thanks to the high sensitivity of entropy measures to minimal distributional changes, the hierarchical system of decomposition allows one to:

- have a precise diagnosis of localization, together with its two empirical dimensions
- account for the complexity in the evolution of localization at the different spatial scales, disentangling within-country regional agglomeration tendencies from national structural changes.

All the analytical constructions can be referred to as a location model of localization, in which perfect localization is associated with the theoretical case of complete specialization of regions, and complete concentration of industries.

Perfect localization occurs when each industry is completely concentrated in just one region, and when each region in which an industry is concentrated is completely specialized in that industry. In the case of perfect localization, agglomeration economies within industries (localization economies) and industry-specific natural advantages are deemed to be fully exploited.

By contrast, *perfect regularity* (a nil value of the *L-index*) is associated with the absence of specialization of regions ($T^{\circ}_{ij}=0$, for all ij) and the absence of geographic concentration of industries ($T_k=0$, for all k). In this case, since all regions in the area have the same allocation of employment across industries, they are characterised by an identical economic structure. The spatial distribution of industries is inevitably symmetrical, since all the industries exhibit perfectly-overlapping distributions across spatial units. Therefore the *L-index* can be conceived as an aggregate measure of localization economies within industries, and of regional specialization based on industry-specific advantages.

The *L-index* makes it possible to ascertain the extent to which two macro-areas (Europe, USA) are different, in terms of localization economies, and whether it is a matter of national or regional specialization (in other words, concentration *across* or *within* countries).

Moreover, the entropy measure of overall localization proposed complies with several baseline principles, outlined by Combes and Overman (2004) and Duranton and Overman (2005) as desirable properties:

(1) It is comparable across spatial units and scales (additively decomposable by geographical subgroups);
(2) It specifies an unambiguous and meaningful null hypothesis of absence of localization ($L=0$);
(3) It is suitable for statistical testing.

However, like all measures based on aggregate regional data, the *L-index* is affected by the modifiable areal unit problem (MAUP) and the checkerboard problem (Arbia, 1989). A line of methodological development, based on spatial disproportionality measures of concentration, has been recently introduced to deal with the checkerboard and the MAUP problems (Bickenbach and Bode, 2006).

The following section is devoted to the spatial decomposition of the *L-index*.[10]

4.2. Within- and between-country components of overall localization

Despite the equivalence between geographical concentration and regional specialization, the evolution of overall localization *within* countries may not proceed in parallel with localization *between* countries. Separating overall localization into its different components (*within* and *between* countries) allows one to disentangle the contribution of national comparative advantages to overall localization patterns from the magnitude of regional agglomeration forces (which

[9] The present analysis confirms insights yielded by recent studies which acknowledge that, when relative measures are adopted, the two phenomena cannot diverge (Aiginger and Davies, 2004; Aiginger and Rossi-Hansberg, 2006). There is a statistical reason for this. If one relies on relative measures, identity at the aggregate level should always apply because, when constructing each relative measure, the researcher makes use of information from the entire matrix, where the rows refer to industries and the columns refer to countries. Conversely, absolute measures are constructed by using the limited information provided by each row (column) of a matrix. In this case, the only benchmark is the theoretical uniform distribution of each row (column) having any connection with the rest of the matrix distribution. As a matter of fact, in the literature, the expected identity has been found only when unweighted relative entropy measures (Aiginger and Davies, 2004) or the commonly-used relative mean deviation (Mulligan and Schmidt, 2005) have been adopted. Therefore, the relevant question is which measure (absolute or relative) is the appropriate one. Indubitably, this depends upon the purpose of the analysis, although relative indices are more appropriate than absolute ones — if the intention is to ascertain the strength of within-industry agglomeration economies.

[10] The entropy index of overall localization is also additively decomposable by industrial subgroups.

may be the result of external economies or intra-firm increasing returns-to-scale) at the industry level.

In what follows, the spatial decomposition of the *L-index*, introduced in the previous section, is presented, and it is shown that specialization and concentration still conceptually and analytically underpin each component of overall localization.

The decomposition of the *L-index* is obtained by means of the following relation between the location quotients, applied thus far:

$$LQ^{*}_{ijk} = LQ_{ijk}LQ_{ik}. \tag{14}$$

Substituting Eq. (14) into Eq. (11) the *L-index* becomes:

$$L = \sum_{k=1}^{n}\sum_{i=1}^{m}\sum_{j=1}^{r_i}\frac{x_{ijk}}{x}\ln\left(LQ_{ijk}LQ_{ik}\right). \tag{15}$$

And, since $\sum_{j=1}^{r_i}\frac{x_{ijk}}{x}=\frac{x_{ik}}{x}$, the decomposition yields:

$$L = \sum_{k=1}^{n}\sum_{i=1}^{m}\sum_{j=1}^{r_i}\frac{x_{ijk}}{x}\ln\left(LQ_{ijk}\right) + \sum_{k=1}^{n}\sum_{i=1}^{m}\frac{x_{ik}}{x}\ln\left(LQ_{ik}\right). \tag{16}$$

The *between-country* component (L^b) of overall localization is:

$$L^b = \sum_{k=1}^{n}\sum_{i=1}^{m}\frac{x_{ik}}{x}\ln\left(LQ_{ik}\right). \tag{17}$$

The two-fold definition of the concept of overall localization still holds at the between-country level. Substituting $\frac{x_{ik}}{x}=\frac{x_i}{x}\cdot\frac{x_{ik}}{x_i}=\frac{x_k}{x}\cdot\frac{x_{ik}}{x_k}$ in Eq. (17) yields:

$$L^b = \sum_{k=1}^{n}\sum_{i=1}^{m}\frac{x_i}{x}\cdot\frac{x_{ik}}{x_i}\ln\left(LQ_{ik}\right) = \sum_{k=1}^{n}\sum_{i=1}^{m}\frac{x_k}{x}\cdot\frac{x_{ik}}{x_k}\ln\left(LQ_{ik}\right). \tag{18}$$

Substituting Eqs. (6) and (3) in the left and right sides of Eq. (18), respectively:

$$L^b = \sum_{i=1}^{m}\frac{x_i}{x}T_i^{\circ} = \sum_{k=1}^{n}\frac{x_k}{x}T_k^b. \tag{19}$$

The within-country component (L^w) of overall localization is:

$$L^w = \sum_{k=1}^{n}\sum_{i=1}^{m}\sum_{j=1}^{r_i}\frac{x_{ijk}}{x}\ln\left(LQ_{ijk}\right). \tag{20}$$

Again, the two-fold definition of localization is confirmed at the within-country level. Since $\frac{x_{ijk}}{x}=\frac{x_{ij}}{x}\cdot\frac{x_{ijk}}{x_{ij}}=\frac{x_k}{x}\cdot\frac{x_{ijk}}{x_k}$, Eq. (20) becomes:

$$L^w = \sum_{k=1}^{n}\sum_{j=1}^{r}\frac{x_{ij}}{x}\cdot\frac{x_{ijk}}{x_{ij}}\ln\left(B_{ijk}\right) = \sum_{k=1}^{n}\sum_{j=1}^{r}\frac{x_k}{x}\cdot\frac{x_{ijk}}{x_k}\ln\left(B_{ijk}\right). \tag{21}$$

Substituting Eq. (5) in the left side and Eq. (2) in the right side of Eq. (21) yields:

$$L^w = \sum_{j=1}^{r}\frac{x_{ij}}{x}T_{ij}^c = \sum_{k=1}^{n}\frac{x_k}{x}T_k^w. \tag{22}$$

The overall localization index (L) is an averaged dissimilarity index, and it is interpretable as a summary statistic of regional specialization indices (T_{ij}°), weighted by the regional manufacturing shares ($\frac{x_{ij}}{x}$):

$$L = \sum_{j=1}^{r}\frac{x_{ij}}{x}T_{ij}^{\circ} = \sum_{i=1}^{m}\frac{x_i}{x}T_i^{\circ} + \sum_{j=1}^{r}\frac{x_{ij}}{x}T_{ij}^c. \tag{23}$$

Similarly, on the concentration side, overall localization can be seen as a summary statistic of relative concentration Theil indices (T_k), weighted by the industry shares ($\frac{x_k}{x}$):

$$L = \sum_{k=1}^{n}\frac{x_k}{x}T_k = \sum_{k=1}^{n}\frac{x_k}{x}T_k^b + \sum_{k=1}^{n}\frac{x_k}{x}T_k^w. \tag{24}$$

To conclude, both Eqs. (23) and (24) correspond to Eq. (16) and to:

$$L = L^b + L^w. \tag{25}$$

The *L-index*, and each single component, is non-negative. Perfect regularity ($L=0$) implies that $L^b=0$ and $L^w=0$. Any departure from perfect regularity ($L>0$) means that some localization economies are at work within countries ($L^w>0$), or some comparative advantage between countries exists ($L^b>0$), or both. Usually, overall localization is jointly explained by international and intra-national components.

To thoroughly understand the economic interpretation of the spatial decomposition analysis, it is helpful to discuss the two extreme cases.

a) Localization occurs only *between* countries ($L=L^b$; $L^w=0$). In this case, internal regions of each country are not specialized, since their economic structures mirror the reference, national, economic structure ($T_{ij}^c=0$ for all ij); therefore, the spatial distributions of all industries within each country overlap with the spatial distribution across internal regions of the national aggregate activity ($T_k^w=0$, for all k). However, national economic structures are different from the supranational area, and the spatial distribution of industries across countries follows different patterns (so that some industries do not overlap with the spatial distribution of economic activity as a whole, across countries). In this case, localization is fully explained by national comparative advantages, and no room is left for industry-specific, regional, agglomeration economies.
b) Localization occurs only *within* countries ($L=L^w$; $L^b=0$). In this case, the internal regions of each country are specialized, because their economic structures are different from the reference, national, economic structure, and the spatial concentrations of industries within each country are dissimilar from the spatial distribution of the national aggregate activity across internal regions of each country. However, national economic structures are exactly proportional ($T_i^{\circ}=0$ for all i), and the international distribution of each industry overlaps with the allocation of economic activity as a whole, across countries ($T_k^b=0$, for all k).

5. Localization patterns in Europe: an illustrative example

This section describes the pattern of the European manufacturing sector, in terms of overall localization, geographic concentration of industries, and specialization of economies, in 2001. Table 1 summarizes the levels of relative concentration of 12 manufacturing industries and the regional specialization of nine countries in Europe, depicting their symmetry, with respect to the overall localization measures introduced.

Data are drawn from the Structural Business Statistics database (Eurostat), which provides employment data by manufacturing industries. Analysis is confined to two-digit manufacturing industries, according to the NACE revision 1 classification, and encompasses 145 nuts–2 regions (the observed unit of analysis) belonging to the following countries: Belgium and Luxembourg (consolidated), Finland, France, Western Germany, Greece, Italy, Netherlands, Spain, United Kingdom.

To provide an idea of the different levels of agglomeration, it is interesting to note that, in 2001, *textiles*, *transport equipment*, *wood*, *non-metal products*, and *chemicals* recorded the highest levels of geographical concentration. Their comparatively higher agglomeration may be the outcome of strong within-industry spillovers, natural advantages, or increasing returns-to-scale at the level of each industrial plant. To understand this result, it is also important to note that smaller industries usually appear to be more concentrated than larger industries

Table 1
Two-fold decomposition of the entropy index of overall localization, 2001

	$\frac{x_k}{x}$	T_k		T_k^b		T_k^w	
	(a)	(b)	(a)*(b)	(c)	(a)*(c)	(d)	(a)*(d)
Food	0.119	0.112	0.013	0.031	0.004	0.081	0.010
Textiles	0.074	0.287	0.021	0.140	0.010	0.147	0.011
Wood	0.027	0.167	0.004	0.062	0.002	0.105	0.003
Paper	0.081	0.128	0.010	0.041	0.003	0.087	0.007
Chemicals	0.064	0.160	0.010	0.021	0.001	0.139	0.009
Rubber and plastics	0.054	0.073	0.004	0.010	0.001	0.063	0.003
Other non-metallic mineral products	0.046	0.167	0.008	0.036	0.002	0.132	0.006
Basic metals and fabric. metal products	0.150	0.070	0.010	0.008	0.001	0.062	0.009
Machinery and equipment	0.114	0.088	0.010	0.040	0.005	0.048	0.006
Electrical and optical equipment	0.117	0.090	0.011	0.027	0.003	0.063	0.007
Transport equipment	0.101	0.169	0.017	0.051	0.005	0.118	0.012
Manufacturing nec	0.052	0.095	0.005	0.029	0.002	0.065	0.003
Overall localization	1.000		0.124		0.038		0.086
	$\frac{x_i}{x}$	aRS_i^o		T_i^o		aRS_i^c	
	(a)	(b)	(a)*(b)	(c)	(a)*(c)	(d)	(a)*(d)
Belgium and Luxembourg	0.030	0.168	0.005	0.035	0.001	0.133	0.004
Germany	0.266	0.115	0.031	0.051	0.014	0.064	0.017
Spain	0.110	0.156	0.017	0.045	0.005	0.111	0.012
Finland	0.013	0.137	0.002	0.090	0.001	0.047	0.001
France	0.159	0.096	0.015	0.015	0.002	0.081	0.013
Greece	0.009	0.316	0.003	0.166	0.002	0.150	0.001
Italy	0.194	0.129	0.025	0.049	0.009	0.080	0.016
Netherlands	0.039	0.098	0.004	0.054	0.002	0.044	0.002
United Kingdom	0.181	0.128	0.023	0.012	0.002	0.115	0.021
Overall localization	1.000		0.124		0.038		0.086

Source: EUROSTAT Region-SBS data.

because there is greater uncertainty about their spatial distribution when compared to overall manufacturing activity.

Similarly, regional size negatively affects the degree of overall specialization. Smaller regions and countries are likely to be more specialized than larger ones, since their manufacturing structures are deemed to be more dissimilar to the European structure than those of larger regions, given that the latter may constitute an important part of the reference distribution.

In fact, turning from industries to regions, the magnitude of region-based specialization, relative to Europe, was found to be higher in smaller and Cohesion countries, notably in Greece and Spain, but also in Belgium and Luxembourg, and Finland. Specialization of some smaller and peripheral countries – namely Greece and Finland – was mainly a "national fact". Instead, in larger and older member-states of the European Union, specialization, in 2001, is mainly a "regional fact". Indeed, a within-country component higher than the between-country one is recurrent in the specialization of Germany, Belgium and Luxembourg, Italy, France, and the United Kingdom. This result is certainly the outcome of the larger size of the national economy, which implies higher internal differences. This does not portend the end of the State; but, European integration has certainly reduced the importance of national borders, in defining appropriate units of economic analysis.

It is important to note that the *L-index* is less biased, with respect to spatial scale and industrial size, than the specific measures of concentration and specialization. Within-country components explain a large part of the localization of industries and the specialization of countries. Bearing in mind the importance of potential MAUP biases, it is nevertheless conceivable that the regional grid is the prominent geographical level, in assessing concentration and specialization. In other word, national borders seems to have scant significance in defining localization, therefore greater attention should be paid to sub-national economies, which represent the realms of localization economies, increasing returns-to-scale and natural advantages, as sources of agglomeration.

6. Conclusion and further developments

This paper has introduced an analytical framework with which to consider concentration and specialization as twin manifestations of the same concept: overall localization — both across and within countries. By means of relative entropy indices and their partition properties, it is possible to employ two hierarchical spatial grids, assessing the evolution of their relative importance in the pattern of localization of economic activities.

First, a taxonomy of raw concentration and specialization measures, derived from the original dissimilarity Theil index, was developed. Here, the concept of overall localization – with its two-fold connotation – found an explicit statistical counterpart. In fact, it was argued that, when dissimilarity logic is adopted, specialization and concentration can be seen as two sides of the same coin, both across and within countries. The index of overall localization allows one to encompass the full structure of industries and regions within a single analysis, because it can be conceived as the aggregation of the specific indices used to assess regional (and country) specialization, on one hand, and, the concentration of industries within (and across) countries, on the other.

Entropy measures are likely to be widely implemented in regional economics, following their numerous applications in the past, in several scientific disciplines and social sciences. However, the methodology could be improved with various refinements which deal with the MAUP and the checkerboard problems, in both industrial and spatial dimensions. Moreover, given the high relevance of statistical testing of results, further work should be done to develop appropriate counterfactuals, embedded in sound economic models of industrial location.

Acknowledgments

This work has been partly supported by the Italian Ministry of University and Research Grants PRIN2005. The author is grateful to Eckhardt Bode, Riccardo Lucchetti, Giulio Palomba and the participants at the International Workshop on Spatial Econometrics and Statistics held in Rome (May 25–27, 2006) and the Kiel Institute staff seminar (December 7, 2006) for their useful comments and suggestions on earlier drafts. The paper has benefited greatly from the comments of two anonymous reviewers. The usual disclaimers apply.

Appendix A. Decomposing the dissimilarity Theil index for geographic concentration

As previously noted, relative concentration refers to the dissimilarity in the localization of each industry, k, with respect to the spreading of overall manufacturing industry across the spatial units considered (countries, regions). If an industry, k, spreads in exact proportion to total manufacturing employment, the relative concentration index will exhibit a nil value.

$$T_k = \sum_{i=1}^{m} \sum_{j=1}^{r_i} \frac{x_{ijk}}{x_k} \ln\left(\frac{\frac{x_{ijk}}{x_k}}{\frac{x_{ij}}{x}}\right). \tag{A.1}$$

Adding and subtracting the term $\sum_{i=1}^{m} \frac{x_{ik}}{x_k} \ln\left(\frac{x_{ik}}{x_i}\right)$ to Eq. (A.1) yields:

$$T_k = \sum_{i=1}^{m} \sum_{j=1}^{r_i} \frac{x_{ijk}}{x_k} \ln\left(\frac{x_{ijk}}{x_{ij}}\right) - \sum_{i=1}^{m} \sum_{j=1}^{r_i} \frac{x_{ijk}}{x_k} \ln\left(\frac{x_k}{x}\right) + \sum_{i=1}^{m} \frac{x_{ik}}{x_k} \ln\left(\frac{x_{ik}}{x_i}\right) + \sum_{i=1}^{m} \frac{x_{ik}}{x_k} \ln\left(\frac{x_{ik}}{x_i}\right) \tag{A.2}$$

and, since, $\sum_{i=1}^{m} \sum_{j=1}^{r_i} \frac{x_{ijk}}{x_k} = \sum_{i=1}^{m} \frac{x_{ik}}{x_k}$, then:

$$T_k = \sum_{i=1}^{m} \sum_{j=1}^{r_i} \frac{x_{ijk}}{x_k} \ln\left(\frac{x_{ijk}}{x_{ij}}\right) - \sum_{i=1}^{m} \frac{x_{ik}}{x_k} \ln\left(\frac{x_k}{x}\right) + \sum_{i=1}^{m} \frac{x_{ik}}{x_k} \ln\left(\frac{x_{ik}}{x_i}\right) - \sum_{i=1}^{m} \sum_{j=1}^{r_i} \frac{x_{ijk}}{x_k} \ln\left(\frac{x_{ik}}{x_i}\right). \quad (A.3)$$

Combining the second and third elements, the *between-country* component is obtained:

$$T_k^b = \sum_{i=1}^{m} \frac{x_{ijk}}{x_k} \ln\left(\frac{\frac{x_{ik}}{x_k}}{\frac{x_i}{x}}\right). \quad (A.4)$$

Instead, the *within-country* component is obtained by combining the first element of Eq. (A.3) with the fourth one:

$$T_k^w = \sum_{i=1}^{m} \sum_{j=1}^{r_i} \frac{x_{ijk}}{x_k} \ln\left(\frac{\frac{x_{ijk}}{x_{ik}}}{\frac{x_{ij}}{x_i}}\right), \quad (A.5)$$

so that:

$$T_k = T_k^b + T_k^w. \quad (A.6)$$

The Theil index within countries (T_k^w) is a weighted average of the relative Theil indices of industry k between regions in each country (T_{ik}^{br}), where the weights are the countries' shares in total employment in industry k ($\frac{x_{ik}}{x_k}$).

$$T_k^w = \sum_{i=1}^{m} \frac{x_{ik}}{x_k} T_{ik}^{br}, \quad (A.7)$$

where $T_{ik}^{br} = \sum_{j=1}^{r_i} \frac{x_{ijk}}{x_{ik}} \ln\left(\frac{\frac{x_{ijk}}{x_{ik}}}{\frac{x_{ij}}{x_i}}\right)$, which, again, can be thought of as a dissimilarity Theil index.

References

Aiginger, K., Davies, S.W., 2004. Industrial specialization and geographic concentration: two sides of the same coin? Not for the European Union. Journal of Applied Economics 7 (2), 231–248.

Aiginger, K., Pfaffermayr, M., 2004. The Single Market and geographic concentration in Europe. Review of International Economics 12 (1), 1–11.

Aiginger, K., Rossi-Hansberg, E., 2006. Specialization and concentration: a note on theory and evidence. Empirica 33 (4), 255–266.

Amiti, M., 1999. Specialization patterns in Europe. Weltwirtschaftliches Archiv 135, 573–593.

Aoki, M., 1996. New Approaches to Macroeconomic Modelling, Evolutionary Stochastic Dynamics, Multiple Equilibria, and Externalities as Field Effects. Cambridge University Press.

Arbia, G., 1989. Spatial Data Configuration in Statistical Analysis of Regional Economic and Related Problems. Kluwer.

Bickenbach, F., Bode, E., 2006. Disproportionality measures of concentration, specialization, polarization. Kiel Working Paper N. 1276. Kiel Institute for the World Economy.

Brülhart, M., 2001. Evolving geographical concentration of European Union. Weltwirtschaftliches Archiv 137 (2), 215–243.

Brülhart, M., Traeger, R., 2005. An account of geographic concentration patterns in Europe. Regional Science and Urban Economics 35, 597–624.

Combes, P.P., Overman, H.G., 2004. The spatial distribution of economic activities in the European Union. In: Thisse, J.-F. and Henderson, V. (Eds.), Handbook of Urban and Regional Economics, vol. 4. Elsevier-North Holland, Amsterdam, pp. 2845–2910.

Devereux, M.P., Griffith, R., Simpson, H., 2004. The geographic distribution of production activity in the UK. Regional Science and Urban Economics 34 (5), 533–564.

Duranton, G., Overman, H.G., 2005. Testing for localization using micro-geographic data. Review of Economic Studies 72 (4), 1077–1106.

Ellison, G., Glaeser, E.L., 1997. Geographic concentration in U.S. manufacturing industries: a dartboard approach. Journal of Political Economy 105 (5), 889–927.

Haaland, J.I., Kind, H.J., Midelfart-Knarvik, K.H., Torstensson, J., 1999. What determines the economic geography of Europe? CEPR Discussion Paper N. 2072.

Hallet, M., 2000. Regional specialization and concentration in the EU. Economic Papers N. 141. The European Commission, Directorate General for Economic and Financial Affairs.

Hill, M.O., 1973. Diversity and evenness: a unifying notation and its consequences. Ecology 54 (2), 427–432.

Hoover, E.M., 1936. The measurement of industrial localization. Review of Economics and Statistics 18 (4), 162–171.

Jacobs, J., 1969. The Economy of Cities. Vintage, New York.

Jaynes, E.T., 1957. Information theory and statistical mechanics. Physical Review 106 (4), 620–630.

Jost, L., 2006. Entropy and diversity. Oikos 103 (2), 363–375.

Kim, S., 1995. Expansion of markets and the geographic distribution of economic activities: the trends in US regional manufacturing structure, 1860–1987. Quarterly Journal of Economics 110 (4), 881–908.

Krugman, P., 1991. Geography and Trade. MIT Press, Cambridge.

Kullback, S., Leibler, R.A., 1951. On information and sufficiency. Annals of Mathematical Statistics 22, 79–86.

Maasoumi, E., 1993. A compendium to information theory in economics and econometrics. Econometric Reviews 12 (2), 137–181.

Marcon, E., Puech, F., 2003. Evaluating the geographic concentration of industries using distance-based methods. Journal of Economic Geography 3 (4), 409–428.

Marshall, A., 1964. Principles of Economics, an Introductory Volume. Macmillan, London. (First edition published 1890.).

Mattioli, E., 1978. Una Nuova Valutazione delle Previsioni Input–Output con Differenti Gradi di Aggregazione. Statistica, anno XXXVIII, n. 1.

Maurel, F., Sédillot, B., 1999. A measure of the geographic concentration in French manufacturing industries. Regional Science and Urban Economics 29 (5), 575–604.

Midelfart, K.H., Overman, H.G., Redding, S.J., Venables, A.J., 2004. The location of European industry. In: Dierx, A., Ilzkovitz, F., Sekkat, K. (Eds.), European Integration and the Functioning of Product Markets. Edward Elgar.

Mori, T., Nishikimi, K., Smith, T.E., 2005. A divergence statistic for industrial localization. Review of Economics and Statistics 87 (4), 635–651.

Mulligan, G.F., Schmidt, C., 2005. A note on localization and specialization. Growth and Change 36 (4), 565–576.

Ohlin, B., 1933. Interregional and International Trade. Harvard Economic Studies, vol.39. Harvard University Press, Cambridge, MA.

Reardon, S.F., Firebaugh, G., 2002. Measures of multi-group segregation. Sociological Methodology 32, 33–67.

Reardon, S.F., O'Sullivan, D., 2004. Measures of spatial segregation. Sociological Methodology 34, 121–162.

Renyi, A., 1961. On measure of entropy and information. In: Neymann, J. (Ed.), Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability 1960, pp. 547–561.

Sala-i-Martin, X., 2002. The disturbing "rise" of global income inequality. NBER Working Paper N. 8904.

Sapir, A., 1996. The effects of Europe's internal market program on production and trade: a first assessment. Weltwirtschaftliches Archiv 132, 457–475.

Shannon, C.E., 1948. A mathematical theory of communication. Bell System Technical Journal 27, 379–423 July.

Shannon, C.E., Weaver, W., 1949. The Mathematical Theory of Communication. University of Illinois Press, Urbana.

Stutzer, M., 1995. An information theoretic index of risk in financial markets. In: Berry, Donald A., Chaloner, Kathryn M., Geweke, J.K. (Eds.), Bayesian Analysis in Statistics and Economics: Essays in Honor of Arnold Zellner. Wiley Interscience.

Stutzer, M., 1994. The statistical mechanics of asset prices. In: Elworthy, K.D., Everett, W. N., Lee, E.B. (Eds.), Differential Equations, Dynamical Systems, and Control Science. Lecture Notes in Pure and Applied Mathematics Series, vol. 152, pp. 321–342.

Taneja, I., 1989. On generalized information measures and their applications. Advances in Electronics and Electronic Physics 76, 327–413.

Theil, H., 1967. Economics and Information Theory. North-Holland, Amsterdam.

Tsallis, C., 1988. Possible generalization of Boltzmann–Gibbs statistics. Journal of Statistical Physics 52 (1–2), 479–487.

Zani, S., 1974. Valutazione delle Previsioni Economiche mediante la Teoria dell'Informazione. Rivista Internazionale di Scienze Sociali, Cattolica del Sacro Cuore, anno LXXXII, fasc. I–II, 256–273.

Please cite this article as: Cutrini, E., Using entropy measures to disentangle regional from national localization patterns, Regional Science and Urban Economics (2008), doi:10.1016/j.regsciurbeco.2008.08.005