WORKING PAPER

14-02

Geographical Agglomeration: the Case of Belgian Manufacturing Industry

Federal
Planning Bureau
Economic analyses and forecasts

Avenue des Arts 47-49
B-1000 Brussels
Tel.: (02)507.73.11
Fax: (02)507.73.73
E-mail: contact@plan.be
URL: http://www.plan.be

L. Bertinelli[1], J. Decrop

November 2002

1. Center for Operations Research and Econometrics (CORE), UCL.

Geographical Agglomeration: the Case of Belgian Manufacturing Industry

L. Bertinelli[1], J. Decrop

November 2002

1. Center for Operations Research and Econometrics (CORE), UCL.

Federal Planning Bureau

The Belgian Federal Planning Bureau (FPB) is a public agency under the authority of the Prime Minister and the Minister of Economic Affairs. The FPB has a legal status that gives it an autonomy and intellectual independence within the Belgian federal public sector.

FPB's activities are primarily focused on macroeconomic forecasting, analyzing and assessing policies in economic, social and environmental fields.

Internet

URL: http://www.plan.be

E-mail: contact@plan.be

Publications

Publications:

The Economic Forecasts
The "Short Term Update"

Planning Papers (recent publications)

giving the results of policy analyses and assessments

91 *Verkenning van de financiële evolutie van de sociale zekerheid 2000-2050 - De vergrijzing en de leefbaarheid van het wettelijk pensioensysteem*
M. Englert, N. Fasquelle, M.-J. Festjens, M. Lambrecht, M. Saintrain, C. Streel, S. Weemaes - January 2002

92 *De administratieve lasten in België voor het jaar 2000 - Eindrapport*
Greet De Vil, Chantal Kegels - January 2002

Working Papers (recent publications)

10-02 Towards E-Gov in Belgium
H. Van Sebroeck - September 2002

11-02 Monetary policy in the euro area - Simulations with the NIME model
E. Meyermans - October 2002

12-02 Constructing Productive ICT Capital Stock Series for Belgium
T. Pamukçu, W. Van Zandweghe - October 2002

13-02 Stedelijke woondynamiek van de Belgische bevolking en haar gezinnen
D. Devogelaer - November 2002

Citations should refer to the publication, mentioning the author(s) and the date of publication.

Responsible Editor Henri Bogaert
Legal Depot: D/2002/7433/28

Table of Contents

Executive summary

The aim of this working paper is to carry out a descriptive analysis of the geographical concentration of the manufacturing sector in Belgium. This work is part of the study on the impact of information and communication technologies (ICT), on the geographical concentration of economic activities in cities, which is being financed by the SSTC[1]. This paper is closely related to a first working paper on this subject, which analyzes the geographical agglomeration patterns of all economic activities[2], but takes a more technical approach and gives a more detailed picture of industrial agglomeration.

This paper is focused on the manufacturing sector, from which geographical agglomeration of sectors of activity is analyzed at a fine industrial level, i.e. NACEBEL 4-digit industries. To assess the degree of concentration, individual plant data on wage and salary earners is used, with a relatively fine geographical breakdown (districts and townships). The short period of investigation (1997-2000) is informative from a descriptive point of view, but one must remain very cautious when it comes to addressing policy issues, which is not at all the primary aim of this paper.

We know from literature that most countries' economic activities are concentrated in a few areas, including cities. This is also the case in Belgium. When comparing manufacturing with overall employment patterns, it seems that the manufacturing sector as a whole is much more concentrated. Using manufacturing employment as a benchmark, we compute the geographical concentration for each NACE′ industry. The major findings of this paper are:

- about one third of all NACEBEL 4-digit industries can be considered as highly concentrated. We find that the activities showing high levels of concentration are not specifically high-tech industries, but also traditional industries;
- evolution of the concentration patterns of industries from 1997 to 2000 is rather stable, which can be seen as the inertia of geographical concentration patterns in most sectors;
- some sectors do show large variations in their concentration index for the period under investigation, including high-tech industries, but these are rather small sectors with a small number of plants.

We further analyze a number of specific problems, which arise when studying concentration patterns at specific geographical and industrial disaggregated levels. First, we try to assess whether broad industries (NACEBEL 2-digit) are concentrated because of the concentration within or across their sub-industries (NACEBEL 4-digit). For most industries, concentration within sub-industries is

1. Service federal des affaires scientifiques, techniques et culturelles. Contract S2/64/01.
2. See Decrop (2002).

prevalent, which is an argument in favour of the presence of localization economies. Four industries, however, do highlight significant concentration across sub-industries: textile, clothing, printing and publishing and precision instruments. The results for these industries are consistent with the importance of diversity in clusters of activity, as suggested by some authors[1].

We also study the relation between agglomeration and plant scale, which appears to be positive: “hard-core” manufacturing, which generally takes place in large plants, seems to be more geographically concentrated than smaller plants within the same industry.

When we consider regional outcomes, it can be shown that Flanders has a more dispersed pattern than Wallonia, where manufacturing activity is concentrated in a few hot spots. Part of the explanation for this lies in the 'history' of the two regions and their different pathways to industrialization. This is reflected in the diverging mix of industries between the two regions, as well as the different degrees of concentration of the same industries.

Finally we study the issue of spatial autocorrelation on a fine geographical scale, namely townships. Autocorrelation tries to assess the degree of value similarity between neighbouring geographical units. It can be shown that a large majority of industries (95%) exhibit positive spatial autocorrelations, 30% of which are significant. The geographical scope of agglomeration within these industries does not seem to be captive within small geographical units, but extends beyond them (at distrilocalizationct and even maybe larger level).

All in all, this paper provides some descriptive features with respect to agglomeration in the manufacturing industry. Although the study gives some hints about potential explanatory factors, the intention is not to build an explanatory framework[2] or to suggest policy measures. Further research into these fields is to be expected.

1. See Feldman and Audretsch (1999).
2. For example, the question of the source of agglomeration, natural advantages *vs* dynamic externalities, is not investigated here.

Introduction

About a decade ago, Krugman (1991) mentioned the 'New Economic Geography'. Although many of Krugman's ideas were actually not as new as he pretended, at least, he presented them in a coherent way and made them popular among professional economists (see for instance Neary (2001) for a synthesis). The basic idea starts from a model in which economic activity is uniformly distributed across space. Krugman shows, in a monopolistic competition setting, that through market forces (what he calls forward/backward linkages), a tiny modification in the perfectly symmetrical equilibrium may foster a snowball effect, leading to a highly skewed distribution of economic activity across regions. More generally, these forward/backward linkages are part of what are generally referred to as Marshallian externalities. Local interaction between producers (knowledge externalities), between workers (labor market pooling) and between intermediary and final producers, on the one hand, and final producers and consumers, on the other hand (forward/backward linkages) may lead to gains from proximity.

There is plenty of evidence of concentration of economic activities, reaching from case studies (i.e. Route 128 and Silicon Valley in the United States (Saxenian (1996), the Blue Banana in Europe etc.) to empirical work, such as Glaeser et al. (1992), Ciccone and Hall (1996), Rauch (1993) and many more. Studies supporting the existence of agglomeration economies were largely carried out in the 1970s and 1980s (Tabuchi (1986), Moomaw (1981), Sveinkauskas (1975),…). In relation to these agglomeration economies, an interesting debate emerged in this context in the 1990s: does specialization or diversity favour agglomeration economies? This is a very old question indeed. Hoover (1937) referred to localization and urbanization economies to characterize both these phenomena. The issue has already been addressed in McLaughlin (1930), but was formally treated in Glaeser et al. (1992) and Henderson et al. (1995) in a dynamic framework. This debate on specialization vs. diversity has some direct relevance to policy. Indeed, agglomerations of people and economic activity, besides the gains mentioned above, also give rise to certain costs, which are generally referred to as congestion costs. Thus, if agglomeration economies mostly arise from intra-industry interactions, political decision- makers should favour the rise of specialized clusters, so as to maximize these agglomeration economies in comparison with the congestion cost. On the other hand, when inter-industry cross-fertilization prevails, diversified clusters of economic activities should be fostered.

From a theoretical point of view, the major conclusion in models of economic geography predicts that agglomeration forces will increase when transportation costs decline. One way of interpreting this result is to say that economic integration fosters regional inequalities. Hence, in view of the amounts of money spent by countries and at European level to promote spatial equity, the need to measure geographical concentration is of prime importance. The refined methods, used to give empirical content to spatial concentration, therefore extend far beyond theoretical considerations and should provide political decision-makers with a map of what geographical concentration actually is, before considering any measures of spatial equity.

This study will therefore not provide any guidelines in relation to the determinants of agglomeration or its consequences, but only a description of the actual situation as for the geographical concentration of manufacturing activity in Belgium. It is a preliminary but necessary step to obtain a positive analysis. The rest of the paper is drawn up as follows: section 1 provides some measures of geographical concentration, followed by a brief description of the data in section 2. Section 3 provides our main results of geographical concentration as well as some extensions.

I Measures of geographical concentration

Recently, Duranton & Overman (2001) have shown that it is possible to treat space continuously instead of using an arbitrary collection of geographical units. As such, they are able to precisely develop distance-based measures of localization and to assess their statistical significance. To do this sort of analysis, it is necessary to have spatial X and Y coordinates of plants. Unfortunately, in Belgium, the smallest geographical units for which plant and employment data are available are the townships, making it impossible to apply the methodology of both authors. We are thus forced to use predefined geographical units, as in many other studies. In the following, we present the main indices. However, contrary to other papers, the aim of this paper is not to present a new measure of geographical concentration.

In the empirical literature, many measures of economic concentration have been proposed. We can classify them by means of three criteria (see Devereux et al., 1999):

- Is it a gross or net measure of geographical concentration? Other things being equal, geographical concentration of an industry tends to be higher if the industry is composed of a few large plants. Gross indices of geographical concentration do not take industrial concentration of industries (i.e. plant size distribution within industries) into account, while net indices of geographical concentration neutralize industrial concentration. The former ones are typically Gini-like indices, whereas the latter are used by Ellison-Glaeser and Maurel-Sédillot. Henceforth, we speak of geographical concentration for gross indices, and geographical agglomeration for net indices;
- Do the indices rely on a theoretically or statistically based model? This criterion allows us to distinguish two net indices of agglomeration: the Ellison-Glaeser index is based on a theoretical model of natural advantages and spillovers, while the Maurel-Sédillot one relies on a statistical approach;
- What is the geographical benchmark to compare the spatial distribution of industries? In most empirical work on industry location, the index is measured relative to the geographical distribution of total manufacturing. Hence the value is generally zero if the industry's employment is located in the same proportion as total manufacturing. Decrop (2002) shows that this may cause some misunderstanding about geographical concentration of industries, since industries that are more uniformly spread across geographical units than total activity can reveal relatively high concentration indices.[1]

1. Note that Decrop (2002) uses total employment data in Belgium rather than only manufacturing employment.

The main problem with the locational Gini is its sensitiveness to the industrial concentration of employment in firms: two industries, with the same number of workers but distributed in a few *vs* many plants, will exhibit different Gini indices even if the location of plants is chosen randomly and independently in the two cases. To take this into account, some authors have tried to neutralize the effect of industrial concentration on geographical concentration. Two types of agglomeration indices (*i.e.* net geographical concentration) will be discussed here: the Ellison-Glaeser and Maurel-Sédillot indices.

A. Ellison-Glaeser index

Ellison and Glaeser (1997) present an agglomeration index of plants, based on a test of comparison between the observed geographical distribution of plants and a random distribution. The randomness of a geographical distribution is defined as the expected distribution in the absence of agglomeration and centrifugal forces, like natural advantages, intra-industrial spillovers, transport costs. They first define an index of gross geographical concentration:

$$G_{EGk} = \frac{\sum_i (S_{ki} - S_i)^2}{1 - \sum_i S_i^2}$$

where *k* stands for the k^{th} industry of activity (k = 1, …, K), *i* is the i^{th} geographical unit (i = 1, …, m) and *s* is the share of a geographical unit i in industry employment k (s_{ki}) and total employment (s_i). This index has the attractive property of being equal in average to the industrial concentration when plants localize independently from each other. Industrial concentration of an industry *k* is measured with the Herfindahl index[1] which we call H_k:

$$H_k = \sum_{f_k = 1}^{F_k} Z_{f_k}^2$$

where f_k (f_k = 1, …, F_k) corresponds to the f^{th} plant belonging to industry *k*, and z_{fk} is the share of employment taken by the f^{th} plant in industry *k*. The value of H_k is a function of the number and the size distribution of plants in industry *k*. For a industry with *F* plants, the index has a minimum value of $1/F_k$ when plants have the same size. Indices are generally high for industries with a small number of plants and with an uneven size distribution. The inverse of the Herfindhal index has an interesting meaning: it corresponds to the number of plants we would obtain if they were all of the same size. The smaller this number, the higher the industrial concentration in the industry. H_k indices equal to 0.1 and 0.01 are equivalents to industries with respectively 10 and 100 equally sized plants.

1. It is important to notice that the Herfindahl index used in the economic geography literature is derived from the field of industrial economy, but is not the same. In comparison with the classical Herfindahl index used in industrial economy, business units considered in economic geography are plants and not firms, and the market shares are employment shares and not shares of turnover.

If there were no agglomeration economies and if the geographical units are equally attractive (no natural advantages), the gross geographical concentration G_{EGk} of an industry k should be exactly equal to its industrial concentration H_k. Both authors show that the expected mean of G_{EGk} is $H_k + \gamma$ $(1-H_k)$, from which they derive an estimator of excess-concentration, called agglomeration index γ:

$$\hat{\gamma}_{EG_k} = \frac{G_{EG_k} - H_k}{1 - H_k}$$

Two important remarks are worth considering when using the E-G index:

- The index is a measure of excess-concentration with respect to industrial concentration, but does not indicate where this excess-concentration comes from: natural advantages, agglomeration economies or other factors.
- The index attempts to control differences in overall size across geographical units, measured by total manufacturing employment. While this reference is relevant for large geographical units as in many empirical studies, it can become a problem in small countries (like Belgium), when using smaller geographical units like districts or townships. Decrop (2002) has shown that in this case, the Ellison-Glaeser index is more an index of regional specificity of industries than one of geographical agglomeration.

B. Maurel-Sédillot index

Maurel and Sédillot (1999) propose an estimator of externality that is slightly different from Ellison and Glaeser. Instead of a theoretical model, they start from a probabilistic model of geographical concentration, also taking industrial concentration into account. The sole difference with the Ellison-Glaeser index is the measure of gross geographical concentration, G_{MSk}:

$$G_{MS_k} = \frac{\sum_i S_{ki}^2 - \sum_i S_i^2}{1 - \sum_i S_i^2}$$

The estimator of excess-concentration, *i.e.* the index of agglomeration is the following:

$$\hat{\gamma}_{MS_k} = \frac{G_{MS_k} - H_k}{1 - H_k}$$

To be exhaustive, it is worth noticing that Devereux et al. (1999) also propose an agglomeration index, based on an intuitive approach. The results in Belgium are very close to those obtained by the Maurel-Sédillot index; it will therefore not be further investigated.

Whatever the model chosen (Ellison-Glaeser *vs* Maurel Sédillot), the interpretation of the parameter γ is fairly simple:

- If plants choose to locate randomly and independently of each other, the value of γ should on average be zero; neither externalities nor natural advantages are influencing location and the non-zero gross geographical concentration observed is only due to industrial concentration (measured by the Herfindahl index);
- A high value of γ for a particular industry has to be seen as excess-concentration, i.e. spatial concentration that is higher than the expected concentration in the case of random location choice. The industry is seen as 'agglomerated'.

In principle, both indices take the Herfindahl index into account in order to neutralize the effect of industrial concentration on geographical concentration. However, Holmes and Stevens (2000) notice that the γ_{EG} index significantly varies with plant size. In most manufacturing industries, the Ellison-Glaeser index increases when deleting small sized plants in the sample of observations. In order to explain this fact, both authors put forward two suggestions:

- Industries can be broadly defined and can encompass sub-industries that have totally different location behaviours according to their size, e.g. small plants that carry out retail or service functions and that are hence geographically dispersed, and major plants doing hard-core manufacturing that are geographically concentrated.
- Plants located in regions with high concentration of economic activities have some advantages (due to agglomeration externalities or natural advantages) that enhance productivity. These plants will tend to grow more rapidly than those situated outside such areas.

II Description of the data and geographical units

A. Data description

The database, which has been used for the present study covers all plants established in Belgium, and spans a 4-year period from 1997 to 2000.[1] For each plant, we have the number of jobs, the industry it belongs to (up to NACEBEL-5 digit classification), and the township it is located in. Employment data come from the national office for social security (ONSS), which collects employment data for all wage earners in Belgium.

In Belgium, the notion of plant is clearly distinguished from the one of employer, the latter corresponding to the notion of firm in a general sense. If the employer has only one activity at one location, then it is considered as a plant. But if the employer carries out its activity in two or more locations (industries or operation units) and/or carries out two different types of activity, each operation unit is seen as a plant. However, if several operation units of the same firm are located in the same township, only one plant is taken into account.

Concerning employment data, it is worth noticing that it corresponds to the number of jobs and not to the number of workers.[2] If a worker is working for two or more employers, he is counted several times. Another limit of the database is that it does not take the difference between part-time and full-time workers within plants into account.

B. Some descriptive features

The number of observations in the database totals 1 002 935 year-plant observations (resp. 13 326 665 year-jobs), with the following yearly breakdown:

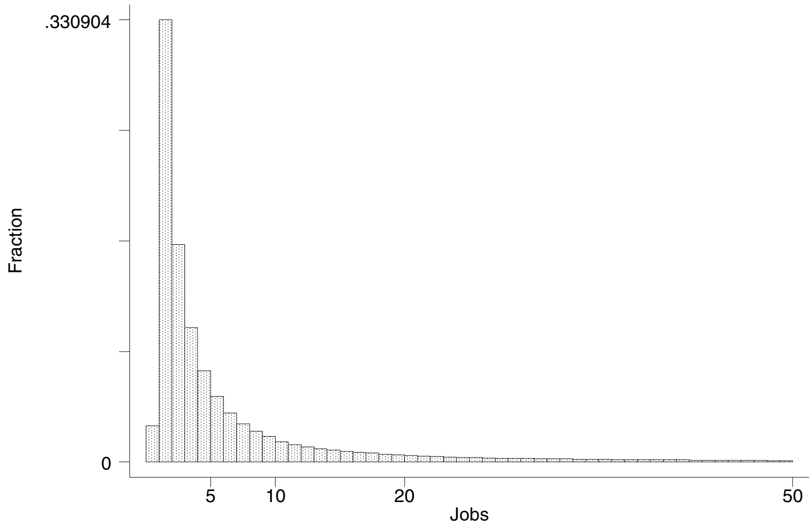
TABLE 1 - Evolution of total number of plants, employment and average plant size

Year	Number of plants	Total employment	Average plant size
1997	247 485	3 248 231	13.1
1998	249 117	3 308 622	13.3
1999	252 105	3 357 027	13.3
2000	254 228	3 412 785	13.4

1. The database comes from the 'Banque Carrefour pour la Sécurité Sociale', that collects all types of information in the field of social security.
2. The number of jobs occupied on June 30th of each year.

These raw figures, however, include some important features. A first important point is the highly skewed *distribution of plants* according to their size. In fig. 1, we have plotted the distribution of plants with less than 50 employees. A striking feature is that about a third of plants smaller than 50 employees actually occupies 2 persons. More generally, taking the whole database into account, two thirds of the observations are comprised in a class ranging from 1 to 5 workers.

FIGURE 1 - Firm size distribution for firms smaller than 50 workers



Although the majority of plants are very small, they by far do not represent the major part of employment. About 80% of employment is concentrated in plants larger than 20 workers, representing hardly more than 10% of the total number of plants.

A second important feature is the industrial distribution of plants and employment. In 2000, manufacturing represented 18.5% of total employment (9.67% of plants), the rest of employment being distributed between agriculture (0.9%), services (43.9%) and public utilities (36.7%).[1]

In the sequel we shall only focus on the manufacturing industries. The major reason to do so, comes from our final aim, which consists in providing some hard facts to support developments in the economic geography literature: are economic activities agglomerated, and if so, which ones? Service sectors are governed by rationales which substantially differ from manufacturing, and hence, we exclude them from the present study for reasons of consistency.[2]

Manufacturing employment grew by 0.97% from 1997 to 2000 (relative to 5.1% for total employment growth in Belgium) and represented more than 600 000 jobs in the period under consideration (see table 2).

1. The breakdown is detailed in appendix 1. Note that construction, as well as water, gas and electricity supply were added to the service sector.
2. Note that though some similar studies include extractive industries, we left them aside. The reason is that such sectors are heavily dependent on the location of natural resources rather than on market forces.

TABLE 2 - Evolution of total manufacturing employment

Year	Jobs
1997	624 254
1998	630 758
1999	624 654
2000	630 287

Out of the 237 NACE4 manufacturing industries, the 10 largest industries are represented in table 3 and represent 26.37% of total manufacturing employment.[1] The largest NACE4 industry is *construction and assembling of automotive vehicles* (industry 3410), representing 5.93% of total manufacturing employment in 2000, whereas industry 2721 (*Manufacture of cast iron tubes*) has the smallest share (close to 0%). Note, however, that the more aggregated NACE2 industries may entail a more or less large number of NACE4 industries, which in turn explains part of the results found in table 3. This is particularly true for food industry (NACE2 15) which has 31 NACE4 sub-industries. For further details, see table 4.

TABLE 3 - Top 10 largest manufacturing industries in Belgium with respect to employment

NACE4	Name	Jobs (2000)
3410	Manufacture of motor vehicles	37 395
1581	Manufacture of bread; manufacture of fresh pastry goods and cakes	21 756
2710	Manufacture of basic iron and steel and of ferro-alloys (ECSC)	19 075
2222	Printing n.e.c.	16 112
2442	Manufacture of pharmaceutical preparations	15 248
2852	General mechanical engineering	12 112
3430	Manufacture of parts and accessories for motor vehicles and their engines	11 723
1751	Manufacture of carpets and rugs	11 709
2414	Manufacture of other organic basic chemicals	11 334
3120	Manufacture of electricity distribution and control apparatus	9 717

The size of each industry can be decomposed in terms of *share of employment* and *share of plant number*. Table 4 does so at the NACE2-digit level.

1. The 20 largest NACE4 sectors represent 39.40% of total manufacturing employment.

TABLE 4 - Industry size by plants and jobs

NACE2	% plants (2000)	% jobs (2000)	Number of NACE4 industries
15 - Food products and beverages	27.06%	13.83%	31
16 - Tobacco products	0.14%	0.47%	1
17 - Textiles	5.24%	6.79%	21
18 - Wearing apparel; dressing and dyeing of fur	3.32%	2.03%	6
19 - Leather and leather products, incl. footwear	0.54%	0.37%	3
20 - Wood and wood products	3.74%	1.84%	6
21 - Pulp, paper and paper products	1.19%	2.57%	6
22 - Publishing, printing and reproduction of media	9.97%	5.11%	13
23 - Coke, refined petroleum products and nuclear fuel	0.12%	0.62%	3
24 - Chemicals and chemical products	3.05%	11.03%	20
25 - Rubber and plastic products	2.74%	3.93%	7
26 - Other non-metallic mineral products	5.15%	5.33%	25
27 - Basic metals	1.03%	6.26%	17
28 - Fabricated metal products	14.96%	9.18%	16
29 - Machinery and equipment n.e.c.	5.07%	6.74%	20
30 - Office machinery and computers	0.07%	0.07%	2
31 - Electrical machinery and apparatus	1.87%	3.64%	7
32 - Radio, television and communication equipment and apparatus	0.45%	3.07%	3
33 - Medical, precision and optical instruments	2.53%	1.15%	5
34 - Motor vehicles, trailers and semi-trailers	1.53%	8.74%	3
35 - Other transport equipment	0.77%	2.81%	8
36 - Furniture; manufacturing n.e.c	8.23%	3.90%	13
37 - Recycling	1.22%	0.54%	2

A striking feature of this table is the relative discrepancy between *share of employment* and *share of plants*. And the finer the industrial breakdown, the higher this disconnection for each industry. Another way of assessing this disconnection can be done by measuring the *industrial concentration*, with the Herfindahl index (H_k) (see Appendix 2 for NACE4 H_k-indices).

The aim of the present study is, however, *not* to analyze the *industrial concentration* of manufacturing activity. Nonetheless, as will be detailed below, studying the spatial concentration of activity requires “some” prior knowledge on industrial concentration. The reason can be illustrated very simply: suppose one industry, entailing 100 jobs. Moreover, suppose that all employment in this industry is concentrated in one plant. Hence, all employment will be concentrated in one area and classical indices of spatial concentration would consider this industry as highly concentrated. But this result is highly dependent on industrial concentration.

Table 4 as well as Appendix 2 lead us to treat carefully the industrial concentration issue prior to analyzing the spatial concentration of industries in Belgium. For instance, the H_k-indices span from 0.0013 to 1, which theoretically corresponds to the case of perfectly identically sized plants going from 1 to 750 jobs. Hence, great caution is called for concerning industrial concentration problem when analyzing the spatial distribution of manufacturing activity in the following sections.[1]

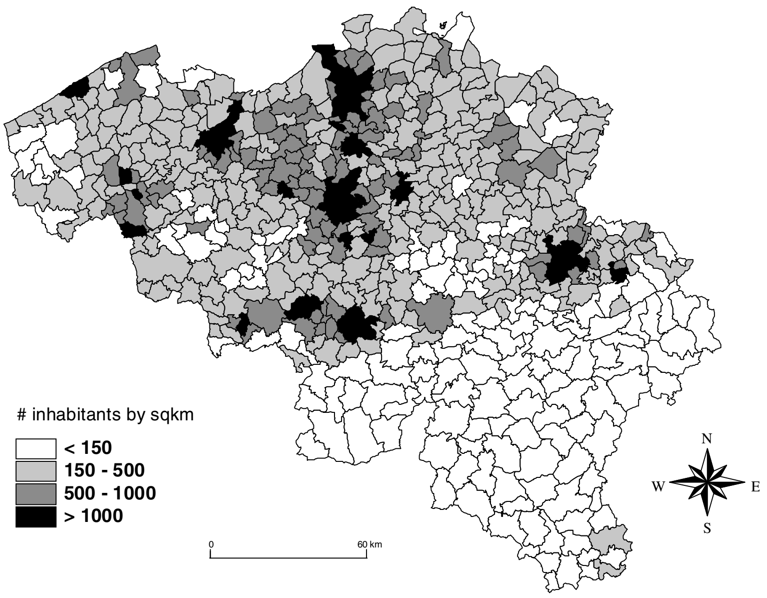
C. Geographical units

So far, we have provided some hints on the distribution of employment across industries and plant size classes. The aim of the present study being to present a map of the spatial distribution of manufacturing activity across Belgium, we should first assess some general features. Belgium is a federal state, divided into 3 regions, 10 provinces, 43 districts and 589 townships. In order to measure spatial concentration of manufacturing activity, we shall mainly concentrate on districts and townships. Districts are about 700 km^2 on average, which is smaller than French departments. By contrast, townships have much smaller areas (52 km^2 on average). It should be noted that both of these spatial units are administratively set and do *not* correspond to any economic rationale whatsoever, contrary for instance to metropolitan areas. But unlike metropolitan areas, districts as well as townships cover the whole Belgium territory, which is a necessary requirement when studying the spatial distribution of activities.

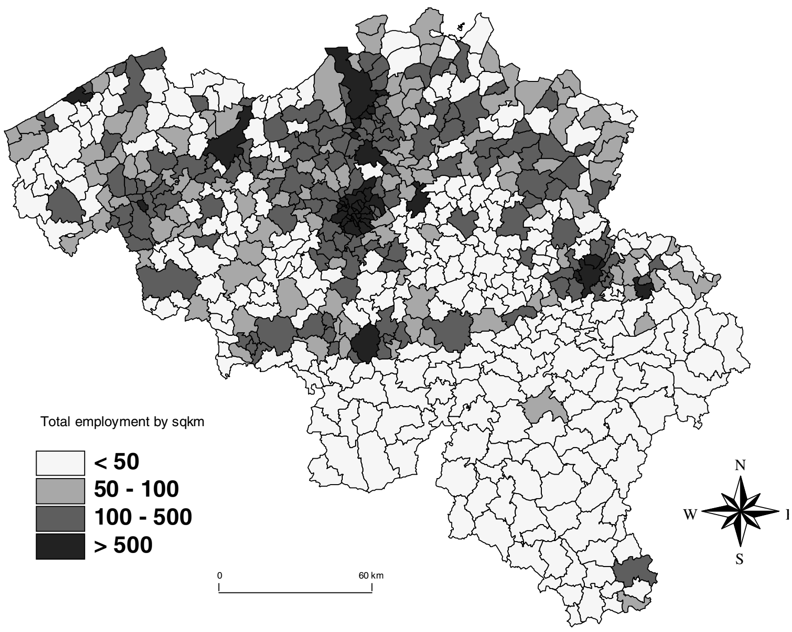
Map 1 draws population density for 2000 (see appendix 5 for a map of Belgian regions). Average densities are 198 inhabitants/km^2 in Wallonia against 439 inhabitants/km^2 in Flanders and 5 922 inhabitants/km^2 for Brussels, indicating that the northern part of the country (Flanders) and Brussels are more densely populated than the Walloon region. A similar map can be drawn for employment density. Differences appear above all for the Brussels-Capital region. The map slightly changes when considering only manufacturing employment rather than total employment.

1. It is worth noting that in Krugman type of models, industrial concentration is the main driving force of geographical agglomeration. Hence, the E-G index totally drops this type of explanation.

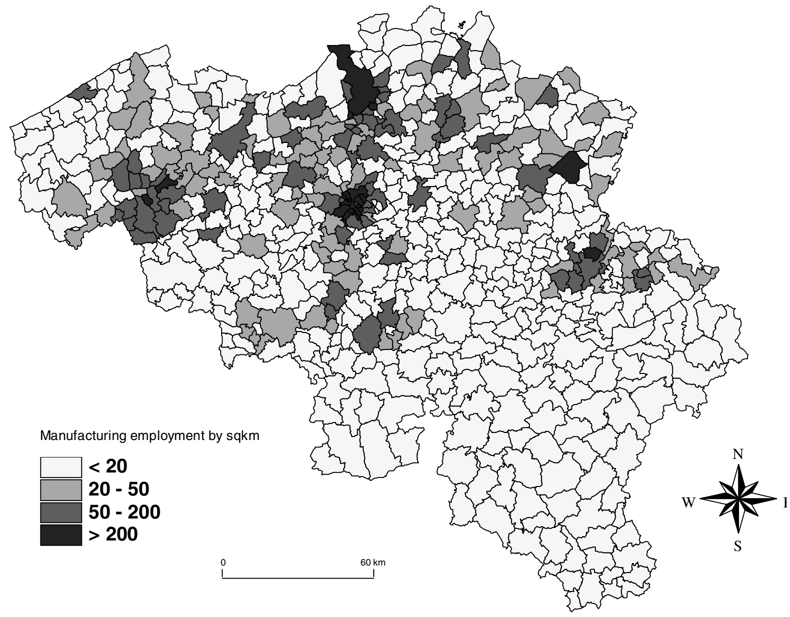
MAP 1 - Population density (2000)



MAP 2 - Employment density (2000)



MAP 3 - Manufacturing employment density (2000)



Two remarks are worth noting:

(i) Total employment is much more concentrated than population and, manufacturing employment is even more concentrated than total employment.

(ii) Wallonia has a more clustered pattern than Flanders. And this is true for total as well as manufacturing employment. See section 3.D for further details.

III Geographical concentration of manufacturing in Belgium

A. Main results

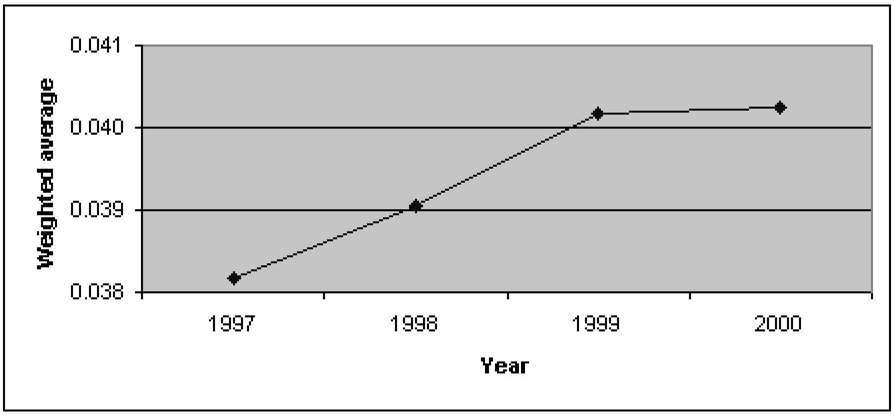
As discussed above, there is a whole series of measures of geographical concentration. Here we shall mainly concentrate on the Ellison and Glaeser (1997) index (henceforth E-G). Appendix 2 provides comparative tables between E-G indices and other measures such as Maurel and Sédillot (1999) for NACE4 industries. There are two major reasons for doing so: first, most studies use such a disaggregation level. By doing so, international comparisons can be made. Second, NACE4 industries seem to be more self-contained than more aggregated levels such as NACE2 or NACE3.[1]

Finally, results presented in this section were computed by using *districts* as the reference spatial units. The E-G index treats spatial units perfectly symmetrically, i.e. it does not distinguish between two neighbouring districts and two districts at both ends of the country. Stated differently, E-G indices totally dismiss the *spatial autocorrelation* issue. Hence, the finer the spatial unit, the more exacerbated this issue is. In section 3.E, we cope with the issue of spatial autocorrelation.

Fig. 2 depicts the evolution of the average E-G index, weighted by the employment share of the NACE4 industries. This index has increased by less than 10% in four years. This rather high figure shall, however, be moderated. Indeed, referring to Ellison and Glaeser's proposed benchmark, industries are on average neither dispersed nor significantly concentrated.[2] Consequently, all that can be said is that agglomeration has increased, but to which extent this increase is significant remains uncertain.

1. In Decrop (2002), NACE2 and 3 disaggregation results are provided for all sectors of activity. NACE3 results, as well as E-G results for townships rather than districts are available upon request.
2. Ellison and Glaeser's (1997) index is built so that values below 0 characterize dispersed industries, whereas values above 0 stand for agglomerated industries (i.e. more agglomerated than random location would suggest). However, in their empirical part, they suggest that industries with values below 0.02 cannot be considered as significantly concentrated, and propose the value of 0.05 as a threshold above which significant spatial concentration can reasonably be claimed.

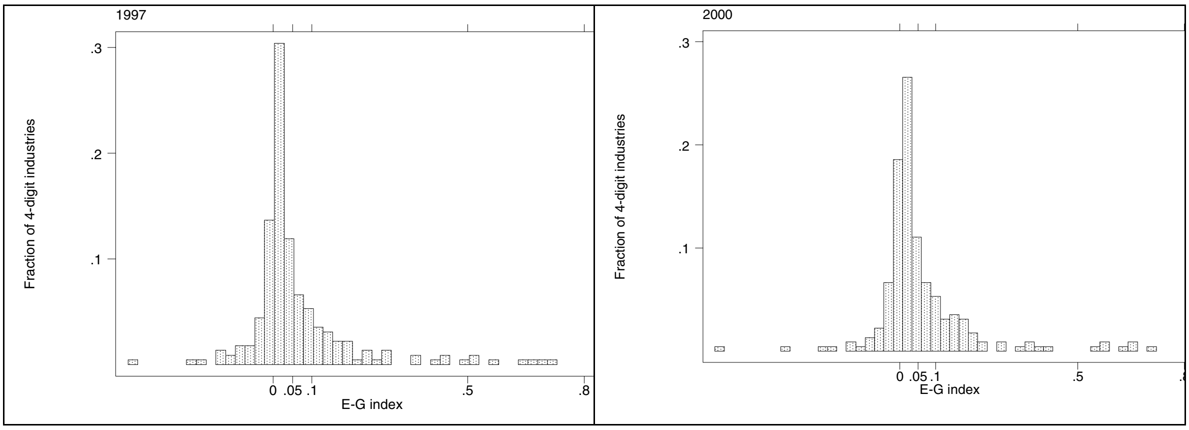
FIGURE 2 - Average E-G index*



* Outlyers (-1 < E-G < 1) are not included in the computation.

Looking at the industrially disaggregated results, fig. 3 plots the distribution of the E-G indices across the NACE4 industries in 1997 and 2000.

FIGURE 3 - E-G indices distribution (1997 and 2000)



As has already been alluded above, most industries are in the range of E-G values which shows no excess concentration. Nonetheless, in 1997, 77 of all 4-digit industries could be considered as excessively concentrated and this figure remained unchanged for 2000. All in all, the distribution is characterized by a strong inertia as, can be inferred from table 5.

TABLE 5 - E-G class distribution

	Number of 4-digit industries	
	1997	2000
E-G < 0	23.1%	24.1%
0 = E-G < 0.02	23.6%	24.6%
0.02 = E-G < 0.05	19.7%	17.5%
E-G > 0.05	33.6%	33.8%

The values of the indices have increased during these four years, but have mostly remained inside the intervals of the critical threshold values. This is mainly the result of the short period of observation. An interesting issue that would deserve closer attention relates to the sources of variation of the indices by industry. Indeed, these E-G indices may vary for two reasons: firstly, entry of new plants, respectively exit of old ones, and secondly through relocation. Unfortunately, this kind of study is out of our scope, as we do not have any plant identifiers at our disposal.[1]

In table 6, we now examine the most localized 4-digit industries in 1997 (initial year).

TABLE 6 - 20 most concentrated NACE4 manufacturing industries (1997 and 2000)

NACE4	Name	Employment share1997	E-G 1997	E-G 2000	Rank 2000
2465*	Manufacture of prepared unrecorded media	0.04%	1.179971	0.045609	82
2233	Reproduction of computer media	0.00%	0.880805	-0.21719	226
2653	Manufacture of plaster	0.02%	0.726968	0.626721	6
1722	Carded wool weaving	0.02%	0.698011	0.703953	3
2615	Manufacture and processing of other glass, including technical glassware	0.13%	0.672499	0.075822	59
1724	Weaving of silk	0.12%	0.647099	0.567289	8
2464*	Manufacture of chemical products for photography	1.00%	0.576459	0.581766	7
3621	Manufacture of coins and medals	0.03%	0.513142	0.419537	10
2330**	Treatment of nuclear fuel	0.17%	0.51001	0.657515	4
2753	Foundry of light metals	0.14%	0.491351	0.656758	5
2611	Manufacture of flat glass	0.44%	0.45183	0.389529	11
2626	Manufacture of refractory ceramics	0.13%	0.447705	0.370092	13
3001**	Manufacture of office machinery	0.01%	0.407472	-2.78E-06	174
2320	Petroleum refining	0.45%	0.374138	0.552565	9
2960**	Manufacture of arms and ammunition	0.27%	0.360152	0.33789	14
1723	Weaving of wool	0.03%	0.291652	0.271878	16
2872	Manufacture of light metal packing	0.24%	0.289532	0.37548	12
1591	Production distilled alcohol beverages	0.01%	0.283065	0.160899	31
1542	Manufacture of refined oils and fats	0.15%	0.256841	0.17546	25
2921*	Manufacture of industrial furnaces and burners	0.02%	0.248493	0.160967	30

* Medium-high tech industries.

** High-tech industries.

The E-G index is a model-based index. Ellison and Glaeser (1997) have started from the fact that excess-agglomeration of economic activity may result from either the presence of natural advantage or spillovers. However, they establish "something of an observational equivalence result between the effects of natural advantages and spillovers on expected concentration levels", hence "any estimated [E-G index] is compatible with a pure natural advantage model, a pure spillover model, or models with various combinations of the two factors" (pp.

1. Data is presented by firm identifiers, hence questions concerning multiplant firms can be tackled.

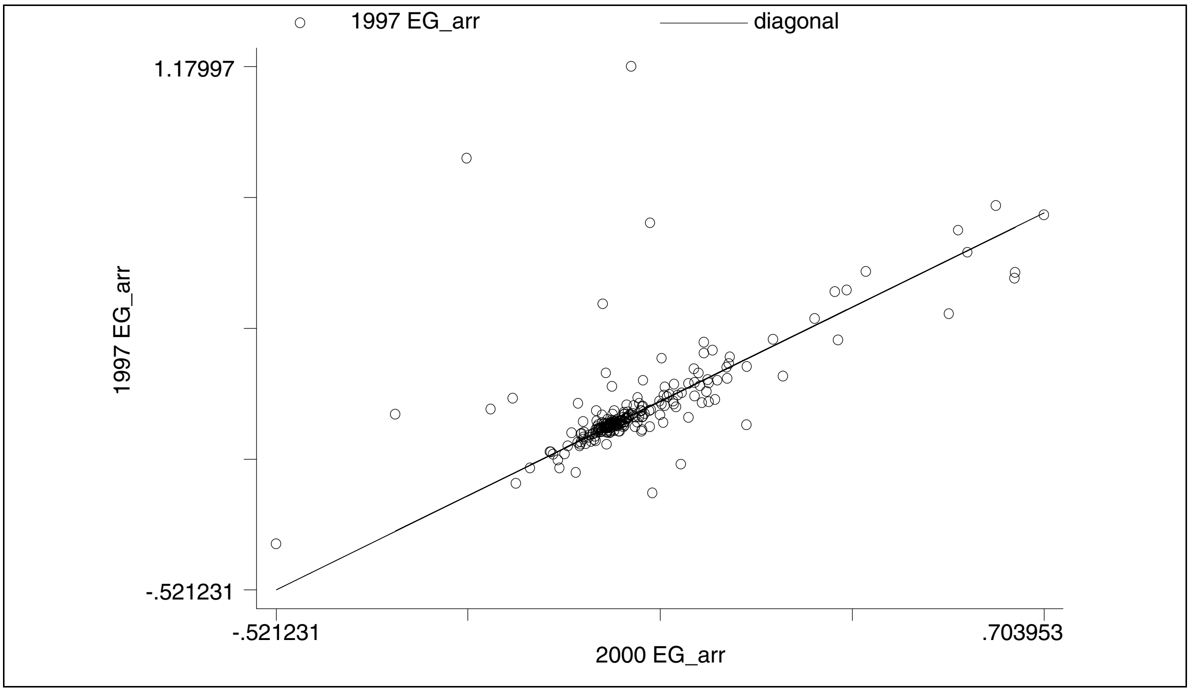
896-7). As such, at this stage of the analysis, only conjectures on the determinants of agglomeration may be put forward.[1] A first comment, deriving from table 6, concerns the comparison between the results in 1997 and 2000. Out of the 20 most localized industries in 1997, 7 did not belong anymore to this top 20 in 2000, among which 4 were ranked above position 31. One has to be very careful when considering such large variations in the E-G index. Indeed, altogether, these 4 industries represented less than 0.2% of total employment. In other words, at least for some of these industries, entry or exit of one single plant may have a dramatic impact on the absolute value of the E-G index. This is for instance the case for industry 2233, with 2 plants in 1997 and 4 in 2000.[2]

A closer look at the results support evidence found in other countries, and can directly be related to historical reasons. For instance textile industry is found to be agglomerated in most studies of this type. Industry 2611 (Flat glass mills) is found to be excessively agglomerated for France too. A second typical result concerns refining of petroleum, which is normally strongly agglomerated near the coast for obvious reasons of high transportation cost. Furthermore, a striking feature in our pattern of agglomeration relates to the fact that we do find some high tech industries in the present top 20 classification, as can be expected. Similar results hold for France (Maurel and Sédillot, 1999) and Portugal (Teixeira, 2002).[3] A potential driving force for the geographical concentration of some high-tech sectors may be their close link to more traditional industries that happen to be concentrated due to natural advantages. One may think that this is the case for industries 2464 and 2921 (Manufacture of chemical products for photography, Manufacture of industrial furnaces and burners). But further analysis is required to assess this explanation.

Except for the very small industries, there is a relative inertia of the E-G indices across this 4-year period. This can be further assessed by the following figure, where x-axis (y-axis) represent the E-G index in year 2000 (1997). Besides some outliers that have already been mentioned, most industry observations are concentrated around the diagonal.

1. Some studies have led such kind of analysis: Dumais *et al*. (2002), Rosenthal and Strange (2001), Teixeira (2002).
2. The same remark applies to sector 3001, 2615, 2465 with respectively 3, 18, 4 plants in 1997 and 2, 17,5 plants in 2000.
3. High-tech classification stems from Hatzichronoglou (1997).

FIGURE 4 - Evolution of E-G indices between 1997-2000



So far the results tend to support that for about one third of the industries (cf. table 5), there is a high correlation between the plants' location decisions. However, this result should be interpreted with great caution as it only represents a picture at a certain moment in time, but does not account for past events having led to such a situation. A certain degree of concentration at the present time may hide movements towards deconcentration in highly concentrated industries and the reverse for relatively dispersed industries. Considering the threshold values suggested by Ellison and Glaeser, we represent the 4-digit industries that were not concentrated in 1997 (E-G < 0.02) and became excessively agglomerated in 2000 (E-G > 0.05), and vice versa, in order to give some hints about this static *versus* dynamic issue.

TABLE 7 - Upward/downward movements of E-G index

Industries that were not considered as localized in 1997 and became excessively localized in 2000		Industries that were considered as localized in 1997and not any more in 2000	
NACE4	Name	NACE4	Name
1572	Manufacture of domestic pets food	1543	Production of margarine
1717	Preparation and spinning of other fibers	1752	Manufacture of strings, ropes and nets
2223	Binding and completion	2215	Other publishing activities
2466*	Manufacture of various chemical products	2221	Newspaper printing
2662	Manufacture of plaster elements for construction	2233	Reproduction of computer recordings
3162*	Manufacture of other electrical materials	2420*	Manufacture of agro-chemical products
3530**	Aeronautical and spatial construction	2732	Cold rolling of strip iron
		3001**	Production of office machinery
		3230**	Production of reception, recording or reproduction of sounds and images apparatus

* Medium-high tech industries.

** High-tech industries.

There is no straightforward interpretation of this upward respectively downward movements in the agglomeration space by these 4-digit industries. Note that high-tech industries are over-represented in table 7, which could indicate that these industries tend to be more footloose. Note, however, that except for two industries, there is no overlapping between table 6 and table 7. Furthermore, a potential explanation of these large movements may come from a very low number of plants in each of these industries, and hence, there can be no economic rationale behind these figures.[1]

B. Scope of spillovers across industries

A further issue that has been investigated in this type of literature is the scope of spillovers across industries. A prior to this study is to analyze how 4-digit industries, belonging to the same 2-digit industry, tend to behave, *i.e.* if their pattern of geographical concentration is correlated. Appendix 3 summarizes the distribution of 4-digit industries within 2-digit industries. The intervals were chosen in order to account for dispersed industries (E-G<0), not significantly localized (0≤E-G<0.02), intermediate values (0.02≤E-G<0.05) and excessively localized (E-G≥0.05) patterns.

In about half of the 2-digit industries, 50% or more of the 4-digit industries are concentrated in one sole interval. Moreover, industries that were found to be excessively concentrated at the 4-digit level can mostly be recovered here. This is for instance the case for the textile industry as well as coking, refining and nuclear industry. Contrary to the United States (Ellison and Glaeser, 1997), but similar to Portugal (Teixeira, 2002), the tobacco industry cannot be considered as significantly agglomerated. Finally, the high-tech industry 30 is found to be excessively localized, compared to 4-digit results. This result has also been highlighted in the French case (Maurel and Sédillot, 1999).

Although these results provide information about patterns of agglomeration for 4-digit industries belonging to the same 2-digit industries, no conclusion can be drawn upon intra-industry spillovers. For 2-digit industries, high levels of geographical concentration may result from within NACE4 sub-industry concentration (localization economies) and/or from between NACE4 sub-industry co-agglomeration (urbanization economies). To cope with this issue, we have computed Ellison and Glaeser's co-agglomeration indices for 4-digit industries belonging to the same 2-digit industries.[2] This allows us to tackle the specialization *versus* diversity issue mentioned above. We also provide results for E-G indices at NACE2 level to discriminate between high and low level concentration industries.

1. This is actually the case for instance for industries 2233 and 3001 which are common in both tables, but represent no more than 0.01% of total employment.
2. Table 8 actually represents the normalized co-agglomeration index. The interpretation is hence straightforward: values of the co-agglomeration index equal to 0 indicate that any spillovers/natural advantages found within the industry group are completely industry-specific, whereas values of 1 are consistent with spillovers benefiting plants in all the industries and/or natural advantages are perfectly correlated.

TABLE 8 - Co-agglomeration of 4-digit industries among 2-digit industries

NACE2		E-G 1997	Co-aggl.1997	E-G 2000	Co-aggl.2000
15	Food industry	0.004	0.164	0.004	0.193
16	Tobacco industry[1]	0.033	/	0.008	/
17	Textile industry	0.098	0.645	0.103	0.644
18	Clothes and fur industry	0.022	0.767	0.026	0.664
19	Leather and footwear industry	0.012	-0.049	0.016	-0.052
20	Wood, cork, basket and esparto industry	0.019	0.125	0.021	0.096
21	Paper and paperboard industry	0.009	0.196	0.011	0.311
22	Publishing, printing and reproduction	0.024	0.934	0.024	1.087
23	Coking, refining and nuclear industries	0.125	-0.275	0.219	-0.148
24	Chemical industry	0.023	0.151	0.02	0.122
25	Rubber and plastic industry	0.012	0.490	0.012	0.498
26	Production of other non-metallic minerals	0.010	0.046	0.007	0.012
27	Metallurgy	0.006	-0.043	0.012	-0.057
28	Metal transformation	0.005	0.170	0.005	0.216
29	Production of machines and equipment	0.004	-0.041	0.003	0.043
30	Production of office machines and computer materials	0.060	0.037	0.056	0.753
31	Production of electrical machines and apparatus	-0.006	0.434	-0.005	-2.866
32	Production of radio, broadcast and communication equipment	-0.035	5.944	-0.023	-3.562
33	Production of medical, precision, optical and clock instruments	0.019	3.141	0.01	0.689
34	Production and assembling of motor vehicle, tow and semi-trailer	-0.038	0.003	-0.022	-0.280
35	Production of other transport materials	0.005	0.843	0.009	0.016
36	Furniture production; various industries	0.009	0.026	0.007	0.040
37	Recycled material recovery	0.006	0.335	0.002	0.203

1 For the tobacco industry (NACE2 16), there is only one NACE4 sub-industry.

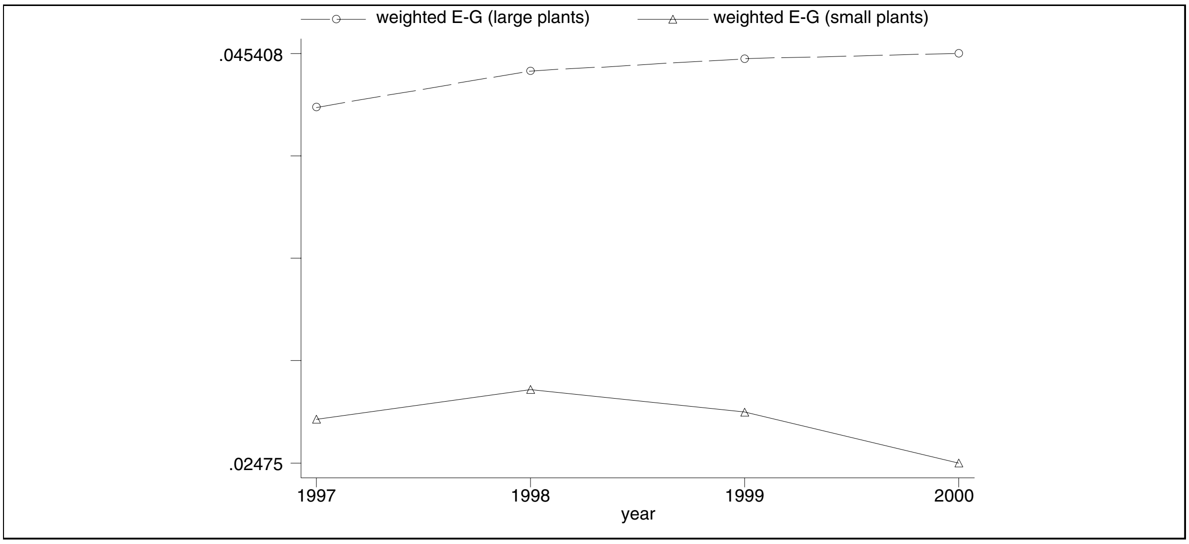
A first striking remark is that the scope of spillovers is largely intra-industrial rather than inter-industrial. Out of 22 two-digit industries, only 4 do highlight high values of co-agglomeration (stable through the time): textile industry, clothes and fur industry, publishing, printing and reproduction, production of medical, precision, optical and clock instruments. The latter has the highest values of co-agglomeration and can be considered as a high-tech industry. Hence this is consistent with the concept of science-based diversity suggested by Feldman and Audretsch (1999), i.e. diversity may favour cross fertilization in R&D intensive activities. Moreover, high values of co-agglomeration for industries 17 and 22 coincide with US results. Finally, no co-agglomeration can be highlighted for 7 two-digit industries: 19, 23, 26, 27, 29, 34 and 36. No horizontal spillovers, respectively no natural advantage correlation exists for these industries. These results hence tend to suggest that localization rather than urbanization economies prevail. However, note that only horizontal spillovers are taken into account when aggregating 4-digit industries in 2-digit ones. Input-output linkages are totally left aside here, thus leaving open the existence of hypothetical forward/backward linkages (Fujita et al., 1999).

C. Agglomeration and plant scale

In two recent contributions, Holmes and Stevens (2000) have raised the issue of the plant size when measuring the geographical concentration of industries. "Do plants located in areas where an industry concentrates tend on average to be larger, the same as, or smaller than establishments located outside such areas?" (p.1). This is the question addressed in the sequel. More specifically, we have shown that the distribution of plant size is excessively skewed: single job plants represent 35% of the total number of plants in 2000, plants with less than or equal to 5 jobs represent 72%, and plants less than or equal to 20 jobs 93% of the total number of plants. Nevertheless, note that the latter class only represents 34% in terms of employment. Thus considering erroneously that location decisions follow the same rationale, independently of plant size, may induce large distortions. Why erroneously? It has been shown that plant productivity in agglomerated areas is higher than outside these areas (Ciccone and Hall, 1996). If we consider that productivity is related to the concentration of the "hard-core" manufacturing plants in an industry, rather than total industry employment, than standard measures of geographical concentration may tend to understate this concentration (Holmes and Stevens, 2000).

Fig. 5 clearly depicts the difference in geographical concentration between large plants (i.e. plants with employment above 20 persons) and small plants (i.e. plants with 20 employees or less). This result is hence consistent with Holmes and Stevens who find that plants located in areas where an industry concentrates, are on average larger than plants in the same industry outside such areas, and this is particularly striking in the case of manufacturing. In terms of spillovers, this means that larger plants are more prone to the absorption of external effects than smaller firms.

FIGURE 5 - Evolution of average E-G indices according to plant size



Looking now at the disaggregated 4-digit level, we see that over our 4 years of observation, in two thirds of the cases, the E-G index for the large plants manufacturing industry was higher than in its small counterparts.[1] Even though the aggregate result is hence qualified by the disaggregated result, it remains important to be aware of these differences which may substantially bias the conclusions.

When one is mainly interested in the ranking of manufacturing industries in terms of geographical concentration, rather than the absolute value of concentration, an interesting issue is to analyze whether large plant E-G indices rank roughly similar to small plant E-G indices, even though, on average, the absolute value of the E-G index is smaller in the latter case. To tackle this issue intuitively, we have computed the Spearman rank correlation indices for the 4 years under observation, as well as the simple correlation coefficient.

TABLE 9 - Correlation of E-G indices according to plant size

	Spearman	Correlation coefficient
1997	0.1897	0.0425
1998	0.2221	0.1398
1999	0.2820	0.0743
2000	0.2454	0.1222

Not only do small firms have lower levels of geographical concentration than larger ones, but, their ranking in terms of levels of concentration also largely differ. This result leads us to great caution in interpreting results upon indices computed on all plants. As Holmes and Stevens (2000) suggest, classical industry definitions (e.g. NACE) may poorly measure what plants are doing. “A four-digit industry might include small, geographically diffuse establishments that perform a retail or service function, as well as, geographically concentrated establishments that do hard-core manufacturing.” (pp. 11-12).

Finally, note that out of the top 20 industries (in terms of geographical concentration), only 2 coincide. However, comparing results with the E-G indices computed for all plants, there are 2 correspondences with the small plants sample, and 11 with the large plant sample. Thus, this may lead us to believe that E-G index reflects large plant rather than small plant locational outcomes.

1. Moreover, note that average differences between E-G indices are in absolute value larger for [E-G(large plants)]-[E-G(small plants)], rather than for [E-G(small plants)]-[E-G(large plants)].

TABLE 10 - Top 20 most localized industries according to plant size

NACE4	Name	E-G (large plants)	NACE4	Name	E-G (small plants)
2753	Casting of light metals	0.7242	1583	Manufacture of sugar	0.8710
2330	Processing of nuclear fuel	0.7145	3543	Manufacture of invalid carriages	0.8276
1724	Silk-type weaving	0.6108	1725	Other textile weaving	0.4604
2464	Manufacture of photographic chemical material	0.5957	2960	Manufacture of weapons and ammunition	0.4490
2320	Manufacture of refined petroleum products	0.5700	1721	Cotton-type weaving	0.4032
2626	Manufacture of refractory ceramic products	0.5285	1751	Manufacture of carpets and rugs	0.3385
1717	Preparation and spinning of other textile fibres	0.4415	1714	Preparation and spinning of flax-type fibres	0.2652
2611	Manufacture of flat glass	0.4093	3615	Manufacture of mattresses	0.2385
2622	Manufacture of ceramic sanitary fixtures	0.4082	3622	Manufacture of jewellery and related articles n.e.c.	0.2188
2872	Manufacture of light metal packaging	0.3968	3662	Manufacture of brooms and brushes	0.2125
1723	Worsted-type weaving	0.3851	1711	Preparation and spinning of cotton-type fibres	0.2106
2615	Manufacture and processing of other glass, including technical glassware	0.3735	1772	Manufacture of knitted and crocheted pullovers, cardigans and similar articles	0.2069
2960	Manufacture of weapons and ammunition	0.2987	2954	Manufacture of machinery for textile, apparel and leather production	0.1986
2225	Other activities related to printing	0.2976	3511	Building and repairing of ships	0.1801
2020	Manufacture of veneer sheets; manufacture of plywood, laminboard, particle board, fibre board and other panels and boards	0.2639	3621	Striking of coins and medals	0.1718
1824	Manufacture of other wearing apparel and accessories n.e.c.	0.2344	2052	Manufacture of articles of cork, straw and plaiting materials	0.1663
3162	Manufacture of other electrical equipment n.e.c.	0.1981	2123	Manufacture of paper stationery	0.1646
1760	Manufacture of knitted and crocheted fabrics	0.1944	1600	Manufacture of tobacco products	0.1281
1562	Manufacture of starches and starch products	0.1932	2461	Manufacture of explosives	0.1280
1772	Manufacture of knitted and crocheted pullovers, cardigans and similar articles	0.1923	1920	Manufacture of luggage, handbags and the like, saddlery and harness	0.1130

D. Regional outcomes

So far, we have analyzed the geographical concentration of manufacturing industry exclusively at national level. This implies that we consider countries as "self-contained" in terms of agglomeration outcomes, *i.e.* trans-border effects are negligible. There are at least two reasons for doing so: data availability and consistency. If available, consistent data can at best be found at national level. Collecting international data in order to compute E-G indices is a tremendous task. One needs consistent industrial as well as spatial units to collect data, but also a similar definition of what a *plant* is, and how *employment* is accounted for. Following this consistency argument, national data probably constitutes the least worse way to neutralize political, social and historical divergences across countries.

However, this issue can also be reversed: why do we measure geographical concentration at the national rather than sub-national level? Indeed, there might be substantial heterogeneity inside countries, justifying such an approach. This is particularly true for Belgium. Belgium is a federal state with three regions: Brussels, Flanders and Wallonia. Computing concentration indices for Brussels makes no sense due to its excessively small extent. However, one may be interested in different patterns of concentration in Flanders and Wallonia. These two regions have faced very different paths of development. In the past, Wallonia used to be the richest part of the country, partly due to its iron and steel industry. But after the crises in this industry in the second half of the 20th century and the industrialization of Flanders at the same time, Flanders caught up and eventually leapfrogged Wallonia in terms of development, with, however, very different areas of specialization, as can be seen in table 11.

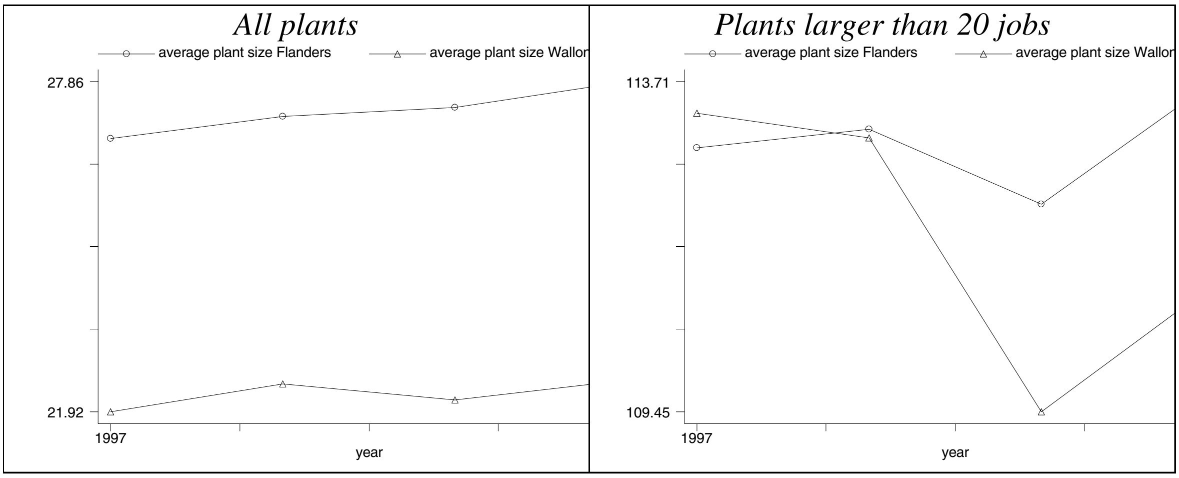
TABLE 11 - Top 10 most important industries according to region (year 2000) - Share in total manufacturing employment

NACE4	Name	Employment share Flanders	NACE4	Name	Employment share Wallonia
3410	Manufacture of motor vehicles	6.712%	2710	Manufacture of basic iron and steel and of ferro-alloys	7.992%
1581	Manufacture of bread; manufacture of fresh pastry goods and cakes	3.027%	1581	Manufacture of bread; manufacture of fresh pastry goods and cakes	4.508%
2222	Printing n.e.c.	2.500%	2442	Manufacture of pharmaceutical preparations	4.504%
1751	Manufacture of carpets and rugs	2.418%	2952	Manufacture of machinery for mining, quarrying and construction	3.370%
2414	Manufacture of other organic basic chemicals	2.238%	2852	General mechanical engineering	2.381%
3430	Manufacture of parts and accessories for motor vehicles and their engines	2.161%	3530	Manufacture of aircraft and spacecraft	2.328%
2524	Manufacture of other plastic products	1.804%	3520	Manufacture of railway and tramway locomotives and rolling stock	2.215%
2852	General mechanical engineering	1.773%	2222	Printing n.e.c.	1.997%
2710	Manufacture of basic iron and steel and of ferro-alloys	1.708%	2112	Manufacture of paper and paperboard	1.870%
2811	Manufacture of metal structures and parts of structures	1.599%	2612	Shaping and processing of flat glass	1.744%

Belgium is one of the leading car manufacturers in Europe (although it has no national make). This clearly appears in table 11 in the Flemish column. On the Walloon side, the burden of history is clearly reflected by the iron and steel industry in the first position with almost 8% of total manufacturing employment. Apart from this latter industry, three more common industries are in the top 10 of the most important regional activities: printing n.e.c., bread, pastry and cake industry, and general mechanical engineering. These may be industries that are generally important in all countries, above all the second one. Note that for Wallonia, there are more heavy industries, *i.e.* iron and steel, machinery for mining, railway, which are absent in the Flemish top 10. A more interesting feature is that large industries are slightly larger in Wallonia compared to Flanders, tending to support a larger specialization.

In terms of average plant size, there has been almost no variation across the period of observation, but Flanders has been steadily above Walloon average.

FIGURE 6 - Average plant size in Flanders and Wallonia



The same observation holds when we restrict our sample to plants larger than 20 jobs. Actually the ratio of small to large plants is more or less similar between both these regions. In 2000, there were indeed 15920 manufacturing plants in Flanders and 6381 in Wallonia, with the following breakdown:

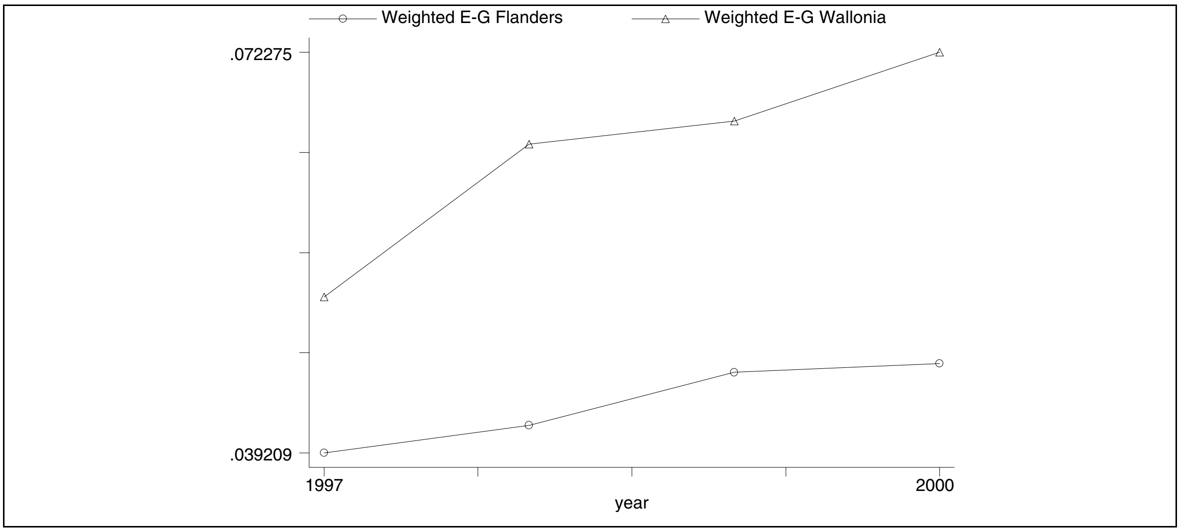
TABLE 12 - Number of manufacturing plants by size and region

	Flanders	Wallonia
0 - 20 jobs	12 595	5 320
21 - 100 jobs	2 536	807
More than 100 jobs	789	254

Moving now to the concentration issue, a first crude observation is the difference in average concentration indices between both regions. Fig. 7 depicts a higher degree of concentration in Wallonia than in Flanders. Indeed, the Walloon curve always exceeds the 0.05 threshold value, whereas Flanders always stays beneath it, although converging towards this value. This feature is already highlighted in map 3, i.e. Wallonia has a more clustered pattern for total manufacturing employement.

A second feature is related to the increase of the E-G index during the period of observation. This is consistent with fig. 2 where the E-G index was also increasing during the period of observation. This might not have been the case, as the Brussels-capital region is totally left aside in this sub-section.

FIGURE 7 - Flanders and Wallonia average E-G index



NB: Outliers (E-G>1 and E-G<-1) were dropped to plot Fig.7

When measuring Spearman's rank correlation between the E-G of Flanders and Wallonia by industry, one finds figures between 0.03 and 0.16 according to the period. This points to rather different patterns of agglomeration according to the region we observe. And indeed, when observing the top 10 E-G indices for Flanders and Wallonia, there are no common industries. But, one can observe that except for two industries, no other industries ranks in the top 20 of the other region's top ten industries. Note, however, that we do find sub-industries of the same NACE2 industry in both regions, especially the textile industry (NACE2 17).

TABLE 13 - Top 10 E-G indices according to region (year 2000)

NACE4	Name	E-G Flanders	Rank Wallonia	NACE4	Name	E-G Wallonia	Rank Flanders
1722	Woollen-type weaving	1.190	/	1714	Preparation and spinning of flax-type fibres	6.650	24
2330	Processing of nuclear fuel	0.926	188	2452	Manufacture of perfumes and toilet preparations	1.863	170
3530	Manufacture of aircraft and spacecraft	0.874	183	2470	Manufacture of man-made fibres	1.739	179
2613	Manufacture of hollow glass	0.849	52	2861	Manufacture of cutlery	1.372	25
2960	Manufacture of weapons and ammunition	0.737	152	1771	Manufacture of knitted and crocheted hosiery	1.355	43
2626	Manufacture of refractory ceramic products	0.737	17	1760	Manufacture of knitted and crocheted fabrics	1.310	62
1910	Tanning and dressing of leather	0.604	19	1721	Cotton-type weaving	1.242	29
2753	Casting of light metals	0.582	/	1830	Dressing and dyeing of fur; manufacture of articles of fur	1.165	125
1724	Silk-type weaving	0.578	/	3161	Manufacture of electrical equipment for engines and vehicles n.e.c.	1.002	55
2464	Manufacture of photographic chemical material	0.519	190	3650	Manufacture of games and toys	0.964	215

E. Agglomeration and spatial autocorrelation

So far, our approach consisted in analyzing the spatial agglomeration pattern across Belgium. In doing so, an implicit assumption on the spatial unit has been made continuously, namely the unit we are considering is the most appropriate scale to measure agglomeration. Among geographers, there is a long tradition of determinating the optimal spatial unit, and the choice of this unit is far from being neutral when it comes to analyzing the spatial distribution of manufacturing activity. Indeed, referring to the theoretical literature on spatial economics, at least two types of mechanisms may lead to different agglomeration patterns. Theory tells us that *pecuniary externalities* are much broader than *technological externalities*, implying that a much denser agglomeration shall result from the latter than from the former, i.e. Manufacturing Belt in the US for the former vs. Akihabara (Electric City) in Tokyo for the latter. A result of this is that fine spatial units of observation will capture more technological externalities related industries whereas rougher spatial units are more prone to highlight pecuniary externalities related industries. A way to circumvent somehow this issue is to have recourse to a measure of spatial autocorrelation.

Exploratory Spatial Data Analysis is a set of techniques aimed at describing and visualizing spatial distributions, identifying atypical localization or spatial outliers, detecting patterns of spatial association, clusters or hot spots, and suggesting spatial regimes or other forms of spatial heterogeneity. Central to this conceptualization is the notion of spatial autocorrelation or spatial association, i.e., the phenomenon where *locational similarity* is matched by *value similarity* (Anselin and Bera, 1998). This means that observations are not statistically independent. Clusters of events, people and facilities are referred to as positive spatial autocorrelation, whereas negative spatial autocorrelation refers to arrangements where people, events or facilities are dispersed.

There are a number of formal statistics that attempt to measure spatial autocorrelation. Among those, Moran's *I* statistic (Moran, 1950) is probably the most popular one. This statistic compares the value of a continuous variable at any location with the value of the same variable at surrounding locations. Formally, it is defined as:

$$I = \frac{N}{S} \cdot \frac{\Sigma_i \Sigma_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\Sigma_i (x_i - \bar{x})^2}$$

with $S = \Sigma_i \Sigma_j w_{ij}$, and x_i representing the value of the observation in region *i*. N is the total number of observations, $\bar{x}$ is the mean of the variable across all observations and w_{ij} is a weight between region *i* and region *j*. Usually w_{ij} is an inverse function of distance between both these regions or may take value 1 when regions *i* and *j* are contiguous and 0 elsewhere. Moran's *I* is similar to a correlation coefficient because it compares the sum of the cross-products of values at different locations, two at a time. However, I is not restricted by the interval [-1,1]. When nearby points have similar values, the cross-product is high. Conversely, when nearby points have dissimilar values, the cross-product is low. More precisely, values of *I*, significantly larger than the expected value of the Moran statistic, E[I] =-1/(N-1), indicate positive spatial association, whereas significantly smaller values indicate negative spatial association. Inference can be based upon two approaches. Values of *I* significantly, which on higher than the expectation of Moran's statistic, will denote positive spatial autocorrelation. In other words, similar values, either high values or low values, are more spatially clustered than could be caused purely by chance. The converse is true for values of *I* which are significantly lower than the expectation of *I*.

A first way of roughly detecting the presence of spatial autocorrelation can be done by comparing Ellison and Glaeser indices at two different spatial unit levels. This has been done for the Belgian case for *districts* and *townships*. Whereas the Spearman rank correlation between the indices is about 0.63 and highly significant, the unweighted E-G averages are 0.031 and 0.058 for the townships respectively the districts, meaning that districts tend to capture more thoroughly hypothetical clusters.[1]

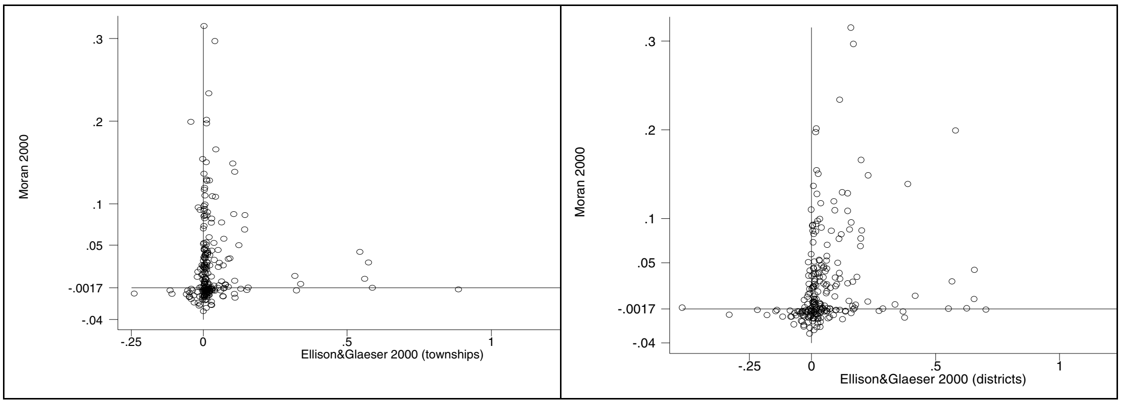
In order to refine this analysis, we have computed Moran indices of spatial autocorrelation for the finest spatial units at our disposal, namely townships. Results are provided in Appendix 4. First and foremost, 95% of the indices indicate positive spatial autocorrelation for the four years under scrutiny. Hence, positive spatial autocorrelation can be considered as a rule rather than as an exception in the present setting. However, only less than one third of these results do highlight statistically significant spatial autocorrelation (5% criterion). So in a spillover perspective, this means that about 30% of all our industries do highlight across township-border externalities. When taking a significant threshold of 10%, industries entering the positive spatial autocorrelation category increase to 35%.

On the other hand, industries highlighting a negative spatial autocorrelation would be more prone to an interpretation where townships are *self-contained* in terms of spillovers, *i.e.* spillovers are intra- rather than inter towns (e.g. a center and its hinterland). Note, however, that none of these industries do highlight statistically significant negative autocorrelation.

Finally, in figure 8 we have plotted Moran indices (at township level) against E-G indices (at the district and township levels). On the left figure, most E-G indices at township level are concentrated around the zero value, pointing to no excessive concentration. Conversely, on the right figure, most dots are concentrated either around the zero value for the E-G index and around the expectation (-0.0017) for the Moran, or in the upper right quadrant. This clearly points to what has already been mentioned above: when measured at township level, only very few clustered industries emerge. A reason for this may be spatial autocorrelation between townships. Thus, recomputing the E-G indices at a higher degree of spatial aggregation, will partly internalize positive spatial autocorrelation, leading to a higher concentration of dots in the north-east part of our right figure.[2]

1. Results refer to the year 2000. Average values for the E-G indices at township respectively district levels for 1997 are 0.035 and 0.072, leading to similar relative values as for the year 2000. Note that outliers (E-G<2 and E-G<-2) where eliminated to compute those unweighted averages.
2. When dropping outlying values of the E-G index, the simple correlation coefficient between Moran's I and E-G(townships) is 0.0085, whereas it increases to 0.21 when taking the E-G(districts) instead of E-G(townships).

FIGURE 8 - E-G index vs. Moran's



Further steps

The aim of this paper is to carry out a thorough descriptive analysis of the distribution of manufacturing activity across Belgium. In doing so, emphasis has been placed on measures of agglomeration. The intention of this study is therefore not to provide straightforward policy prescriptions.

Describing the spatial patterns of manufacturing activity may, however, be relevant to policymakers. Governments devote quite significant amounts of money to encourage spatial equity. This goal can be achieved in many different ways, one of which is to promote the development of less favoured regions, e.g. by means of incentive policies granting special advantages to firms that locate in the designated areas.

Although such policies are necessary and valuable, they imply a trade-off against efficiency. As we have mentioned, the measures of spatial agglomeration which are used in this paper, i.e. the Ellison-Glaeser and the Maurel-Sédillot indices, are model-based indices. In these models, location is the result of natural advantages and/or externalities. Having confined ourselves to manufacturing industries, a further step would now be to analyze the determinants of agglomeration, i.e. natural advantages or externalities. As a result one could focus on specific policies for industries for which externalities play a greater part (such as stimulating specific clustering of linked activities). As we find very different patterns of agglomeration for the manufacturing industries, externalities may be especially dominant in some industries. In this respect it seems reasonable to think that a more equitable spatial policy should take the differential benefits and costs of agglomeration between industries into account.

Bibliography

Ciccone, Antonio and Robert Hall (1996), Productivity and the Density of Economic Activity, *American Economic Review*, 86(1), pp. 54-70

Decrop, Jehan (2002), Agglomération et dynamique des activités économiques dans les villes belges: Une approche spatiale et sectorielle, *Federal Planning Bureau of Belgium working paper* 09-02

Devereux Michael, Rachel Griffith and Helen Simpson (1999), The geographical Distribution of Production Activity in the UK, *IFS working paper* n°26/99

Dumais Guy, Glenn Ellison and Edward Glaeser (2002), geographical Concentration as a Dynamic Process, Review of Economics and Statistics, 84(2), pp. 193-204

Duranton, Gilles and Henry Overman (2001), localizationgeographical in the UK Manufacturing Industries, mimeo LSE

Ellison, Glenn and Edward Glaeser (1997), geographical Concentration in U.S. Manufacturing Industries: A Dartboard Approach, *Journal of Political Economy*, 105(5), pp. 889-927

Ellison, Glenn and Edward Glaeser (1999), The geographical Concentration of Industry: Does Natural Advantage Explain Agglomeration?, *American Economic Review*, 89(2), May 1999, pages 311-16

Feldman, Maryann and David Audretsch (1999), Innovation in Cities: Science-Based Diversity, Specialization and Localized Competition, *European Economic Review*, 43(2), pp. 409-29

Fujita, Masahisa, Paul Krugman and Anthony Venables (1999), The Spatial Economy: Cities, Regions, and International Trade, MIT Press

Glaeser, Edward, Heidi Kallal, José Scheinkman and Andrei Shleifer (1992), Growth in Cities, *Journal of Political Economy*, 100, pp. 1126-52

Hatzichronoglou Thomas (1997), Revision of High-Technology Sector and Product Specification, STI working paper 1997/2, OECD, Paris

Henderson, Vernon, Ari Kuncoro and Matt Turner (1995), Industrial Development in Cities, *Journal of Political Economy*, 103(5), pp. 1067-90

Holmes, Thomas and John Stevens (2000), geographical Concentration and Establishment Scale, forthcoming *Review of Economics and Statistics*

Holmes, Thomas and John Stevens (2002), The Role of Cities: Evidence from the Placement of Sales Officies, *Federal Reserve Bank of Minneapolis* Staff Report #298

Hoover, Edgar (1937), Location Theory and the Shoe and Leather Industry, Cambridge, Mass.: Harvard University Press

Krugman, Paul (1991), Increasing Returns and Economic Geography, *Journal of Political Economy*, 99(3), pp. 483-99

Maurel, Françoise and Béatrice Sédillot (1999), A Measure of the geographical Concentration in French Manufacturing Industries, *Regional Science and Urban Economics*, 29(5), pp. 575-604

McLaughlin, Glenn (1930), Industrial Diversification in American Cities, *Quarterly Journal of Economics*, 45(1), pp. 131-49

Moomaw, Ronald (1981), Productivity and City Size: a Crtique of the Evidence, *Quarterly Journal of Economics*, 96, 675-88

Neary, Peter (2001), Of Hype and Hyperbolas: Introducing the New Economic Geography, *Journal of Economic Literature*, 39(2), pp. 536-61

Rauch, James (1993), Productivity Gains from geographical Concentration of Human Capital: Evidence from the Cities, *Journal of Urban Economics*, 34(3), pp. 380-400

Rosenthal, Stuart and William Strange (2001), The Determinants of Agglomeration, Journal of Urban Economics, 50(2), pp. 191-229

Saxenian, AnnaLee (1996), Regional Advantage: Culture and Competition in Silicon Valley and Route 128, Harvard University Press

Sveikauskas, Leo (1975), The Productivity of Cities, *Quarterly Journal of Economics*, 89, pp. 393-413

Tabuchi, Takatoshi (1986), Urban Agglomeration, Capital Augmenting Technology, and Labor Market Equilibrium, *Journal of Urban Economics*, 20, pp. 211-28

Teixeira, Antonio (2002), The geographical Distribution of Production Activity in Portugal and its Determinants, *mimeo* CORE, Université catholique de Louvain

Appendix

APPENDIX 1 - Percent of plants and jobs by NACEBEL 2-digit in 2000

NACE2	Percent plants (2000)	Percent jobs (2000)	NACE2	Percent plants (2000)	Percent jobs (2000)
1	2.49	0.86	41	0.06	0.21
2	0.09	0.02	45	10.6	5.66
5	0.05	0.02	50	3.9	1.64
10	0.01	0	51	8.45	5.23
14	0.07	0.12	52	14.38	5.85
15	2.55	2.53	55	8.33	3.17
16	0.01	0.08	60	2.41	3.1
17	0.5	1.21	61	0.06	0.06
18	0.29	0.31	62	0.05	0.36
19	0.05	0.06	63	1.21	1.26
20	0.35	0.34	64	0.6	2.36
21	0.12	0.5	65	1.81	2.64
22	0.98	0.94	66	0.14	0.78
23	0.01	0.11	67	1.48	0.41
24	0.29	2.06	70	3.04	0.54
25	0.27	0.79	71	0.39	0.18
26	0.49	0.96	72	1.02	1.09
27	0.1	1.12	73	0.15	0.36
28	1.48	1.76	74	8.61	8.39
29	0.51	1.24	75	2.77	10.93
30	0.01	0.01	80	3.07	10.18
31	0.19	0.72	85	5.61	10.96
32	0.04	0.57	90	0.26	0.43
33	0.24	0.19	91	3.13	1.09
34	0.16	1.63	92	2.3	1.6
35	0.08	0.52	93	2.78	0.73
36	0.76	0.69	95	0.9	0.09
37	0.11	0.1	98	0.02	0.55
40	0.08	0.56	99	0.1	0.09
			Total	100	100

APPENDIX 2 - Herfindahl, Ellison-Glaeser and Maurel-Sédillot indices for the 76 industries with E-G>0.05 in 2000

NACE4	Herfindahl index (Hk)	E-G_arr2000	MS_arr2000	E-G_arr1997	MS_arr1997
2310	0.985	5.895	-0.051	0.128	-0.052
1722	0.907	0.704	-0.051	0.698	-0.052
2330	0.222	0.658	0.698	0.510	0.536
2753	0.444	0.657	0.889	0.491	0.789
2653	0.874	0.627	-0.051	0.727	-0.052
2464	0.381	0.582	0.784	0.577	0.802
1724	0.408	0.567	0.573	0.647	0.651
2320	0.287	0.553	0.724	0.374	0.548
3621	0.574	0.420	0.521	0.513	0.573
2611	0.16	0.390	0.345	0.452	0.396
2872	0.264	0.376	0.525	0.290	0.428
2626	0.287	0.370	0.272	0.448	0.356
2960	0.444	0.338	0.329	0.360	0.351
2622	0.781	0.288	-0.048	0.172	0.027
1723	0.169	0.272	0.219	0.292	0.236
2020	0.074	0.230	0.189	0.204	0.162
1717	0.792	0.229	-0.050	0.015	-0.044
1772	0.019	0.203	0.158	0.234	0.186
1721	0.05	0.201	0.179	0.214	0.193
1714	0.068	0.199	0.189	0.165	0.152
3622	0.016	0.198	0.262	0.202	0.276
2752	0.205	0.183	0.164	0.160	0.138
2052	0.144	0.179	0.149	0.098	0.060
1542	0.613	0.176	0.099	0.257	0.110
1751	0.042	0.169	0.133	0.151	0.110
1753	0.754	0.169	-0.051	0.090	-0.052
2441	0.113	0.167	0.133	0.162	0.117
1715	0.133	0.166	0.100	0.122	0.079
2921	0.288	0.161	0.214	0.249	0.304
1591	0.283	0.161	0.199	0.283	0.304
2225	0.018	0.159	0.174	0.085	0.110
3511	0.054	0.155	0.194	0.142	0.185
3662	0.226	0.153	0.076	0.183	0.110
1725	0.102	0.146	0.083	0.108	0.048
1711	0.114	0.146	0.103	0.152	0.109

NACE4	Herfindahl index (Hk)	E-G_arr2000	MS_arr2000	E-G_arr1997	MS_arr1997
1760	0.207	0.146	0.059	0.196	0.097
1562	0.676	0.137	-0.050	0.149	-0.052
1810	0.567	0.137	-0.047	0.037	-0.010
1771	0.101	0.126	0.067	0.120	0.049
3162	0.12	0.125	0.081	-0.113	-0.027
2630	0.231	0.120	0.149	0.112	0.129
2613	0.264	0.117	0.021	0.072	0.002
1730	0.026	0.114	0.091	0.146	0.124
2954	0.167	0.113	0.073	0.082	0.033
2612	0.066	0.112	0.097	0.092	0.076
2414	0.072	0.106	0.169	0.113	0.187
2932	0.214	0.103	0.038	0.107	0.039
1552	0.138	0.100	0.096	0.137	0.134
3512	0.324	0.099	0.213	0.077	0.226
2651	0.139	0.098	0.052	0.112	0.057
2913	0.089	0.096	0.109	0.022	0.026
2442	0.08	0.094	0.086	0.083	0.069
2213	0.068	0.094	0.116	0.232	0.279
2212	0.034	0.091	0.108	0.046	0.055
1713	0.11	0.090	0.054	0.091	0.057
2662	0.302	0.080	0.177	-0.207	-0.052
1910	0.334	0.077	0.027	0.063	0.032
2615	0.757	0.076	-0.048	0.673	0.611
1572	0.197	0.075	0.014	0.008	-0.049
3161	0.421	0.074	-0.038	0.060	-0.045
1597	0.173	0.067	0.023	0.051	0.006
2914	0.153	0.065	0.056	0.076	0.067
1520	0.047	0.064	0.031	0.160	0.104
2461	0.376	0.063	-0.042	0.041	-0.042
3530	0.186	0.063	0.062	-0.001	-0.002
2710	0.147	0.063	0.067	0.040	0.041
1594	0.393	0.062	0.058	0.073	0.064
1754	0.038	0.062	0.037	0.060	0.023
2466	0.044	0.062	0.065	-0.006	-0.002
1716	0.464	0.060	0.041	0.061	0.026
3002	0.179	0.059	-0.002	0.085	0.021
1920	0.49	0.058	-0.024	0.040	-0.016

NACE4	Herfindahl index (Hk)	E-G_arr2000	MS_arr2000	E-G_arr1997	MS_arr1997
1823	0.1	0.055	0.017	0.103	0.062
2862	0.072	0.055	0.062	0.024	0.028
3630	0.059	0.052	0.031	0.040	0.008
2223	0.055	0.051	0.053	0.009	0.028
2214	0.134	0.050	0.036	0.077	0.086

Spearman	E-G (1997)	E-G (2000)	MS (1997)	MS (2000)
E-G (1997)	1	/	/	/
E-G (2000)	0.7648	1	/	/
MS (1997)	0.6871	0.4702	1	/
MS (2000)	0.5016	0.5805	0.8152	1

Correlation	E-G (1997)	E-G (2000)	MS (1997)	MS (2000)
E-G (1997)	1	/	/	/
E-G (2000)	0.2456	1	/	/
MS (1997)	0.7038	0.138	1	/
MS (2000)	0.5001	0.2091	0.7921	1

APPENDIX 3 - 4-digit classification among 2-digit industries

NACE2	Name	E-G 2000	NACE2 rank	NACE4				
				Obs.	E-G≤0	0<E-G ≤0.02	0.02<E-G ≤0.05	E-G>0.05
15	Food industry	0.004	18	31	19.4%	25.8%	25.8%	29.0%
16	Tobacco industry	0.008	14	1	0.0%	100.0%	0.0%	0.0%
17	Textile industry	0.103	2	21	0.0%	4.8%	9.6%	85.7%
18	Clothes and fur industry	0.026	4	6	0.0%	16.7%	50.0%	33.3%
19	Leather and footwear industry	0.016	8	3	0.0%	33.3%	0.0%	66.7%
20	Wood, cork, basket and esparto industry	0.021	6	6	0.0%	50.0%	16.7%	33.3%
21	Paper and paperboard industry	0.011	11	6	0.0%	33.3%	66.7%	0.0%
22	Publishing, printing and reproduction	0.024	5	13	30.8%	7.7%	23.1%	38.5%
23	Coking, refining and nuclear industries	0.219	1	3	0.0%	0.0%	0.0%	100.0%
24	Chemical industry	0.02	7	20	35.0%	5.0%	30.0%	30.0%
25	Rubber and plastic industry	0.012	10	7	0.0%	28.6%	71.4%	0.0%
26	Production of other non-metallic minerals	0.007	15	25	28.0%	12.0%	4.0%	56.0%
27	Metallurgy	0.012	9	17	29.4%	29.4%	5.9%	35.3%
28	Metal transformation	0.005	17	16	25.0%	50.0%	12.5%	12.8%
29	Production of machines and equipment	0.003	19	20	40.0%	15.0%	10.0%	35.0%
30	Production of office machines and computer materials	0.056	3	2	50.0%	0.0%	0.0%	50.0%
31	Production of electrical machines and apparatus	-0.005	21	7	28.6%	42.9%	0.0%	28.6%
32	Production of radio, broadcast and communication equipment	-0.023	23	3	0.0%	100.0%	0.0%	0.0%
33	Production of medical, precision, optical and clock instruments	0.01	12	5	40.0%	40.0%	20.0%	0.0%
34	Production and assembling of motor vehicle, tow and semi-trailer	-0.022	22	3	33.3%	66.7%	0.0%	0.0%
35	Production of other transport materials	0.009	13	8	37.5%	12.5%	0.0%	50.0%
36	Furniture production; various industries	0.007	16	13	30.8%	38.5%	0.0%	30.8%
37	Recycled material recovery	0.002	20	2	50.0%	0.0%	50.0%	0.0%

APPENDIX 4 - Moran's I index of spatial autocorrelation (townships) for the 76 industries with E-G>0.05 in 2000

NACE4	EG_arr2000	Moran I (2000)	p-value (2000)	Moran I (1997)	p-value (1997)
2310	5.8952	-0.0014	0.7827	-0.0041	0.8854
1722	0.704	-0.0028	0.6075	-0.0028	0.615
2330	0.6575	0.0415	0	0.0231	0
2753	0.6568	0.0088	0	0.0058	0
2653	0.6267	-0.0016	0.9563	-0.0016	0.9517
2464	0.5818	0.1992	0	0.208	0
1724	0.5673	0.0286	0	0.0286	0
2320	0.5526	-0.0017	0.9861	-0.0017	0.9949
3621	0.4195	0.0125	0.0004	0.0659	0
2611	0.3895	0.1386	0	0.1415	0
2872	0.3755	-0.0121	0.4798	-0.0134	0.5026
2626	0.3701	-0.0051	0.6858	0.0156	0.0008
2960	0.3379	0.003	0.5394	0.0029	0.4664
2622	0.2878	-0.0018	0.9815	-0.0033	0.9255
1723	0.2719	-0.0043	0.8779	-0.005	0.8368
2020	0.2299	0.1485	0	0.1583	0
1717	0.2286	0.006	0.0126	0.0047	0.5758
1772	0.2029	0.0864	0	0.0695	0
1721	0.2014	0.1658	0	0.1797	0
1714	0.1989	0.0773	0	0.0475	0.0143
3622	0.1981	0.0685	0	0.0596	0
2752	0.1831	0.0339	0.0111	0.0113	0.2877
2052	0.1793	0.0027	0.8312	0.0036	0.7988
1542	0.1755	-0.0011	0.9295	-0.0021	0.9174
1751	0.1692	0.2966	0	0.2773	0
1753	0.1687	-0.0027	0.8602	-0.0031	0.8569
2441	0.1671	0.0285	0.1527	0.0336	0.0784
1715	0.1659	0.0011	0.8859	-0.0046	0.8891
2921	0.161	-0.0068	0.7226	-0.0063	0.6968
1591	0.1609	0.0956	0	0.0894	0
2225	0.1588	0.3151	0	0.1196	0
3511	0.1553	0.0875	0	0.0852	0
3662	0.1526	0.0066	0.644	0.0196	0.2382
1725	0.1462	0.1083	0	0.0886	0
1711	0.146	0.1284	0	0.1079	0
1760	0.1455	0.0267	0.1829	-0.0067	0.7961
1562	0.1368	-0.0052	0.4799	-0.005	0.4626
1810	0.1368	-0.0044	0.6376	-0.0153	0.4461
1771	0.126	-0.0138	0.5325	-0.009	0.6009
3162	0.1246	0.1293	0	0.0008	0.7987

NACE4	EG_arr2000	Moran I (2000)	p-value (2000)	Moran I (1997)	p-value (1997)
2630	0.1197	0.0819	0	0.0934	0
2613	0.1166	0.0316	0.0676	0.0488	0.0101
1730	0.1138	0.2338	0	0.2784	0
2954	0.1128	0.0382	0.0201	0.0154	0.2577
2612	0.112	0.0771	0	0.0711	0.0004
2414	0.1059	0.0496	0	0.048	0
2932	0.103	0.0253	0	0.0239	0
1552	0.0997	-0.0033	0.9375	-0.0069	0.7999
3512	0.099	-0.0035	0.8632	-0.0049	0.673
2651	0.0976	0.0233	0.2228	0.0233	0.2144
2913	0.0963	0.0331	0.0218	0.0042	0.732
2442	0.0942	0.0118	0.5197	0.0131	0.4668
2213	0.0939	0.1091	0	0.1167	0
2212	0.0911	0.1191	0	0.1292	0
1713	0.0904	-0.0093	0.7134	-0.0079	0.7549
2662	0.0796	0.086	0	-0.0022	0.9672
1910	0.077	0.0216	0.1554	0.0443	0.0105
2615	0.0758	-0.0019	0.9203	-0.0018	0.9675
1572	0.0747	-0.0117	0.5663	-0.0094	0.6622
3161	0.0744	-0.0022	0.944	-0.0049	0.844
1597	0.0673	-0.0129	0.6005	-0.0159	0.5084
2914	0.0645	0.0699	0.0004	0.0854	0
1520	0.0643	-0.0052	0.868	0.029	0.0837
2461	0.0632	-0.0046	0.8405	-0.0063	0.7912
3530	0.0631	0.0574	0.0026	0.0615	0.0014
2710	0.063	0.0176	0.3008	0.0015	0.8698
1594	0.0624	0.0256	0.0902	0.0263	0.0652
1754	0.0622	0.0432	0.0294	0.0209	0.2874
2466	0.0615	0.0524	0.0043	0.0155	0.4467
1716	0.0602	0.0156	0.1275	0.0316	0.041
3002	0.0586	0.0239	0.1835	0.0413	0.0216
1920	0.0579	-0.0038	0.6912	-0.0037	0.8346
1823	0.0554	0.0407	0.014	0.044	0.0035
2862	0.0545	0.0014	0.7971	0.0003	0.8436
3630	0.0524	-0.0051	0.8782	0.0115	0.5698
2223	0.0511	0.0452	0.0193	-0.0011	0.978

APPENDIX 5 - Belgium and its three regions (the thin strokes indicate the limits of the districts)