IDENTIFYING THE GEOGRAPHICAL AGGLOMERATIONS OF MANUFACTURING INDUSTRIES

Giovanni Iuzzolino[*]

1. Introduction and main conclusions

The tendency of manufacturing firms to cluster together has long been observed, and the theoretical discussion of the advantages of producing in an "industrial district" also dates far back.[1]

From the textiles and metallurgic agglomerations observed by Marshall in England at the end of the nineteenth century, to the industrial districts of the North and Centre of Italy in the 1970s and to the modern clusters of small or large high-tech firms in California or Texas, over the past decades different patterns of agglomeration have had great success in various institutional and technological contexts.

The presence of localized industrial aggregation, also from the viewpoint of regional and country growth theories,[2] can be therefore regarded as a potential competitive advantage for regions or nations.

However, while there is an abundance of qualitative analysis of the clustering of individual industries and the way in which spatial concentration changes into economies of agglomeration,[3] for a long time it was difficult to undertake quantitative work on the same subject because of the lack of any econometrically operational definition of industrial agglomerations.

It has only recently become possible, thanks to the work of Ellison and Glaeser,[4] to measure the geographical concentration of production in a statistically correct way, excluding from the study of agglomeration cases

---

* Bank of Italy, Naples branch, regional research unit. I am grateful to Luigi Cannari, Massimo Omiccioli, Guido Pellegrini and Luigi Federico Signorini for their suggestions and comments. I retain responsibility for remaining errors.

1 Marshall, 1890.

2 Pellegrini, 2000.

3 For theoretical work on these subjects see Fujita *et al.* 1999.

4 Ellison and Glaeser, 1997.

of industrial concentration that are due to the random distribution of a small number of large plants.

Following Ellison and Glaeser, many measures of agglomeration have been recently proposed.[5] The empirical evidence provided by this literature confirms that most industries are, in a large number of countries, more intensely clustered than a model of random spatial distribution of firms would predict.

If industrial agglomeration is more the rule than the exception, and if we can now measure its intensity in a more rigorous way, then it is natural to ask whether it is possible to draw a map of agglomerations based on these new measures. This is the main focus of this paper.

We believe that such a map represents a necessary instrument for carrying out a reliable econometric analysis of the existence and intensity of agglomeration advantages.

Indeed, whether such advantages actually exist, how large they are, and how they change with differences in the structural features of agglomerations, are all empirical questions that require the use of econometric techniques to compare, for instance, the performance of agglomerated areas with alternative forms of organization of the productive processes. In this regard a crucial role is played by the availability of a map through which one can count how many industrial agglomerations exist in a country/region, see where they are located and control for their structural features.

In the next two paragraphs, starting from the work of Ellison and Glaeser, we derive an algorithm that can be used to select and map industrial agglomerations. We propose, specifically, a test of the industrial specialization of an area based on the null hypothesis of absence of agglomeration advantages: in this way we can select the areas where such advantages are most likely to exist, and build around them a set of neighbouring areas that represent the space of probable diffusion of proximity advantages.

Then (par. 4) we apply the algorithm to the geographical distribution of manufacturing activities in Italy and give a description of the main

---

5 Maurel and Sedillot, 1999; Devereux *et al.*, 1999; Duranton and Overmars, 2002. See Pagnini (in this volume) for a survey.

features of agglomerated areas in the country. This enables us, among other things, to verify the high variability that exists among Italian agglomerations concerning size, specialization and presence of large firms. The considerable dissimilarities found among agglomerations suggest that, in the econometric analysis of the advantages of producing in an industrial agglomeration, a strictly dichotomous approach can be misleading.

Finally (par. 5), we discuss the conceptual differences between the industrial agglomerations obtained with our algorithm and the Marshallian notion of "industrial district", which represents a key analytical category in the literature on local development in Italy.

2. Raw geographical concentration and proximity advantages: in search of an operational definition of industrial agglomeration

Geographical concentration and agglomeration phenomena. – Until a few years ago, the geographical concentration of production was measured by comparing the share of employment in a given industry across regions with the share of aggregate employment in the same regions.[6]

Writing z^{ip} for the number of employees in the i^{th} region and in the p^{th} industry, and Z^p for the employment in the whole country in the same industry, a frequently used index of geographical concentration is given by:

$$G = \sum_{i=1}^{n} \left(s_i^{\,p} - x_i\right)^2 \text{, with } s_i^{\,p} = \frac{z_i^{\,p}}{Z^p} \text{ and } x_i = \frac{\sum_p z_i^{\,p}}{\sum_p Z^p} .$$

However, the G index, as well as all the indexes that summarize inequality and concentration, is not a good measure of agglomeration, that is, of the tendency of firms to locate near each other in a particular area. Indeed, the value of G depends not only on the presence of agglomeration advantages, but also on the inequality of employee distribution among plants.

In particular, in the random location choice model proposed by Ellison and Glaeser the expected value of the G index is:

6 Krugman, 1991.

$$E(G)=\left(1-\sum_{i=1}^{n}x_i^{\,2}\right)[\gamma+(1-\gamma)H]\ ,$$

where H is the Herfindahl index of the distribution of the industry plant size, and γ is a parameter that is positively correlated with the intensity of agglomerating forces in the industry.[7] If such forces are negligible (γ=0), we have:

$$\tilde{G}^{\gamma=0} \sim \Phi(\mu,\sigma^2);\ \mu=(1-\sum x_i^2)H>0\ .$$

Therefore G is not a correct measure of agglomeration, because a spatial random distribution of a few plants may lead to a high value of G (through the H effect) even if agglomerating forces are not in action. Thus G can be viewed as a raw concentration index that must be corrected for the share of concentration that would be expected to arise randomly.

In this paper we propose an algorithm to find industrial agglomerations based on this intuition of Ellison and Glaeser.[8]

As a first step, where no agglomeration advantages exist, we can split the random variable G into mutually independent local components ($\tilde{G}^{\gamma=0}=\sum_{i=1}^{n}\tilde{G}_i^{\gamma=0}$) so that:

$$\tilde{G}_i^{\gamma=0} \sim \Phi(\mu_i,\sigma^2{}_i) \text{ with } \mu=\sum_{i=1}^{n}\mu_i\,;\sigma^2=\sum_{i=1}^{n}\sigma^2{}_i\ .$$

In such a way we build an analytical link between the geographical concentration of an industry and the industrial specialization of a region. G_i can be viewed as being both the contribution of the i^{th} region to the "raw

---

[7] Ellison and Glaeser, 1997. In the Ellison-Glaeser model, in which a firm's profits are affected (also) by the characteristics of the localization area, we have: $\gamma = \gamma^n + \gamma^s - \gamma^n\gamma^s$, where γ^n captures the importance of natural advantages to the industry and γ^s represents the probability that valuable spillovers (that is, any kind of benefit generated by proximity) exist between pairs of firms located near other firms in the same industry.

[8] See Iuzzolino, 2004, for details.

concentration” in the industry, and the degree of “raw specialization” of this region in that industry.

Available census data allow us to estimate for each area (down to the level of the municipality) both G_i and the mean and variance parameters; we can therefore test if the level of raw specialization of any local area is statistically consistent with the absence of agglomeration economies.

In particular, the basic step of the algorithm we propose is the selection of the areas where the difference between raw specialization (G_i) and the sample mean, under the null of γ_i=0, is larger than twice its standard deviation:

(1a) $G_i > \mu_i + 2 * \sigma_i$.

The ratio of the two members of (1a),

(1b) $$\frac{G_i}{\mu_i + 2 * \sigma_i},$$

can be used as a measure of the intensity of the agglomerating forces in each area.[9]

Since $\sum_i s_i^2 h_i = H$,[10] where h_i is the Herfindahl index of plant employment shares in the i^{th} area, a few calculations show that (1a) corresponds to:[11]

---

9 Of course the threshold on the right side of (1a) is somewhat arbitrary, but we can consider it a minimum level under which agglomeration is unlikely to be found. Then we can concentrate our attention on the most agglomerated areas, that is, on those with high values of (1b).

10 Writing z_j^i for the number of employees in the j^{th} plant located in the i^{th} area, and k_i for the number of plants in the same area, we have:

$$H = \frac{\sum_{j=1}^{k1} (z_j^1)^2 + \ldots + \sum_{j=ki-1+1}^{ki} (z_j^i)^2 + \ldots + \sum_{j=kn-1+1}^{m} (z_j^n)^2}{Z^2}$$

Then multiplying and dividing each addendum by (Z_i^2) we have:

$$H = \sum_{j=1}^{k1} \frac{(z_j^1)^2}{Z_1^2}\left(\frac{Z_1^2}{Z^2}\right) + \ldots + \sum_{j=ki-1+1}^{ki} \frac{(z_j^i)^2}{Z_i^2}\left(\frac{Z_i^2}{Z^2}\right) + \ldots + \sum_{j=kn-1+1}^{m} \frac{(z_j^n)^2}{Z_n^2}\left(\frac{Z_n^2}{Z^2}\right)$$ and finally:

(continues)

$$(s_i - x_i)^2 > s_i^2 h_i \left(1 - \sum_{i=1}^{n} x_i^2\right) + 2\left\{s_i^2 h_i Hk - s_i^4 \sum_{j=1}^{k_i} \frac{z_{ij}^4}{Z_i^4} y\right\}$$

It is interesting to note that, raw specialization being equal (that is, holding the value of s_i and x_i fixed), the difference between the two sides of inequality (1a) depends on the h_i's and therefore on the number and relative size of the firms located in the area. Therefore, sorting the areas by descending levels of (1b), we expect to find at the top of the list those areas whose specialization is due to the presence of a large number of homogeneous firms: a condition that fits well with the intuitive notion of agglomeration.

Agglomeration points and agglomerated areas. – With (1a) we can select the municipalities with the strongest industrial specialization, and we assume that these can be ascribed to agglomeration advantages. The next question is how to build up, around those municipalities, an agglomerated area (AA) that could approximate the space of probable diffusion of such advantages.

Cross-border spillover mechanisms from an agglomeration point are likely to determine relatively high levels of specialization (at least in the sense of $\gamma_i>0 \Leftrightarrow G_i > \mu_i$) in neighbouring areas. Denoting with $d(i,j)$ a dichotomic variable that equals zero if the i^{th} and j^{th} areas are neighbouring, an AA can be defined by:

(2) $$AA = \bigcup_{i=1}^{s} i : \gamma_i > 0 \text{ and } \exists\, j \in \text{AA}: \forall i,\ d(i,j)=0.$$

In this way, an AA expands until it borders only with non-specialized areas. Such a property can be interpreted as an example of the process of spatial correlation that is common to several economic

---

$$H = h_1 s_1^2 + \ldots + h_i s_i^2 + \ldots + h_n s_n^2 = \sum_i h_i s_i^2 .$$

11 The variables k and y are defined as follows:

$$k = 2\left[\sum_{i=1}^{n} x_i^2 - 2\sum_{i=1}^{n} x_i^3 + \left(\sum_{i=1}^{n} x_i^2\right)^2\right]; y = 2\left[\sum_{i=1}^{n} x_i^2 - 4\sum_{i=1}^{n} x_i^3 + 3\left(\sum_{i=1}^{n} x_i^2\right)^2\right]$$

phenomena. An area with a large value of G_i (beyond the (1a) threshold) may represent a centre of agglomeration, so that firms will locate their plants in the neighbourhood. Therefore, the likelihood of finding other specialized areas around the "centre" will be high. Moreover, if we suppose that the intensity of agglomeration advantages falls off with distance, this likelihood will decrease as we move away from the "centre". The presence of non-specialized areas that enclose the AA will then signal the exhaustion of localization advantages.

After we have built the most agglomerated area, starting from the municipality with the maximum value of G_i fulfilling the condition (1a), we look for other possible agglomerations, repeating the process up to a complete exploration of the country. The outcome of this process will be a partition of the country into three sets: non-specialized areas (γ_i<0), weakly specialized areas (γ_i>0, but $i \notin$AA) and strongly specialized areas (γ_i>0 and $i \in$AA).

In conclusion, from (1a) and (2) we can define an AA as a continuum of municipalities, all specialized in a given industry, at least one of which significantly exceeds the degree of specialization expected under the null hypothesis of absence of agglomeration advantages.

3. Agglomeration and co-agglomeration phenomena: which groups of industries are eligible?

In the previous paragraph we propose a criterion for defining the geography of agglomerations in a given industry. But we have yet to define the boundaries of such an industry, that is, the exact definition of the industries for which we will be seeking agglomeration phenomena.

In effect, if we suppose that agglomerative advantages are narrowly industry-specific, then the search for AAs will be reduced to the most narrowly-defined industries. But this condition would appear too restrictive: we can imagine many circumstances in which plants try to locate near other plants in related but not identical lines of business. For example, Ellison and Glaeser find evidence that "*industries ... appear to* *coagglomerate both with important upstream suppliers and with important downstream customers*".[12]

Of course, the selection of such groups of industries is a very complicated matter, partly because of insufficient industrial detail in the available data. To define a partition of manufacturing activities suitable for our purpose we have therefore adopted an empirical and partially data-driven solution.

The first step in our construction relies on data from the input-output matrix of Italian manufacturing activities. We use these data, which include detailed information on flows of manufactures built and exchanged among 49 industries, to determine, as far as possible, the main *filières* that characterize the industrial structure of the country.

By looking at the ratio between flows of intermediate goods exchanged within each industry and the value of all intermediate goods used by the same industry, one sees that the 49 branches of activity are in fact not very "self-contained", as the median value of the ratio is only equal to 23.7 per cent. We then aggregate some branches, essentially moving from the n-digit to the (n–1)-digit level, and including in the same aggregate all the industries (such as specialized machinery) that show clear upstream-downstream relationships. In this way we identify 7 aggregates of industries with a share of intermediate goods supplied and used inside the same aggregate equal to or greater than two-thirds (Table 1).

We then proceed to look, within each of the 7 aggregates, for strong locational complementarities. Local interdependencies can occur not only inside a *filière* relationship, i.e. between phases inside the same production process, but also for many other reasons, such as the use of a common technology or a specialized input. The second step in our construction is therefore a data-driven search for such local interdependencies between industries belonging to the same aggregate.

In other words, since the available data do not allow us to group all the industries that are in some way complementary or interdependent in any way, we suppose that a similarity in the geographical distribution of firms must be the outcome of such interdependence.

---

[12] Ellison and Glaeser, 1997, p. 892.

Table 1

Flows of manufactures used as intermediate goods by industries
(percentages)

Purchases by (aggregate of industries)	Sales by: The same aggregate	Sales by: Other industries	Total
I Food and beverages	69.5	30.5	100
II Textiles, clothing, leather and footwear	77.1	22.9	100
III Paper, printing and publishing	73.4	26.6	100
IV Wood, furniture and non-metallic mineral products	70.0	30.0	100
V Petroleum, chemicals, pharmaceuticals and rubber products	65.5	34.5	100
VI Metal goods, mechanical, electrical and electronic engineering	82.2	17.8	100
VII Transport equipment	64.6	35.4	100

Source: Based on Istat, input-output matrix 1992.

For this purpose, and within each of the 7 aggregates, we build a matrix of "locational coefficients", where the generic element a_{ij} represents the share of employers of the i^{th} industry (at 5-digit level) in the j^{th} Italian municipality (j: 1...8.101). Then, carrying out a cluster analysis of these coefficients, we identify 16 clusters of industries that represent homogenous sectors. These we use to find industrial agglomerations.[13]

In Table 2 we report the main features of these clusters in terms of size and geographical concentration. It should be noted that, moving from the raw concentration index G to the agglomeration index, the ranking of industries changes significantly. For instance, the transport equipment industry moves from first to thirteenth position, while clothing and furniture move in the opposite direction.

[13] Devereux *et al.* (1999) suggest that Ellison and Glaeser's γ index overestimates the intensity of agglomeration in an industry when the number of plants is less than the number of areas in which the plants might be located. We have therefore selected industrial clusters with a number of plants equal to or greater than the number of Italian municipalities.

Table 2

Size and geographical concentration of the 16 industry clusters

	Plants	Employment	Concentration and agglomeration indexes	
			G	$\frac{G}{\mu + 2 * \sigma}$
I.1 Beverages, milk-based products, pasta and confectionery	62,297	329,538	1.8	3.5
I.2 Tinned food, prepared meats and machinery for food industry	14,648	137,470	6.3	5.5
II.1 Textile products and related machinery	18,150	234,363	14.0	22.3
II.2 Clothing industry	57,328	399,274	3.6	16.2
II.3 Knitwear	19,714	133,588	6.9	14.8
II.4 Leather, footwear and related machinery	18,668	196,003	13.4	31.1
III Paper, printing, publishing and related machinery	33,924	269,132	2.6	4.3
IV.1 Cement and glass products	16,176	135,730	5.1	4.8
IV.2 Wood, furniture and related machinery	91,406	396,885	4.6	36.3
IV.3 Jewels, musical instruments and toys	17,866	90,048	13.2	20.5
IV.4 Ceramics and stone products; related machinery	17,391	140,039	10.9	12.3
V.1 Chemicals and petrochemical products; related machinery	8,958	236,925	6.3	3.3
V.2 Rubber, plastic and related machinery	15,921	211,656	4.6	9.0
VI.1 Electronic and electrical equipment, machine tools	109,452	775,951	1.2	3.4
VI.2 Primary metal industries and industrial machinery	80,856	838,859	4.2	12.0
VII Transport equipment	8,355	330,316	20.5	4.2

Source: Based on Istat, *Censimento intermedio dell'industria e dei servizi* 1996.

4. Patterns of geographical concentration in Italy

We have seen that with our algorithm three classes of areas can be identified: non-specialized (γ_i<0), weakly specialized (γ_i>0, but $i \notin$AAs), and agglomerated or strongly specialized (γ_i>0 and $i \in$AAs).

In a given country and at a given time, the relative weight of these classes depend on the localization processes followed by firms over time and, with respect to agglomerated areas, on the intensity of centripetal forces that have arisen from the most specialized points because of some kinds of proximity advantages.

Thus, if we agree with the Ellison-Glaeser model and go back from the effects to the causes, a region or a country with many agglomerated areas can be viewed as a place in which peculiar sources of competitiveness are operating or have operated.

We are therefore interested in building a map that would enable us to see how many agglomerations there are in a country, where they are located and what are their dimensions and structural features. We present the map of Italian industrial agglomeration built using the 1996 census data at the municipality level.

Applying our algorithm to each of the 16 clusters of industries, we identified 156 industrial agglomerations spread over most of Italy (Figure 1): the municipalities belonging to AAs are 2,209, or 29 per cent. The share of manufacturing employment in the AAs is a similar: of the 4.9 million manufacturing employees in the country, 2.5 million are located in non-specialized municipalities, 0.9 million in weakly specialized municipalities and 1.5 million (30.6 per cent of total employment) in agglomerated areas. While it is difficult to say whether this figure is high or low because of the lack of comparable data from other countries, we can observe that the same figure varies across regions or industrial clusters.

The spatial variance of agglomeration is indeed high (Table 3). In 8 regions out of 20, less than 10 per cent of manufacturing employment is in industrial agglomerations, while in 5 regions (Piedmont, Lombardy, Veneto, Marche and Tuscany) the share is greater than one third. In the latter regions 68 per cent of total agglomeration employment is concentrated: the same regions (which are among the richest in Italy) account for 58 per cent of total industrial employment and for less than 40 per cent of the population.

By contrast, the less developed regions, those of southern Italy, with over 36 per cent of the population and 15 per cent of manufacturing employment, account for only 8 per cent of employment in industrial agglomerations.

The geographical distribution of employment in the weakly specialized areas is much more balanced, with a share of about 19 per cent in both the North and the South of Italy. The latter is the only area where employment in weakly specialized municipalities exceeds that of AAs, pointing to the diffusion of cases of industrial specialization that still have a low intensity in comparison with the rest of the country.

Figure 1

Italian municipalities belonging to areas of industrial agglomeration in 1996

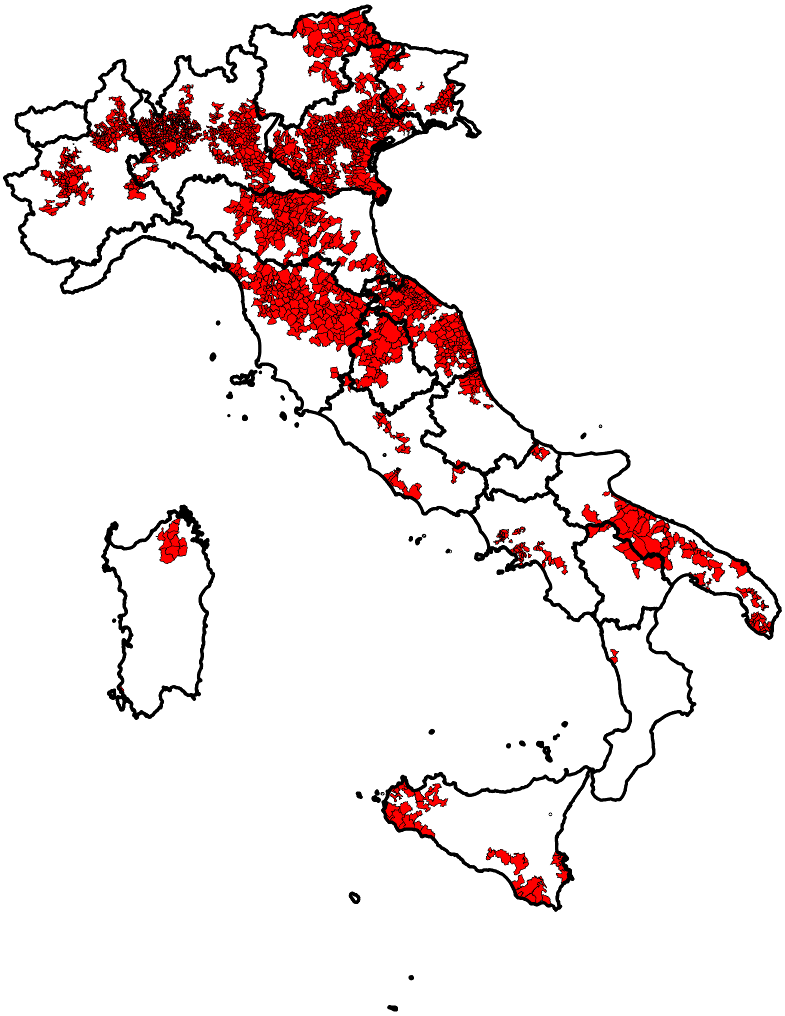

Table 3

Size and geographical distribution of industrial agglomerations in Italy

(units and percentages)

	No. of AAs (1)	Munici-palities in AAs	Employment in AAs		Employment in weakly specialized municipalities	
			Units	Share of employ-ment in the area	Units	Share of employ-ment in the area
North-West	**41**	**872**	**620,559**	**32.8**	**361,918**	**19.1**
Liguria	0	2	462	0.6	12,089	14.5
Lombardy	31	664	443,071	34.8	245,289	19.3
Piedmont	10	206	177,026	33.2	103,090	19.4
Valle d'Aosta	0	0	0	0.0	1,450	23.6
North-East	**43**	**714**	**414,154**	**30.4**	**264,101**	**19.4**
Emilia Romagna	14	157	151,805	29.6	93,176	18.2
Friuli	4	73	35,283	27.8	25,474	20.1
Trentino	1	51	6,388	8.9	16,943	23.7
Veneto	24	433	220,678	34.0	128,508	19.8
Centre	**39**	**394**	**324,902**	**37.9**	**94,496**	**11.0**
Lazio	5	23	47,164	21.8	20,878	9.7
Marche	10	181	92,901	48.4	26,692	13.9
Tuscany	19	156	170,795	45.0	36,780	9.7
Umbria	5	34	14,042	20.1	10,146	14.5
South and Islands	**33**	**229**	**127,094**	**17.1**	**143,438**	**19.3**
Abruzzo	3	59	18,719	17.9	25,674	24.5
Basilicata	0	4	1,584	6.0	9,899	37.4
Calabria	1	2	272	0.8	5,739	16.2
Campania	11	56	41,345	19.3	34,859	16.2
Molise	1	3	741	4.4	3,416	20.2
Puglia	11	73	54,491	30.2	29,781	16.5
Sardinia	1	9	1,165	2.3	14,884	29.2
Sicily	5	23	8,777	7.6	19,186	16.7
Italy	**156**	**2,209**	**1,486,709**	**30.6**	**863,953**	**17.8**

Source: Based on Istat, *Censimento intermedio dell'industria e dei servizi 1996*.

(1) Number of municipalities that are "central areas" of AAs.

Table 4

Size and sectoral distribution of industrial agglomerations in Italy

(units and percentages)

	No. of AAs (1)	Employment in AAs		Employment in weakly specialized municipalities	
		Units	Share of employ-ment in the cluster	Units	Share of employ-ment in the cluster
I.1 Beverages, milk-based products, pasta and confectionery	1	2,053	0.6	46,119	14.0
I.2 Tinned food, prepared meats and machinery for food industry	6	16,105	11.7	29,020	21.1
II.1 Textile products and related machinery	7	130,742	55.8	27,406	11.7
II.2 Clothing industry	22	178,460	44.7	50,882	12.7
II.3 Knitwear	13	58,219	43.6	27,400	20.5
II.4 Leather, footwear and related machinery	21	151,909	77.5	9,090	4.6
III Paper, printing, publishing and related machinery	6	84,626	31.4	45,973	17.1
IV.1 Cement and glass products	6	15,043	11.1	38,660	28.5
IV.2 Wood, furniture and related machinery	15	164,079	41.3	48,203	12.1
IV.3 Jewels, musical instruments and toys	14	41,225	45.8	6,755	7.5
IV.4 Ceramics and stone products; related machinery	14	53,713	38.4	22,123	15.8
V.1 Chemicals and petrochemical products; related machinery	4	69,162	29.2	50,612	21.4
V.2 Rubber, plastic and related machinery	8	46,882	22.2	64,013	30.2
VI.1 Electronic and electrical equipment, machine tools	7	97,350	12.5	186,381	24.0
VI.2 Primary metal industries and industrial machinery	10	290,899	34.7	152,385	18.2
VII Transport equipment	2	86,242	26.1	58,931	17.8
Total	**156**	**1,486,709**	**30.6**	**863,953**	**17.8**

Source: Based on Istat, *Censimento intermedio dell'industria e dei servizi* 1996.

(1) Number of AAs specialized in the cluster.

With regard to the different clusters of industries (Table 4), the incidence of agglomerations in total employment largely exceeds 40 per cent in clusters belonging to the fashion sector (up to 77.5 per cent in the leather and footwear industries) and to the wood, furniture and jewellery industries. The density of agglomerated areas in ceramic goods and in the metal and engineering sector is also high (38.4 and 34.7 per cent, respectively), while it is very low in the food and electronic industries.

Dissimilarities inside the agglomerated areas. – Differences in agglomeration intensity among regions or industries are useful to describe the pattern of industrial concentration. Another interesting question is whether or not the agglomerated areas are similar in their structural features.

In fact, the 156 agglomerations we have found turn out to be quite varied (Table 5). First, they have very different sizes: in terms of number of employees, around a median value of 3,905 we have an inter-quartile distance of 7,733. In terms of extensions of the AAs, we found 39 “small agglomerations” (in the lower quartile) with 4 or less municipalities and 39 “large agglomerations” which include 21 or more (up to 138) municipalities.

Table 5

Structural features of industrial agglomerations in Italy
(units and percentages)

Percentiles	Employment	Number of municipalities	Extension (1)	Industrial diversification (2)	Share of large firms (3)
0 (minimum)	259	1	0.0	11.8	0.0
15	986	2	7.0	20.0	0.0
25	1,366	4	10.5	28.6	0.0
50 (median)	3,905	9	20.4	40.6	0.0
75	9,099	21	39.0	66.7	15.6
85	16,663	40	45.7	78.8	31.0
100 (maximum)	72,693	138	207.5	100.0	99.1

Source: Based on Istat, *Censimento intermedio dell’industria e dei servizi* 1996.
(1) Distance in kilometres between the farthest municipalities in the AAs.
(2) Number of specializations in the AAs on the whole of the industries in the cluster.
(3) Share of employment concentrated in plants with 250 or more employees.

Table 6

Intensity of agglomerations in AAs

	Main municipality	Area	Employment	Value of (1b) index: $\frac{G_i}{\mu_i + 2 * \sigma_i}$
		10 most agglomerated areas		
Textile products and related machinery	Prato	Centre	39,426	224.5
Jewels, musical instruments and toys	Valenza	North-West	6,807	144.6
Leather, footwear and related machinery	Santa Croce sull'Arno	Centre	32,960	77.0
Leather, footwear and related machinery	Porto Sant'Elpidio	Centre	40,234	72.2
Jewels, musical instruments and toys	Vicenza	North-East	8,124	56.8
Knitwear	Carpi - Novi di Modena	North-East	11,558	47.1
Jewels, musical instruments and toys	Arezzo	Centre	10,101	43.0
Wood, furniture and related machinery	San Giovanni al Natisone	North-East	15,321	37.6
Clothing industry	Carpi - Reggiolo	North-East	8,627	37.6
Leather, footwear and related machinery	Vigevano	North-West	6,649	33.1
		10 least agglomerated areas		
Rubber, plastic and related machinery	Oderzo	North-East	2,881	1.1
Rubber, plastic and related machinery	Ciserano	North-West	5,304	1.1
Ceramics and stone products; related machinery	Caltagirone	South and Islands	259	1.1
Clothing industry	Como	North-West	3,231	1.0
Electronic and electric equipment, machine tools	Marcianise	South and Islands	5,324	1.0
Electronic and electric equipment, machine tools	Longarone	North-East	8,308	1.0
Ceramics and stone products; related machinery	Rezzato	North-West	765	1.0
Rubber, plastic and related machinery	Battipaglia	South and Islands	983	1.0
Tinned food, prepared meats and machinery for food industry	Cremona	North-West	1,285	1.0
Wood, furniture and related machinery	Mosciano Sant'Angelo	South and Islands	773	1.0

Second, we note considerable dissimilarities in the degree of industrial diversification. In this regard, consider that the 16 clusters of industries include a median number of 6 sub-industries at the 3-digit level: we observe that the share of sub-industries in which AAs are strongly specialized[14] varies considerably, from a minimum of about 12 per cent to a maximum of 100 per cent.

Finally, industrial agglomerations are very different with regard to the presence of large firms. The share of employment in firms with 250 or more employees is very low (between 0 and 16 per cent) for three-quarters of AAs but exceeds 30 per cent in the areas above the 85th percentile.

Moreover, we must stress the strong variability of AAs with respect to the value of the index (1b), which is a measure of the agglomeration intensity that summarizes many characteristics of spatial concentration (extreme values are reported in Table 6). In 60 AAs out of 156, the value of the index is less than 2; in 31 cases this value is greater than 10, that is, the raw concentration in these areas is more than ten times the threshold (1a).

To show that such an index is correlated with the main features of the AAs, we report in Table 7 the median value of the index calculated in a number of structurally different agglomerated areas. It is clear that the intensity of agglomeration declines with the share of employment in large firms and rises with the size and the industrial diversification of AAs.

We feel that all these differences suggest an analytical approach that takes appropriate account of them. Many papers try to evaluate the significance of an "agglomeration effect" (in firms' performances or in the operation of some markets) by estimating equations with a simple "agglomerated area" dummy. But if we agree that different typologies of agglomeration can account for very different sources (and intensities) of proximity advantages, then a simple dichotomy between agglomerated and non-agglomerated areas could be misleading.[15] To overcome such a difficulty we need to introduce elements of graduality in the measure of agglomeration phenomena, like the index (1b) that our algorithm produces.

---

[14] Recall that the threshold (1a) is calculated with regard to the aggregation of all the industries belonging to a cluster. The "degree of industrial diversification" here is simply the share of industries in the cluster that pass the test (1a).

[15] Cannari and Signorini, 2000.

Table 7

Intensity of agglomerations in groups of AAs
(median value of index (1b))

Structural features of AAs				
Size (2)	Share of employment in large firms (1)			
	High	Medium	Low	**Total**
High	2.1	8.5	26.5	**9.4**
Medium	1.7	4.6	7.5	**3.0**
Low	1.2	2.8	2.3	**2.1**
Total	**1.4**	**3.0**	**3.3**	**2.7**
Industrial diversification (3)	High	Medium	Low	**Total**
High	9.9	8.5	12.9	**10.6**
Medium	2.1	2.6	3.3	**2.9**
Low	1.3	3.1	2.3	**1.9**
Total	**1.4**	**3.0**	**3.3**	**2,7**

(1) "High" if almost 25 per cent of employment is concentrated in plants with 250 or more employees; "low" if such figure is less than 5 per cent; "medium" otherwise.

(2) "High" if the AAs has almost 10,000 employees; "low" if such figure is less than 3,000; "medium" otherwise.

(3) "High" if the number of specialization in the AAs (among the industries in the cluster) is greater than 66 per cent; "low" if such figure is less than 33 per cent; "medium" otherwise.

5. Industrial agglomeration and industrial districts in Italy

The Italian industrial structure is characterized by many territorial systems of small firms specialized in so-called traditional sectors. In some of these systems, commonly denominated industrial districts (ID), in addition to the usual external economies due to spatial concentration, firms seem to enjoy peculiar competitive advantages in connection with the favourable social environment of the areas where they are located.[16] The combination of such advantages may explain the remarkable growth that

[16] Brusco and Paba, 1997.

IDs achieved in the last thirty years, for example in employment or exports.[17]

Establishing which areas are to be qualified as ID is a more complex task than identifying AAs, because most of the district's socio-economic features (i.e. cooperation, trust, system of values and views) are very difficult to translate into quantitative variables.[18]

Most quantitative papers on Italian IDs are based on a two-step statistical criterion developed by Fabio Sforzi and Istat.[19]

In the first step of this algorithm the country is divided into "local labour systems" (LLS), that is, aggregates of municipalities that are relatively self-contained in terms of daily journey flows from residence to work places. In particular, using the 1991 population census data, Italy is divided into 799 LLS. Interpreting LLS as "*spatially-coherent systems of interacting localities*" they "*are put forward as the first element for the empirical definition of Marshallian industrial districts*".[20]

In the second step, each LLS is classified as either a district or non-district if the share of manufacturing employment, the level of concentration in a single industry and the share of small and medium firms are greater or smaller than the country average.[21] In this way, out of 799 LLSs, 199 are identified as "Sforzi-Istat industrial districts" (SIDs).

Although based on a reasonable criterion, this algorithm produces a rather crude identification of district areas, owing to both the dichotomous classification and the largely arbitrary determination of the thresholds.

Some authors, even if they agree with the basic premises of the Sforzi-Istat partition, have proposed finer classifications of districts and other industrialized areas, defining several categories.[22] But the reality of the IDs is so varied that it is difficult to capture its variety just by imposing

---

17 Signorini, 2000.

18 Bellandi, 1979.

19 Sforzi, 1985; Istat, 1997.

20 Sforzi, 1990.

21 The small and medium firms are those with less than 250 employees.

22 Cannari and Signorini, 2000.

quantitative criteria of selection. Thus the value of the various maps of districts so far proposed has been questioned time and again.[23]

The problem of these maps and of the algorithms that produce them is not failure to select the best-known industrial areas: textiles at Prato, ceramics at Sassuolo, hides at Arzignano and Santa Croce, glasses at Belluno or machinery at Reggio Emilia. These are easily classified as IDs by any reasonable algorithm. The problem rather is at what distance from Prato or from Sassuolo are the enterprises that produce clothes or textiles no longer in an ID? Or can an ID be strongly specialized over a different local or sectoral range than those that are conventionally considered? Which combination of criteria, for example, can capture the district of blown glass in Venice, that of footwear in the Brenta Valley or that of the publishing business in Milan? And again, is the presence of a few plants of very large size enough to exclude the presence of an ID in the system of hundreds of small and medium firms that for several decades have been producing motor vehicle parts near Turin?

That said, the question now is what relation can be established between the notion of ID and the definition of AA proposed in this paper.

In the next two paragraphs we will address this question, firstly from a quantitative point of view, that is, by comparing the map of SIDs with that of AAs; then by discussing to what extent the two concepts are compatible.

Industrial district and agglomeration: a quantitative comparison. – Even if the number of municipalities belonging to SIDs or to AAs is similar (2,479 and 2,209 respectively), there is a considerable geographical mismatch between the two maps: more than one third of municipalities in AAs do not belong to SIDs, and more than 41 per cent of SID municipalities do not belong to AAs.

A large disparity exists also in terms of employment: in the case of the industrial specialization sectors, while AAs account for 30.6 per cent of the total, the figure for SIDs is much lower (20.1 per cent).

These results are not surprising if we keep in mind the very different criteria adopted to identify SIDs and AAs.

[23] Tattara, 2001.

In the first place, the clusters of industries we have selected represent, at least partially, vertically integrated groups of activities, while the specialization criterion for SIDs refers to single-digit industries. Thus, if firms belonging to a same *filière* tend to locate near each other, the density of specialized firms tends to be larger in AAs than in SIDs.

In the second place, with our algorithm each municipality belonging to an AA must be specialized in a given cluster of industries (as defined in paragraph 3), while the Sforzi-Istat criterion requires that specialization should emerge only in the average of municipalities in the LLS. Our method makes it possible to identify AAs inside a given LLS, or in different but neighbouring LLSs.

Third, it is possible that an area with a considerable presence of large plants represents an AA with our method.[24] Such an area could never be an SID. This is a crucial difference between the two criteria: of the 156 AAs we have found, 61 do not fulfil (if only by a few points) the requirement of a prevalence of small and medium-sized firms imposed by the Sforzi-Istat method.[25]

In conclusion, we can assert that our algorithm captures a more variegated typology of agglomerations than SDIs, essentially because it has more “degrees of freedom” in the definition of industrial, local and dimensional thresholds.

But does such flexibility entail some confusion between the concepts of industrial agglomeration and industrial district?

Agglomeration economies and district advantages. – Our approach for selecting agglomerations has the advantage of making use of some of the methods and terms of modern spatial economics. The price to pay for that is that we cannot include *explicitly* in our algorithm non-operational variables, such as most of the socio-economic ones.

This point deserves to be discussed further. Both environmental and socio-economic variables on which most of a district’s peculiarities depend, and which account for the productive advantages over other

---

[24] If the number of plants is large enough and/or the plants dimensions are similar enough.

[25] It is important to stress that most of these 61 AAs are not agglomerations dominated by large firms: the share of employment in plants with 250 or more employees exceeds 50 per cent only in 15 AAs and is lower than one third in 40 cases.

agglomerations of specialized firms, are in fact taken into account in our model, albeit implicitly. They are part of the set of location and agglomeration advantages that are included in the parameter γ. Due to the infeasibility of separating the direct impact of each component included in this parameter, our algorithm cannot distinguish clearly between firm agglomerations that enjoy advantages *mainly* based on social and institutional factors and those for which economic and technological factors prevail.

But our algorithm does respect the standard *necessary* condition for identifying an ID, that is, an agglomeration of firms, mostly small or medium-sized, specialized in a few industries in a bounded area. Of course, since such a definition only partially captures the more complex concept of ID (which also includes conditions linked to the endowment of social capital), our algorithm tends to identify districts *and* different types of agglomerations. In other words, Italian IDs are, by and large, included in the 156 agglomerated areas we have identified, but such areas also include clusters of firms that have no district specificity.

However, we are convinced that the measure of agglomeration intensity (1b) can be used, to an extent, as an instrument for identifying IDs among the agglomerated areas in the upper part of the list. This is an important question and it needs clarification. Does the proposed algorithm bring us to a *degree* or *specie* distinction between IDs and other agglomerated areas?

The answer is not a trivial one. As we have shown, the algorithm partitions the economic territory into three parts, identifying *strongly* agglomerated areas, *weakly* agglomerated areas and areas not specialized in any industrial production. In the last two kinds of areas, the presence of economies of agglomeration is not proven: for firms in those areas, location choices do not seem to depend strongly on the characteristics of the region.

With regard to *strongly* agglomerated specialized areas, there are important structural differences between them as to the number and relative size of firms, raising the question of whether they are, strictly speaking, industrial districts, according to the definition commonly accepted in the literature.

The issue can, we believe, be addressed by assuming that the existence of certain *special* traits of the districts become more likely as the degree of agglomeration (γ_i) of the specialized area grows. We already know that this quantity, under the same degree of raw specialization (G_i), grows with the number of firms and with the degree of homogeneity in their size (see paragraph 3.1). In this regard, levels of agglomeration largely exceeding the probabilistic threshold, as can be found in the most specialized areas, would signal the existence of conditions that better fit the standard description of industrial districts.

It is worth recalling that some externalities found in industrial agglomerations are not peculiar to districts; they can be the result of other kinds of spatial concentration, e.g. those associated with the location of large plants. On the other hand, other advantages, mainly due to the presence of social capital, are peculiar to districts. In the model proposed, the possibility of properly identifying districts relies heavily on the hypothesis that agglomeration advantages *add up* over the economic area. In other words, if we call *districts in a strict sense* not those productive clusterings whose structure "is *solely* due to peculiar social interrelation",[26] but the regions in which these interrelations work together with other kinds of externalities, then districts, other things equal, can be identified as those areas that most benefit from clustering advantages and therefore tend to show higher values of agglomeration advantages among specialized areas.

[26] Signorini, 2000, p. XXIII.

REFERENCES

Bellandi, G. (1979), “Dal “settore” industriale al “distretto” industriale. Alcune considerazioni sull’unità di indagine dell’economia industriale ”, *Rivista di economia e politica industriale*, No. 1.

Brusco, S. and Paba, S. (1997), “Per una storia dei distretti industriali italiani dal secondo dopoguerra agli anni novanta”, in F. Barca (ed.) *Storia del capitalismo italiano dal dopoguerra ad oggi*, Rome, Donzelli.

Cannari, L. and Signorini, L.F. (2000), “Nuovi strumenti per la classificazione dei sistemi locali”, in L.F. Signorini (ed.), *Lo sviluppo locale*, Rome, Donzelli.

Devereux, M.P., Griffith, R. and Simpson, H. (1999), “The Geographic Distribution of Production Activity in the United Kingdom ”, The Institute for Fiscal Studies, Working Paper No. 26.

Duranton, G. and Overman, H. (2002), “Testing for Localization Using Micro-Geographic Data ”, CEPR Discussion Paper No. 3379.

Ellison, G. and Glaeser, E.L. (1997), “Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach”, *Journal of Political Economy*, Vol. 105, No. 5, pp. 889-927.

Fujita, M., Krugman, P. and Venables, A.J (1999), *The Spatial Economy*, Cambridge, MA, MIT Press.

Istat (1997), *I sistemi locali del lavoro 1991*, Rome.

Iuzzolino (2004), “Costruzione di un algoritmo di identificazione delle agglomerazioni territoriali di imprese manifatturiere”, in L.F. Signorini (ed.) *Economie locali, modelli di agglomerazione e apertura internazionale*, Rome, Banca d’Italia.

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

Marshall, A. (1890), *Principles of Economics*, London, Macmillan.

Maurel, F. and Sedillot, B. (1999), “A Measure of Geographic Concentration in French Manufacturing Industries”, *Regional Science and Urban Economics*, Vol. 29, pp. 575-604.

Pellegrini, G. (2000), “I fattori strutturali dello sviluppo locale nelle recenti analisi teoriche ed empiriche della crescita”, in E. Ciciotti and A.

Spaziante (eds.) *Economia, territorio e istituzioni. I nuovi fattori delle politiche di sviluppo locale*, Milan, Franco Angeli.

Sforzi, F. (1990), "The Quantitative Importance of Marshallian Industrial Districts in the Italian Economy", in F. Pyke, G. Becattini, and W. Sengenberger (eds.), *Industrial Districts and Inter-firm Co-operation in Italy*, Geneva, International Institute for Labour Studies.

Signorini, L.F. (2000), "L'"effetto distretto": motivazioni e risultati di un progetto di ricerca", in L.F. Signorini (ed.), *Lo sviluppo locale*, Rome, Donzelli.

Tattara, G. (2001), "L'efficienza dei distretti industriali: una ricerca condotta dal Servizio Studi della Banca d'Italia", *Economia e società regionale*, Vol. 76, No. 4, pp.110-145.