



HAL
open science

Concentration of Economic Activity: An Analytical Framework

Olga Alonso-Villar, Coral del Río

► **To cite this version:**

Olga Alonso-Villar, Coral del Río. Concentration of Economic Activity: An Analytical Framework. Regional Studies, 2011, pp.1. 10.1080/00343404.2011.587796 . hal-00724178

HAL Id: hal-00724178

<https://hal.science/hal-00724178>

Submitted on 19 Aug 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Concentration of Economic Activity: An Analytical Framework

Journal:	<i>Regional Studies</i>
Manuscript ID:	CRES-2010-0044.R2
Manuscript Type:	Main Section
JEL codes:	R12 - Size and Spatial Distributions of Regional Economic Activity < R1 - General Regional Economics < R - Urban, Rural, and Regional Economics, D63 - Equity, Justice, Inequality, etc. and Measurement < D6 - Welfare Economics < D - Microeconomics
Keywords:	Geographic Concentration, Inequality Measures

SCHOLARONE™
Manuscripts

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57

Concentration of Economic Activity: An Analytical Framework

Olga Alonso-Villar[#] and Coral del Río^{*}

(Received February 2010; in revised form May 2011)

[#] Universidade de Vigo; Facultade de CC. Económicas; Departamento de Economía Aplicada; Campus Lagoas-Marcosende s/n; 36310 Vigo; Spain. E-mail: ovillar@uvigo.es

^{*} Universidade de Vigo; Facultade de CC. Económicas; Departamento de Economía Aplicada; Campus Lagoas-Marcosende s/n; 36310 Vigo; Spain. E-mail: crio@uvigo.es

1
2
3 Abstract
4

5 The goal of this paper is to offer an analytical framework within which *relative*
6 concentration, including both the concentration of each sector and aggregate
7 concentration, can be analyzed. By borrowing properties from the literature on income
8 inequality and segregation and adapting them to a location context, this paper
9 characterizes the generalized entropy family of concentration indexes and shows the
10 properties of the L-index. In addition, it offers other measures taken from the
11 segregation literature. All these tools are used to analyze the spatial patterns of
12 manufacturing industries in Spain from 1977 to 2008, paying special attention to their
13 technological intensity.
14
15
16
17
18
19
20
21

22 JEL Classification: R12; D63
23
24

25 Keywords: Geographic concentration; Inequality measures
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

La Concentración de la Actividad Económica: Un Marco Analítico

Resumen

El objetivo de este trabajo es ofrecer un marco analítico dentro del cual examinar tanto la concentración *relativa* de un sector como la del conjunto de sectores. Tomando propiedades de la literatura de distribución de la renta y de segregación y adaptándolas a un contexto de localización, este artículo caracteriza las medidas de concentración derivadas de la familia de índices de entropía generalizada y muestra las propiedades del índice L. Además, ofrece también otras medidas derivadas de la literatura de segregación. Todas estas herramientas son usadas para analizar los patrones espaciales de la industrias manufactureras españolas entre 1977 y 2008, prestando especial atención a sus intensidades tecnológicas.

1. Introduction

In recent years, the study of production location patterns has received increasing interest in the field, both empirically and theoretically. This flourishing interest is motivated in part by a general concern with the effects of economic integration processes on industrial localization, especially in Europe, where the creation of the Single Market has stimulated the debate (HAALAND et al., 1999; BRÜLHART, 2001; AINGINGER and PFAFFERMAYR, 2004; RESMINI, 2007, inter alia).

In quantifying the spatial concentration of a sector, most measures follow a *relative* notion so that the spatial distribution of the sector is compared with that of the whole set of sectors (ELLISON and GLAESER, 1997; AMITI, 1999; BRÜLHART and TRAEGER, 2005). If economic activity is measured in terms of employment, as is traditionally done, and the focus is on manufacturing industries, the distribution of overall manufacturing employment is usually considered the distribution of reference against which to compare that of any single sector. Thus, concentration rises in a sector as long as its employment distribution among locations departs from that of overall manufacturing employment.

By following this approach, many studies have calculated the concentration level of each manufacturing sector (which can be labeled partial concentration) in different economies; however, there has been no formal discussion of how to aggregate this information to calculate *relative* concentration for the entire manufacturing industry. Certainly, interest in measuring overall concentration is not new in the field, and complementary perspectives have been offered. Thus, some studies have proposed comparing distribution of employment in the manufacturing industry across locations with distribution of employment in the whole economy (BRÜLHART and TRAEGER, 2005). In other words, in measuring concentration in the manufacturing industry, the benchmark considered is outside that industry. Others have proposed instead the use of the average concentration of the sectors involved, making use of an external benchmark unnecessary (AINGINGER and DAVIS, 2004; MULLIGAN and SCHMIDT, 2005; CUTRINI, 2009, 2010). However, as far as we know, there has been no debate on overall concentration measurement in terms of basic properties.

1
2
3 Among the spatial concentration measures existing in the literature, those borrowed
4 from the literature on income inequality are some of the most widely used. In this
5 regard, the Gini index has traditionally been used for analyzing spatial location patterns
6 of manufacturing industries (KRUGMAN, 1991; AMITI, 1999; BRÜLHART, 2001;
7 SUEDEKUM, 2006, inter alia). More recently, the generalized entropy family of
8 indexes (GE, henceforth) has been used as well because of its advantages in terms of
9 decomposability (BRÜLHART and TRAEGER, 2005; BRAKMAN et al., 2005;
10 PÉREZ-XIMÉNEZ and SANZ-GRACIA, 2007; CUTRINI, 2009). The properties of the
11 inequality measures on which these concentration indexes are based are well known
12 because the literature on income distribution has dealt with inequality from an
13 axiomatic perspective; however, the corresponding geographical concentration
14 measures are being applied without explicitly stating the consequences of using them.
15
16
17
18
19
20
21
22
23
24

25
26 To close these gaps somewhat, this paper offers a set-up within which *relative*
27 concentration, including both the concentration of each sector and aggregate
28 concentration, can be analyzed in terms of basic properties. This set-up allows us to
29 characterize the GE family of concentration indexes of a sector (BRÜLHART and
30 TRAEGER, 2005) and show the properties of the aggregate concentration index, the L-
31 index (CUTRINI, 2009, 2010), resulting from the weighted sum of the one of them (the
32 Theil index). For that purpose, some properties borrowed from the literature on income
33 inequality and segregation are adapted to our context (ALONSO-VILLAR and DEL
34 RÍO, 2010a; FRANKEL and VOLIJ, 2010).¹ In addition, other partial and overall
35 measures borrowed from the segregation literature are proposed to quantify
36 concentration. All these tools are finally used to analyze the spatial patterns of
37 manufacturing industries in Spain during its democratic period (1977-2008).
38
39
40
41
42
43
44
45
46
47

48 Our empirical analysis suggests that high-tech industries tend to concentrate at a higher
49 extent than industries with a lower technological intensity. In addition, we find that
50 overall concentration slightly increased until 1981 while it tended to decrease from
51 1985 onward. This evolution seems to be mainly shaped by concentration changes in
52 the sectors involved and not by shifts in the sectoral structure of the economy.
53
54
55
56
57

58 The paper is structured as follows. Section 2 offers an overview of recent concentration
59 measures and characterizes the partial concentration measures derived from the GE
60 family in terms of basic properties. Section 3 introduces several aggregate concentration

1
2
3 indexes related to the partial measures described above. These partial and aggregate
4 measures are used in Section 4 to analyze the manufacturing industry in Spain, paying
5 attention to the contribution of technological groups to overall concentration. Finally,
6
7 Section 5 presents the main conclusions.
8
9

10 11 2. The spatial concentration of an industry 12

13 14 2.1 An overview of recent concentration measures 15

16
17 The literature offers a variety of measures to assess the distribution of economic activity
18 across geographic units (counties, regions, countries, etc.). The terminology used is not
19 always consistent since they are indiscriminately called spatial/geographical
20 concentration measures, localization measures, and agglomeration measures. In this
21 paper, we prefer the term *geographical concentration*, which seems more general, since
22 the other labels are often used when these measures are intended to check the existence
23 of externalities as the main force driving the spatial concentration of firms.²
24
25
26
27
28
29

30
31 Some of these measures are formally derived from location models, such as the popular
32 index proposed by ELLISON and GLAESER (1997) (EG, henceforth), which measures
33 the extent of geographic concentration once the size of establishments (via the
34 Herfindahl index) and inherent randomness in the concentration of firms are accounted
35 for. Following the same spirit, MAUREL and SÉDILLOT (1999) and GUIMARÃES et
36 al. (2007) developed new indexes. Other indexes are instead borrowed from the
37 literature on statistics, as is the case with the *D*-index proposed by MORI et al. (2005),
38 which quantifies the divergence between the observed distribution of establishments
39 and a uniform distribution across physical space according to the Kullback-Leibler
40 distance (also called relative entropy). A similar index is that used by AINGINGER and
41 DAVIES (2004) and AINGINGER and PFAFFERMAYR (2004), who borrowed the
42 concept of entropy from information theory, rather than that of relative entropy
43 (COVER and THOMAS, 1991).³ Other measures derived from the literature on
44 statistics include those put forward by MARCON and PUECH (2003) and
45 DURANTON and OVERMAN (2005). As opposed to previous measures, these tools
46 are distance-based methods that consider the distribution of distances between pairs of
47 establishments. This allows taking into account various spatial scales at the same time,
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 avoiding the problem of arbitrary geographic scales (usually administrative). An
4 important drawback of this approach is that it requires data that are often not available.
5
6

7
8 Apart from the EG index, which has been applied in a large number of empirical
9 studies, some of the most widely used geographic concentration measures are derived
10 from the literature on income inequality. This is the case with the Gini coefficient and
11 the Theil index (also related to information theory).⁴ Several versions exist depending
12 on the view of spatial concentration; thus, the *absolute* concentration adaptations of
13 these measures are used when the no-concentration benchmark against which to
14 compare the distribution of a sector is given by a uniform distribution across locations
15 (BRAKMAN et al., 2005). There are also *topographic* concentration measures
16 (BRAKMAN et al., 2005; BRÜLHART and TRAEGER, 2005) according to which no-
17 concentration exists if the sector is evenly spread over physical space (which is also the
18 approach followed in the *D*-index, which can actually be interpreted as the topographic
19 version of the Theil index).
20
21
22
23
24
25
26
27
28
29

30 However, most concentration measures have followed a different approach. Suppose,
31 for example, that economic activity is measured in terms of employment, as
32 traditionally assumed. Focusing on manufacturing industries, the distribution of overall
33 manufacturing employment is usually considered the distribution of reference against
34 which to compare that of any single sector, so that no spatial concentration exists in the
35 sector so long as its employment distribution across locations is equal to that of the
36 industrial aggregate. This notion is labeled *relative* concentration and has been
37 extensively used in empirical research (ELLISON and GLAESER, 1997; AMITI, 1999;
38 BRÜLHART and TRAEGER, 2005). This is the approach followed in this paper when
39 unveiling, in this section, the properties that scholars in regional science are implicitly
40 assuming when using these inequality-based measures (which do not account for either
41 inherent randomness in the concentration of firms or distances between locations).
42
43
44
45
46
47
48
49
50
51

52 The measures mentioned so far assess the concentration of a sector, labeled here partial
53 concentration; however, one can also be concerned with aggregate or overall
54 concentration. This issue will be analyzed in Section 3.
55
56
57
58
59
60

2.2 Characterizing partial measures

Since the literature on income distribution has tackled inequality from an axiomatic perspective, there is general agreement on the basic properties that any inequality measure should satisfy. The approach followed by the literature on geographic concentration has been rather different, since such an axiomatization does not exist (COMBES and OVERMAN, 2004). The aim of this paper is not to propose an axiomatic framework for measuring geographical concentration, but to unveil some of the main properties we implicitly assume when using inequality measures in this context.

For the sake of simplicity, assume that economic activity is measured in terms of employment. Consider that there are T workers in the economy allocated among $L > 1$ locations according to $t \equiv (t_1, t_2, \dots, t_L)$, where $T = \sum_l t_l$. Vector t represents the distribution of reference against which that of any sector is compared. Let us denote by $x^s \equiv (x_1^s, x_2^s, \dots, x_L^s)$ the distribution of sector s , where x_l^s is the number of workers in sector s in location l . Therefore, the total number of workers in location l is $t_l = \sum_s x_l^s$, while the total number of workers in sector s is $X^s = \sum_l x_l^s$.

In this paper, an index of partial concentration is a function $I_c : D \rightarrow \mathbb{R}$, where $D = \bigcup_{L>1} \{(x^s; t) \in \mathbb{R}_+^L \times \mathbb{R}_+^L : x_l^s \leq t_l \forall l\}$, such that $I_c(x^s; t)$ represents the concentration level of sector s , which is distributed across locations according to x^s , when comparing it with the distribution of reference t .

The GE family comprises indexes frequently used to measure the concentration of an industry according to a *relative* notion and can be written as:

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

where α is a sensitivity parameter.⁵ If sector s is distributed across locations in the same way as aggregate employment, i.e., if $x_l^s / X^s = t_l / T \quad \forall l$, any index of this class is equal to zero.⁶ An advantage of these measures is that they are additively decomposable, which is helpful for empirical analysis.

Even though in the literature on income distribution this family of indexes has been characterized in terms of basic properties, to our knowledge, such a characterization does not exist in the field of spatial concentration. In the next proposition, we show that five properties borrowed from the literature on inequality and segregation completely characterize the GE family of concentration indexes:⁷

- 1) *Symmetry in locations* (the partial concentration index is unaffected by the order in which locations are enumerated). This property implies that the spatial distribution of locations is irrelevant.
- 2) *Movement between locations* (when a location with a lower employment level in the sector of study than another, but with the same aggregate employment, loses employment in the sector in favor of the other location, the concentration of the sector must increase). This guarantees that a disequalizing movement of workers from one location to another having the same total employment level fosters concentration.
- 3) *Scale invariance I* (concentration should not change when the employment level of the aggregate distribution and/or that of the sector varies, so long as the weight that each location represents in distributions t and x^s , $\frac{t_l}{T}$ and $\frac{x_l^s}{X^s}$, respectively, remains unaltered). This implies that it is only employment shares that matter, not employment levels.
- 4) *Insensitivity to proportional divisions of locations I* (subdividing a location into several units of equal size, in terms of both aggregate employment and employment in the sector of study, does not affect the concentration level of the sector). This property means that the geographical scale used in the analysis (usually administrative units) is irrelevant as long as employment distributions within locations are uniform.⁸
- 5) *Aggregation* (when classifying locations into two mutually exclusive classes, the concentration of the sector can be written as a function of the concentration of

the sector in each class of locations, the employment level in each class, and the employment share of the sector in each class of locations). This is a helpful property since it allows the indexes to decompose by subgroups.

Proposition. Let I_c be a continuous concentration index that takes a zero value when the distribution of the sector of study among locations coincides with the distribution of reference (i.e., when $\frac{x_l^s}{X^s} = \frac{t_l}{T}$). I_c is a concentration index satisfying properties 1-5 if and only if it can be written as an increasing monotonic transformation of $\Psi_\alpha(x^s; t)$.

Proof: See Appendix A.

Another relative measure widely used is the locational Gini coefficient:

$$G^s = \frac{\sum_{l,l'} \frac{t_l}{T} \frac{t_{l'}}{T} \left| \frac{x_l^s}{t_l} - \frac{x_{l'}^s}{t_{l'}} \right|}{2 \frac{X^s}{T}}, \quad (2)$$

which satisfies properties 1-4 but not property 5. An additional discrepancy between Ψ_α and G^s rests on their different sensitivity toward movements of workers between locations. Thus, Ψ_α (with $\alpha \leq 1$) is more sensitive to movements of workers between locations in which the relative presence of the sector, $\frac{x_l^s}{t_l}$, is low.⁹ G^s attaches, instead, more weight to movements affecting locations around the mode of distribution $\left(\frac{x_1^s}{t_1}, \dots, \frac{x_L^s}{t_L} \right)$.

Given the parallels between spatial concentration of a sector across locations and segregation of a population group across organizational units, a variation of the index of dissimilarity proposed by DUNCAN and DUNCAN (1955) can be adapted to measure partial concentration:

$$D^s = \frac{1}{2} \sum_l \left| \frac{x_l^s}{X^s} - \frac{t_l}{T} \right|. \quad (3)$$

1
2
3 It is easy to see that this index satisfies properties 1, 3, and 4, but not property 2 since
4
5 D^s is not sensitive to disequalizing movements of workers between locations on the
6
7 same side of the mean of distribution $\left(\frac{x_1^s}{t_1}, \dots, \frac{x_L^s}{t_L}\right)$. Therefore, as opposed to Ψ_α and
8
9
10
11 G^s , D^s does not always increase when disequalizing movements of workers occur.¹⁰
12

13
14 So far, we have focused on the properties analyzed in inequality/segregation contexts
15 when using GE measures since, as far as we know, they have not been unveiled in a
16 location context. These properties are of a different nature from those presented in
17 COMBES and OVERMAN (2004). In any case, note that the GE family of
18 concentration indexes also satisfies some of those properties. They are comparable
19 across spatial scales, and some are also comparable across activities since they are
20 decomposable by locations and subsectors, respectively (BRÜLHART and TRAEGER,
21 2005). In addition, statistical significance of the results is possible by using
22 bootstrapping (BRÜLHART and TRAEGER, 2005) and the index takes a unique value
23 (zero) under the null hypothesis of no concentration. Note also that all the partial
24 measures included in this section are affected by the “modifiable areal unit problem”
25 (MAUP, ARBIA, 1989), since in quantifying concentration they do not take into
26 account the spatial distribution inside observed regions (as discussed by BRÜLHART
27 and TRAEGER, 2005, in the GE case). This means that they are sensitive to changes in
28 spatial scale. Moreover, since these measures neglect distances across regions, they are
29 affected by the “checkerboard problem” as well (WHITE, 1983).
30
31
32
33
34
35
36
37
38
39
40
41
42

43 3. Overall concentration measures

44
45
46 In this section, we first show that the weighted sum of the Theil index ($\Psi_1(x^s; t)$) for
47 each of the mutually exclusive sectors into which the economy can be partitioned (i.e.,
48 the L -index proposed by CUTRINI, 2009) is actually the mutual information index (M ,
49 henceforth) derived from information theory. Next, by analyzing the properties satisfied
50 by the M -index in other contexts, we unveil the properties that one is implicitly
51 assuming when using it to measure geographic concentration. Finally, we propose two
52 more aggregate concentration measures derived from the segregation literature to use in
53 the empirical section to analyze the robustness of our results.
54
55
56
57
58
59
60

3.1 The M -index

In a recent paper, CUTRINI (2009) proposes using the L -index, $L = \sum_s \frac{X^s}{T} \Psi_1(x^s; t)$, to

measure overall localization of manufacturing industries in Europe. Given that

$\sum_s \frac{X^s}{T} \sum_l \frac{x_l^s}{X^s} \ln\left(\frac{t_l}{T}\right) = \sum_l \frac{\sum_s x_l^s}{T} \ln\left(\frac{t_l}{T}\right) = \sum_l \frac{t_l}{T} \ln\left(\frac{t_l}{T}\right)$, it is easy to see that this index

can be rewritten as $L = \sum_s \frac{X^s}{T} \sum_l \frac{x_l^s}{X^s} \ln\left(\frac{x_l^s}{X^s}\right) - \sum_l \frac{t_l}{T} \ln\left(\frac{t_l}{T}\right)$. Note that, on the one hand

the second component of the above expression is the entropy, $H(V)$, of a random

variable, V , whose frequency distribution is $\left(\frac{t_1}{T}, \dots, \frac{t_L}{T}\right)$. On the other hand, the first

component is equal to minus the conditional entropy of $V|Z$, $-H(V|Z)$, where the

distribution of Z is $\left(\frac{X^1}{T}, \dots, \frac{X^S}{T}\right)$, and the conditional distribution $V|Z = z$ is

$\left(\frac{x_1^s}{X^s}, \dots, \frac{x_L^s}{X^s}\right)$. From all of the above, it follows that the L -index turns out to be equal to

the M -index of random distributions V and Z , since M is defined as

$M(V; Z) = H(V) - H(V|Z)$, where $H(V) = -\sum_v p(v) \ln p(v)$,

$H(V|Z) = -\sum_z p(z) \sum_v p(v|z) \ln p(v|z)$, $p(v)$ and $p(z)$ denote the probability mass

functions of V and Z , respectively, and $p(v|z)$ is the probability distribution of

conditional distribution $V|Z = z$ (see COVER and THOMAS, 1991). In other words,

$$L = \underbrace{-\sum_l \frac{t_l}{T} \ln\left(\frac{t_l}{T}\right)}_{H(V)} + \underbrace{\sum_s \frac{X^s}{T} \sum_l \frac{x_l^s}{X^s} \ln\left(\frac{x_l^s}{X^s}\right)}_{-H(V|Z)} = M(V; Z). \quad (4)$$

Therefore, M can also be written as

$$M = \sum_s \frac{X^s}{T} \Psi_1(x^s; t). \quad (5)$$

The M -index has been axiomatically characterized by FRANKEL and VOLIJ (2010) for analyzing overall school segregation in a multiracial context.¹¹ The parallels between

1
2
3 measurement of overall segregation across organizational units and measurement of
4 overall geographic concentration is evident. The former involves comparisons among
5 the distributions of racial groups across schools, while the latter requires comparing
6 distributions of industries across locations. Given the good properties of this index in a
7 segregation context (actually considered axioms, see JAMES and TAEUBER, 1985), it
8 also seems reasonable to analyze its suitability for quantifying overall concentration.
9

10
11
12 In what follows, we adapt these properties to our context (for more technical definitions,
13 see Appendix A). Some of these properties are similar to those previously defined for
14 partial measures except that now all sectors are involved, which implies adding specific
15 requirements for them.
16
17

18
19
20 The *M*-index satisfies *continuity* and is invariant to:
21

- 22 a) Any reordering of sectors and locations (*symmetry in sectors and locations*);
- 23 b) Proportional changes in all sectors and locations (*scale invariance II*);
- 24 c) Splitting one location into two if both have the same sectoral structure
25 (*insensitivity to proportional divisions of locations II*).

26
27
28 In addition,

- 29 d) If a sector is split into two subsectors both having the same distribution across
30 locations, concentration remains unaltered (*group division*). This implies that
31 concentration is unbiased to the industrial classification so long as there are no
32 internal differences among the spatial distributions of subsectors included in
33 each sector.
- 34 e) Given two economies having the same total employment and sizes of sectors
35 (locations), if an economy, representing new locations (sectors), is adjoined to
36 each of them, the ranking between them according to the index does not change
37 (*independence*). This is a separability property that means that if workers of a
38 sector (location) are reallocated among a subset of locations (sectors),
39 concentration rises if and only if concentration in these locations (sectors) rises.
40
41
42
43
44
45
46
47
48
49
50
51
52
53

54 An advantage of this index is that it can be decomposed additively by groups
55 (CUTRINI, 2009). Thus, if the manufacturing sectors are classified into several
56 mutually exclusive groups, it is possible to determine whether aggregate manufacturing
57 concentration is due mainly to the *between-group* (*B*) component (i.e., to differences in
58 the spatial distribution of employment between these groups) or to the *within-group* (*W*)
59
60

component (i.e., to differences in the spatial distributions of the sectors included in each group compared with employment distribution of the whole group). Let us assume that the manufacturing sectors are classified into (say) two groups: low-tech (G) and high-tech industries (H). Then,

$$M = M^B + M^W,$$

where:

$$M^B = \frac{X^G}{T} \sum_l \frac{x_l^G}{X^G} \ln \left(\frac{\frac{x_l^G}{X^G}}{\frac{t_l}{T}} \right) + \frac{X^H}{T} \sum_l \frac{x_l^H}{X^H} \ln \left(\frac{\frac{x_l^H}{X^H}}{\frac{t_l}{T}} \right), \text{ and}$$

$$M^W = \frac{X^G}{T} \sum_{s \in G} \frac{X^s}{X^G} \sum_l \frac{x_l^s}{X^s} \ln \left(\frac{\frac{x_l^s}{X^s}}{\frac{x_l^G}{X^G}} \right) + \frac{X^H}{T} \sum_{s \in H} \frac{X^s}{X^H} \sum_l \frac{x_l^s}{X^s} \ln \left(\frac{\frac{x_l^s}{X^s}}{\frac{x_l^H}{X^H}} \right),$$

where x_l^G (x_l^H) represents the low-tech (high-tech) employment in location l , and X^G (X^H) is the total low-tech (high-tech) employment in the economy.

The M -index can also be used to compare overall concentration in two different years. To do that, let us calculate the weighted average of index Ψ_1 in year 1 according to demographic weights in year 2: $\sum_s \frac{X^{s(2)}}{T^{(2)}} \Psi_1(x^{s(1)}; t^{(1)})$. Next, add and subtract that expression in $M^{(2)} - M^{(1)}$. It is straightforward to show that an intertemporal change between years (2) and (1) can be decomposed in two terms, one showing the gap due to changes in the spatial concentration of sectors and another due to changes in the sectoral structure of the economy:

$$M^{(2)} - M^{(1)} = \underbrace{\sum_s \frac{X^{s(2)}}{T^{(2)}} \left(\Psi_1(x^{s(2)}; t^{(2)}) - \Psi_1(x^{s(1)}; t^{(1)}) \right)}_{\text{concentration factor}} + \underbrace{\sum_s \Psi_1(x^{s(1)}; t^{(1)}) \left(\frac{X^{s(2)}}{T^{(2)}} - \frac{X^{s(1)}}{T^{(1)}} \right)}_{\text{sectoral structure factor}}.$$

The first component (concentration factor) aggregates the differences between concentration index Ψ_1 in years 2 and 1 ($\Psi_1(x^{s(2)}; t^{(2)}) - \Psi_1(x^{s(1)}; t^{(1)})$) according to the demographic weights in year 2. The second component (sectoral structure factor) aggregates the differences due to changes in the demographic weight of each sector ($(X^{s(2)} / T^{(2)}) - (X^{s(1)} / T^{(1)})$).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

M also satisfies several of the properties discussed in COMBES and OVERMAN (2004), as mentioned in CUTRINI (2009) with respect to the L -index. The within-between decomposition allows M to be comparable across activities at different scales.¹² In addition, since M is equal to the weighted average of Ψ_1 , which is decomposable by groups of locations, M is also comparable across spatial scales. For the same reason, statistical significance of the results is possible by using bootstrapping, and the index takes a unique value (zero) under the null hypothesis of no concentration.

3.2 Other overall concentration indexes

The unbounded Gini index, G , proposed by REARDON and FIREBAUGH (2002) to measure overall segregation can also be used to quantify aggregate concentration since it can be expressed as the weighted mean of index G^s for each sector:

$$G = \sum_s \frac{X^s}{T} G^s . \quad (6)$$

SILBER (1992) also offers an overall segregation index that extends the popular index of dissimilarity proposed by DUNCAN and DUNCAN (1955). This index can be adapted to measure overall concentration as follows:

$$IS = \frac{1}{2} \sum_s \sum_l \left| \frac{x_l^s}{T} - \frac{X^s}{T} \frac{t_l}{T} \right| .$$

It is easy to prove that this modified version can also be written as the weighted sum of index D^s for each sector into which the manufacturing industry can be partitioned:

$$IS = \sum_s \frac{X^s}{T} D^s . \quad (7)$$

Consequently, G and IS are naturally related to the aforementioned partial measures since they are weighted averages of G^s and D^s , respectively.¹³ This implies, first, that they can be used to determine the contribution of each sector to overall concentration, as happens in the case of M . And second, the differences among M , G , and IS rest on the disparities among the partial measures on which they are based. These indexes can also be decomposed like the M -index to undertake intertemporal comparisons.

In the empirical section, we show not only the results with the M -index but also with the above indexes. This allows us to check the robustness of our results since M , G , and IS

1
2
3 are based on partial concentration measures satisfying different properties (as shown in
4 Section 2.2), which means that these indexes do not necessarily lead to the same
5 outcomes.
6
7

8 9 10 4. Concentration of manufacturing industries in Spain: An 11 illustration 12

13
14
15 This section illustrates the use of the above tools in the case of Spain, showing the
16 usefulness of the decompositions of these indexes both by groups of sectors (classified
17 by technological intensity) and across time. The data used in this paper come from the
18 Labor Force Survey (*EPA*) conducted by the Spanish Institute of Statistics (*INE*)
19 following EUROSTAT guidelines. Our data correspond to the second quarter of each
20 year from 1977 to 2008. Manufacturing industries are considered at a two-digit level in
21 the National Classification of Economic Activities (*CNAE*), and the territorial scale is
22 that of provinces (nuts III).
23
24
25
26
27
28

29 30 4.1 Manufacturing concentration in 2008 by technological intensity 31

32
33 When analyzing spatial patterns of manufacturing industries in 2008, we find important
34 similarities in the rankings of these industries according to partial indexes $\Psi_1(x^s; t)$,
35 G^s , and D^s .¹⁴ Moreover, all these indexes coincide in detecting concentration in both
36 low-tech industries at a two-digit level (tobacco, leather, and recycling) and high-tech
37 industries (office and IT equipment; aircraft and other transport material).
38
39
40
41
42

43
44 To analyze in more detail whether substantial differences exist among industries
45 depending on their technological intensity, we group manufacturing industries by
46 following the OECD and *INE* classifications (four groups of sectors have been
47 considered, see Table 5 in Appendix B). Thus, we calculate the contribution of each
48 sector to the overall concentration at a two-digit level and then aggregate these
49 contributions according to the technology group to which each sector belongs. Three
50 overall measures are used to obtain the concentration of the whole manufacturing
51 industry: M , G , and IS (see Table 1 and expressions (5)-(7)).¹⁵ Note that the
52 contribution of a group of industries to overall concentration depends on two factors:
53 the concentration of its industries and the employment share of the group. If the
54 contributions of all groups are similar to their demographic weights, one can conclude
55
56
57
58
59
60

1
2
3 that there are barely differences among the concentration levels of the groups. So long
4 as the contribution of a group is higher than its demographic weight, the concentration
5 of the group is rather intense (and, therefore, higher than that of other groups). The
6 opposite happens when the contribution is lower than its demographic weight.
7
8
9

10
11 We find that the contribution of high-tech industries to overall concentration when
12 using M more than doubles their demographic weight (12.8% versus 5.1%) and it also
13 exceeds their demographic weight, although at a lower extent, when using G and IS
14 (8.6% versus 5.1%). Regarding medium-high-tech industries, the results do depend on
15 the index being used. On the contrary, in the case of medium-low-tech industries, their
16 contribution to overall concentration is remarkably lower than expected (21-22% versus
17 27%) while the contribution of low-tech sectors to overall concentration is similar to the
18 weight this group represents in terms of manufacturing employment.
19
20
21
22
23
24
25

26 [Insert Table 1 around here]
27
28

29
30 Consequently, high-tech industries seem more concentrated than industries with a lower
31 technological intensity; however, can we say that technological intensity is a relevant
32 variable to explain concentration in Spain? To answer this, we use the corresponding
33 within-between decomposition of index M . We find that the *between-group* component
34 explains around 28.5% of overall manufacturing concentration at a two-digit level.
35 Therefore, there are indeed important differences in the spatial distributions of these
36 four groups. Moreover, concentration increases with technological intensity (see Table
37 2). Thus, index Ψ_1 indicates that partial concentration in the high-tech group (jointly
38 considered) triples that of the medium-high-tech group, while that of the latter doubles
39 the value of the groups with lower technological intensity.¹⁶ Consequently, the spatial
40 pattern of the high-tech group considered as a whole (i.e., neglecting the disparities
41 among its sectors) clearly departs from the spatial distribution of total employment,
42 which leads to a high concentration level.
43
44
45
46
47
48
49
50
51
52

53 [Insert Table 2 around here]
54
55

56
57 From all of the above, it seems that the technological intensity of industries plays an
58 important role in explaining the concentration of manufacturing employment in Spain
59 since the high-tech industries, jointly considered, have a higher concentration value than
60 the remaining groups. This result is in line with that obtained by ALONSO-VILLAR et

1
2
3 al. (2004), who used the index proposed by MAUREL and SÉDILLOT (1999) and a
4 different Spanish dataset for 1999. Therefore, the results for the Spanish economy seem
5 rather robust.
6
7

10 4.2 Evolution of manufacturing concentration, 1977-2008

12 Figure 1 shows the evolution of the overall concentration of the manufacturing industry
13 at a two-digit level from 1977 to 2008.¹⁷ We can establish three periods of change in the
14 evolution of spatial concentration. First, we observe a slight increase until 1981, a
15 noteworthy decrease from 1985 to 1990, and finally, a remarkable decline from 2001
16 onward.¹⁸ The bootstrap analysis for the M index based on 10,000 replications shows
17 that the three changes are significant at the 99% level of significance (the null
18 hypothesis being $\Delta M = 0$).¹⁹
19
20
21
22
23
24
25

26 [Insert Figure 1 around here]

27
28 By using the temporal decompositions of the above overall measures (Table 3), we find
29 that the overall concentration rise in the first years of democracy is due entirely to the
30 increase in concentration of the manufacturing industry (concentration factor).
31 Decreases in the other two periods are also due mainly to changes in the concentration
32 levels of the manufacturing industries (at least 60% of total change), even though the
33 sectoral structure factor has an important weight, as well.
34
35
36
37
38
39

40 [Insert Table 3 around here]

41
42 [Insert Figure 2 around here]

43
44 Figure 2 shows the contributions of industries, grouped by technological intensity, to
45 overall concentration at a two-digit level throughout the entire period.²⁰ We observe a
46 decreasing tendency in the contribution of low-tech industries to overall concentration
47 from 1999 onward, together with an increase in the contribution of medium-high-tech
48 industries from 2004 onward. We should note, however, that while the decrease in the
49 low-tech group could be explained by the evolution of employment in this group, the
50 evolution of the medium-high-tech group seems to be explained only partially by that
51 factor (see Figure 3 in Appendix B).
52
53
54
55
56
57
58
59
60

5. Final comments

The measurement of population segregation across organizational units (occupations, schools, neighborhoods, etc.) and the measurement of spatial concentration share much in common. Thus, while the former focuses on the distribution of racial groups across schools, for example, the latter addresses the distribution of economic sectors across location units; however, each field has dealt with measurement from a different perspective. On the one hand, while the segregation literature has tackled this matter mainly from an axiomatic point of view, the literature on spatial concentration has not. On the other hand, the former has focused mostly on the measurement of overall segregation, whereas the latter has dealt with the concentration of any single sector (labeled here partial concentration) rather than with overall concentration; exceptions are AINGER and DAVIES (2004), MULLIGAN and SCHMIDT (2005), and CUTRINI (2009, 2010).

Given the parallels between the two phenomena, this paper has proposed two overall concentration measures which are adapted from segregation measures (REARDON and FIREBAUGH, 2002, and SILBER, 1992). In addition, using FRANKEL and VOLIJ's (2010) characterization of the mutual information index, used to quantify school segregation, this paper has shown the properties of the L-index proposed by CUTRINI (2009). The partial concentration index in which the L-index is based, the Theil index, together with the remaining members of the GE family of concentration indexes, has been also characterized in terms of basic properties adapted from the literature on income distribution and segregation to our case. Consequently, this paper brings analytical support to partial and overall concentration measures existing in the literature (BRÜLHART and TRAEGER, 2005; CUTRINI, 2009, 2010).

Finally, these indexes have been used to measure overall concentration of Spanish manufacturing industries over the last three decades. The analysis reveals that the process of economic integration of Spain into the EU, signed in 1986, together with improvements in transport infrastructure, has not fueled the spatial concentration of the Spanish manufacturing industry. On the contrary, this process seems to be accompanied by a decreasing, though intermittent, trend in the concentration level that continues today. We have also found that the technological intensity of an industry is a relevant variable to explain geographic concentration in Spain.

Acknowledgements

Financial support from the *Xunta de Galicia* (INCITE08PXIB300005PR) and from FEDER is gratefully acknowledged.

Appendix A

Properties of partial measures

Symmetry in locations. If $(x^s'; t')$ is obtained from $(x^s; t)$ by a permutation of locations, then $I_c(x^s'; t') = I_c(x^s; t)$.

Movement between locations. If $(x^s'; t)$ is obtained from $(x^s; t)$ in such a way that: (i) $x_i^s' = x_i^s - d$, $x_h^s' = x_h^s + d$ ($0 < d \leq x_i^s$), where i and h are two locations satisfying that $t_i = t_h$, and $x_i^s < x_h^s$; and (ii) $x_l^s' = x_l^s \quad \forall l \neq i, h$; then $I_c(x^s'; t) > I_c(x^s; t)$.

Scale invariance I. If $a, b \in R^{++}$ ($ax_i^s \leq bt_i$), then $I_c(ax^s; bt) = I_c(x^s; t)$.

Insensitivity to proportional subdivisions of locations I. If $(x'; t')$ is obtained from $(x^s; t)$ after splitting location l into two in such a way that the size of the sector in each of them is $x_l/2$ and the size of the distribution of reference in each of them is $t_l/2$, then $I_c(x^s'; t') = I_c(x^s; t)$.

Aggregation. Let us classify locations into two mutually exclusive groups so that $(x^s; t) = (x^{s1}, x^{s2}; t^1, t^2)$, where the aggregate employment level in locations included in group 1 (2) is denoted by T_1 (T_2), while X_1^s (X_2^s) represents the employment level of the sector of study in the corresponding group of locations. Concentration index I_c is defined as aggregative if there exists a continuous aggregator function A such that $I_c(x^s; t) = A\left(I_c(x^{s1}; t^1), \frac{X_1^s}{T_1}, T_1, I_c(x^{s2}; t^2), \frac{X_2^s}{T_2}, T_2\right)$, where A is strictly increasing in the first and fourth arguments.

Properties of overall measures

In order to formally define the corresponding properties of overall measures, we denote overall index by OI_c and define matrix E :

$$E = \begin{bmatrix} x_1^1 & \cdots & x_L^1 \\ \vdots & & \vdots \\ x_1^s & \cdots & x_L^s \end{bmatrix}$$

Symmetry in sectors and locations. If E' is obtained from E by a permutation of sectors/locations, then $OI_c(E') = OI_c(E)$.

Scale invariance II. If the size of all sectors in all locations is multiplied by $\lambda \in \mathbb{R}^{++}$, overall concentration does not change, i.e., $OI_c(\lambda E) = OI_c(E)$.

Insensitivity to proportional subdivisions of locations II. If matrix E' results from E after splitting a location l into two in such a way that the proportion of each sector s (with respect to t) in each new location is the same, then $OI_c(E) = OI_c(E')$.

Group division. If E' results from E by splitting sector s into two subsectors, s' and s'' , such that both have the same spatial distribution $(\frac{x_l^{s'}}{X^{s'}} = \frac{x_l^{s''}}{X^{s''}} \quad \forall l)$, then $OI_c(E') = OI_c(E)$.

Independence. Consider two economies, E and \tilde{E} , having the same total employment and the same employment level in each sector (location). If E' is obtained from E and \tilde{E}' from \tilde{E} after adjoining to each of them a common economy with new locations (sectors), then $OI_c(E') \geq OI_c(\tilde{E}')$ if and only if $OI_c(E) \geq OI_c(\tilde{E})$.

Proof of proposition

In the first step, we establish a formal relationship between geographic concentration and inequality (i.e., we show that any concentration index I_c satisfying properties 1-4 leads to a inequality index, I , satisfying scale invariance, symmetry, the Pigou-Dalton

principle, and replication invariance). In the second step, we make use of this relation, together with the results from the inequality literature, to show that any concentration index satisfying properties 1-5 is a strictly increasing monotonic transformation of Ψ_α . In the third step, we show that any strictly increasing monotonic transformation of Ψ_α satisfies properties 1-5.

First step.

If the concentration index I_c satisfies properties 1-4, then index I evaluated at the

hypothetical distribution $y \equiv \left(\underbrace{\frac{x_1^s}{t_1}, \dots, \frac{x_1^s}{t_1}}_{t_1}, \dots, \underbrace{\frac{x_L^s}{t_L}, \dots, \frac{x_L^s}{t_L}}_{t_L} \right)$ as $I(y) := I_c(x^s; t)$ works as an

inequality index since it satisfies scale invariance, symmetry, the Pigou-Dalton principle, and replication invariance (these are axioms proposed in the literature on income distribution and we maintain their original labels).

a) I is well defined. Note that several vectors $(x^s; t)$ can be reached after grouping individuals in the hypothetical distribution who belong to the same location depending on how many locations are considered. However, by property 4, all of these vectors have the same spatial concentration level since they can be obtained from each other by proportional subdivisions.

b) Scale invariance. This property is satisfied by index I since

$$I\left(\theta \frac{x_1^s}{t_1}, \dots, \theta \frac{x_1^s}{t_1}, \dots, \theta \frac{x_L^s}{t_L}, \dots, \theta \frac{x_L^s}{t_L}\right) = I_c(\theta x^s; t), \text{ which is equal to } I_c(x^s; t) \text{ because}$$

I_c satisfies property 3 (case where $a > 0$, $b = 1$).

c) Symmetry. It requires that individuals play symmetric roles in the inequality index. This is satisfied by I since I_c satisfies properties 1 and 4.

d) The Pigou-Dalton transfer principle. According to property 4, any regressive transfer in this hypothetical economy can be expressed as a sequence of disequalizing employment movements in an economy constructed from the original one by proportional subdivisions of locations so that the distribution of

reference becomes $\left(\underbrace{1, \dots, 1}_T\right)$. Since I_c satisfies property 2, the second situation

leads to a higher concentration index and, therefore, to a higher value of I .

- e) Replication invariance. It means 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 not altered. This property is satisfied here since a k -replication of the hypothetical distribution leads to a k -replication of vector $(x^s; t)$, and I_c satisfies property 3 (case where $a = b$).

Second step.

Following SHORROCKS (1984) and FOSTER (1985), any continuous inequality measure I taking a zero value at the egalitarian distribution and satisfying scale invariance, replication invariance, the Pigou-Dalton transfer principle, symmetry, and aggregation can be written as $I(y) = F^{-1}(I_\alpha(y))$ for some parameter α , where F is a strictly increasing function such that $F: [0, \infty) \rightarrow \mathbb{R}$ with $F(0) = 0$ and I_α is the well-known GE family of inequality indexes:

$$I_\alpha(y) = \begin{cases} \frac{1}{n\alpha(\alpha-1)} \sum_i \left[\left(\frac{y_i}{\frac{1}{n} \sum_k y_k} \right)^\alpha - 1 \right] & \text{if } \alpha \neq 0, 1 \\ \frac{1}{n} \sum_i \left[\frac{y_i}{\frac{1}{n} \sum_k y_k} \ln \left(\frac{y_i}{\frac{1}{n} \sum_k y_k} \right) \right] & \text{if } \alpha = 1 \\ \frac{1}{n} \sum_i \ln \left(\frac{\frac{1}{n} \sum_k y_k}{y_i} \right) & \text{if } \alpha = 0 \end{cases}$$

In Step 1, we proved that any concentration index I_c satisfying properties 1-4 can be regarded as an inequality index I satisfying scale invariance, symmetry, the Pigou-Dalton transfer principle and replication invariance. It is easy to see that if I_c is a continuous function, so too is I . If we additionally show that I is aggregative and also

that it is equal to zero at the egalitarian distribution, we can use Shorrocks's result to characterize inequality index I .

An inequality index I is defined as aggregative if $I(y) = A(I(y^1), \mu(y^1), n(y^1), I(y^2), \mu(y^2), n(y^2))$, where A is a continuous function that is strictly increasing in the first and fourth arguments, y^i represents the income distribution corresponding on the individuals' group i , $\mu(\cdot)$ is the average of the corresponding distribution, and $n(\cdot)$ is the number of individuals in the corresponding group. In our case, $y \equiv \left(\underbrace{\frac{x_1^s}{t_1}, \dots, \frac{x_i^s}{t_i}}_{t_1}, \dots, \underbrace{\frac{x_L^s}{t_L}, \dots, \frac{x_L^s}{t_L}}_{t_L} \right)$, and the average of that distribution is

equal to $\frac{X^s}{T}$. In what follows, we show that our I is an aggregative inequality index.

For the sake of simplicity, assume that class 1 includes locations $l = 1, \dots, i$, while class 2 is the complementary. By definition

$$I \left(\underbrace{\frac{x_1^s}{t_1}, \dots, \frac{x_i^s}{t_i}}_{\text{class 1}}, \dots, \underbrace{\frac{x_{i+1}^s}{t_{i+1}}, \dots, \frac{x_L^s}{t_L}}_{\text{class 2}} \right) = I_c(x^s; t).$$

According to property 5, I_c is an aggregative concentration index:

$$I_c(x^s; t) = I_c(x^{s1}, x^{s2}; t^1, t^2) = A \left(I_c(x^{s1}; t^1), \frac{X_1^s}{T_1}, T_1, I_c(x^{s2}; t^2), \frac{X_2^s}{T_2}, T_2 \right).$$

Note that $I_c(x^{s1}; t^1) = I \left(\frac{x_1^s}{t_1}, \dots, \frac{x_i^s}{t_i} \right)$, and $I_c(x^{s2}; t^2) = I \left(\frac{x_{i+1}^s}{t_{i+1}}, \dots, \frac{x_L^s}{t_L} \right)$.

Therefore, the inequality index I is aggregative:

$$I \left(\underbrace{\frac{x_1^s}{t_1}, \dots, \frac{x_i^s}{t_i}}_{\text{class 1}}, \dots, \underbrace{\frac{x_{i+1}^s}{t_{i+1}}, \dots, \frac{x_L^s}{t_L}}_{\text{class 2}} \right) = A \left(I \left(\frac{x_1^s}{t_1}, \dots, \frac{x_i^s}{t_i} \right), \frac{X_1^s}{T_1}, T_1, I \left(\frac{x_{i+1}^s}{t_{i+1}}, \dots, \frac{x_L^s}{t_L} \right), \frac{X_2^s}{T_2}, T_2 \right).$$

Finally, note that I is equal to zero at the equalitarian distribution (i.e., when all locations have the same employment shares in the sector).

Therefore, by using Shorrocks's result, it follows that $I(y) = F^{-1}(I_\alpha(y))$ for $\alpha \neq 0, 1$ or $\alpha = 1$.²¹ On the other hand, $I_c(x^s; t) = I(y)$ and $F^{-1}(I_\alpha(y)) = F^{-1}(\Psi_\alpha(x^s; t))$, which completes the proof of step two.

Third step.

To prove that $F^{-1}(\Psi_\alpha)$ is a concentration index satisfying properties 1-5, it suffices to show that Ψ_α satisfies them, which is done in what follows. It is easy to prove that Ψ_α verifies *scale invariance I*, *symmetry in locations*, and *insensitivity to proportional subdivisions I*. To demonstrate that Ψ_α satisfies the property of *movement between locations*, note that any disequalizing movement from location i to h , where $t_i = t_h$ and

$x_i^s < x_h^s$, implies moving from distribution $y = \left(\frac{x_1^s}{t_1}, \dots, \frac{x_1^s}{t_1}, \dots, \frac{x_i^s}{t_i}, \dots, \frac{x_i^s}{t_i}, \dots, \frac{x_h^s}{t_h}, \dots, \frac{x_h^s}{t_h}, \dots, \frac{x_L^s}{t_L}, \dots, \frac{x_L^s}{t_L} \right)$ to distribution

$y' = \left(\frac{x_1^s}{t_1}, \dots, \frac{x_1^s}{t_1}, \dots, \frac{x_i^s - d}{t_i}, \dots, \frac{x_i^s - d}{t_i}, \dots, \frac{x_h^s + d}{t_h}, \dots, \frac{x_h^s + d}{t_h}, \dots, \frac{x_L^s}{t_L}, \dots, \frac{x_L^s}{t_L} \right)$. Note that

$I_\alpha(y) = \Psi_\alpha(x^s; t)$ and $I_\alpha(y') = \Psi_\alpha(x^s; t')$. 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 $\Psi_\alpha(x^s; t') > \Psi_\alpha(x^s; t)$.

Next, we prove that Ψ_α is aggregative. Using simple calculations Ψ_α can be written as

$$\Psi_\alpha(x^{s1}, x^{s2}; t^1, t^2) = \begin{cases} -\frac{1}{\alpha(\alpha-1)} + \left[\left(\frac{T_1}{T} \right)^{1-\alpha} \left(\frac{X_1^s}{X^s} \right)^\alpha \left(\Psi_\alpha(x^{s1}; t^1) + \frac{1}{\alpha(\alpha-1)} \right) + \left(\frac{T_2}{T} \right)^{1-\alpha} \left(\frac{X_2^s}{X^s} \right)^\alpha \left(\Psi_\alpha(x^{s2}; t^2) + \frac{1}{\alpha(\alpha-1)} \right) \right] & \text{for } \alpha \neq 0, 1 \\ \frac{X^{s1}}{X^s} \left[\Psi_\alpha(x^{s1}; t^1) + \ln \left(\frac{X_1^s T}{T_1 X^s} \right) \right] + \frac{X^{s2}}{X^s} \left[\Psi_\alpha(x^{s2}; t^2) + \ln \left(\frac{X_2^s T}{T_2 X^s} \right) \right] & \text{for } \alpha = 1 \end{cases}$$

On the other hand, $T = T_1 + T_2$ and $X^s = X_1^s + X_2^s$. Therefore, Ψ_α can be written as

$$\Psi_\alpha(x^{s1}, x^{s2}; t^1, t^2) = A\left(\Psi_\alpha(x^{s1}; t^1), \frac{X_1^s}{T_1}, T_1, \Psi_\alpha(x^{s2}; t^2), \frac{X_2^s}{T_2}, T_2\right),$$

which completes the proof. □

Appendix B

[Insert Table 4 around here]

[Insert Table 5 around here]

[Insert Figure 3 around here]

References

- AINGINGER, K. and DAVIES, S.W. (2004) Industrial specialization and geographic concentration: two sides of the same coin? Not for the European Union, *Journal of Applied Economics* VII(2), 231-248.
- AINGINGER, K. and PFAFFERMAYR, M. (2004) The single market and geographic concentration in Europe, *Review of International Economics* 12(1), 1-11.
- ALONSO-VILLAR, O. (2011) Measuring geographic concentration: Lorenz curves and their decompositions, *Annals of Regional Science*, forthcoming (doi:10.1007/s00168-010-0369-5).
- ALONSO-VILLAR, O., CHAMORRO-RIVAS, J.M., and GONZÁLEZ-CERDEIRA, X. (2004) Agglomeration economies in manufacturing industries: the case of Spain, *Applied Economics* 36(18), 2103- 2116.
- ALONSO-VILLAR, O. and DEL RÍO, C. (2010a) Local versus overall segregation measures, *Mathematical Social Sciences* 60, 30-38.
- ALONSO-VILLAR, O. and DEL RÍO, C. (2010b) Concentration of economic activity: inequality-based measures, Documento de Trabajo 1006, Departamento de Economía Aplicada, Universidade de Vigo.

1
2
3 AMITI, M. (1999) Specialization patterns in Europe, *Weltwirtschaftliches Archiv*
4 135(4), 573-593.
5

6
7
8 ARBIA, G. (1989): *Spatial data configuration in statistical analysis of regional*
9 *economic and related problems*. Kluwer, Dordrecht.
10

11
12 ATKINSON, A.B. (1970) On the measurement of inequality, *Journal of Economic*
13 *Theory* 2, 244-263.
14

15
16
17 BICKENBACH, F. and BODE, E. (2008) Disproportionality measures of concentration,
18 specialization, and localization, *International Regional Science Review* 31, 359-388.
19

20
21 BOTAZZI, D., DOSI, G., FAGIOLO, G., and SECCHI, A. (2008) Sectoral and
22 geographical specificities in the spatial structure of economic activities, *Structural*
23 *Change and Economic Dynamics* 19, 189-202.
24
25

26
27
28 BRAKMAN, S., GARRETSEN, H., GORTER, J., VAN DER HORST, A. and
29 SCHRAMM, M. (2005) New economic geography, empirics, and regional policy, CPB
30 (Netherlands Bureau for Economic Policy Analysis) Special Publication 56.
31
32

33
34 BRÜLHART, M. (2001) Evolving geographical concentration of European
35 manufacturing industries, *Weltwirtschaftliches Archiv* 137(2), 215-243.
36
37

38
39 BRÜLHART, M. and TRAEGER, R. (2005) An account of geographic concentration
40 patterns in Europe, *Regional Science and Urban Economics* 35(6), 597-624.
41

42
43 COMBES, P.-P. and OVERMAN, H. (2004) The spatial distribution of economic
44 activities in the EU, in HENDERSON J.V. and THISSE J.-F. (Eds), *Handbook of*
45 *Regional and Urban Economics* 4, pp. 2845-2909. Elsevier-North Holland, Amsterdam.
46
47

48
49 COVER, T.M. and THOMAS, Y.A. (1991): *Elements of Information Theory*. Wiley
50 Interscience, New York.
51

52
53
54 CUTRINI, E. (2009) Using entropy measures to disentangle regional from national
55 localization patterns, *Regional Science and Urban Economics* 39, 243-250.
56

57
58
59 CUTRINI, E. (2010) Specialization and concentration from a twofold geographical
60 perspective: evidence from Europe, *Regional Studies* 44(3), 315-336.

1
2
3 DUNCAN, O.D. and DUNCAN, B. (1955) A methodological analysis of segregation
4 indexes, *American Sociological Review* 20(2), 210-217.
5
6

7
8 DURANTON, G. and OVERMAN, H. (2005) Testing for localization using micro-
9 geographic data, *Review of Economic Studies* 72, 1077-1106.
10
11

12
13 ELLISON, G. and GLAESER, E. (1997) Geographic concentration in U.S.
14 manufacturing industries: a dartboard approach, *Journal of Political Economy* 105, 889-
15 927.
16
17

18
19 FOSTER, J.E. (1985): Inequality measurement, in PEYTON JOUND H. (Ed), *Fair*
20 *Allocation, American Mathematical Society. Proceedings of Simposia in Applied*
21 *Mathematics* 33, pp. 31-68, Providence.
22
23

24
25 FRANKEL, D.M. and VOLIJ, O. (2010) Measuring School Segregation, *Journal of*
26 *Economic Theory*, forthcoming (doi: 10.1016/j.physletb.2003.10.071).
27
28

29
30 GUIMARÃES, P., FIGUEIREDO, O. and WOODWARD, D. (2007) Measuring the
31 localization of economic activity: a parametric approach, *Journal of Regional Science*
32 47(4), 753-774.
33
34

35
36 HAALAND, J., KIND, H., MIDELFART KNARVIK, K., and TORSTENSSON, J.
37 (1999) What determines the economic geography of Europe?, CEPR Discussion Paper
38 2072.
39
40

41
42 JAMES, D.R. and TAEUBER, K.E. (1985) Measures of segregation, *Sociological*
43 *Methodology* 15, 1-32.
44
45

46
47 KAPUR, J.N. and KESSAVAN, H.K. (1992) *Entropy Optimization Principles with*
48 *Applications*. Academic Press, San Diego.
49
50

51
52 KRUGMAN, P. (1991) *Geography and Trade*. MIT Press, Cambridge, Massachusetts.
53

54
55 MARCON, E. and PUECH, F. (2003) Evaluating the geographic concentration of
56 industries using distance-base methods, *Journal of Economic Geography* 3, 409-428.
57
58

59
60 MAUREL, F. and SÉDILLOT, B. (1999) A measure of the geographic concentration in
French manufacturing industries, *Regional Science and Urban Economics* 29, 575-604.

1
2
3 MORA, R. and RUIZ-CASTILLO, J. (2010) A Kullback-Leibler measure of
4 conditional segregation, Working Paper-Economic Series 10-15, Department of
5 Economics, Universidad Carlos III de Madrid.
6
7

8
9
10 MORI, T., NISHIKIMI, K., and SMITH, T. (2005) A divergence statistic for industrial
11 localization, *The Review of Economics and Statistics* 87(4), 635-651.
12

13
14 MULLIGAN, G.F. and SCHMIDT C. (2005) A note on localization and specialization,
15 *Growth and Change* 36(4), 565-575.
16
17

18
19 OECD (2007): *OECD Science, Technology and Industry. Scoreboard 2007*.
20

21
22 PALUZIE, E., PONS, J., and TIRADO, D. (2004) The geographical concentration of
23 industry across Spanish regions, 1856-1995, *Review of Regional Research*, 24(2), 143-
24 160.
25
26

27
28 PÉREZ-XIMÉNEZ, D. and SANZ-GRACIA, F. (2007) Geographical concentration of
29 service activities across U.S: states and counties. 1969-2000. Paper presented at the
30 XVII International RESER (European Research Network on Services and Space)
31 Conference in Tempere (Finland).
32
33

34
35 REARDON, S.F. and FIREBAUGH, G. (2002) Measures of multigroup segregation,
36 *Sociological Methodology* 32, 33-76.
37
38

39
40 RESMINI, L. (2007) Regional patterns of industry location in transition countries: does
41 economic integration with the European Union matter?, *Regional Studies* 41(6), 747-
42 764.
43
44

45
46 SHANNON, C.E. (1948) A mathematical theory of communication, *The Bell System*
47 *Technical Journal* 27, 379-423, 623-656.
48
49

50
51 SHORROCKS, A.F. (1984) Inequality decompositions by population subgroups,
52 *Econometrica* 48, 1369-1386.
53
54

55
56 SILBER, J. (1992) Occupational segregation indices in the multidimensional case: a
57 note, *The Economic Record* 68, 276-277.
58
59
60

1
2
3 SUEDEKUM, J. (2006) Concentration and specialization trends in Germany since re-
4 unification, *Regional Studies* 40(8), 861-873.
5
6

7
8 WHITE, M.J. (1983) The measurement of spatial segregation, *American Journal of*
9 *Sociology* 88 (5), 1008-1018.
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review Only

Technological groups 2008	Contribution to <i>M</i> (%)	Contribution to <i>G</i> (%)	Contribution to <i>IS</i> (%)	Employment share (%)
High-tech	12.8	8.7	8.6	5.1
Medium-high-tech	23.7	26.8	26.8	25.3
Medium-low-tech	21.3	22.1	21.4	27.2
Low-tech	42.2	42.4	43.2	42.3

Table 1. Contribution of each group of sectors to overall concentration at a two-digit level and their employment shares.

Technological groups 2008	$\Psi_1(x^s; t)$	G^s	D^s
High-tech	0.336	0.438	0.330
Medium-high-tech	0.101	0.250	0.190
Medium-low-tech	0.044	0.160	0.119
Low-tech	0.037	0.151	0.109

Table 2. Partial concentration indexes of each technological group.

	Concentration/structure factor decomposition <i>M</i>	Concentration/structure factor decomposition <i>G</i>	Concentration/structure factor decomposition <i>IS</i>
1977-1981	109.98% -9.98%	109.17% -9.17%	109.16% -9.16%
1985-1990	64.21% 35.79%	72.73% 27.27%	73.89% 26.11%
2001-2008	59.86% 40.14%	75.82% 24.18%	75.34% 24.66%

Table 3. Intertemporal decompositions of overall concentration indexes.

For Peer Review Only

High-technology group	<i>Office and computing machinery</i> (33) <i>Electronics</i> (35) <i>Medical, precision and optical instruments</i> (39) <i>Manufacture of other transport material (including aircraft)</i> (38)
Medium-high-technology group	<i>Chemicals</i> (25) <i>Machinery and equipment</i> (32) <i>Electrical machinery and apparatus</i> (34) <i>Motor vehicle and trailers</i> (36) <i>Shipping building</i> (37)
Medium-low technology group	<i>Refined petroleum products</i> (13) <i>Rubber and plastic products</i> (48) <i>Non-metallic mineral products</i> (24) <i>Fabricated metal products(except machinery and transport material)</i> (31) <i>Production and first transformation of metals</i> (22)
Low-technology group	<i>Food products, beverages, and tobacco</i> (41, 42) <i>Textile industry</i> (43) <i>Leather industry</i> (44) <i>Footwear and clothing</i> (45) <i>Wood, cork industry, and furniture</i> (46) <i>Paper industry, publishing, and graphic arts</i> (47) <i>Other manufacturing industries n.e.c.</i> (49)

Table 4. Classification of two-digit industries by technological intensity: 1977-1992.

High-technology group	<i>Office, accounting, and computing machinery</i> (30) <i>Radio, TV, and communications equipment</i> (32) <i>Medical, precision and optical instruments</i> (33) <i>Manufacture of other transport material (including aircraft)</i> (35)
Medium-high-technology group	<i>Chemicals</i> (24) <i>Machinery and equipment n.e.c.</i> (29) <i>Electrical machinery and apparatus n.e.c.</i> (31) <i>Motor vehicles, trailers and semi-trailers</i> (34)
Medium-low-technology group	<i>Coke, refined petroleum products, and nuclear fuel</i> (23) <i>Rubber and plastic products</i> (25) <i>Other non-metallic mineral products</i> (26) <i>Metallurgy</i> (27) <i>Fabricated metal products (machinery and equipment excluded)</i> (28)
Low-technology group	<i>Food products and beverages</i> (15) <i>Tobacco</i> (16) <i>Textile industry</i> (17) <i>Clothing and fur industry</i> (18) <i>Leather and footwear</i> (19) <i>Wood and cork industry, except furniture</i> (20) <i>Paper industry</i> (21) <i>Publishing, graphic arts, and reproduction of recorded supports</i> (22) <i>Manufacture of furniture and other manufacturing industries n.e.c.</i> (36) <i>Recycling</i> (37)

Table 5. Classification of two-digit industries by technological intensity: 1993-2008.²²

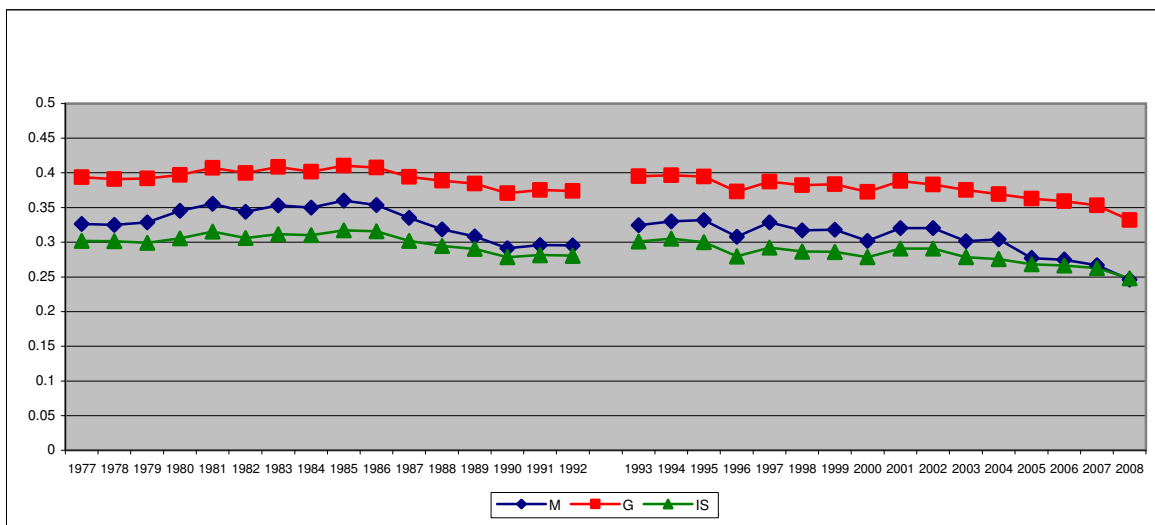


Figure 1. Overall concentration of the manufacturing industry over the period 1977-2008.

Peer Review Only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

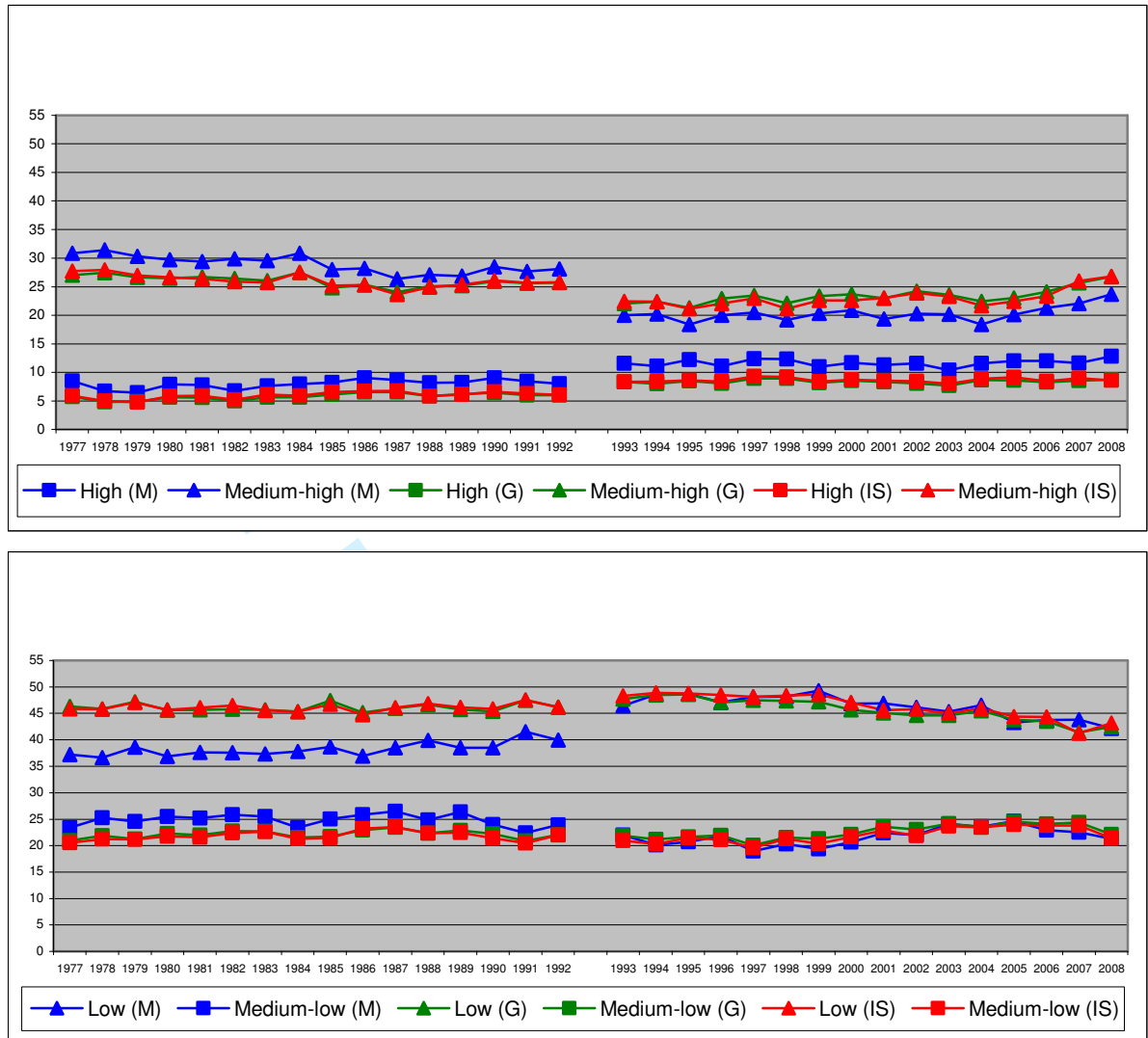


Figure 2. Contribution of each technological group to three overall concentration indexes.

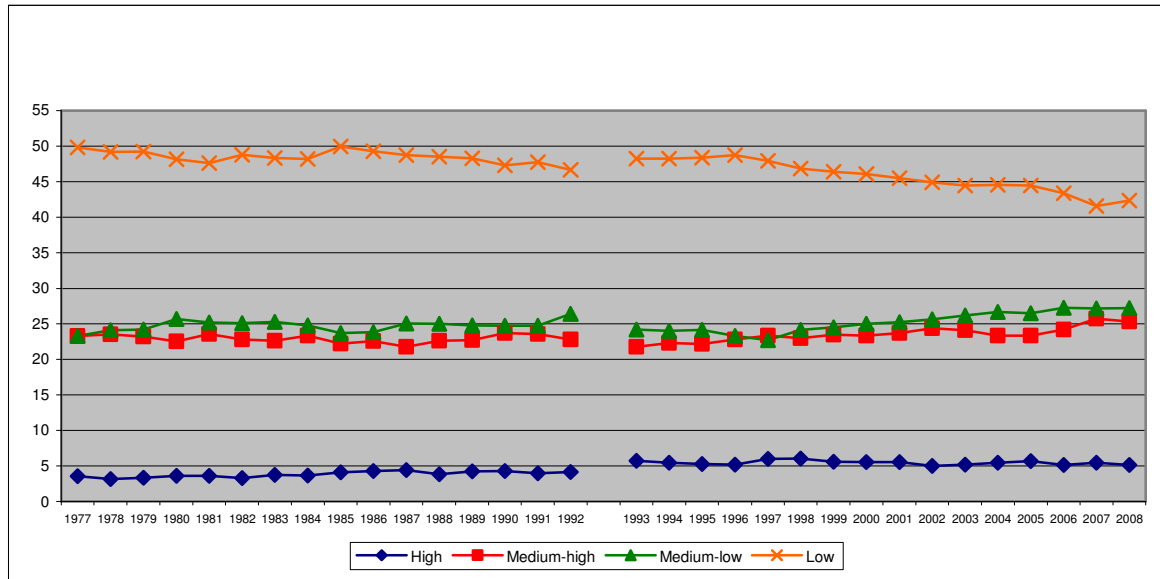


Figure 3. Manufacturing employment share along the period 1977-2008.

Peer Review Only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

¹ Some of these properties have been used in ALONSO-VILLAR (2011) to characterize employment Lorenz curves.

² For an econometric approach to the roles played by natural advantages and technological spillovers, see BOTTAZZI et al. (2008).

³ The properties of entropy measures outside economics can be seen in SHANNON (1948) and KAPUR and KESAVAN (1992).

⁴ For other proposals, see BICKENBACK and BODE (2008) and ALONSO-VILLAR and DEL RÍO (2010b).

⁵ The GE family also includes $\Psi_0(x^s; t) = \sum_l \frac{t_l}{T} \ln \left(\frac{t_l / T}{x_l^s / X^s} \right)$ if $x_l^s \neq 0 \quad \forall l$.

⁶ The values of these indexes are unbounded.

⁷ For technical definitions see Appendix A.

⁸ Properties 2 and 4 imply aversion to inequality in distribution $\left(\frac{x_1^s}{t_1}, \dots, \frac{x_L^s}{t_L} \right)$.

⁹ This is related to a property proposed in the literature of income distribution according to which, the greater inequality increase, the lower the income of the poorer individual involved in a regressive transfer from the poorer to the richer (see ATKINSON, 1970). In our context, this means that disequalizing movements of workers between locations lead to higher concentration increases in the sector when these movements occur between locations in which the relative presence of this sector, as compared with overall employment there, is low.

¹⁰ G^s and D^s take values in the interval $[0,1)$.

¹¹ For additional properties, see MORA and RUIZ-CASTILLO (2010).

¹² If grouping (say) three-digit industries into two-digit industries, overall concentration at a three-digit level can be expressed as aggregate concentration at a two-digit level (the between-group component) plus the concentration inside each two-digit industry considered separately (the distribution of each three-digit industry is compared with that of the two-digit industry to which it belongs rather than with distribution t).

¹³ As opposed to the *IS* index, which is based on comparisons between $\frac{x_l^s}{X^s}$ and $\frac{t_l}{T}$ in each location, in the global concentration index proposed by MULLIGAN and SCHMIDT (2005), $G(L) = \sum_s \frac{X^s}{T} \left(\frac{1}{2} \sum_l \left| \frac{x_l^s}{t_l} - \frac{X^s}{T} \right| \right)$, $\frac{x_l^s}{t_l}$ is contrasted with $\frac{X^s}{T}$.

¹⁴ In fact, the Spearman correlation coefficient between these indexes ranges between 0.94 and 0.98.

¹⁵ The values of these indexes for 2008 are shown in Figure 1, which includes not only these figures but also those of the whole period.

¹⁶ This result is not a consequence of the small weight of the high-tech industry in the economy since the GE family is unaffected by the size of the sector (see property 3).

¹⁷ The change from 1992 to 1993 is due to use of a different classification of manufacturing industries: From 1977 to 1992, *EPA* gathered information at a two-digit level according to classification *CNAE-1974*, while from 1993 to 2008, the classification used was *CNAE-93*.

¹⁸ This finding corroborates that obtained by PALUZIE et al. (2004), who used alternative indexes and datasets for the period 1955-1995.

¹⁹ For arguments in favor of using this method in the case of entropy measures, see BRÜLHART and TRAEGER (2005) and CUTRINI (2010).

²⁰ Since classification of the manufacturing industry at a two-digit level between 1977 and 1992 is different from the one used between 1993 and 2008, we have used another technological grouping (see Tables 4 and 5 in Appendix B). In any case, a broad criterion similar to that used in 2008 has been used for the whole period.

²¹ The case where $\alpha = 0$ is discarded, because when the sector has no employment in location l (i.e., when $x_l^s = 0$) and $\alpha = 0$, the index value would be infinite and, therefore, it makes no sense. The case where $\alpha = 1$ does not have the same problem since

$$\lim_{x_l^s \rightarrow 0} \frac{x_l^s / X^s}{t_l / T} \ln \left(\frac{x_l^s / X^s}{t_l / T} \right) = 0.$$

²² Since in this study industries are considered at a two-digit level and the OECD (2007) and *INE* classifications consider both two- and three-digit industries, we have introduced some changes with respect to them. In particular, the *INE* includes one of the subsectors of sector **24** in the high-tech group and the remaining subsectors in the medium-high-tech group. Here, we have decided to include the whole sector in the latter

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

group. On the other hand, the *INE* classifies part of sector **35** (i.e. aircraft) in the former group and part in the latter. We have decided to include the whole sector in the former.

For Peer Review Only