Universidade de Vigo
Departamento de Economía Aplicada

Documento de Traballo
0804

An alternative inequality-based concentration measure

Olga Alonso-Villar
Coral del Río

Documentos de Traballo

Novembro 2008

Departamento de Economía Aplicada
Universidade de Vigo
As Lagoas Marcosende S/N, 36310 –Vigo
Tfno: +34 986 812500 - Fax: +34 986 812401
http://webs.uvigo.es/x06/
E-mail: depx06@uvigo.es

An alternative inequality-based concentration measure*

Olga Alonso-Villar# and Coral del Río

Universidade de Vigo

Abstract

The aim of this paper is twofold. First, it shows the properties that regional economics is implicitly assuming when "relative" inequality measures, such as the Gini coefficient and the generalized entropy family of indexes, are used to quantify the geographic concentration of economic activity. Second, it proposes a new geographic concentration index that is based on an "absolute" inequality measure. This measure, which is additively decomposable, is also analyzed from an axiomatic point of view. Comparisons between these concentration indexes are illustrated by using manufacturing employment data in Spain.

JEL Classification: R12; D63

Keywords: Inequality measures; Segregation; Geographic concentration; Axioms

* Financial support from the *Xunta de Galicia* (*Programa de Estruturación de Unidades de Investigación en Humanidades e Ciencias Sociais* 2006/33 and INCITE08PXIB300005PR) and from FEDER is gratefully acknowledged.

Correspondence address: Universidade de Vigo; Facultade de CC. Económicas; Departamento de Economía Aplicada; Campus Lagoas-Marcosende s/n; 36310 Vigo; Spain. Tel.: +34 986812507; fax: +34 986812401; e-mail: ovillar@uvigo.es

1. Introduction

In recent years, researchers have shown increasing interest in the study of spatial concentration patterns of economic activity, both empirically and theoretically. This heightened interest is in part motivated by the general concern with the effects of economic integration processes on production location patterns, especially in Europe where the creation of the Single Market has stimulated the debate (Amiti, 1999; Haland et al., 1999; Brülhart, 2001; and Aiginger and Pfaffermayr (2004), inter alia).[1] Among the spatial concentration measures existing in the literature, those borrowed from the literature on income inequality are some of the most widely used.[2] In this regard, the locational Gini coefficient has been traditionally used for analyzing the spatial location patterns of manufacturing industries (Krugman, 1991; Amiti, 1999; Brülhart, 2001; and Suedekum, 2006, among many others), and lately, some of the indexes included in the generalized entropy family have been used as well (Brülhart and Traeger, 2005; Mori et al., 2005; Brakman et al., 2005).

There is a wide consensus in the literature on income distribution about the properties an inequality measure has to satisfy when it is used to compare income distributions having the same mean. Basically, one must invoke the symmetry axiom, which guarantees anonymity among individuals, and the Pigou-Dalton principle of transfers, which requires a transfer of income from a poorer to a richer person to increase inequality. However, if one is interested in comparing two income distributions that have different means, an additional property has to be specified, the one regarding the type of mean-invariance.[3] This requires introducing another judgment value into the analysis, and scholars have reached no agreement with respect to this matter. Some opt to invoke the scale invariance axiom, which stipulates that the inequality of a distribution remains unaffected when all incomes increase (or decrease) by the same proportion. This property gives rise to "relative" inequality measures such as the Gini index and the generalized entropy family of indexes, which are consistent with the Lorenz criterion. Others prefer, instead, to call on the translation invariance axiom, under which

---

[1] From a theoretical perspective, the literature of the new economic geography has contributed extensively to this debate. A review of this literature can be seen in Fujita et al. (2000), Neary (2001), and Ottaviano and Thisse (2004), among others.

[2] Other concentration measures proposed in the literature are formally derived from location models (Ellison and Glaeser, 1997; Maurel and Sédillot, 1999; Guimarães et al., 2007). There are also distance-based measures related to the literature on spatial statistics (Marcon and Puech, 2003; Duranton and Overman, 2005).

[3] Properties such as normalization, continuity, differentiability, and replication invariance are also commonly invoked, but they are more technical.

inequality remains unaltered if all incomes are augmented (or diminished) by the same amount, thereby giving rise to “absolute” inequality measures.

Certainly, in a context of income distribution, the properties of these inequality measures are well-known since this literature has tackled inequality measurement from an axiomatic perspective (Atkinson, 1970; Kolm, 1976; Shorrocks 1980; Cowell, 2000; and Zheng, 2007, inter alia). However, some of these measures have been extensively used to quantify the geographic concentration of the economic activity without an analysis of their properties in the new context. For this reason, it seems timely to reflect about the consequences of using inequality indexes to measure the spatial concentration of production depending on whether they satisfy one mean-invariance condition or the other.

The aim of this paper is twofold. First, it shows the properties that regional economics is implicitly assuming when using “relative” inequality measures to quantify the geographic concentration of economic activity. Second, it proposes a new geographic concentration index based on an “absolute” inequality measure, which is additively decomposable. These measures are analyzed from both an axiomatic and an empirical point of view. In doing so, a *relative* approach is followed, so that to measure the spatial concentration of a sector, its distribution across location units is compared with that of aggregate activity.[4]

For this purpose, we put on display the basic axioms behind inequality-based concentration measures. As far as we know, this is the first time that these axioms, borrowed from the literature on both income distribution and occupational segregation, are formally established in a location context. This approach allows us not only to show the properties that scholars in regional science are actually assuming when using the Gini index and the generalized entropy family for spatial analyses but also to propose a new geographic concentration measure that is additively decomposable. In particular, it reveals that according to the two former measures, concentration does not change when the level of aggregate activity and/or the level of the sector under consideration vary, so long as the weight that each location represents in both distributions remains unaltered. Consequently, any change in the level of economic activity of a sector does not affect its concentration degree, insofar as each location still accounts for the

[4] It is important to note that the labels “relative” and “absolute” used in the literature on income distribution do not have the same meaning as the labels *relative* and *absolute* in the literature on spatial concentration (Brülhart and Traeger, 2005).

original share of the sector. However, the new measure is concentration-invariant against any increase (or decrease) in the level of the sector so long as it is allocated among locations in such a way that each of them receives an amount proportional to the share that that location represents in terms of aggregate activity. We should be aware of the fact that the conclusions reached can be substantially affected by the type of concentration invariance assumed by the indexes. Consider for example that the economy consists of two locations, one with 500 jobs and the other with 1000. According to the locational Gini coefficient and the generalized entropy family, the concentration level of a sector having 1 worker in the first location and 10 in the second is the same as it would be if there were 10 workers in the first location and 100 in the second. Nevertheless, according to the new measure proposed in this paper, concentration in the latter case is higher than in the former. In order to show the empirical relevance of these discrepancies, comparisons between both types of indexes are illustrated by using Spanish manufacturing employment data for 2008.

The paper is structured as follows. Section 2 reflects, from an axiomatic perspective, on the implications of using inequality-based concentration indexes. Section 3 introduces a new concentration measure that satisfies a concentration-invariance condition that differs from that assumed by the locational Gini index and the generalized entropy family. As with the latter, this new concentration index is additively decomposable, which is a helpful property for empirical analysis. Using employment data of manufacturing industries in Spain, Section 4 first illustrates the differences between the aforementioned concentration measures, and, second, it analyzes the Spanish case in more detail by using the decompositions of the new index. Finally, Section 5 presents the main conclusions.

2. Geographic concentration: an axiomatic approach

2.1 Notation

Consider an economy with $L > 1$ location units (counties, regions, countries, etc.) across which aggregate employment, denoted by T, is distributed. Let $t \equiv (t_1, t_2, ..., t_L)$ denote this distribution, where $T = \sum_l t_l$. This distribution represents the benchmark against which the distribution of any sector is compared. This concentration notion is labeled *relative* and has been extensively used in empirical research. If a researcher is concerned, for example, with the geographic concentration of manufacturing industries, t could represent the distribution of manufacturing employment among regions (as in Amiti, 1999; and Brülhart, 2001). If a researcher is concerned with a broader perspective, t could instead represent the distribution of overall employment, services included (as in Brülhart and Traeger, 2005).

Let us denote by $x \equiv (x_1, x_2, ..., x_L)$ the employment distribution of the sector in which we are interested and by X its employment level ($X = \sum_l x_l$). In this paper, an index of geographic concentration is a function $I_c : D \to \mathbb{R}$ such that $I_c(x;t)$ represents the concentration level of the sector having distribution x when comparing it with the distribution of reference t, where $D = \bigcup_{L>1} \left\{ (x;t) \in \mathbb{R}^L{}_+ \times \mathbb{R}^L{}_{++} : x_l \le t_l \forall l \right\}$.

In order to better understand the axioms proposed below, first of all, we formally establish the relationship between the measurement of spatial concentration of economic activity and the measurement of income inequality. For that purpose, a hypothetical "income" distribution derived from vector $(x;t)$ is obtained. In doing so, in each location l, x_l is allocated in equal amounts among t_l workers. In other words, in each location, the variable of study (employment in the sector of study) is equally split among all individuals (both those working in the sector of study and those in the remaining sectors). This per capita employment level, $\frac{x_l}{t_l}$, represents the employment in the sector of study that corresponds to each individual in location l, and it plays the role of individual "income". Namely, the fictitious "income" distribution is constructed as follows: there are t_1 persons with an individual "income" of $\frac{x_1}{t_1}$, t_2 persons with an individual "income" of $\frac{x_2}{t_2}$, and so on. Therefore, we have built "income" distribution $y \equiv (\underbrace{\frac{x_1}{t_1}, ..., \frac{x_1}{t_1}}_{t_1 \text{ individuals}}, ..., \underbrace{\frac{x_L}{t_L}, ..., \frac{x_L}{t_L}}_{t_L \text{ individuals}})$ in a world of $T = \sum_l t_l$ individuals where total "income" is $X = \sum_l t_l \frac{x_l}{t_l}$.

Suppose, for example, that we want to measure the geographic concentration of the chemical sector by comparing its employment distribution across regions with that of manufacturing employment. Consider that the economy has three locations and that the employment distribution of the chemical industry among them is $(3,2,5)$, while the distribution of manufacturing workers is $(30,10,30)$. In other words, $(x;t)=(3,2,5;30,10,30)$. Therefore, our fictitious "income" distribution would be one with 70 people having a total income of 10 units: there are 30 people with an individual "income" of 0.1, 10 people with an individual "income" of 0.2, and 30 people with an individual "income" of 0.6, i.e., the "income" distribution is equal to $y \equiv \left(\underbrace{\frac{3}{30},...,\frac{3}{30}}_{30}, \underbrace{\frac{2}{10},...,\frac{2}{10}}_{10}, \underbrace{\frac{5}{30},...,\frac{5}{30}}_{30} \right)$.

The parallelism between employment distribution $(x;t)$ and hypothetical "income" distribution y will be helpful for understanding the axiomatic framework presented in what follows, where some basic axioms, borrowed from the literature on income distribution and occupational segregation, are adapted to analyze spatial concentration measures.

2.2 Basic properties

We can start our list by proposing a symmetry axiom in the spatial concentration context. We call this axiom *symmetry in locations,* which means that if locations are enumerated in a different order, the concentration index should remain unchanged.[5]

Axiom 1: *Symmetry in locations*. If $(\Pi(1),...,\Pi(L))$ represents a permutation of locations, then $I_c(x\Pi;t\Pi)=I_c(x;t)$, where $x\Pi=\left(x_{\Pi(1)},...,x_{\Pi(L)}\right)$ and $t\Pi=\left(t_{\Pi(1)},...,t_{\Pi(L)}\right)$.

As mentioned earlier, another basic axiom of any inequality measure is the Pigou-Dalton principle, which requires a transfer of income from a poorer to a richer person to increase inequality. This property gives rise to our next axiom: *movement between locations.* If we focus again on the chemical sector, this property requires that when a region with a lower employment level in chemicals than another (but with the same manufacturing employment)

[5] In the income distribution literature this axiom requires that the inequality index does not change when individuals' incomes swap. In the occupational segregation literature, this axiom is called "symmetry in groups" and requires anonymity among occupations (see Hutchens, 1991).

loses employment in chemicals in favor of the other location, concentration in the chemicals sector must increase.[6]

Axiom 2: *Movement between locations*. If $\left(x';t\right)\in D$ is obtained from $\left(x;t\right)\in D$ in such a way that:

(i) location *i* loses employment in the sector of study, while the opposite happens to location *h*, i.e., $x'_i = x_i - d$, $x'_h = x_h + d$ $(0 < d \le x_i)$, where *i* and *h* are two locations with the same aggregate employment level, $t_i = t_h$, but with different shares in the sector of study since $x_i < x_h$;

(ii) the employment level of the sector of study does not change in the remaining locations, i.e., $x'_l = x_l \quad \forall l \ne i,h$;

then $I_c\left(x';t\right) > I_c\left(x;t\right)$.

The next axiom we present, *insensitivity to proportional subdivisions of locations,* is not borrowed from the literature on income distribution but from that on occupational segregation.[7] This axiom requires that subdividing a location into several units of equal size, both in terms of aggregate employment and in terms of employment in the sector of study, does not affect the concentration level of the sector. Without loss of generality, the subdivision in the next axiom is written for the last location in order to make notation easier.

Axiom 3: *Insensitivity to proportional subdivisions of locations.* If $\left(x';t'\right)\in D$ is obtained from $\left(x;t\right)\in D$ in such a way that:

(i) all locations except the last one remain unaltered both in terms of aggregate employment and employment in the sector of study, i.e., $t'_l = t_l$ and $x'_l = x_l$ for any $l = 1,...,L-1$;

(ii) the last location is subdivided in M location units without introducing any differences among them in terms of employment shares, i.e., $x'_l = x_L/M$, $t'_l = t_L/M$ for any $l = L,...,L+M-1$,

then, $I_c\left(x';t'\right) = I_c\left(x;t\right)$.

---

[6] This axiom has also been adapted to measure occupational segregation, where it is called “movement between groups” (see Hutchens, 2004; and Alonso-Villar and Del Río, 2008b).
[7] The corresponding axiom is named “insensitivity to proportional divisions” (see Hutchens, 2004).

In order to understand the relevance of the above axiom, we go back to the example given at the beginning of Section 2. Note that "income" distribution $y \equiv \left(\frac{3}{30},...,\frac{3}{30},\frac{2}{10},...,\frac{2}{10},\frac{5}{30},...,\frac{5}{30}\right)$ can be obtained from different $(x;t)$ vectors, depending on how the "income" data are grouped. We could, for example, group the "income" data into three groups, $\left(\underbrace{\frac{3}{30},...,\frac{3}{30}}_{group1(30)},\underbrace{\frac{2}{10},...,\frac{2}{10}}_{group2(10)},\underbrace{\frac{5}{30},...,\frac{5}{30}}_{group3(30)}\right)$, so that the 30 individuals having an "income" level equal to 0.1 are in group 1, the 10 individuals having an "income" of 0.2 are in group 2, and the 30 individuals having an "income" of 0.6 are in group 3. In this case, we would obtain former vector $(x;t)=(3,2,5;30,10,30)$. But we could also group individuals into five groups, $\left(\underbrace{\frac{3}{30},...,\frac{3}{30}}_{group1(10)},\underbrace{\frac{3}{30},...,\frac{3}{30}}_{group2(10)},\underbrace{\frac{3}{30},...,\frac{3}{30}}_{group3(10)},\underbrace{\frac{2}{10},...,\frac{2}{10}}_{group4(10)},\underbrace{\frac{5}{30},...,\frac{5}{30}}_{group5(30)}\right)$, so that 10 of the individuals having an "income" of 0.1 are included in the first group, 10 are in group 2, and the remaining 10 are in the third group, while those having an "income" of 0.2 are included in the fourth group, and those with an "income" of 0.6 are in the fifth group. In this case, $(x';t')=(1,1,1,2,5;10,10,10,10,30)$. Note that, according to axiom 3, both $(x;t)$ and $(x';t')$ have the same concentration level since the latter can be obtained from the former by a proportional subdivision of locations. As a consequence of axiom 3, the value of the concentration index evaluated at any employment distribution $(x;t)$ does not change so long as the corresponding hypothetical "income" distribution y remains unaltered.[8]

Another standard assumption of inequality indexes is the replication invariance or population principle, which allows comparisons among income distributions having different population sizes. This axiom requires that when replicating the economy k-times, so that for every individual in the previous economy there are now *k* identical individuals, income inequality is

---

[8] Note that axiom 2 together with axiom 3 allows one to conclude that if $\frac{x_i}{t_i}<\frac{x_h}{t_h}$, $x'_i = x_i - d$, $x'_h = x_h + d \quad (0<d\le x_i)$, and $x'_l = x_l \quad \forall l \neq i,h$, then $I_c(x';t) > I_c(x;t)$, even when $t_i \neq t_h$. The explanation is that, as a consequence of axiom 3, locations can be subdivided, without affecting concentration, in such a way that the distribution of reference becomes eventually equal to $\left(\underbrace{1,...,1}_{T}\right)$. Therefore, the movements of the sector among locations involve new locations having the same aggregate employment.

not altered. This principle is adapted to our context in order to make it possible to compare economies with different numbers of location units.

Axiom 4: *Population Principle*. If $(x';t')$ is a k-replication of distribution $(x;t)$ so that for any initial location unit l there are now k identical locations where the employment level of the sector of study and the aggregate employment in each of them is equal to x_l and t_l, respectively, then $I_c(x';t') = I_c(x;t)$.

Note that axioms 3 and 4 imply that $I_c(ax;at) = I_c(x;t)$ for any $a > 0$. Therefore, any concentration index satisfying these two properties is cardinally unaffected by the unit of measurement (if employment is measured either in hundreds or thousands of individuals, the index does not change).[9]

The next axiom, scale invariance, is not satisfied by all inequality indexes, since there is no consensus in the literature with respect to this matter. Only the "relative" inequality measures (such as the Gini index and the generalized entropy family of indexes) satisfy it. As mentioned above, it requires that inequality remains constant when multiplying all incomes by the same positive scalar. This axiom is adapted to our context as follows.

Axiom 5: *Scale Invariance*. If the distribution of the sector of study, x, is multiplied by a positive scalar, a, and the distribution of reference, t, is multiplied by another positive scalar, b, in such a way that $ax_l \leq bt_l \;\; \forall l$, then the concentration level of the sector does not change, i.e., $I_c(ax;bt) = I_c(x;t)$.

This property means that the value of the concentration index should not change when the employment level of the distribution of reference and/or that of the sector under consideration vary, so long as the weight that each location represents in distributions t and x remains unaltered. In other words, if manufacturing employment in each location doubles and that of chemicals triples, provided that both facts are compatible, the spatial concentration of the

[9] A weaker property, where rankings of distributions are ordinally unaffected by the unit of measurement, has been recently invoked in the literature on income distribution where it is called unit consistency (Zheng, 2007).

chemical industry should not change. Therefore, this axiom means that in measuring spatial concentration, it is only employment shares that matter, not employment levels.[10]

However, as mentioned above, some scholars opt to invoke the translation invariance axiom rather than the scale invariance axiom so that inequality remains, instead, unaltered if all incomes are augmented (or diminished) by the same amount. This property gives rise to "absolute" inequality measures. This axiom is formally defined in our context in what follows.

Axiom 5': *Translation Invariance*. If employment in the sector of study increases (or decreases) in such a way that the change, a, is distributed across locations according to their employment weights in the distribution of reference, i.e., $(x';t)=\left(x_1+a\frac{t_1}{T},...,x_L+a\frac{t_L}{T};t\right)$, and $x_l+a\frac{t_l}{T}\le t_l\ \forall l$, then the concentration level of the sector should not change, i.e., $I_c\left(x';t\right)=I_c\left(x;t\right)$.[11]

As a consequence of the above axiom, if employment in the chemicals industry increases, and this surplus is distributed among locations in such a way that if in a location aggregate manufacturing employment doubles that of another location, the former location receives twice as much of the extra employment in chemicals as the latter, then, the spatial concentration of the chemical industry should not change.

It follows, therefore, that the *translation invariance* axiom and the *scale invariance* axiom differ regarding the type of increments in the sector of study that are considered to be concentration-invariant. Consider, for instance, that the distribution of aggregate manufacturing employment among the three locations of the economy is $t\equiv\left(100,200,500\right)$,

[10] In a context of occupational segregation, a similar axiom has been proposed by Alonso-Villar and Del Río (2008b).

[11] Equivalently, the concentration index does not vary if distribution t, rather than x, varies in such a way that $(x;t')=\left(x;t_1-\frac{a\frac{t_1}{T}}{\frac{x_1}{t_1}+\frac{a}{T}},...,t_L-\frac{a\frac{t_L}{T}}{\frac{x_L}{t_L}+\frac{a}{T}}\right)$.

and the distribution of the chemical industry is $x \equiv (1,10,29)$. Any concentration index satisfying axiom 5 would reach the same value at distributions x and $x' \equiv (10,100,290)$, the latter resulting from allocating the employment surplus of the chemical industry (360 jobs) according to the employment shares in distribution x (i.e., $\left(\frac{x_1}{X},\frac{x_2}{X},\frac{x_3}{X}\right) = (0.025, 0.25, 0.725)$). However, when using any index satisfying axiom 5', the extra employment of the chemical industry must be allocated across locations according to employment shares in distribution t (i.e., $\left(\frac{t_1}{T},\frac{t_2}{T},\frac{t_3}{T}\right) = (0.125, 0.25, 0.625)$) in order to keep concentration unaltered. In other words, in this case, distribution x has the same concentration level as distribution $x'' \equiv (46,100,254)$.[12]

It is important to know which type of concentration invariance we prefer using and, therefore, which measure we should single out in order to measure the spatial concentration of production, since results can vary considerably.

3. Inequality-based concentration measures: A new proposal

In this section, we first show that the locational Gini coefficient and the generalized entropy family of indexes used in the literature to analyze geographic concentration satisfy axioms 1-5. Second, we propose a new concentration measure which satisfies axioms 1, 2, 3, 4, and 5'.

[12] Note that in order to keep concentration unaltered when moving from distribution x to x'', axiom 5' requires that the employment shares in the latter distribution are closer to those of distribution t than employment shares in the former. In our example, this means that $x'' \equiv (46,100,254)$ Lorenz dominates distribution $x \equiv (1,10,29)$, but this does not necessarily have to be the case. If the distribution of reference were instead $\tilde{t} \equiv (10,200,590)$, which is Lorenz dominated by $x \equiv (1,10,29)$, we would obtain that x Lorenz dominates $\tilde{x}'' \equiv (5.5,100,294.5)$, which implies the latter has higher "relative" inequality. This is an important difference with respect to "absolute" inequality indexes, since the invariance condition in that context implies that if average income increases, "relative" inequality must necessarily to decrease when keeping "absolute" inequality as constant.

3.1 Previous indexes

The locational Gini index

The Gini index of a given sector can be written as the sum of the differences between the employment shares of the sector in each pair of locations weighted by their demographic weights, divided by twice the employment share of the sector in the whole economy:

$$G = \frac{\sum_{l,l'} \frac{t_l}{T}\frac{t_{l'}}{T}\left| \frac{x_l}{t_l} - \frac{x_{l'}}{t_{l'}} \right|}{2\frac{X}{T}}.$$

It is easy to see that this index satisfies axioms 1, 3, 4 and 5. In order to see that it also satisfies axiom 2, note that if $x_i < x_h$ and $t_i = t_h$, any disequalizing movement of employment in the sector of study from location *i* to location *h* would make the index increase since:

a) $\left| \frac{x_i - d}{t_i} - \frac{x_h + d}{t_h} \right| = \frac{x_h + d}{t_h} - \frac{x_i - d}{t_i} > \frac{x_h}{t_h} - \frac{x_i}{t_i} = \left| \frac{x_i}{t_i} - \frac{x_h}{t_h} \right|$;

b) If l $(l \neq i,h)$ is such that $\frac{x_l}{t_l} < \frac{x_i}{t_i}$ or $\frac{x_l}{t_l} > \frac{x_h}{t_h}$, then any change in component $\left| \frac{x_l}{t_l} - \frac{x_h + d}{t_h} \right|$ is exactly offset by a change of the same magnitude and opposite direction in component $\left| \frac{x_i - d}{t_i} - \frac{x_l}{t_l} \right|$;

c) If l $(l \neq i,h)$ is such that $\frac{x_i}{t_i} \leq \frac{x_l}{t_l} \leq \frac{x_h}{t_h}$, then no component decreases.

The generalized entropy family

The generalized entropy family of inequality indexes gives rise to the next family of concentration measures:

$$\Psi_\alpha(x;t) = \begin{cases} \frac{1}{\alpha(\alpha-1)} \sum_l \frac{t_l}{T}\left[\left(\frac{x_l/X}{t_l/T}\right)^\alpha - 1\right] & \text{if } \alpha \neq 0,1 \\ \sum_l \frac{x_l}{X} \ln\left(\frac{x_l/X}{t_l/T}\right) & \text{if } \alpha = 1 \end{cases}$$

where α is a sensitivity parameter.[13] It is straightforward to prove that this family satisfies axioms 1, 3, 4 and 5. To demonstrate that Ψ_α satisfies the axiom of *movement between locations*, note that:

a) $$I_\alpha(y)=\begin{cases}\dfrac{1}{n\alpha(\alpha-1)}\sum_i\left[\left(\dfrac{y_i}{\frac{1}{n}\sum_k y_k}\right)^\alpha-1\right] & \text{if } \alpha\neq 0,1\\ \dfrac{1}{\mathrm{n}}\sum_i\left[\dfrac{y_i}{\frac{1}{n}\sum_k y_k}\ln\left(\dfrac{y_i}{\frac{1}{n}\sum_k y_k}\right)\right] & \text{if } \alpha=1\end{cases}$$ are inequality indexes belonging to the generalized entropy family.

b) Any disequalizing movement from location *i* to *h*, where $t_i=t_h$ and $x_i<x_h$, implies moving from "income" distribution $y=\left(\frac{x_1}{t_1},...,\frac{x_1}{t_1},...,\frac{x_i}{t_i},...,\frac{x_i}{t_i},...\frac{x_h}{t_h},...,\frac{x_h}{t_h},...,\frac{x_L}{t_L},...,\frac{x_L}{t_L}\right)$ to $y'=\left(\frac{x_1}{t_1},...,\frac{x_1}{t_1},...,\frac{x_i-d}{t_i},...,\frac{x_i-d}{t_i},...,\frac{x_h+d}{t_h},...,\frac{x_h+d}{t_h},...,\frac{x_L}{t_L},...,\frac{x_L}{t_L}\right)$.

c) $\Psi_\alpha(x;t)=I_\alpha(y)$.

Since I_α is an inequality measure satisfying the Pigou-Dalton transfer principle and *y'* can be obtained from *y* by a finite sequence of regressive transfers, it follows that $\Psi_\alpha(x';t')>\Psi_\alpha(x;t)$.

An advantage of these concentration indexes is that they can be additively decomposed, which is useful for undertaking detailed analyses of spatial patterns (Brülhart and Traeger, 2005; Alonso-Villar and Del Río, 2008a). Consider, for example, that our location units are the European regions, and that they are grouped by country. Given a classification of regions into *K* countries, we can write $(x;t)=(x^1,...,x^K;t^1,...,t^K)$, where x^k denotes the employment distribution of the sector across regions in country *k*, and t^k is the distribution of aggregate

[13] If we had considered concentration indexes defined on the space of employment distributions $(x;t)$ where all components of vector x were strictly positive, rather than positive, then another index could be defined: $\Psi_\alpha(x;t)=\sum_l\frac{t_l}{T}\ln\left(\frac{t_l/T}{x_l/X}\right)$ if $\alpha=0$.

employment across regions in that country ($k = 1,...,K$). Then, index $\Psi_\alpha(x;t)$ can be decomposed as follows:

$$\Psi_\alpha(x;t) = \sum_k \left(\frac{X_k}{X}\right)^\alpha \left(\frac{T_k}{T}\right)^{1-\alpha} \Psi_\alpha(x^k;t^k) + \Psi_\alpha(X_1,...,X_K;T_1,...,T_k)$$

where X_k is the employment level of the sector in country *k*, T_k is the aggregate employment level in that country, and $(X_1,...,X_K;T_1,...,T_K)$ represents the distributions of the sector of study and that of aggregate employment across countries. The first addend of the above formula represents the *within* component, i.e., the weighted sum of regional concentration inside each country, while the second addend reflects the *between* component, i.e., the concentration that would exist if there were no regional disparities within countries but only among countries.[14]

3.2 A new concentration index

We should notice that the above concentration indexes are not the only decomposable indexes that can be defined by extending inequality measures. In fact, the variance is another inequality measure that can be additively decomposed.[15] However, we should keep in mind that the value judgments behind this inequality measure differ from those of the generalized entropy family. In fact, the variance is an inequality index that does not satisfy the scale invariance axiom but the translation invariance axiom, which means that inequality remains unchanged if all incomes are augmented (or diminished) by the same amount and not if all incomes increase proportionally. This index can be adapted to measure concentration as follows:

$$\Phi\left(x;t\right) = \sum_l \frac{t_l}{T}\left[\frac{x_l}{t_l} - \frac{X}{T}\right]^2 .$$

It is easy to see that concentration index $\Phi\left(x;t\right)$ satisfies axioms 1, 3, and 4. In order to prove that it satisfies axiom 2, note that function $f(z) = z^2$ is an increasing convex function, which implies that $\frac{t_i}{T}\left[\frac{x_i - d}{t_i} - \frac{X}{T}\right]^2 + \frac{t_h}{T}\left[\frac{x_h + d}{t_h} - \frac{X}{T}\right]^2 > \frac{t_i}{T}\left[\frac{x_i}{t_i} - \frac{X}{T}\right]^2 + \frac{t_h}{T}\left[\frac{x_h}{t_h} - \frac{X}{T}\right]^2$ if *i* and *h* are

[14] Index Ψ_2 can also be decomposed by subsectors, so that the contribution of each subsector to the concentration of the sector can be determined (see Brülhart and Traeger, 2005).

[15] For other good properties of this index, see Chakravarty (2001) and Zheng (2007).

two locations such that $x_i < x_h$ and $t_i = t_h$. Therefore, if $(x';t')$ is obtained from $(x;t)$ through a disequalizing movement of the type described in axiom 2, then $\Phi(x';t') > \Phi(x;t)$.

Finally, note that $\Phi(x;t)$ satisfies axiom 5' since $\Phi\left(x_1 + a\frac{t_1}{T},...,x_L + a\frac{t_L}{T};t\right) = \sum_l \frac{t_l}{T}\left[\frac{x_l + a\frac{t_l}{T}}{t_l} - \frac{X+a}{T}\right]^2 = \sum_l \frac{t_l}{T}\left[\frac{x_l}{t_l} - \frac{X}{T}\right]^2 = \Phi(x;t)$.

Observe that even though index $\Phi(x;t)$ has been obtained by extending an "absolute" inequality measure, it is actually a *relative* concentration measure since it quantifies how much the distribution of the sector across locations, x, departs from the distribution of reference, t.

Since index $\Phi(x;t)$ can be rewritten as

$$\Phi(x;t) = \left(\frac{X}{T}\right)^2 \sum_l \frac{t_l}{T}\left(\frac{\frac{x_l}{X} - \frac{t_l}{T}}{\frac{t_l}{T}}\right)^2,$$

this measure not only depends on how much $\frac{x_l}{X}$ departs from $\frac{t_l}{T}$ (as the entropy family) but also on the weight that the sector of study represents in the economy, $\frac{X}{T}$. In other words, this measure takes into account not only the distribution of the sector across locations but also its level.

An advantage of the aforementioned index is that it is additively decomposable. Following the decomposition used in the literature on income distribution (Chakravarty, 2001), index $\Phi(x;t)$ can be decomposed as

$$\Phi(x;t) = \sum_k \frac{T_k}{T}\Phi(x^k;t^k) + \Phi(X_1,...,X_K;T_1,...,T_k), \quad (1)$$

when location units (regions, for example) are grouped into K classes (countries, for example). This decomposition of total concentration in the within (first addend) and between (second addend) components is analogous to the one corresponding to the generalized entropy family. Therefore, the concentration of the sector can be written as the weighted sum of regional concentration inside each country (according to its demographic weight) plus the concentration that would exist if there were no regional disparities within countries but only among countries.

On the other hand, if the sector of study is decomposed into several mutually exclusive subsectors, $s = 1,...,S$, then $\Phi(x;t)$ can also be decomposed by subsectors since

$$\Phi(x;t) = \sum_{s} \Phi(x^{s};t) + 2\sum_{s=1}^{S}\sum_{s'>s} \tilde{\Phi}(x^{s};x^{s'};t), \qquad (2)$$

where $\tilde{\Phi}(x^{s};x^{s'};t) = \sum_{l} \frac{t_l}{T}\left(\frac{x_l^{s}}{t_l} - \frac{X^{s}}{T}\right)\left(\frac{x_l^{s'}}{t_l} - \frac{X^{s}}{T}\right)$ represents the covariance between the distributions of subsectors s and s' across locations.

The first addend of the former decomposition is a summary of the internal contribution of each subsector to the overall concentration of the sector, while the second addend involves the concentration due to locational interdependencies among subsectors.

From all of the above, it follows that apart from the concentration measures derived from the generalized entropy family, other decomposable measures satisfying alternative invariance properties can be used to determine the spatial concentration of economic activity. It is important to reflect about the type of concentration-invariance condition the researcher finds more suitable and, therefore, which measure she/he choses to quantify the spatial concentration of production, since empirical conclusions can substantially differ.

4. The geographic concentration of manufacturing in Spain

The data used in this paper comes from the Labor Force Survey (*EPA*) conducted by the Spanish Institute of Statistics (*INE*) by following EUROSTAT’s guidelines. Our data corresponds to the second quarter of the year 2008. Manufacturing industries are considered at a two- and three- digit level of the National Classification of Economic Activities (*CNAE-1993 Rev1*).

4.1 Comparisons between Φ and other inequality-based indexes

Table 1 shows the concentration level of the manufacturing industries in 2008 at a two-digit level (23 industries) by comparing the employment distribution of each sector across provinces with the distribution of total manufacturing employment. For this purpose several *relative* concentration measures are used: Φ, Ψ_α with $\alpha = 1, 2$, and G. Industries are ranked in descending order according to index Φ.

MANUFACTURING INDUSTRIES (two-digit) 2008	**Number of three-digit industries (*S*)**	Φ	Ψ_1	Ψ_2	G	$\frac{X}{T}$ **(%)**
15 *Manufacture of food products and beverages*	9	0.00626	0.11	0.12	0.26	15.90
26 *Manufacture of other non-metallic ore products (Glass, ceramic products, bricks, tiles, cement, etc.)*	8	0.00412	0.25	0.43	0.33	6.88
34 *Manufacture of motor vehicles, trailers and semi-trailers*	3	0.00393	0.39	0.35	0.47	7.45
24 *Chemical industry*	7	0.00197	0.22	0.22	0.37	6.66
27 *Metallurgy*	5	0.00196	0.46	0.68	0.49	3.80
19 *Preparation, tanning and dressing of leather; manufacture of leather goods and luggage articles*	3	0.00195	1.63	3.72	0.83	1.62
35 *Manufacture of other transport material (Ships, railway material, aircraft, bicycles, motorcycles, etc.)*	5	0.00164	0.68	1.17	0.59	2.65
29 *Machinery and mechanical equipment construction industry*	7	0.00155	0.11	0.11	0.24	8.23
22 *Publishing, graphic arts, and reproduction of recorded supports*	3	0.00153	0.21	0.22	0.36	5.85
36 *Manufacture of furniture; other manufacturing industries (Jewelry, musical instruments, sport articles, toys, etc.)*	6	0.00110	0.09	0.10	0.23	7.45
28 *Manufacture of metal products, except machinery and equipment*	7	0.00090	0.03	0.03	0.13	12.53
18 *Clothing and fur industry*	3	0.00080	0.32	0.48	0.39	2.90
17 *Textile industry*	7	0.00076	0.42	0.45	0.48	2.90
20 *Wood and cork industry, except furniture; basket making and wickerwork*	5	0.00071	0.21	0.26	0.35	3.71
25 *Manufacture of rubber and plastic products*	2	0.00050	0.23	0.21	0.35	3.43
31 *Manufacture of electrical machinery and material*	6	0.00036	0.20	0.20	0.34	2.98
23 *Manufacture of coke, refinement of petroleum and treatment of nuclear fuels*	3	0.00013	1.20	2.02	0.76	0.57
21 *Paper industry*	2	0.00013	0.28	0.30	0.37	1.47
32 *Manufacture of electronic material; manufacture of radio, television and communications apparatus*	3	0.00011	0.51	0.50	0.52	1.06
33 *Manufacture of medical-surgical, precision and optical equipment and instruments, and clocks and watches*	5	0.00011	0.44	0.39	0.49	1.16
16 *Tobacco industry*	1	0.00008	2.67	14.95	0.93	0.17
30 *Manufacture of office machines and IT equipment*	1	0.00002	1.22	1.64	0.76	0.26
37 *Recycling*	2	0.00002	0.70	0.67	0.59	0.36

Table 1: *Relative* concentration indexes and weights of two-digit sectors in 2008.

We observe that according to index Φ (which satisfies axiom 5'), *Manufacture of food products and beverages* (**15**), *Manufacture of other non-metallic ore products* (**26**), and *Manufacture of motor vehicles, trailers and semi-trailers* (**34**) are the most concentrated sectors, followed at a certain distance by the *Chemical industry* (**24**), *Metallurgy* (**27**), and *Manufacture of leather goods* (**19**). Therefore, among the highly concentrated sectors, we find both high- and medium-high-technology industries (**34** and **24**), and low- and medium-low-technology industries (**15**, **19**, **26** and **27**).[16]

On the other hand, *Recycling* (**37**), *Manufacture of office machines and IT equipment* (**30**), the *Tobacco industry* (**16**), *Manufacture of medical-surgical, precision and optical equipment and instruments, and clocks and watches* (**33**), *Manufacture of electronic material; manufacture of radio, television and communications apparatus* (**32**), *Paper industry* (**21**)**,** and *Manufacture of coke, refinement of petroleum and treatment of nuclear fuels* (**23**) are among the less concentrated sectors. Therefore, low- (**16**, **21,** and **37**), medium-low- (**23**) and high-technology sectors (**30**, **32**, and **33**) appear as little concentrated.

Regarding the locational Gini coefficient and the members of the entropy family, we see that they behave similarly to each other, which seems reasonable since these indexes satisfy the same concentration-invariance condition (axiom 5). The results obtained with these measures substantially differ, however, from those obtained before. The most concentrated sectors are now the *Tobacco industry* (**16**), *Manufacture of coke, petroleum, and nuclear fuels* (**23**), *Manufacture of office machines and IT equipment* (**30**), and a sector also considered highly concentrated according to index Φ: *Manufacture of leather goods* (**19**).

Note that sectors **16, 23,** and **30** are some of the less concentrated industries according to index Φ. If we plot the distribution of sector **16** across locations and compare it with that of sector **19** (Figure 1)**,** we can understand why index Φ classifies them in different concentration levels, while for the remaining indexes both sectors behave rather similarly: Absolute levels are relevant for concentration measures satisfying axiom 5', while they are not important for measures satisfying axiom 5. In fact, according to the Gini index, in sector **16** concentration is a little higher than in sector **19**, in spite of the fact that the latter shows much higher discrepancies among locations in terms of employment levels.

[16] For a classification of manufacturing industries based on technology, see OECD (2007), annex 1.1.

Observe, however, that index Φ not only depends on the weight of the sector but also on its distribution across locations, as mentioned in Section 3. In fact, the *Manufacture of metal products, except machinery and equipment* (**28**) is a sector with an important weight in the economy. (It represents around 12.5% of total manufacturing employment; see last column in Table 1). However, index Φ does not classify it among those that are the most concentrated. On the other hand, in spite of the small weight that sector **19** represents in terms of manufacturing employment (1.6%), its concentration level is rather high. The fact that indexes that focus on different aspects of the employment distribution classify this sector in the group of the most concentrated sectors suggests that the agglomeration of this industry is a rather robust result.

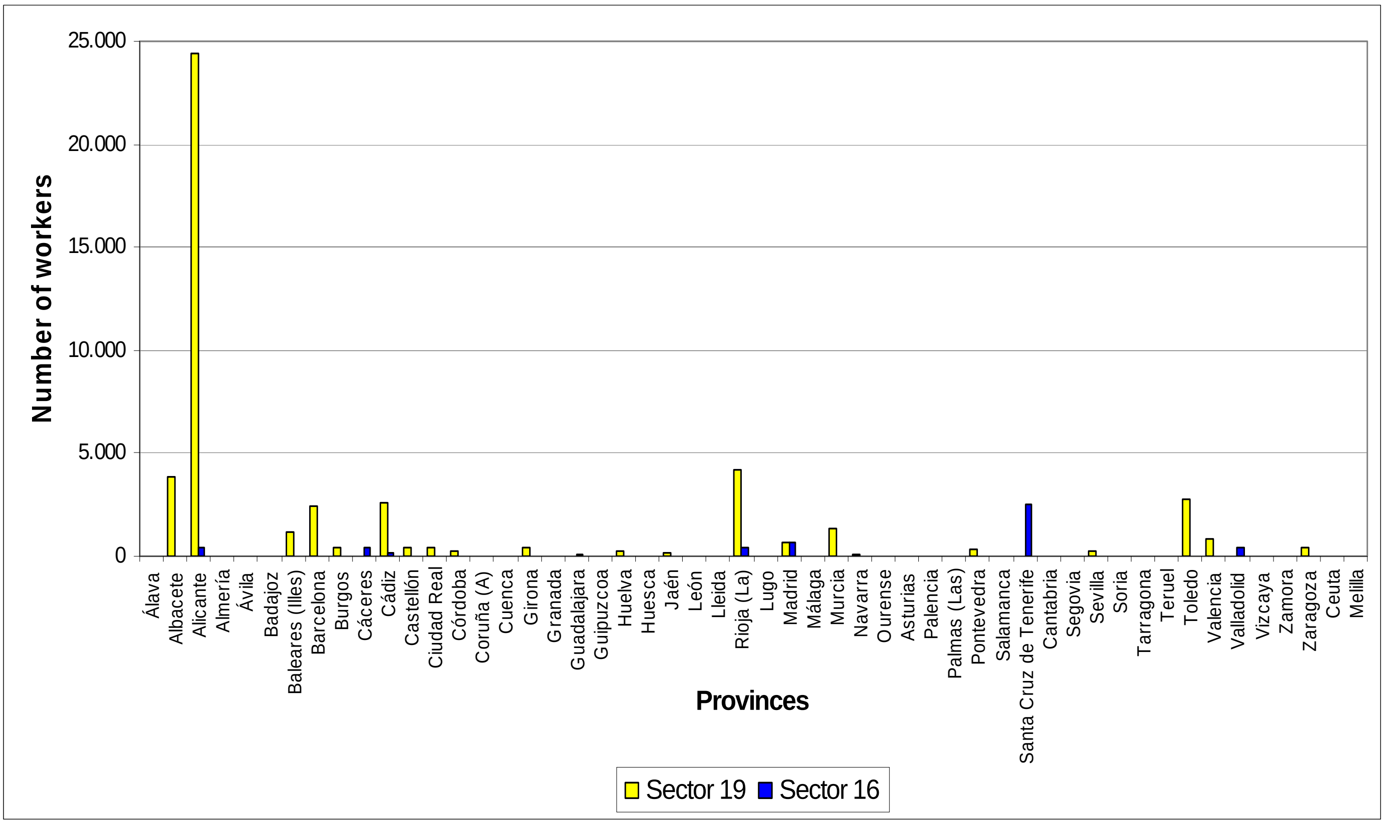

Figure 1. Employment distributions of sectors **19** and **16** across Spanish provinces in 2008.

One would expect to find a higher concentration level in high-tech industries, since knowledge spillovers can be an important source of agglomeration externalities in this kind of sector. However, from the information shown above, we cannot conclude that this happens in Spain, neither according to index Φ nor according to the locational Gini coefficient and the generalized entropy family. A similar conclusion was pointed out by Devereux et al. (2004) in the UK by using the index proposed by Maurel and Sédillot (1999).

Next, we calculate the concentration level of manufacturing industries at a three-digit level (103 industries) in order to see if there are important internal discrepancies. Tables 2 and 3 show the 20 most-concentrated and the 20 least-concentrated subindustries, respectively.

MANUFACTURING INDUSTRIES (Three-digit) **The 20 most-concentrated sectors** **2008**	Φ	Ψ_1	Ψ_2	G	$\frac{X}{T}$ **(%)**
263 *Manufacture of ceramic tiles and flags*	0.0030735	2.65	13.44	0.90	1.07
193 *Footwear*	0.0016416	2.31	5.66	0.92	1.20
341 *Manufacture of motor vehicles*	0.0015915	0.68	0.69	0.60	3.39
151 *Meat industry*	0.0015835	0.42	0.66	0.47	3.46
351 *Building and repairing of ships and boats*	0.0012945	1.57	3.33	0.83	1.39
343 *Manufacture of non-electric parts, pieces and accessories of motor vehicles*	0.0010268	0.45	0.39	0.47	3.61
182 *Tailoring of clothing and accessories*	0.0008374	0.41	0.62	0.45	2.61
361 *Manufacture of furniture*	0.0008288	0.09	0.10	0.24	6.35
271 *Iron and steel industry*	0.0007253	0.62	1.14	0.55	1.78
281 *Manufacture of metal products for construction*	0.0007253	0.07	0.08	0.21	6.85
153 *Preparation and preservation of vegetables and fruits*	0.0006151	1.04	1.38	0.73	1.50
158 *Other food products* *(Bread and pastry; sugar, cocoa and sweet products, etc.)*	0.0005674	0.08	0.09	0.22	5.60
152 *Preparation and preservation of fish*	0.0005520	1.34	2.94	0.77	0.97
244 *Manufacture of pharmaceutical products*	0.0005460	0.54	0.53	0.55	2.28
221 *Publishing*	0.0005048	0.42	0.45	0.49	2.37
222 *Graphic arts*	0.0004513	0.19	0.19	0.33	3.46
203 *Manufacture of wooden structures and carpentry pieces for construction*	0.0004385	0.31	0.48	0.41	2.15
252 *Manufacture of plastic products* *(Rubber inner tubes and plates, plastic packing items, etc.)*	0.0003226	0.27	0.24	0.39	2.58
241 *Manufacture of basic chemical products* *(Industrial gases, pigments, fertilizers, and nitrogenous compounds, etc.)*	0.0003186	0.40	0.49	0.48	1.80
155 *Dairy industries*	0.0002767	0.52	0.75	0.53	1.36

Table 2: *Relative* concentration indexes of the 20 most-concentrated sectors according to Φ at three-digit level and weight of each of them in 2008.

We observe that the values of index Φ when considering the manufacturing industries at a three-digit level are lower than those obtained at a two-digit level (compare Tables 2 and 3 with Table 1). On the other hand, the most-concentrated subindustries belong mainly to industries with either high levels of concentration (as is the case in sectors **15**, **26**, **34**, **24**, **27**, and **19**) or intermediate levels (**35**, **22**, **36**, **28**, **18,** and **20**). In particular, we find that five out of the nine subindustries included in sector **15** are among the most concentrated ones, while in sector **34**, this happens for two out of three subindustries. In some sectors, we find, however, remarkable heterogeneity regarding the concentration levels of their respective subindustries. Thus, sectors **19**, **24**, **26,** and **35** have some of the most- and the least-concentrated subindustries. The former are, respectively, *Footwear* (**193**); *Manufacture of pharmaceutical* *and basic chemical products* (**244** and **241**);[17] *Manufacture of ceramic tiles and flags* (**263**); and *Building and repairing of ships and boats* (**351),** while the latter include *Preparation, tanning and dressing of leather* (**191**); *Manufacture of pesticides and other agro-chemical products* and *Manufacture of man-made and synthetic fibres* (**242** and **247**); *Manufacture of several non-metallic ore products* (**268**); and *Manufacture of other transport material* (**355)**.

MANUFACTURING INDUSTRIES (Three-digit) **The 20 least-concentrated sectors** **2008**	Φ	Ψ_1	Ψ_2	G	$\frac{X}{T}$ **(%)**
371 *Recycling of scrap metal and metal waste*	0.0000074	1.34	2.07	0.78	0.13
365 *Manufacture of games and toys*	0.0000071	0.89	0.84	0.66	0.21
173 *Textile finishing*	0.0000070	0.80	0.86	0.63	0.20
314 *Manufacture of accumulators and electric batteries*	0.0000061	2.06	4.58	0.89	0.08
323 *Manufacture of receivers, sound and image reproduction, and recording equipment*	0.0000061	0.82	0.79	0.64	0.20
283 *Manufacture of steam generators*	0.0000047	2.03	3.55	0.89	0.08
268 *Manufacture of several non-metallic ore products*	0.0000042	1.13	1.64	0.75	0.11
364 *Manufacture of sports articles*	0.0000039	1.02	1.33	0.68	0.12
334 *Manufacture of optical instruments and photographic equipment*	0.0000032	1.91	2.90	0.86	0.07
242 *Manufacture of pesticides and other agro-chemical products*	0.0000031	1.25	1.63	0.77	0.10
191 *Preparation, tanning, and dressing of leather*	0.0000027	1.11	1.30	0.74	0.10
355 *Manufacture of other transport material*	0.0000026	2.30	4.83	0.91	0.05
247 *Manufacture of man-made and synthetic fibres*	0.0000023	1.33	1.83	0.79	0.08
223 *Reproduction of recorded supports*	0.0000017	2.35	12.51	0.88	0.03
335 *Manufacture of watches and clocks*	0.0000014	3.17	13.14	0.97	0.02
233 *Treatment of nuclear fuels*	0.0000008	5.31	101.16	1.00	0.01
363 *Manufacture of musical instruments*	0.0000008	2.10	3.63	0.88	0.03
176 *Manufacture of knitted textiles*	0.0000007	1.52	1.78	0.79	0.05
231 *Manufacture of coke*	0.0000003	3.94	25.10	0.98	0.01
181 *Tailoring of leather clothes*	0.0000002	3.49	15.85	0.97	0.01

Table 3: *Relative* concentration indexes of the 20 least-concentrated sectors, according to Φ at three-digit level, and weight of each of them in 2008.

4.2 Decompositions of index Φ

By subindustries

In the following section, we analyze whether the concentration level of each manufacturing industry in Spain is mainly due to the concentration of each subindustry at the three-digit

[17] Note that the *Manufacture of pharmaceutical products* (**244**) is a high-technology subindustry. Another high-technology subsector that appears as highly concentrated (among the 30 most-concentrated ones) is *Manufacture of aircraft and spacecraft* (**353**) (where $\Phi = 0.00018$). This subsector, as opposed to the former, is also classified as concentrated by the other indexes ($G = 0.78$), which suggests that the concentration of sector **353** is a robust result.

level taken separately or whether, on the contrary, there are strong spatial interdependencies among them.

Table 4 illustrates the decomposition of index Φ by subsectors (see Section 3, expression (2)) so that for each manufacturing industry at a two-digit level two components are given: one showing the summary contribution of its subindustries at a three-digit level to the concentration of the industry and the other showing the concentration due to the locational interdependencies among subindustries.

Manufacturing industries **2008**	**Number of subindustries** (*S*)	$\Phi(x;t)$	$\frac{\sum_s \Phi(x^s;t)}{\Phi(x;t)}$ **(%)**	$\frac{2\sum_{s=1}^{S}\sum_{s'>s}\tilde{\Phi}(x^s;x^{s'};t)}{\Phi(x;t)}$ **(%)**
15	9	0.00626	63.29	36.71
16	1	0.00008	100.00	0.00
17	7	0.00076	40.95	59.05
18	3	0.00080	107.06	-7.06
19	3	0.00195	87.67	12.33
20	5	0.00071	84.39	15.61
21	2	0.00013	93.02	6.98
22	3	0.00153	62.79	37.21
23	3	0.00013	100.69	-0.69
24	7	0.00197	55.47	44.53
25	2	0.00050	115.14	-15.14
26	8	0.00412	94.79	5.21
27	5	0.00196	57.92	42.08
28	7	0.00090	136.81	-36.81
29	7	0.00155	50.76	49.24
30	1	0.00002	100.00	0.00
31	6	0.00036	85.51	14.49
32	3	0.00011	63.00	37.00
33	5	0.00011	75.40	24.60
34	3	0.00393	67.72	32.28
35	5	0.00164	93.60	6.40
36	6	0.00110	88.20	11.80
37	2	0.00002	105.10	-5.10

Table 4: Concentration index Φ of each manufacturing industry at a two-digit level in 2008 and decomposition by subsectors at a three-digit level.

Focusing now on the most concentrated industries, we observe that the contributions of both factors are rather similar in the case of sectors **24** and **27** (55.5%-44.5% in the former case and 57.9%-42.1% in the latter; see Table 4, columns 3 and 4). Consequently, the spatial interdependencies among subsectors in the *Chemical industry* and in *Metallurgy* are remarkably high (perhaps due to knowledge spillovers in the former case and to input-output linkages in the latter). Something similar happens in the case of *Manufacture of food products and beverages* (**15)** and *Manufacture of motor vehicles, trailers and semi-trailers* (**34)**, even though at a lower degree**.** However, in sectors **19** and **26**, the spatial interdependencies among subsectors are much lower, which seems consistent with the internal heterogeneity of these subsectors mentioned in the previous section.

With respect to the industries with low concentrations, we see that in most of them, the spatial interdependencies among subsectors are low or even negative. Exceptions to this pattern are two high-tech industries: *Manufacture of electronic material; manufacture of radio, television and communications apparatus* (**32**), and *Manufacture of medical-surgical, precision and optical equipment and instruments, and clocks and watches* (**33**), which present important internal interdependencies. In fact, this pattern is generally shared with the remaining high-technology industries, which suggests that externalities among firms may play a role that exceeds the finer industrial classification.

By location groups

We now raise another question. Is it relevant to classify Spanish provinces according to their per capita GDP levels in order to explain the spatial concentration of manufacturing industries? In order to find an answer, provinces were partitioned into three groups of similar sizes: one including the poorer provinces (those having a per capita GDP lower than 85% of the national average), another including those with an intermediate level (between 85% and 105%), and another including the richest provinces.[18]

Table 5 shows the decomposition of index Φ in the within and between components, as explained in Section 3, expression (1). We can see that this categorization of provinces is especially relevant to explaining the concentration of two highly concentrated sectors: **15** and **24**, since the between components explain about 30-35% of the concentration of the

[18] The latest available data at provincial level corresponds to 2005, and they are also offered by the *INE*. The first group includes the poorest provinces: Cádiz, Córdoba, Granada, Huelva, Jaén, Málaga, Sevilla, Ávila, Salamanca, Zamora, Albacete, Ciudad Real, Cuenca, Toledo, Badajoz, Cáceres, Lugo, Ourense, and Pontevedra. Almería, Huesca, Teruel, Asturias, Las Palmas, Santa Cruz de Tenerife, Cantabria, León, Palencia, Segovia, Soria, Valladolid, Guadalajara, Alicante, Valencia, A Coruña, Murcia, Ceuta, and Melilla are included in the second group while Zaragoza, Illes Balears, Burgos, Barcelona, Girona, Lleida, Tarragona, Castellón, Madrid, Navarra, Álava, Guipúzcoa, Vizcaya, and La Rioja are in the third group.

corresponding sector (see column four). Note, however, that while the former, *Manufacture of food products and beverages,* is overrepresented in the poorest provinces, the latter, *Chemical industry,* is mostly found in the richest ones (see columns five and seven). In the remaining concentrated sectors, this classification is less important. In fact, in the case of sectors **26** and **27**, the between component is irrelevant.

Finally, note that sector **29** (which has an intermediate concentration level) is the one having the highest between component (over 42%), mainly due to the overrepresentation of the sector in the richest provinces. Even though in the remaining high- and medium-technology sectors (**24**, **30**, **31**, **32, 33** and **34**) the between component is much lower than in sector **29**, in all of them we find overrepresentation of the sector in the richest provinces.

Manufacturing industries (two-digit) 2008	Φ	$\frac{X}{T}$ %	**Within component** %	**Between component** %	$\frac{X_k}{T_k}$		
					Poor %	**Intermediate** %	**Rich** %
15	0.00626	15.9	64.62	35.38	24.15	17.75	12.02
16	0.00008	0.17	97.24	2.76	0.1	0.42	0.07
17	0.00076	2.9	97.97	2.03	2.69	3.55	2.66
18	0.00080	2.9	90.33	9.67	3.8	3.9	2.09
19	0.00195	1.62	93.20	6.80	1.92	3.35	0.65
20	0.00071	3.71	75.44	24.56	4.58	5.48	2.53
21	0.00013	1.47	95.41	4.59	1.16	1.25	1.69
22	0.00153	5.85	86.06	13.94	4.66	4.03	7.19
23	0.00013	0.57	97.10	2.90	0.97	0.44	0.49
24	0.00197	6.66	70.86	29.14	3.68	4.37	8.87
25	0.00050	3.43	90.28	9.72	2.12	4.16	3.54
26	0.00412	6.88	98.38	1.62	7.68	7.82	6.13
27	0.00196	3.8	99.10	0.90	2.95	4.06	3.98
28	0.00090	12.53	98.91	1.09	12.46	13.04	12.3
29	0.00155	8.23	57.83	42.17	4.49	6.29	10.52
30	0.00002	0.26	81.87	18.13	0.08	0.02	0.45
31	0.00036	2.98	98.31	1.69	2.69	2.73	3.21
32	0.00011	1.06	88.82	11.18	0.5	0.84	1.37
33	0.00011	1.16	97.18	2.82	0.8	1.23	1.25
34	0.00393	7.45	83.16	16.84	4.25	4.99	9.82
35	0.00164	2.65	91.14	8.86	5.01	2.63	1.81
36	0.00110	7.45	94.93	5.07	8.93	7.4	6.95
37	0.00002	0.36	97.49	2.51	0.32	0.27	0.42

Table 5: Concentration index and weight of each manufacturing industry at a two-digit level in 2008 and decomposition by groups of provinces according to their GDP in 2005.

5. Conclusions

This paper has formally presented the axioms underlying several concentration measures derived from the literature on income distribution. As far as we know, this is the first time that this debate has been formally broached since, until now, inequality indexes have been adapted to measure the spatial concentration of economic activity without an axiomatic discussion. This examination has allowed us to unveil the properties that the literature on regional economics is implicitly assuming when using the locational Gini coefficient and the generalized entropy family of indexes, and it has permitted us to propose another inequality-based concentration measure. This research showed that this new concentration index satisfies a concentration-invariance condition that differs from that of the aforementioned indexes. As in the case of the members of the generalized entropy family, this index is also additively decomposable, which seems to be a desirable property for applied research.

This study showed that even though all these concentration indexes quantify the *relative* concentration of an industry, satisfying some basic properties, the results are substantially affected by the type of inequality-based index being used. When measuring the spatial concentration of a sector, those measures based on "relative" inequality indexes focus their attention on the employment share that the sector has in each location unit while measures based on "absolute" inequality indexes take into account not only employment shares but also employment levels. Some researchers may find it more suitable to choose one kind of concentration-invariance condition rather than the other while others may prefer using both types of measures in order to determine the sectors in which the results are robust enough to stand up against changes in the invariance condition.

When analyzing the spatial concentration of manufacturing industries in Spain in 2008, we confirmed the existence of important differences among the outcomes reached according to relative and absolute inequality-based concentration indexes. In fact, according to the new index, the *Manufacture of food products and beverages* is the most concentrated sector, while it occupies one of the lowest positions in the ranking according to the locational Gini coefficient as well as the generalized entropy family. All the indexes coincide, however, in classifying the *Leather industry* among the most concentrated ones, which suggest that the concentration of this sector is a result that is rather robust against changes in the concentration-invariance condition.

On the other hand, both types of measures allow us to conclude that there are important discrepancies among the concentration levels of high-technology industries. In other words, we cannot state that high-tech industries in Spain tend to concentrate at a higher degree than other sectors, which is in line with results obtained by Devereux et al. (2004) in the UK by using the index proposed by Maurel and Sédillot (1999). However, we found that in all high-tech industries, the spatial interdependencies among their corresponding subindustries are remarkably high, as the decomposition of the new concentration index by subindustries shows. Furthermore, the decomposition of this index by location subgroups allows us to conclude that these industries are overrepresented in the richest provinces of the country.

References

Aiginger, K. and Pfaffermayr, M. (2004): “The single market and geographic concentration in Europe”, *Review of International Economics* 12(1), 1-11.

Alonso-Villar, O. and Del Río, C. (2008a): “The geographical concentration of unemployment: A male-female comparison in Spain”, *Regional Studies* 42(3), 401-412.

Alonso-Villar, O. and Del Río, C. (2008b): “Local versus overall segregation measures”, Documento de Traballo 0802, Departamento de Economía Aplicada, Universidade de Vigo.

Amiti, M. (1999): “Specialization patterns in Europe”, *Weltwirtschaftliches Archiv* 135(4), 573-593.

Atkinson, A.B. (1970): “On the measurement of inequality”, *Journal of Economic Theory* 2, 244-263.

Brakman, S., Garretsen, H., Gorter, J., van der Horst, A. and Schramm, M. (2005): “New economic geography, empirics, and regional policy”, CPB (Netherlands Bureau for Economic Policy Analysis) Special Publication 56.

Brülhart, M. (2001): “Evolving geographical concentration of European manufacturing industries”, *Weltwirtschaftliches Archiv* 137(2), 215-243.

Brülhart, M. and Traeger, R. (2005): “An account of geographic concentration patterns in Europe”, *Regional Science and Urban Economics* 35(6), 597-624.

Chakravarty, S. (2001): “The variance as a subgroup decomposable measure of inequality”, *Social Indicators Research* 53, 79-95.

Cowell, F.A. (2000): “Measurement of Inequality”, in A.B. Atkinson and F. Bourguignon (eds.), *Handbook of Income Distribution vol. 1*, Elsevier-North Holland.

Devereux, M., Griffith, R., and Simpson, H. (2004): “The geographic distribution of production activity in the UK”, *Regional Science and Urban Economics* 34, 533-564.

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

Ellison, G. and Glaeser, E. (1997): "Geographic concentration in U.S. manufacturing industries: a dartboard approach", *Journal of Political Economy* 105, 889-927.

Fujita, M., Krugman, P. and Venables, A. (2000): *The Spatial Economy. Cities, Regions and International Trade*, MIT Press, second printing.

Guimarães, P., Figueiredo, O. and Woodward, D. (2007): "Measuring the localization of economic activity: A parametric approach", *Journal of Regional Science* 47(4), 753-774.

Haaland, J., Kind, H., Midelfart Knarvik, K., and Torstensson, J. (1999): "What determines the economic geography of Europe?", CEPR Discussion Paper 2072.

Hutchens, R.M. (1991): "Segregation curves, Lorenz curves, and inequality in the distribution of people across occupations", *Mathematical Social Sciences* 21, 31-51.

Hutchens, R.M. (2004): "One measure of segregation", *International Economic Review* 45(2), 555-578.

Kolm, S.-C. (1976): "Unequal inequalities I", *Journal of Economic Theory* 12, 416-442.

Krugman, P. (1991): *Geography and Trade*, MIT Press.

Marcon, E. and Puech, F. (2003): "Evaluating the geographic concentration of industries using distance-base methods", *Journal of Economic Geography* 3, 409-428.

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

Mori, T., Nishikimi, K., and Smith, T. (2005): "A divergence statistic for industrial localization", *The Review of Economics and Statistics* 87(4): 635-651.

Neary, P. (2001): "Of hype and hyperbolas: introducing the new economic geography", *Journal of Economic Literature*, 39 (2): 536-561.

OECD (2007): *OECD Science, Technology and Industry. Scoreboard 2007*.

Ottaviano, G. and Thisse J.-F. (2004): "Agglomeration and economic geography". In: J.V. Henderson and J.-F. Thisse (eds), *Handbook of Urban and Regional Economics*, vol. 4, Elservier-North Holland.

Shorrocks, A.F. (1980): "The class of additively decomposable inequality measures", *Econometrica* 48(3), 613-626.

Suedekum, J. (2006): "Concentration and specialization trends in Germany since re-unification", *Regional Studies* 40(8), 861-873.

Zheng, B. (2007): "Unit-consistent decomposable inequality measures", *Economica*, 74: 97-111.

Relación de Documentos de Traballo publicados

9901 Philippe Polomé: Experimental E vidence on Voting Manipulation in Referendum Contingent Valuation with Induced Value
9902 Xosé M. González e Daniel Miles: Análisis Envolvente de Datos: Un Estudio de Sensibilidad
9903 Philippe Polomé: Combining contingent valuation and revealed preferences by simulated maximum likelihood
9904 Eva Rodríguez: Social value of health programs: is the age a relevant factor?
9905 Carlos Gradín e Mª Soledad Giráldez: Incorporación laboral de la mujer en España: efecto sobre la desigualdad en la renta familiar
9906 Carlos Gradín: Polarization by sub-populations in Spain, 1973-91
9907 Carlos Gradín: Polarization and inequality in Spain: 1973-91
0001 Olga Alonso e José María Chamorro: How do producer services affect the location of manufacturing firms?. The role of información accesibility
0002 Coral del Río Otero: Desigualdad *Intermedia* Paretiana
0003 Miguel Rodríguez Méndez: Margins, Unions and the Business Cycle in High nd Low Concentrated Industries
0004 Olga Alonso Villar: Large metropolies in the Third World: an explanation
0005 Xulia González e Daniel Miles**:** Wage Inequality in a Developing Country: Decrease of Minimum Wage or Increase of Education Returns
0006 Daniel Miles: Infecuencia de las Compras y Errores de Medida
0007 Lucy Amigo: Integración de los Mercados de Cambio: Análisis rentabilidad-riesgo de la cotización Peseta/Dólar
0008 Eduardo L. Giménez e Manuel González-Gómez: Efficient Allocation of Land Between Productive Use and Recreational Use.
0009 Manuel González-Gómez, P.Polomé e A. Prada Blanco: Sesgo sobre la Información Obtenida y la Estimación de Beneficios en Entrevistas a Visitantes de un Espacio Natural
0010 M. Xosé Vázquez Rodríguez e Carmelo León: Preferencias Imprecisas y Contexto en la Valoración de Cambios en la Salud.
0011 Begoña Alvarez: Can we Identify Fraudulent Behaviour?. An Aplication to Sickness Absence in Spain
0012 Xulia González, Xosé M. González e Daniel Miles: La Transición de la Universidad al Trabajo: una Aproximación Empírica.
0013 Olga Cantó: Climbing out of poverty, Falling back in: Low Incomes' Stability in Spain
0101 Arancha Murillas: Investment and Development of Fishing Resources: A Real Options Approach
0102 Arancha Murillas: Sole Ownership and Common Property Under Management Flexibility: Valuation, Optimal Exploitation and Regulation
0103 Olga Alonso Villar; José-María Chamorro Rivas e Xulia González Cerdeira: An análisis of the Geographic Concentratin of Industry in Spain
0104 Antonio Molina Abraldes e Juan Pinto-Clapés: A Complete Characterization of Pareto Optimality for General OLG Economies
0105 José María Chamorro Rivas: Communications technology and the incentives of firms to suburbanize
0106 Lucy Amigo Dobaño e Francisco Rodríguez de Prado: Incidencia del efecto día en los valores tecnológicos en España

0107 Eva Rodríguez-Míguez; C. Herrero e J. L. Pinto-Prades: Using a point system in the management of waiting lists: the case of cataracts
0108 Xosé M. González e D. Miles: Análisis de los incentivos en el empleo público
0109 Begoña Álvarez e D. Miles: Gender effect on housework allocation: evidence from spanish two-earned couples
0110 Pilar Abad: Transmisión de volatilidad a lo largo de la estructura temporal de swaps: evidencia internacional
0111 Pilar Abad: Inestabilidad en la relación entre los tipos forward y los tipos de contado futuros en la estructura temporal del mercado de swaps de tipos de interés
0112 Xulia González, Consuelo Pazó e Jordi Jaumandreu: Barriers to innovation and subsidies effectiveness
0201 Olga Cantó, Coral del Río e Carlos Gradín: What helps households with children in leaving poverty?: Evidence from Spain in contrast with other EU countries
0202 Olga Alonso-Villar, José María Chamorro-Rivas e Xulia González: Agglomeration eocnomies in manufacturing industries: the case of Spain
0203 Lucy Amigo Dobaño, Marcos Alvarez Díaz e Francisco Rodríguez de Prado: Efficiency in the spanish stock market. A test of the weak hypothesis based on cluster prediction technique
0204 Jaime Alonso-Carrera e María Jesús Freire-Serén: Multiple equilibria, fiscal policy, and human capital accumulation
0205 Marcos Alvarez Díaz e Alberto Alvarez: Predicción no-lineal de tipos de cambio. Aplicación de un algoritmo genético
0206 María J. Moral: Optimal multiproduct prices in differenciated product market
0207 Jaime Alonso-Carrera y Baltasar Manzano: Análisis dinámico del coste de bienestar del sistema impositivo español. Una explotación cuantitativa
0208 Xulia González e Consuelo Pazó: Firms' R&D dilemma: to undertake or not to undertake R&D
0209 Begoña Álvarez: The use of medicines in a comparative study across European interview-based surveys
0210 Begoña Álvarez: Family illness, work absence and gender
0301 Marcos Álvarez-Díaz e Alberto Álvarez: Predicción no-lineal de tipos de cambio: algoritmos genéticos, redes neuronales y fusión de datos
0302 Marcos Álvarez-Díaz, Manuel González Gómez e Alberto Álvarez: Using data-driven prediction methods in a hedonic regression problem
0303 Marcos Álvarez-Díaz e Lucy Amigo Dobaño: Predicción no lineal en el mercado de valores tecnológicos español. Una verificación de la hipótesis débil de eficiencia
0304 Arantza Murillas Maza: Option value and optimal rotation policies for aquaculture explotations
0305 Arantza Murillas Maza: Interdependence between pollution and fish resource harvest policies
0306 Pilar Abad: Un contraste alternativo de la hipótesis de las expectativas en Swaps de tipos de interés
0307 Xulio Pardellas de Blas e Carmen Padín Fabeiro: A tourist destination planning and design model: application to the area around the Miño river in the south of Galicia and the north of Portugal
0308 Lucy Amigo Dobaño e Francisco Rodríguez de Prado: Alteraciones en el comportamiento bursátil de las acciones de empresas tecnológicas inducidas por el vencimiento de derivados

0309 Raquel Arévalo Tomé e José María Chamorro Rivas: A Quality Index for Spanish Housing

0310 Xulia González e Ruben Tansini: Eficiencia técnica en la industria española: tamaño, I+D y localización

0311 Jaime Alonso Carrera e José-María Chamorro Rivas: Environmental fiscal competition under product differenciation and endogeous firm location

0312 José Carlos Álvarez Villamarín, Mª José Caride Estévez e Xosé Manuel González Martínez: Demanda de transporte. Efectos del cambio en la oferta ferroviaria del corredor Galicia-Madrid

0313 José Carlos Álvarez Villamarín, Mª José Caride Estévez e Xosé Manuel González Martínez: Análisis coste-beneficio de la conexión Galicia-Madrid con un servicio de Alta Velocidad.

0401 María José Caride e Eduardo L. Giménez: Thaler's "all-you-can-eat" puzzle: two alternative explanations.

0402 Begoña Álvarez e Daniel Miles: Husbands' Housework Time: Does Wives' Paid Employment Make a Difference?

0403 María José Caride e Eduardo L. Giménez: Leisure and Travel Choice.

0404 Raquel Arévalo Tomé e José María Chamorro-Rivas: Credible collusion in a model of spatial competition.

0405 Coral del Río Otero, Carlos Gradín Lago e Olga Cantó Sánchez: El enfoque distributivo en el análisis de la discriminación salarial por razón de género.

0406 Olga Alonso Villar: Ciudades y globalización en la Nueva Geografía Económica.

0407 Olga Alonso Villar: The effects of transport costs revisited

0408 Xavier Labandeira e Miguel Rodríguez: The effects of a sudden CO_2 reduction in Spain.

0409 Gema Álvarez Llorente, Mª Soledad Otero Giráldez, Alberto Rodríguez Casal e Jacobo de Uña Álvarez: La duración del desempleo de la mujer casada en Galicia.

0410 Jacobo de Uña-Álvarez, Gema Álvarez-Llorente e Mª Soledad Otero-Giráldez: Estimation of time spent in unemployment for married women: An application at regional level.

0411 Mª José Moral: Modelos empíricos de oligopolio con producto diferenciado: un panorama.

0412 Mª José Moral: An approach to the demand of durable and differentiated products.

0501 Raquel Arévalo-Tomé e José-María Chamorro-Rivas: Location as an instrument for social welfare improvement in a spatial model of Cournot competition.

0502 Olga Alonso-Villar: The effects of transport costs within the new economic geography.

0503 Raquel Arévalo Tomé, Mª Soledad Otero Giráldez e Jacobo de Uña Álvarez: Estimación de la duración residencial a partir del periodo de ocupación declarado por los hogares españoles.

0504 Olga Alonso-Villar, Coral de Río e Luis Toharia: Un análisis espacial del desempleo a nivel municipal.

0601 Xulia González, Consuelo Pazó: Do public subsidies stimulate private R&D spending?

0602 Lucy Amigo Dobaño: Anomalías de los mercados financieros. Análisis de las empresas gallegas que cotizan en el mercado de renta variable.

0603 Daniel Miles Touya: Can we teach civic attitudes?

0604 Jacobo de Uña Álvarez, Raquel Arévalo Tomé e Mª Soledad Otero Giráldez: Advances in the estimation of households' duration of residence.

0605 Pilar Abad Romero, Begoña Álvarez García, Eva Rodríguez Míguez e Antonio Rodríguez Sampayo: Una aplicación de los sistemas de puntos en la priorización de pacientes en lista de espera quirúrgica.

0606 Coral del Río, Carlos Gradín e Olga Cantó: Pobreza y discriminación salarial por razón de género en España.

0607 Xulia González : Some empirical regularities on vertical restraints.

0608 José María Chamorro Rivas: Shopping hours and bundling as an entry barrier.

0701 Lucy Amigo Dobaño e Mª Dolores Garza Gil: La flota artesanal gallega. Análisis de la cadena de distribución.

0702 Coral del Río e Olga Alonso-Villar: Diferencias entre mujeres y hombres en el mercado de trabajo: desempleo y salarios.

0703 Coral del Río e Olga Alonso-Villar: Una reflexión sobre las medidas de desigualdad intermedias.

0704 Begoña Álvarez e Marcos Vera-Hernández: Unravelling the determinants of the non-use of health care.

0801 Olga Alonso-Villar: What are we assuming when using inequality measures to quantify geographic concentration?

0802 Olga Alonso-Villar e Coral del Río: Local versus overall segregation measures.

0803 Raquel Arévalo Tomé, Loly Ferrero Martínez, Mª Soledad Otero Giráldez e Jacobo de Uña Álvarez: Movilidad residencial en España: un análisis longitudinal.

0804 Olga Alonso-Villar e Coral del Río: An alternative inequality-based concentration measure.