A Note on New Measures of Agglomeration and Specialization

Francesco Franceschi
(University of Waterloo, Ontario, Canada)
Maurizio Mussoni
(University of Bologna and Rimini Centre for Economic Analysis)
Gianluigi Pelloni*†
(University of Bologna, Wilfrid Laurier University
and Rimini Centre for Economic Analysis)

Abstract

Most measures and indexes of industrial agglomeration, concentration and specialization are variations of the Gini's index and may be limited in scope and structure. Rarely, if ever, they include any reference to territorial dimensions and tend to be measures of volumes adjusted for the task at end. They do not embody any information relative to the industrial density and size of the analyzed regions. Alternatively some authors use just density as a measure of agglomeration ignoring the related dimensionalities of volume and territorial size. These shortcomings may bring about both difficulties in interpreting results and distorted pictures about the actual industrial regional structure of an area.

Regional dimensionality is particularly relevant for researchers focusing on analytical aspects connected with agglomeration and concentration. The present paper introduces measures of agglomeration, concentration and specialization encompassing information about volume, density and region dimensionality. Such an index is built on simple heuristic notion of industrial mass and would provide a more reliable, accurate and flexible instrument than previous measures.

Keywords: Geographical Localization, Industrial Agglomeration, Index.

JEL Classification: C22, C32, C50, E24, E32, J63.

*Corresponding author.
†Gianluigi Pelloni thankfully acknowledges financial support by the *FRP* programme.

1 Introduction

In recent years the fast development of the field of Economic Geography has triggered an increasing interest in indicators of industrial agglomeration, concentration and specialization (Krugman, 1991; Ellison and Glaeser, 1997; Midelfart-Knarvik *et al.*, 2000; Maurel and Sédillot, 1999; Hallet, 2000) measuring industrial clusters . Most of these measures are variations of the Gini's index and may be limited in scope and structure. Rarely, if ever, they include any reference to territorial dimensions and tend to be measures of volumes adjusted for the task at end. They do not embody any information relative to the industrial density and size of the analyzed regions. Alternatively some authors (Ciccone and Hall, 1996) use density as a measure of agglomeration ignoring the related dimensionalities of volume and territorial size. These shortcomings may bring about both difficulties in interpreting results and distorted pictures about the actual industrial regional structure of an area. Regional dimensionality is particularly relevant for researchers focusing on analytical aspects connected with agglomeration and concentration.

We define agglomeration, concentration and specialization in a slightly different way than previous research. In particular, agglomeration aims to measure the total regional industrial mass (i.e. all sectors are included) or, in other words, the geographic concentration of all industries in a specific region; concentration tries to measure the industrial concentration of a specific industry in a given region. Finally, specialization measures how specialized (or diversified) is the economy of a specific region, by measuring how an industrial sector is important for the economy of that region (i.e. it measures the specialization of a specific region in a given industry). The index we propose is applied to all the three concepts. It is evident that an index, able to encompass information about volume and density, should be able to provide a more reliable picture of the industrial structure and a more flexible analytical tool. It is the purpose of the present paper to do so by introducing a notion of industrial mass.

The paper is organized as follows. In Section 2 we discuss previous results in this area while in Section 3 we introduce our new measures and indexes. In Section 4 we calculate our indexes and measures and draw comparisons with some previous indexes using data for Canadian provinces. Section 5 provides conclusions and suggestions for further analyses.

2 Summary of previous results

The TCI Network[1] defines industrial regional clusters as follows:

- An industrial cluster (see Porter 1990) is a set of industries related through buyer-supplier and supplier-buyer relationships, or by common technologies, common buyers or distribution channels, or common labor pools.

[1] See http://www.competitiveness.org/article/view/14/1/5.

- A regional cluster (see Enright 1992, 1993) is an industrial cluster in which member firms are in close geographic proximity to each other. A more inclusive definition would be: regional clusters are geographic agglomerations of firms in the same or closely related industries
- Industrial districts (such as the Italian industrial districts described in Brusco 1992; Piore and Sabel 1984; Becattini 1987, 1989; Goodman and Bamford 1990; and Pyke, Becattini, and Sengenberger 1992) are concentrations of firms involved in interdependent production processes, often in the same industry or industry segment, that are embedded in the local community and delimited by daily travel to work distances (Sforzi 1992).
- A business network consists of several firms that have ongoing communication and interaction, and might have a certain level of interdependence, but that need not operate in related industries or be geographically concentrated in space (see Staber 1996 for discussions of business networks).

It is clear that these terms could have rather different meanings though often they are used interchangeably. Most of the existing indexes of agglomeration, concentration and specialization try to capture one of the above concept and are based on the "*Lorenz curve*" (*LC* henceforth). In fact, the "*Gini coefficient*" (*GC* henceforth) is the most used index to measure the degree of inequality of a distribution. To study geographic concentration of an industry the first step is to calculate the *GC.* As well known, the *LC* was developed to measure the distribution of income and consists in plotting the cumulative density function of a distribution (Lorenz, 1905). The *LC* can also be used to measure the concentration of all statistical variables, thus it may represent the first step to study industrial concentration.

The *GC* is also an inequality measure of a distribution but, differently from the *LC*, it is a number between 0 and 1. Perfect equality corresponds to a zero value, while a value of 1 signals maximal concentration. The *GC* is defined as the area between the *LC* and the uniform distribution line divided by the total area under the uniform distribution. It is one of the most frequently used indexes because of its computing simplicity.

Although its simplicity, the *GC* might fail to properly describe the industrial concentration of a region. In fact, the *GC* lacks completely of any reference to territorial size. The indexes we are proposing in this paper try to improve the existing indexes accounting explicitly for the territorial size. Including the geographical size might facilitate comparisons between the industrial structure of different regions. For instance, it is easier to interpret how Canada and Germany differ in terms of industrial concentration with an index that accounts for their size. Despite the lack of any reference to territorial size, in the wake of Krugman (1991), the *GC* is probably the best known index used in measuring geographic concentration.

Typically, employment and output are the two most commonly used variables in building measures of agglomeration, concentration or specialization. However, though the choice of variable(s) determines the index values, such a choice does not normally affect the intrinsic characteristics of the index itself.

Thus in what follows we focus on the indexes structure and not on variables selection.

In the remaining of this section we present the most important and more frequently used indexes in the literature. For the sake of consistency with the rest of the paper we divide the list into two sub-sets: agglomeration and specialization indexes.

2.1 Agglomeration indexes

Agglomeration indexes measure the geographic concentration of all industries in a specific geographical area, such as a country, a province or a county[2].

As above said, the GC is probably the best known index in this field. It can be used to measure agglomeration taking advantage of the approximated form used when the whole Lorenz curve is not known. In particular, by denoting the cumulative share of sector k employment in the first j regions as $S_j^k = \sum_{m=1}^{j} S_m^k$ and the cumulative share of aggregate employment as $X_j = \sum_{m=1}^{j} x_m$, we can write down the GC as follows[3]:

$$G_k = 1 - \sum_{j=1}^{k} \left(X_j - X_{j-1}\right)\left(S_j^k + S_{j-1}^k\right) \tag{1}$$

When $G_k = 0$ there is no agglomeration, while when $G_k = 1$ sector k is all in one region (maximum of sectoral agglomeration).

The second index in our brief survey is a commonly used index of agglomeration:

$$V_k = \frac{\frac{1}{\overline{y}_k}\sqrt{\frac{\sum_j \left(y_j^k - \overline{y}_k\right)^2}{N}}}{\frac{1}{\overline{y}_j}\sqrt{\frac{\sum_j \left(y_j - \overline{y}_j\right)^2}{N}}} \tag{2}$$

where k is the industry, j is the region, y is the output sectoral share and N is the number of regions.

V_k is a coefficient of variation which gives a measure of spatial dispersion of production. This is a relative index because it compares the industrial agglomeration in a specific region with the average agglomeration of the country. In fact, if the regional agglomeration value is identical to the average national value, then the index value is equal to 1. Instead $V_k < 1$ when in region j there is less agglomeration than in the country. Viceversa $V_k > 1$ if the regional agglomeration is relatively higher.

The so-called "*Clustering index*" is based on gravity models of international trade (Bergstrand, 1985) and is defined:

[2] Midelfart-Knarvik *et al.* (2000) and Hallet (2000) provide summaries and discussions of the wide range of indexes including those we are dealing with in this section. We refer back to these papers for furter details.

[3] See Lafourcade and Mion (2003), p.5.

$$C_k = \frac{\sum_i \sum_j \left(\frac{y_i^k y_j^k}{d_{i,j}}\right)}{\sum_i \sum_j \left(\frac{y_i y_j}{d_{i,j}}\right)} \tag{3}$$

The symbol $d_{i,j}$ denotes the distance between region i and region j. A high C_k should suggest that production of similar goods is taking place in neighboring regions.

There are other indexes that are on the boundary between being agglomeration or specialization indicators. These indexes measure the level of spatial concentration of specific industries. We view these indexes as agglomeration indexes because they do not provide specific information about the characteristics of a specific region.

One of these indexes is the "*Herfindahl-Hirschman*" index (*HH* henceforth). This index is used to measure both the degree of concentration of a market and to estimate the degree of spatial concentration (and agglomeration). The *HH* measures the spatial concentration of industry k across m regions:

$$HH_k = \sum_{i=1}^{m} \left[\frac{E_{k,i}}{E_{k,j}} - \frac{E_i}{E_j}\right]^2 \tag{4}$$

where $E_{k,i}$ is employment in industry k and region i, $E_{k,j}$ is the employment in industry k in country j, E_i is total employment in region i, while E_j is total employment in the country. If the sector k is homogeneous across regions $HH_k = 0$, while in case of maximum agglomeration this index becomes 2[4].

Alternatively this index can be calculated using output instead of employment. The *HH* index can be seen as providing a measure of industrial volume, but it does not carry any information about the dimension of the considered area. Thus it says nothing about density and we do not know if that volume refers to a small or a large area.

The Ellison and Glaeser (1997) is developed within a theoretical framework for business location decisions:

$$EG = \frac{\frac{\sum_{j=1}^{m}\left(\frac{E_{i,j}}{E_j} - \frac{E_j}{E_n}\right)^2}{1 - \sum_{j=1}^{m}\left(\frac{E_j}{E_n}\right)^2} - H}{1 - H} \tag{5}$$

where H measures the distribution of plant size[5].

All these indexes are measures of distance from an average and do not give information about the proper absolute level of agglomeration.

Finally we would present a concentration index that consider some dimensional aspects of the considered area (Spiezia 2002):

[4] See Graham (2003), p.6.
[5] See Devereux, Griffith and Simpson (2002).

$$AGC = \frac{GC}{GC^{\max}} \tag{6}$$

where:

$$GC = \sum_{j=1}^{M} |y_j - a_j|$$

and:

$$GC^{\max} = 2\left(1 - a_{\min}\right)$$

where y_j is the production share of region j with respect to national production and a_j is the area of region j as a percentage of the country area. The AGC index becomes 0 when there is no concentration and 1 when the whole production is in the smallest region. This index considers the size of regions but only with respect to the size of the country, while our index is based on the absolute area.

2.2 Specialization indexes

The characteristics of the traditional specialization indexes are the same of the previously discussed agglomeration indexes. They are measures of (relative) distance from an average.

The first index, S_j, as presented in Hallet (2000) is given by:

$$S_j = \frac{1}{2} \sum_{k=1}^{n} \left| y_j^k - \overline{y^k} \right| \tag{7}$$

where k denotes sectors and j denotes regions. It follows that $S_j \in [0, 1]$ and so in absence of specialization S_j is equal to zero, while it increases towards one as the level of specialization in region j increases. In case of complete specialization S_j becomes 1. Kim (1995) uses a very similar index to study the determinants of agglomeration and specialization for US manufacturing.

In Hallet (2000) output is the variable used to calculate the index S_j. Employment or another suitable variable could have also been used to construct it. The "*Location Quotient*", LQ (aka "*Hoover-Balassa coefficient*")[6] was originally built having employment as reference variable, though again any other suitable variable would have done the job:

$$LQ_{j,k} = \frac{\frac{E_{j,k}}{E_k}}{\frac{E_k}{E}} \tag{8}$$

where the meaning of j and k are the same as in the *Herfindhal-Hirschman* index, E_k denotes the aggregate employment in sector k and E denotes total

[6] See Lafourcade and Mion (2003).

employment. If the Location Quotient is greater than 1, in the region j the industry k is more concentrated then in the average of the country.

The last index we present is due to Duranton and Puga (2001): the "*Relative Diversity Index*" is defined as:

$$RDI_k = \frac{1}{\sum_{j=1}^{n} \left| \frac{E_{j,k}}{E_k} - \frac{E_k}{E} \right|} \tag{9}$$

The value of the RDI_k increases as the regional specialization (in term of employment) approaches the specialization of the national economy. Thus, if a region has exactly the same industrial structure of the average of the country, RDI becomes infinite.

If we refer to industrial agglomeration, i.e. the concept encapsulating the intra/inter industry geographic agglomeration of firms, as a notion capturing the regional industrial "bulk", then the above definitions may often be unsatisfactory as they tend to be expressed either in terms of some measure of volume (often in units of total industrial employment of the area in question) or an indicator of density (e.g. employment per square meter). Once industrial agglomeration stands for something akin to "the quantity of industrial mass" of a region then both volume and density should be taken into account.

From the above short survey it is apparent that the current measures of agglomeration, concentration and specialization are essentially distance measures of volumes (usually employment). These indexes/measures are built with no reference to territorial size and do not have an inbuilt structure to allow for this dimension. It is then possible to calculate the same values of agglomeration/concentration/specialization for different areas of different territorial size. Without a specific reference to information external to the index, sound judgements could be difficult and the risk of drawing wrong conclusions increases. In principle, it is possible to calculate an index of industrial agglomeration with identical values for two areas with remarkably different territorial sizes such as for instance Quebec and Newfoundland in Canada or Bavaria and Liguria in Europe. In principle, these indexes could record interregional similar values for the industrial volumes without picking up the associated different interregional industrial densities. In several research contexts and in more than one dimension the existing measures would be prone to provide distorted information. It is obvious that these drawbacks could be of far reaching consequence.

3 A measure of industrial mass as a measure of industrial agglomeration

Since traditional indexes would not only have a limited explicative power (in terms of carried information) but could also be unreliable measures of the underlying industrial structure, we propose measures of agglomeration, concentration and specialization based on the elementary notion of mass in physics.

In physics mass, in its most elementary form[7], is defined as the quantity of matter of a body, expressed as the product of volume times density:

$$m = \rho V \tag{10}$$

where m = mass, ρ = density and V = volume.

Borrowing this insight from physics, we define by analogy the quantity of industrial matter in a specific geographical area:

$$RIM_j = RID_j \cdot RIV_j \tag{11}$$

This measure, which we call "*Regional Industrial Mass*" (RIM_j), is given by the product of the "*Regional Industrial Density*" (RID_j) and the "*Regional Industrial Volume*" (RIV_j). RIM_j should provide estimates of the level of industrial agglomeration of a specific region j. It is clear that the same amount of RIM_j can be generated alternatively by having a small RIV_j and large RID_j or viceversa by inverting the values of the variables in the same proportions.

With respect to previous measures, RIM_j carries important information about the territorial dimensions of the areas in question. As the volume, RIV_j, is weighted by the density, RID_j, and this density measure is related to the size of territory, RIM_j can take into account the dimension factor and can provide unbiased information on the level of agglomeration. RIM_j can rule out the possibility of estimating identical values for two areas with different territorial sizes.

The RIM_j measure can be easily transformed in an "*Index of Regional Industrial Mass*", $IRIM_j$ (i.e. a measure of distance), by dividing it by the relevant "global" measure:

$$IRIM_j = \frac{RIM_j}{RIM} = \frac{RID_j \cdot RIV_j}{RID \cdot RIV} \tag{12}$$

where RIM = "global" industrial mass (m) ; RID = "global" industrial density (ρ); RIV = "global" industrial volume (V).

The RIM_j-measure would assume precise connotations once the relevant "industrial body" has been selected so that the relevant variables entering the measure can be specified. Once the relevant entities have been chosen, then we can identify the units of measurement of RIM_j. Which variables should go into the definition would depend upon the analytical context and the economic "body" in question.

Let's suppose that we wish to measure the industrial mass of the province of Ontario in Canada. In principle we can do that using as RIV_{ont} the number of firms present on the Ontario territory. Commonly density is defined by the mass per unit volume of a body. In our case RID_{ont} could be given by the number of firms divided by the size of the Ontario territory (A_{ont}). Thus we have:

[7] For an historical study of the development of the concept of mass in physics see Jammer (1961).

$$RIM_{ont} = \frac{RIV_{ont}}{A_{ont}} \cdot RIV_{ont} = RID_{ont} \cdot RIV_{ont} \quad (13)$$

which using data for the year 2004[8] would give an industrial mass for Ontario of:

$$\begin{aligned} RIM_{ont} &= \frac{353,838}{1,076,395} \cdot 353,838 \\ &= 0.328725 \cdot 353,838 = 116,315.4 \end{aligned} \quad (14)$$

The industrial mass of Alberta instead would be defined as:

$$RIM_{alb} = \frac{RIV_{alb}}{\mathbf{A}_{alb}} \cdot RIV_{alb} = RID_{alb} \cdot RIV_{alb} \quad (15)$$

which, always using 2004 data, would be equal to:

$$\begin{aligned} RIM_{alb} &= \frac{147,183}{661,848} \cdot 147,183 \\ &= 0.222382 \cdot 147,183 = 32,730.8 \end{aligned} \quad (16)$$

Using these measures we could then calculate the mass indexes for Ontario and Alberta:

$$\begin{aligned} IRIM_{ont} &= \frac{RIM_{ont}}{RIM} = \frac{RID_{ont} \cdot RIV_{ont}}{RID \cdot RIV} \\ &= \frac{116,315.4}{256,319.86} = 0.45379 \end{aligned} \quad (17)$$

$$\begin{aligned} IRIM_{alb} &= \frac{RIM_{alb}}{RIM} = \frac{RID_{alb} \cdot RIV_{alb}}{RID \cdot RIV} \\ &= \frac{32,730.8}{256,319.86} = 0.1277 \end{aligned} \quad (18)$$

where RIM $(= RID \cdot RIV)$ is Canada's industrial mass.

3.1 Measures of Concentration and Specialization

Having defined agglomeration as the quantity of industrial matter in a region, we can define measures and indexes of concentration and specialization in analogy to the RIM_j measure.

[8] Source: Statistics Canada.

Defining the Regional Industrial Density of sector k in region j by $RID_{j,k}$ and the Regional Industrial Volume of that sector by $RIV_{j,k}$, we have the measure of "*Regional Industrial Concentration*":

$$RIC_{j,k} = RID_{j,k} \cdot RIV_{j,k} \tag{19}$$

and the associated "*Index of Regional Industrial Concentration*" ($IRIC_{j,k}$) providing the share of the "global" industrial mass of sector k (RIC_k), which is concentrated in region j:

$$IRIC_{j,k} = \frac{RIC_{j,k}}{RIC_k} = \frac{RID_{j,k} \cdot RIV_{j,k}}{RID_k \cdot RIV_k} \tag{20}$$

Using (19) we can also define an "*Index of Regional Industrial Sectoral Specialization*" which provides the share of the "global" industrial mass of region j (RIM_j), which is specialized in sector k:

$$IRISS_{j,k} = \frac{RIC_{j,k}}{RIM_j} = \frac{RID_{j,k} \cdot RIV_{j,k}}{RID_j \cdot RIV_j} \tag{21}$$

and an "*Index of Regional Industrial Specialization*":

$$IRIS_j = \sum_k IRISS_{j,k} \cdot RIC_{j,k} = \sum_k \frac{\left(RIC_{j,k}\right)^2}{RIM_j} \tag{22}$$

Equations (20) and (21) would provide the share of the Regional Industrial Mass which pertains to region j and sector k, while equation (22) would tell us how much region j is specialized.

4 Agglomeration, Concentration and Specialization Indexes for Canada

Table 1 summarizes the measures of industrial agglomeration of all Canadian Provinces, using 2004 data and number of firms, i.e. $RID_j = \frac{RIV_j}{A_j}$.

Table 1: Measures of industrial agglomeration of Canadian Provinces (year 2004, $RID_j = \frac{RIV_j}{A_j}$).

Province	$\mathbf{IRIM}_j = \frac{RIM_j}{RIM}$	$\mathbf{RIM}_j = \mathbf{m}_j$	$\mathbf{RID}_j = \boldsymbol{\rho}_j$	$\mathbf{RIV}_j = \mathbf{V}_j$
Ontario	0.45379	$116,315.4$	0.328725	$353,838$
Quebec	0.13653	$34,994.0$	0.150642	$232,299$
Alberta	0.12770	$32,730.8$	0.222382	$147,183$
British Columbia	0.11647	$29,853.3$	0.177763	$167,939$
Nova Scotia	0.07062	$18,100.1$	0.572191	$31,633$
New Brunswick	0.03984	$10,211.8$	0.376349	$27,134$
Prince Edward	0.03373	$8,644.9$	1.235866	$6,995$
Saskatchewan	0.00934	$2,393.9$	0.060639	$39,478$
Manitoba	0.00817	$2,093.8$	0.056853	$36,829$
Newfoundland	0.00383	981.6	0.049219	$19,944$

Alternatively, depending on the analytical content of the analysis, density can be measured using factors of production. While keeping RIV_j equal to the number of firms present on a territory, we could define RID_j either as the ratio $\frac{N_j}{A_j}$ or as the ratio $\frac{K_j}{A_j}$, with $N_j =$ Ontario employment and $K_j =$ Ontario physical capital. In this case the number of firms in Ontario would be weighted by either employment per square meter or by capital per square according to the context. This definition would allow us to connect the volume of firms with the factors of production that they have optimally chosen.

In fact, continuing the analogy with the notion of mass in physics, we can envisage the amount of firms in a specific region as the volume of the industrial body, whereas the density of such a body could be represented by the amount of labour (or physical capital) divided by the size of the region itself. For instance for Ontario and Alberta, using employment to measure density, i.e. $RID_j = \frac{N_j}{A_j}$, and 2004 data, the indexes are:

$$\begin{aligned} RIM_{ont} &= 2,076.3; RIM_{alb} = 390.9 \\ IRIM_{ont} &= 0.528639; IRIM_{alb} = 0.09953 \end{aligned} \tag{23}$$

Table 2 summarizes the measures of industrial agglomeration of all Canadian Provinces, using 2004 data and employment, i.e. $RID_j = \frac{N_j}{A_j}$.

Table 2: Measures of industrial agglomeration of Canadian Provinces (year 2004, $RID_j = \frac{N_j}{A_j}$).

Province	$\mathbf{IRIM}_j = \frac{RIM_j}{RIM}$	$\mathbf{RIM}_j = \mathbf{m}_j$	$\mathbf{RID}_j = \boldsymbol{\rho}_j$	$\mathbf{RIV}_j = \mathbf{V}_j$
Ontario	0.52864	2,076.3	0.005868	353,838
Quebec	0.14135	555.3	0.002390	232,299
Alberta	0.09953	390.9	0.002656	147,183
British Columbia	0.09321	366.1	0.002180	167,939
Nova Scotia	0.06431	252.7	0.007988	31,633
New Brunswick	0.03356	131.9	0.004861	27,134
Prince Edward	0.02096	82.4	0.011780	6,995
Manitoba	0.00834	32.7	0.000888	36,829
Saskatchewan	0.00741	29.1	0.000737	39,478
Newfoundland	0.00269	10.6	0.000531	19,944

If we look at Table 1 and Table 2 we can immediately notice two features. First, independently on how RID_j is calculated, the ranking of the ten Provinces is almost the same with the exception of an inversion in the 8th and 9th position between Saskatchewan and Manitoba. That would mean that, at least in terms of ranking in this specific example, the agglomeration ranking is invariant with respect to the variables entering the definition of RID_j.

Second, if we were to use density RID_j as a measure of industrial agglomeration, either using the number of firms or the number of employed workers to define it, the invariance property will hold as well. However while our measure of mass $IRIM_j$ provides a ranking which is intuitively consonant with the actual industrial structure of the Canadian Provinces, if we were to use density RID_j, instead of $IRIM_j$, as a measure of agglomeration we would get the following ranking:

Table 3: Ranking of Canadian Provinces using densities RID_j as a measure of industrial agglomeration.

$\mathbf{RID}_j = \frac{RIV_j}{A_j}$	$\mathbf{RID}_j = \frac{N_j}{A_j}$
Prince Edward	Prince Edward
Nova Scotia	Nova Scotia
Ontario	Ontario
New Brunswick	New Brunswick
Alberta	Alberta
Quebec	Quebec
British Columbia	British Columbia
Saskatchewan	Manitoba
Manitoba	Saskatchewan
Newfoundland	Newfoundland

Therefore using these two ranking provinces of a small industrial dimension, like Prince Edward, Nova Scotia and New Brunswick, are leading the ranking with only Ontario among the largest economies, in third position. If we were to use blindly such ranking we would likely get a highly distorted picture of the regional industrial agglomeration of the Canadian Provinces.

It is obvious that if we wish to calculate measures of industrial specialization it would be much more reliable using the $IRIC_{j,k}$, $IRISS_{j,k}$, $IRISS_j$ and $IRIS_j$ than trying to use measures based on density alone. Below we present results for $IRISS_j$ and $IRIS_j$ for all Canadian Provinces using the two different definitions of density ($RID_j = \frac{RIV_j}{A_j}$ in Table 4 and $RID_j = \frac{N_j}{A_j}$ in Table 5).

Table 4: Measures of industrial specialization of Canadian Provinces (year 2004, $RID_j = \frac{RIV_j}{A_j}$).

Province	$\mathbf{IRIS}_j = \sum_k \mathbf{IRISS}_{j,k} \cdot \mathbf{RIC}_{j,k}$	$\mathbf{IRISS}_j = \sum_k \mathbf{IRISS}_{j,k}$
Ontario	333.16294	0.119826
Alberta	111.69866	0.121382
Quebec	73.85840	0.108860
British Columbia	67.38895	0.111533
Nova Scotia	33.26544	0.108450
Prince Edward	22.56496	0.114381
New Brunswick	16.98134	0.104806
Saskatchewan	6.99596	0.116207
Newfoundland	3.85745	0.126208
Manitoba	3.42413	0.103717

If instead we use employment to measure density and 2004 data, the measures of industrial specialization of all Canadian Provinces, using 2004 data and employment, i.e. $RID_j = \frac{N_j}{A_j}$, are:

Table 5: Measures of industrial specialization of Canadian Provinces (year 2004, $RID_j = \frac{N_j}{A_j}$).

Province	$\mathbf{IRIS}_j = \sum_k \mathbf{IRISS}_{j,k} \cdot \mathbf{RIC}_{j,k}$	$\mathbf{IRISS}_j = \sum_k \mathbf{IRISS}_{j,k}$
Ontario	3.06992	0.092404
Quebec	0.81113	0.089671
Alberta	0.49828	0.090290
British Columbia	0.47088	0.089833
Nova Scotia	0.31740	0.086474
New Brunswick	0.15069	0.085981
Prince Edward	0.09207	0.083612
Saskatchewan	0.04622	0.092245
Manitoba	0.03324	0.079021
Newfoundland	0.01573	0.087761

It is interesting to note that the region with the highest level of agglomeration Ontario has also by far the highest level of regional industrial specialization. We can also notice that the provinces characterized by a low level of industrial agglomeration are also characterized by extremely low levels of industrial specialization while the three regions with the strongest specialization profile, Ontario, Quebec and Alberta are also the leading agglomeration areas.

5 Conclusions and Suggestions for Further Results

Our RIM_j measure (and so the $IRIM_j$ index) is an heuristic one and not based on economic theory. However, it presents several advantages with respect to previously adopted measures and indexes by reducing the potential biases brought about by looking at only one aspect (either volume or density). In fact, since we construct measures based on the notion of mass in physics, we manage to explicitly consider the territorial size: in this way it is possible to make a comparison between regions of significantly different size. However, it is clear that in principle the RIM_j measure should emerge as a function of the flow of creation and destruction of firms in a specific region and thus of a vector of regional attractors and repellors. Continuing the analogy with physics further research can investigate wether regions with high mass (i.e. high RIM_j) attract more investments than regions with low mass. Therefore not only the industrial mass depends on the process of job creation and destruction, but it could also affect the process itself. Finally, on a purely empirical and pragmatic dimension, further comparisons with other indexes are required.

References

[1] Becattini, G., editor, (1987), *Mercato e forze Locali: il Distretto Industriale*, Bologna, Il Mulino.

[2] Becattini, G., editor, (1989), *Modelli locali di Sviluppo*, Bologna, Il Mulino.

[3] Bergstrand, Jeffrey H. (1985), "The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence", The Review of Economics and Statistics, MIT Press, vol. 67(3), 474-81.

[4] Brusco, S. (1992), "The idea of the Industrial District: Its genesis", in in Pyke et al. (1992), 10-19.

[5] Ciccone, A. and R. Hall, (1996), "Productivity and the Density of Economic Activity", American Economic Review 86, 54-70.

[6] Devereux, M.P., R. Griffith and H. Simpson (1999), "The Geographical Distribution of Production Activity in the UK", CEPR Discussion Papers 3627.

[7] Duranton G. and D. Puga (2001), "Nursery Cities: Urban Diversity, Process Innovation, and the Life-Cycle of Products", American Economic Review 91(5), 1454–1477.

[8] Duranton, G. and D. Puga (2005), "From sectoral to functional urban specialisation", Journal of Urban Economics, Elsevier, vol. 57(2), 343-370.

[9] Ellison, G. and E. Glaeser (1997), "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach", Journal of Political Economy, vol. 105, no. 5, 889-927.

[10] Ellison, G. and Glaeser, E. (1999), "The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration?", American Economic Review, 89.2, 311-316.

[11] Enright, M. J. (1992), "Why Local Clusters are the Way to Win the Game", World Link, 5, July/August, 24-25.

[12] Enright, M. J. (1993), "The Geographic Scope of Competitive Advantage", in E. Dirven, J. Groenewegen, and S. van Hoof, editors, Stuck in the Region?: Changing Scales of Regional Identity. Utrecht: Netherlands Geographical Studies 155, 87-102.

[13] Goodman, E. and J. Bamford, editors, (1990), *Small Firms and Industrial Districts in Italy*, London, Routledge.

[14] Graham, E. (2003), *Reforming Korea's Industrial Conglomerates*, Institute of International Economics: Washington D.C.

[15] Jammer, M. (1961), *Concepts of Mass in Classical and Modern Physics*, Harvard University Press.

[16] Hallet, M. (2000), "Regional specialisation and concentration in the EU", Economic Papers, 141, Brussels: European Commission.

[17] Kim, S. (1995), "Expansion of Markets and the Geographic Distribution of Economic Activities: The Trends in U.S. Regional Manufacturing Structure, 1860-1987", Quarterly Journal of Economics 110:881-908.

[18] Krugman, P. (1991), *Geography and Trade*, Cambridge, MA: MIT Press.

[19] Lafourcade, M. and G. Mion (2003), "Concentration, Spatial Clustering and the Size of Plants: Disentangling the Sources of Co-Location Externalities", CORE Discussion Papers 2003/91, Université catholique de Louvain, Center for Operations Research and Econometrics (CORE).

[20] Lafourcade, M. and G. Mion (2007), "Concentration, Agglomeration and the Size of Plants", Regional Science and Urban Economics, 37, 46-68.

[21] Lorenz, M. O. (1905), "Methods of measuring the concentration of wealth", Publications of the American Statistical Association 9: 209–219.

[22] Maurel, F. and Sédillot, B. (1999), "A Measure of the Geographic Concentration in French Manufacturing Industries", Regional Science and Urban Economics, 29(5), 575-604.

[23] Midelfart-Knarvik, K.H., H.G. Overman, S.J. Redding and A.J. Venables, 2000, "The Location of European Industry", Economic Papers, 142, Brussels: European Commission.

[24] Porter, M. E. (1990), *The Competitive Advantage of Nations*, New York, The Free Press.

[25] Pyke, F., G. Becattini, and W. Sengenberger, editors, (1992), *Industrial Districts and Inter-firm Co-operation in Italy*, Geneva, International Institute for Labour Studies.

[26] Piore, M. and C. Sabel (1984), *The Second Industrial Divide*, New York, Basic Books.

[27] Sforzi, F. (1992), "The quantitative importance of Marshallian industrial districts in the Italian economy", in Pyke et al. (1992), 75-107.

[28] Spiezia, V. (2002), "Geographic Concentration of Production and Unemployment in OECD countries", Cities and Regions, December Issue.

[29] Staber, U., N. V. Schaefer, and B. Sharma, editors (1996), *Business Networks: Prospects for Regional Development*, New York, De Gruyter.

[30] Staber, U. (1996), "Networks and Regional Development", in Staber *et al* (1996), 1-23.

6 Data

Source: Statistics Canada.

Data for Canada and for all Canadian provinces: Alberta, British Columbia, Manitoba, New Brunswick, Newfoundland & Labrator, Nova Scotia, Ontario, Prince Edward Island, Quebec,

Saskatchewan.