CONTRIBUTI DI RICERCA CRENOS

UNIVERSITÀ DI CAGLIARI UNIVERSITÀ DI SASSARI
CENTRE FOR NORTH SOUTH ECONOMIC RESEARCH
CRENoS
CENTRO RICERCHE ECONOMICHE NORD SUD

AGGLOMERATION IN MANUFACTURING AND SERVICES: AN EXPERIMENTAL APPLICATION OF A DISTANCE-BASED MEASURE TO SARDINIA

Alberto Tidu
Stefano Usai
Frederick Guy

WORKING PAPERS

2021/10

arkadia

Centro Ricerche Economiche Nord Sud
(CRENoS)
Università di Cagliari
Università di Sassari

CRENOS was set up in 1993 with the purpose of organising the joint research effort of economists from the two Sardinian universities (Cagliari and Sassari) investigating dualism at the international and regional level. CRENoS' primary aim is to improve knowledge on the economic gap between areas and to provide useful information for policy intervention. Particular attention is paid to the role of institutions, technological progress and diffusion of innovation in the process of convergence or divergence between economic areas. To carry out its research, CRENoS collaborates with research centres and universities at both national and international level. The centre is also active in the field of scientific dissemination, organizing conferences and workshops along with other activities such as seminars and summer schools.
CRENoS creates and manages several databases of various socio-economic variables on Italy and Sardinia. At the local level, CRENoS promotes and participates to projects impacting on the most relevant issues in the Sardinian economy, such as tourism, environment, transports and macroeconomic forecasts.

www.crenos.unica.it
crenos@unica.it

CRENoS – Cagliari
Via San Giorgio 12, I-09124 Cagliari, Italia
tel. +39-070-6756397; fax +39-070- 6756402

CRENoS - Sassari
Via MURONI 25, I-07100 Sassari, Italia
tel. +39-079-213511

Title: AGGLOMERATION IN MANUFACTURING AND SERVICES: AN EXPERIMENTAL APPLICATION OF A DISTANCE-BASED MEASURE TO SARDINIA

Prima Edizione: Dicembre 2021

ISBN: 978 88 68513 795

Viale Bonaria 98 - 09125 Cagliari
Tel. 070/6848663 - info@arkadiaeditore.it
www.arkadiaeditore.it

Agglomeration in manufacturing and services: an experimental application of a distance-based measure to Sardinia

Alberto Tidu
University of Cagliari & CRENoS
Stefano Usai
University of Cagliari & CRENoS
Frederick Guy
Birkbeck College, University of London

Abstract

This paper aims to assess the extent of agglomeration processes across industries and along time within a regional economic system by means of a distance-based measure. Specifically, we compute Marcon & Puech's (2017) M for every industry in Sardinia in 2007 and 2012. This computation allows us to to assess the underlying patterns of agglomeration or dispersion throughout the Great Recession, through a study that is not limited to manufacturing activities but also covers service industries and other sectors.

At the same time, this is the first tentative operationalization of M for an entire region, thanks to an approximation of plant addresses with the centroids of the municipalities where they are located. Such an approximation is aimed to reduce the computational intensity that has prevented M from being actually used for the study of the entire economic activity of areas larger than individual neighborhoods or cities. Preliminary evidence seems encouraging and suggests future developments in this direction.

Keywords: Sardinia; agglomeration; spatial methods; economic geography; distance-based measures.
JEL Classification: R12; L60; L80

1 Introduction

Economic activity tends towards concentration and localization - both on a world scale and locally -and it is therefore undeniable that studying agglomeration is essential to understand all sorts of economic phenomena and patterns of growth, in order *"to explain the riddle of uneven spatial development"* (Garretsen & Martin, 2010).

Traditionally, agglomeration studies have relied on Gini, Theil or Herfindahl indices until the development of the so-called *dartboard approach* by Ellison & Glaeser (1997), which allowed to weight the actual distribution of activities against a theoretical distribution. However, even these measures were still biased by what is commonly referred to as the *Modifiable Areal Unit Problem* (Openshaw & Taylor, 1979). The Modifiable Areal Unit Problem affects all quantitative studies of spatial phenomena that rely on territorial aggregates - such as regions, provinces, municipalities, counties, etc. - as the unit of analysis. Indeed, when territorial aggregates are built on pre-defined borders, we have no way to distinguish spatial associations originating from the simple aggregation of data and real associations actually existing in the individual data irrespective of said spatial aggregation (Openshaw, 1984). A solution to the MAUP was proposed by Duranton & Overman (2005) when they introduced *distance-based methods* - agglomeration indices whose measurement relies not on geographical aggregates but on actual units, such as firms or plants. After Duranton & Overman's (2005) initial application, their innovative index *Kd* has been chosen by many scholars to measure agglomeration in a variety of countries all around the world, but the use of an absolute measure prevents results from being comparable across studies. Marcon & Puech (2010) provided a further improvement along this direction by developing a new cumulative measure, called *M*. M shows many interesting features: it is a relative measure, its units of analysis can be weighted, and it accounts for both local and overall agglomeration. However, a significant drawback limiting its use so far is represented by its computational intensity, which is proportional to the squared number of points under study (Marcon & Puech, 2017): a huge figure when pairing each couple of plants in an entire country, or even within a single region. The other traditionally major difficulty - the availability of much needed micro-geographic data - has been increasingly coped with in recent years, with the big data revolution providing researchers with a wide variety of data, often originating from unconventional sources as well (Piacentino, Arbia & Espa, 2021). It was indeed the availability of a huge dataset – ISTAT's own ASIA – that drove us to explore distance measures with the aim of finding a solution to the afore-mentioned computational intensity, by way of reducing the number of points without excessively sacrificing accuracy. Our proposed solution is to approximate plants' locations to the centroids of the municipalities where they are located, largely reducing the total number of spatial interactions to be accounted for. We believe that – given the multitude of municipalities and their relatively small size – not much reliability is lost with such an approximation, especially when considering the very high accuracy provided by ASIA compared to other similar studies: differently from much existing literature, we are dealing with single establishments instead of entire firms not only as their geographic location is concerned, but also their industry of activity and their number of employees; moreover, we are not limited to manufacturing activities, but we are also able to study services, which are usually excluded for a lack of reliable data. These features convince us that the loss in original accuracy is bearable since said accuracy was originally much higher than usually expected in most previous works. Therefore, we believe that such an approximation might be viable and would still allow to retain sufficient accuracy, while allowing to implement an innovative – and, so far, unexploited – measure such as Marcon and Puech's M, opening up to its application to study even larger territorial units, such as entire countries.

In this paper, we aim at assessing the extent of agglomeration across sectors and along time within a regional economic system by means of a distance based measure. Specifically, we compute Marcon & Puech's M for every industry, in both manufacturing and services, in Sardinia in 2007 and 2012, in order to assess both their underlying patterns through the Great Recession and the reliability of our methodology through the comparison of our results to literature expectations and to general Italian patterns (Tidu, 2021). Sardinia looks like the ideal candidate for such a study. For starters, dealing with an island prevents annoying edge effects that would make everything on the other side of the border disappearing, strongly misrepresenting the actual economic activity of communities located on the outskirts, especially in the context of an open-border economy such as the European Union. However – except for Sicily and, indeed, Sardinia - every Italian island is way too small to have anything that even remotely resembles a real economy with a wide array of industries: the largest minor island – Elba – consists of only 7 municipalities and its slightly over 30.000 inhabitants are disproportionately employed in touristic activities. On the other side, Sicily is the largest and by far the most populous Italian (and Mediterranean, as well) island, but the very fact that it counts over 5 million people provides an extremely high number of pairs when the goal is to relate every plant to each other: such a high figure would probably prevent us from testing the robustness of our approximation method in the upcoming future, by comparing our results to those that will be obtained without any approximation whatsoever. On the other hand, Sardinia – which for our purposes is just as geographically large as Sicily and hosts a significantly lower population, but is still large enough to sustain a diversified economy – is probably more akin to an autonomous economy because of its distance from the mainland, which is far larger than the tiny Messina Strait that separates Sicily from Calabria.

We begin our paper by putting into context what distance measures are and why they are helpful when investigating agglomerations, also referring to previous empirical applications within countries and urban areas. We then proceed to describe our datasets and the methodology we followed for our analysis. Finally, we summarize our results and describe the major changes occurred between 2007 and 2012, comparing our findings to the existing literature in order to estimate whether our methodology is sound and provides figures that are consistent with previous expectations.

2 Theoretical and empirical background

A set of recent evidence for a small group of European (Barlet, Briant & Crusson, 2013, Koh & Riedel, 2014), Asian (Nakajima, Saito & Uesugi, 2012) and American (Klier & McMillen, 2008 and Behrens & Bougna, 2015) countries confirms the widespread prediction that industrial activity exhibits specific location patterns. These findings suggest, therefore, that a high level of concentration in manufacturing can be observed in different countries of the world. Models describing and predicting agglomeration have been developed by economists as diverse as A.C. Pigou and Paul Krugman and such theoretical literature has been also accompanied by a sizeable amount of empirical studies aiming to measure agglomeration as accurately as possible (Tidu, 2021). These studies have relied on different *generations* (Nakajima, Saito & Uesugi, 2012) of indices before arriving to the current wave of distance-based methods.

The first generation corresponds to indices that rely on areal data to measure spatial concentration, such as Gini, Isard, Herfindahl, and Theil, where "*the precise location of firms is not available and the data only consists in aggregated counts over administrative zones*" (Bonneu & Thomas-Agnan, 2015, p. 291). In the study of Italy, three different first-generation measures were used by Pagnini (2003, p. 3) in order to measure agglomeration in manufacturing in 1996, showing "*that for an overwhelmingly majority of sectors centripetal forces prevail over centrifugal ones*". Other studies of concentration in Italy by means of first-generation indices were performed by Lafourcade & Mion (2007) and by De Dominicis, Arbia & De Groot (2013), with both providing a particular focus on the relationship between size and spatial agglomeration patterns.

The second generation started out when Ellison & Glaeser (1997) introduced the so-called *dartboard approach*, by way of comparing the degree of spatial concentration of employment in a given sector with the degree of concentration that would result if every plant in that sector were redistributed randomly across actually existing locations, that is, like darts thrown at the map. Ellison & Glaeser's index (henceforth, the *EG index*) would be used by Rosenthal & Strange (2001) to measure the level of spatial concentration among manufacturing industries at a 4-digit level for different geographic scales (zip code, county, and state) for the fourth quarter of 2000. Their aim was to explain differences in the spatial concentration of industries, by matching it with data on industry characteristics. To purse their aim, they regressed the EG index against those industry characteristics that they had identified as viable proxies for the three Marshallian forces of agglomeration – knowledge spillovers, labor market pooling, and input sharing - also controlling for product shipping costs and natural advantage. The EG index was also used by Kolko (2010), but his focus was also on services, and not limited to manufacturing. He relied on a far deeper level of industrial detail, getting down to 6-digit industries. Studying US firms in 2004, he found that service industries, although more urbanized, are less agglomerated than manufacturing, because transport costs represent an incentive to locate near their customers and also because they are far less reliant on natural resources.

More recently, the necessity to deal with the so-called Modifiable Areal Unit Problem allowed the development of a new assumption - continuous space - leading to the birth of a third generation of indices, dubbed distance-based methods. Distance-based methods are a relatively recent introduction to the field of spatial economics. Proceeding from Ripley's (1976, 1977) seminal works and his K function, Duranton & Overman (2005) put forward a new approach that allowed distance measures to be increasingly used to analyze spatial structures and agglomerations, without the need to rely on an approximation of space as discreet. Indeed - unlike more conventional measures ranging from Gini (1912) to Ellison & Glaeser (1997) indices – distance measures do not rely on any pre-defined zoning (i.e.: neighbourhoods, municipalities, communes, provinces, counties, regions), but on the distance between single points of interest, notwithstanding the geographical aggregation they – maybe only temporarily - belong to. Since they study spatial distribution through the actual position of the target entities (such as individual plants or shops) and not through intermediary aggregates, distance-based methods can be a useful improvement compared to conventional spatial measures. Indeed, they are the only reliable way to overcome those issues that arise from referring to pre-defined zoning: geographic units are not necessarily homogenous, neither geographically nor economically, and therefore final values are dependent on the shape and size of the aggregation unit (since the distribution inside each area is lost through aggregation, and units at the opposite end of the same area are treated the same way as neighbouring units). Such an issue is commonly defined as the *Modifiable Areal Unit Problem*[1]. To overcome the *Modifiable Areal Unit Problem*, Duranton & Overman (2005) treat space as continuous. Precisely, they measure the distribution of geographical distances between pairs of firms in an industry and compared such distributions with a randomly generated distribution of firms. The distance measure they use is the following:

(1) $$\widehat{K}_d(r) = \frac{1}{n\,(n-1)} \sum_{x_i \in R} \sum_{x_j \neq x_i,\, x_j \in \aleph} k\left(\|x_i - x_j\|, r\right)$$

where n denotes the total number of points, x_i are the reference points and x_j are its neighbors, with $k(\bullet)$ as a kernel estimator whose total sum is an estimate of the number of neighbors of x_i at the selected distance r

While researching distance-based methods, Duranton & Overman (2005) proposed five characteristics that sound distance measures should have:

1) It should be comparable across industries;
2) It should control for overall agglomeration trends across industries;
3) It should separate spatial concentration from industrial concentration;
4) It should be unbiased with respect to the degree of spatial aggregation;
5) It should provide an indication of the significance of the results.

A few years after its first introduction, Duranton & Overman's Kd was still the measure of choice when dealing with by then "booming" distance methods and the one that probably respected the largest number of properties listed above (Marcon & Puech, 2010). However, Marcon & Puech (2010) noted that most studies until then had not discussed an essential property of distance-based methods[2]: the difference between probability density functions and cumulative functions. *Density functions* measure agglomeration at a specific distance from a reference point, whereas *cumulative functions* measure it up to a specific distance.

[1] Wong (2004, p. 572) notes that <<*Even though Gehlke and Biehl (1934) discovered certain aspects of the modifiable areal unit problem (MAUP), the term MAUP was not coined formally until Openshaw and Taylor (1979) evaluated systematically the variability of correlation values when different boundaries systems were used in the analysis*>>.

[2] With the exception of a short note by Duranton & Overman (2005) in the conclusion of their paper, where they argue that probability density functions reveal more information than cumulative functions do.

Marcon & Puech (2010) identify another dimension of distance-based methods: they can be *topographic*, *relative* or *absolute* measures, according to the reference value used to compare the distribution. A *topographic* reference uses physical space as a benchmark: the number of neighbors on a disk of radius *r* for a *cumulative function*, or on the ring at distance *r* for a *density function*. Topographic functions might simplify space - treating it as homogenous - or alternatively take into account the lack of homogeneity in the geographical space. A *relative* reference may use any other benchmark that is not physical space (e.g.: the distribution of plants that belong to every industry as a benchmark for the distribution of plants belonging to one specific industry). Finally, in the case of no reference, an *absolute* measure is defined, such as the absolute number of plants located at or within a given distance from a given one.

Marcon & Puech's *M* is a cumulative function that provides the relative frequency of neighbours of a given type (such as firms belonging to the same industry as opposed to the entire population of firms) within a certain distance, compared to the same frequency in the whole space. It is estimated by:

$$\widehat{M}(r) = \sum_i \frac{\sum_{j\neq i} \mathbf{1}\left(\left\|x_i - x_j^c\right\| \leq r\right) w\,(x_j^c)}{\sum_{j\neq i} \mathbf{1}\left(\left\|x_i - x_j\right\| \leq r\right) w\,(x_j)} \Bigg/ \sum_i \frac{W_c - w(x_i)}{W - w(x_i)} \qquad (2)$$

where x_j^c are neighbours of the chosen type, x_j are neighbours of any type, *r* is the selected distance, w is the weight of choice, W_c is the total weight of the first type of points, and W is the total weight of all points.

Duranton & Overman (2005) pioneered the application of distance-based measures for the study of agglomeration across industries in a developed country. They investigated location patterns in the manufacturing sector in the UK, by relying on their newly developed Kd index. They found that 52% of industries exhibited localization at a 5% confidence level, with 24% of them showing dispersion at the same confidence level, corresponding to a non-random distribution across space. This first contribution, which is both methodological and empirical, has been followed by many other studies which rely on this index to assess agglomeration levels across industries and, most importantly, their determinants along the line of Rosenthal & Strange (2001). Nakajima, Saito & Uesugi (2012) focused on Japan and found that about half of the 561 four-digit manufacturing industries they studied can be classified as localized, in contrast with a lower figure of only about 35% for service industries, also concluding that "*industries are becoming neither more concentrated nor more dispersed and the location patterns are stable over time*" (Nakajima, Saito & Uesugi, 2012, p. 18). Barlet, Briant & Crusson (2012) studied the location patterns of business-oriented service and manufacturing industries in France relying on an improved version of the Kd index, which takes into account the number of plants in each industry. They showed that concentration is more present among service industries (61%) than manufacturing industries (42%), especially at short distance. Researching Germany, Koh & Riedel (2014) assessed the agglomeration patterns of four-digit industries in Germany using the Kd index. They found that 71% of manufacturing industries are localized while this ratio shoots up to 97% for the service industries. In line with the results above, Behrens & Bougna (2015, p. 48) found that "*depending on industry definitions and years, 40% to 60% of manufacturing industries are clustered*" and that localization in Canada has generally decreased during recent years. Cainelli, Ganau & Jiang (2020) demonstrated that different statistical techniques produce quite different pictures. In particular, they found that most Italian manufacturing industries experienced spatial dispersion processes during the period of the Great Recession. Finally,

their results indicate that space–time dispersion processes occurred within small spatial distances and a short time horizon, although space–time interactions do not seem statistically significant.
As regards developing countries, the available evidence is scarcer, although some interesting contributions have recently appeared. Brakman, Garretsen & Zhao (2017) examined the location of manufacturing in China and found that around 80% of industries at 4-digit in China are significantly localized. Moreover, they found that localization increased rapidly in the period between 2002 and 2008, especially as a consequence of new entrants. Aleksandrova, Behrens & Kuznetsova (2020) analyzed the agglomeration and co-agglomeration patterns of manufacturing industries in Russia and found that 80% of 3-digit industries are both agglomerated and co-agglomerated. Almeida, Neto & Rocha (2020) found that almost 90% of Brazilian manufacturing at 3-digit have statistically significant localization for 2006 and 2015.
Whereas applications of Duranton & Overman's Kd have been plenty, we have been unable to find a tentative measurement of agglomeration for every industry on a regional – or larger - scale through our measure of choice, Marcon & Puech's M. In order to find some empirical applications of *M*, one could turn to Jensen & Michel (2011) who used it to infer the spatial pattern of stores in Lyon (France), although this could be taken more like a mathematical exercise rather than an economic study[3]. Marcon & Puech (2015) themselves later developed such an application when "releasing" their newest measure, the lower-case *m*, in order to show how this could provide a different type of information in respect to Duranton & Overman's *Kd*[4], when describing the distribution of pharmacies in Lyon weighed against the distribution of non-food retail stores. Two other empirical applications of M were developed by Coll-Martìnez, Moreno-Monroy & Arauzo-Carod (2019) and Méndez-Ortega & Arauzo-Carod (2019) who, respectively, computed both *m* and *M* for creative industries and for software-developing industries in Barcelona metropolitan area, underlining how such measures provide the great advantage of being *relative* and not *absolute* (such as Duranton & Overman's *Kd*), thus comparable between industries and years. Also, Moreno-Monroy & Garcìa-Cruz (2016) used M to assess the degree of spatial agglomeration and co-agglomeration of formal versus informal manufacturing activity within Cali metropolitan area in Colombia. Finally, an interesting contribution has been provided by Zhang, Yao, Sila-Nowicka & Song (2021), who used both *M* and *m* to explore the geographic concentration of five manufacturing industries in the Chinese urban region of Jiangsu, relying on firm-level data. However, each one of the cited contributions is limited to either single industries and/or single urban areas and is unable to study the whole economy of an entire region or large island, such as Sardinia.

[3] Points (firms) were not even weighted by the number of employees working for them.
[4] It must be remembered that they both are *density measures*, not *cumulative measures* such as *M*.

3 Data and methods

Despite the difficulties highlighted above, the accuracy provided by Marcon & Puech's M and the possibility to pioneer such a measure on a national scale led us to select it as our preferred index to study agglomeration in Italy. The exceptional detail provided by ISTAT's ASIA datasets in describing not only every firm, but every single plant in the country, convinced us that we could obtain precise enough results, even when accounting for the slight approximation we were forced to accept concerning the geographical location of each establishment. Our intent was to understand not only the geographical distribution of economic activities per se: we also aimed to infer their patterns of change during the Great Recession caused by the financial crisis that struck the whole world after the bursting of the U.S. housing bubble and the bankruptcy of Lehman Brothers in September 2008. Therefore, we measured agglomeration for two different years[5]: the initial year is 2007, a year that ISTAT at the time described as "*exceptional as concerns firms' birth rate*"[6], showing a dynamicity that would not only be lost the following year, but probably was still unrecovered even a decade later. As concerns the choice of 2012 as the closing year for our analysis, it was the first year since the beginning of the Great Recession that showed an increase both in the number of firms and in the number of employees, although this would have later revealed itself as more of a rebound rather than a real recovery, since both firms and employees would then decrease every following year until 2016[7].

3.1 The dataset

ASIA (Archivio statistico delle imprese attive[8]) is a register established in 1996 in accordance with the provisions of European Council Regulation No. 2816/93 on Community coordination in drawing up business registers for statistical purposes, later replaced by Regulation (EC) No. 177/2008, and according to an harmonized methodology adopted by Eurostat.

Since 1996, ASIA covers every enterprise[9] currently active in Italy and contributing to gross domestic product, in the fields of manufacturing, trade and services, providing name, address, field of activity, number of employees, legal form, turnover class, and dates of creation and cessation.

Economic activities not included in ASIA are: agriculture, forestry and fishing; public administration and defense; compulsory social security; activities of membership organizations; activities of households as employers; undifferentiated goods- and services-

---

[5] It is not a coincidence that those same years were also chosen by Cainelli, Ganau & Jiang (2020), who acknowledged that 2007 <<*is generally regarded as a pre-crisis year*>> and that 2012 <<*corresponds to the first year the Italian economy entered a second wave of downturn after the recovery peak reached in 2011*>>.

[6] https://www.istat.it/it/files//2011/02/testointegrale20091006.pdf.

[7] https://www.istat.it/it/files/2018/12/C14.pdf.

[8] Italian for "Statistical register of active enterprises".

[9] Defined by ISTAT's quality report (https://www.istat.it/it/archivio/216767), in accordance with European Council Regulation No. 696/93, as <<*the smallest combination of legal units that is an organizational unit producing goods or services, which benefits from a certain degree of autonomy in decision-making, especially for the allocation of its current resources. An enterprise carries out one or more activities at one or more locations. An enterprise may be a sole legal unit*>>.

producing activities of households for own use; activities of extraterritorial organizations and bodies; units classified as public institutions and private non-profit institutions.
ASIA is updated every year through a process[10] that integrates several administrative and statistical sources[11], guaranteeing a proper statistical representation of active enterprises and of their identification, demographic and economic information. The register has a central role within economic statistics, and it is used for national accounting estimates.
Since 2004, ISTAT also provides another dataset, called Registro statistico delle Unità Locali (ASIA – UL), whose scope is roughly the same as the original register's and which has been built-up through a specific survey: Indagine sulle Unità Locali delle Grandi Imprese (IULGI). This survey has allowed to locate and define the main variables of each local unit[12].

3.2 M index and methodology

Marcon & Puech (2010) noted how the largest number of properties identified by Duranton & Overman (2005) to define a sound distance measure, were thus far respected by their own measure, that is the *K-density* function (denoted *Kd*). However, since *Kd* is a density measure, they believed there was still the need for a cumulative function, that would be useless in order to evaluate geographic concentration. The authors showed how the two types of functions are not substitutes, but indeed complement each other, and, consequently, they built a new function named M[13], for the measurement of intra- and inter-industry geographic concentration.

(3)
$$\widehat{M}(r) = \sum_i \frac{\sum_{j \neq i} \mathbf{1}\left(\left\|x_i - x_j^c\right\| \leq r\right) w\left(x_j^c\right)}{\sum_{j \neq i} \mathbf{1}\left(\left\|x_i - x_j\right\| \leq r\right) w\left(x_j\right)} \Big/ \sum_i \frac{W_c - w(x_i)}{W - w(x_i)}$$

In a map, two types of points – which, in our case, represent plants - are identified:

a) reference points (in our case, plants belonging to a specific industry);

---

[10] ISTAT's quality report defines it as consisting in:

- Data acquisition;
- Analysis of the appropriateness of the sources;
- Transformation of data to standardize definitions;
- Transformation of data to standardize classifications;
- Record linkage;
- Audit and integration of unusual and/or missing data;
- Standardization, geocodification, de-duplication and validation of address data;
- Evaluation of consistency with previous data from the same elaboration.

[11] Agenzia delle Entrate; INAIL (Istituto Nazionale per l'Assicurazione contro gli Infortuni sul Lavoro); CCIAA (Camere di Commercio, Industria, Agricoltura e Artigianato); Banca d'Italia, INPS (Istituto Nazionale della Previdenza Sociale); Seat – pagine gialle Spa; ISVAP (Istituto per la Vigilanza sulle Assicurazioni Private e di Interesse Collettivo).

[12] Defined by ISTAT's quality report (http://siqual.istat.it/SIQual/visualizza.do?id=8889016), in accordance with European Council Regulation No. 696/93, as <<*an enterprise or part thereof (e.g. a workshop, factory, warehouse, office, mine or depot) situated in a geographically identified place. At or from this place economic activity is carried out for which – save for certain exceptions – one or more persons work (even if only part-time) for one and the same enterprise*>>.

[13] Marcon & Puech (2010, pp. 747-748) <<*called it the M function because it is an extension of the existing cumulative distance-based methods, namely Ripley's K function (1976, 1977) and Besag's L function (1977)*>>.

b) target neighbor points (in our case, plants belonging to the same industry as the reference point).

The average number of target neighbors is compared to a benchmark in order to verify whether they are more or less frequent than they would be if plants were distributed randomly. In order to control for the local density of points, target neighbor points (in our case, the number of plants, belonging to the same industry, located within the selected distance r from the reference point) are normalized by the number of all the neighbors located within the same radius. The average of the resulting ratio for each reference point will then be weighted against the same ratio for the entire area – in our case, the whole country: if the former is higher than the latter – that is, M is greater than 1 - then the industry is somehow concentrated with points showing some degree of mutual attraction that would not be spotted if they were randomly distributed and independent from each other. On the other hand, if the latter ratio is higher than the former, it means that points tend to repel each other, therefore the industry is more dispersed than a random distribution. M also allows to weigh points, for example by – as is our case – employees working at the plant.

One major difficulty was posed by the huge number of interactions required in order to account for every couple of plants located at less than the largest distance range we selected for our analysis. We overcame such an issue by approximating the plants' locations to the centroids of the municipalities where they are located, reducing the number of total spatial interactions to slightly over 10.000 - still a large figure, but far more manageable than the original 9 billion pairs. A similar expedient was found by Brakman, Garretsen & Zhao (2017) when studying spatial concentration of Chinese manufacturing firms: their limit was not computational, but concerned the actual location of the firms, since information was provided only at county level. They offered an interesting justification to such an approximation, by comparing the mean value of intra-county distances (19 kilometers) to the median value of all pair-wise distances between manufacturing firms in China (around 900 kilometers). At first sight, such an approximation might seem counterintuitive when one is handling complex distance-based methods in order to pursue accuracy, but, as well expressed by Marcon & Puech (2017, p. 30) themselves, "*cumulative functions are insensitive to errors at smaller scales than the distance they consider: if the uncertainty is a few hectometers, the number of neighbors up to a few kilometers is known with no error except for the more distant ones, which are a small proportion*". Our expedient shall not be perceived as a simple aggregation of data, since each plant is considered separately from the others. Instead, what we are doing is approximating the geographical position of the plant by no more than a few kilometers: only 2 municipalities out of 377 cover more than 300 km², and the median surface is barely over 40 km²; moreover, it is also easy to presume that most plants actually gravitate closer to the municipality centroid than a random distribution would predict, further reducing the magnitude of our approximation. Therefore, it is irrelevant that our methodology allows to simplify computations by numerically aggregating employees after their location has been registered, since this would occur even with the most pristine and punctual usage of distance methods: in the real world, employees are not piled up one above the other in the exact geo-localized position of the plant, but they move in space and are also separated from each other by at least a few meters, and – more often than not – much more than that, with many plants covering an ample surface (think of airports, harbors, large warehouses). Since nobody would require distance measures to take into account this physiological separation between people in the same working area or even the *exact* position of each one of them in each moment, the real question should be whether the magnitude of our approximation is too large and whether it makes the implementation of distance measures useless compared to more conventional methods relying on aggregation. Such a question is clearly legitimate and will surely find a proper answer in a future study, where we aim to compare results obtained through our methodology with non-approximated computations.

ISTAT provides origin-destination matrices with distances between Italian municipalities, both in meters of road travel and in minutes of time travel. As described in the related methodological note[14], values were computed through GIS tools and TomTom MultiNet 2013 road network, relying on municipalities' centroids (identified as the census section that includes the municipal house) as they were in 2013. Travel times and roads are computed in ideal conditions, not accounting for traffic but only for average travel speed in each road tract[15].

The ASIA dataset for each year was then crossed with the ISTAT distance matrix for the same year through SQL queries that create new columns, showing for each plant *i*:

- the number of employees *w* working in plants of the same industry (x_j^c) within the borders of the municipality it belongs to and in municipalities whose centroid is located within the selected distance range *r*;

- the number of employees *w* working in plants of every industry (x_j) within the borders of the municipality it belongs to and in municipalities whose centroid is located within the selected distance range *r*.

We then proceeded to compute the total number of employees working in each industry (W_c) and the total number of employees working in every industry (*W*).

These numbers allowed us to compute the M for each industry at the selected distance range, through the following formula:

(4) $$\widehat{M}(r) = \sum_i \frac{\sum_{j \neq i} \mathbf{1}\left(\left\|x_i - x_j^c\right\| \leq r\right) w\left(x_j^c\right)}{\sum_{j \neq i} \mathbf{1}\left(\left\|x_i - x_j\right\| \leq r\right) w\left(x_j\right)} \Big/ \sum_i \frac{W_c - w(x_i)}{W - w(x_i)}$$

The process is repeated for each distance (5, 10, 15, 20 and 30 minutes) and for both years (2007 and 2012).

[14] https://www.istat.it/it/files//2015/04/Nota_Tecnica_MatriciDistanza.pdf

[15] Between 15 and 50 km/h for urban roads, roundabouts and interchanges; between 60 and 80 km/h for non urban roads ("*strade statali*", "*strade provinciali*" and "*strade comunali*"); between 85 and 95 km/h for high-speed non urban roads ("*superstrade*"); between 100 and 120 km/h for highways ("*autostrade*").

4 Baseline results

Sardinia is arguably an economically marginal region both because of its small demographic size and its low productivity, with its GDP per capita - amounting to "19,722 Euros, higher than the average of southern Italian regions (17,353) but much lower than the Italian average (25,728.6 Euros)" in 2012 (Sideri & Usai, 2016, p. 176) - precipitating under 80% of the European average during the Great Recession and still falling even after the end of its acute phase, down to 69% in 2017 (Biagi, et al., 2021, p. 168). Furthermore, Sardinia's total GDP (amounting to €31,300 million in 2012) consists for about 75% of consumption expenditure within the regional territory, and the chunk invested in fixed capital formation only amounts to 17% and "has been worryingly decreasing in recent years" (Sideri & Usai, 2016, p. 176). Such an economic marginality – joint with its physical distance from larger markets – has been the main cause of an inflated public sector, which accounts for 24% of total value added and almost half of total compensation; moreover, it is further aggravated by the extremely small size of its firms, which is the lowest in Italy after other peripheral regions such as Sicily, Molise and Calabria, with 97% of firms employing less than 10 people and contributing to 64% of the workforce. Such a precarious state of the economy reflects onto the labor market and produces an unemployment rate that – although significantly lower today than its apex of 18,9% in 2014 – is still worrysome, especially as concerns the youth (age 15-24), afflicted by on the highest unemployment rates in Europe (peaking at 56.4% in 2016), although nowadays below the Mezzogiorno rate especially after featuring a 20% decrease to 35,7% in 2018 (Biagi et al., 2021, p. 173). The primary sector shows a heavier weight than in Italy as a whole, accounting for 5,6% of the total workforce against a national average of 3,7%. On the other hand, the most productive sectors - industry and high-value added service, connected to real estate, professional and services for people – respectively accounted for 8% and 15% of active firms, against a national average of 10% and 22%.

M results for the five computed distance ranges are summarized – with means and standard deviations weighted by the number of plants in each industry - in tables 1 and 2, in addition with descriptive statistics about employees and plants. Several industries, operating in different fields of activity, show zero agglomeration for every distance range, but all of them consist of less than 5 plants each. When each industry is weighted by its number of plants, mean values are remarkably similar between 2007 and 2012. Figure 1 shows the distribution of M results for every distance range and year: it is unsurprising that the largest bulk of results is between 1 and 2, since 1 would be obtained by an industry whose plants were distributed in a pattern exactly mimicking the general distribution of every economic activity within the entire territory analyzed; since most industries tend to show some degree of agglomeration and, moreover, our distance ranges are far shorter than a radius that could include the whole island, it is easy to understand how most industries' M hovers slightly above 1. There also seems to be an increase of dispersed industries and a decrease of agglomerated industries between 2007 and 2012, but this might be true only for shorter distance ranges.

Many of the industries that came out as the most agglomerated are extremely small in terms of both employees and plants, therefore we cannot put much trust in the significance of their results. However, such industries as *Manufacture of metal-forming machinery and machine tools (284)* and *Manufacture of tubes, pipes and hollow profiles and of tube or pipe fittings of cast-iron (242)*, despite their extremely small size, remained consistently on top of the ranking, suggesting that there might be some actual force driving them towards agglomeration. Furthermore, *Manufacture of*

cutlery, hand tools and general hardware (257) and *Manufacture of refractory products (232)* - both quite small in terms of either plants or employees, but not so small as concerns the other measure, thus indicating respectively a relatively large number of single operators and a small number of relatively large firms – seem to show the same consistency and might therefore sustain some actual agglomerating tendency and not result from chance alone. Some larger industries, on the other hand, shall be regarded as most likely being driven by agglomerating forces, and among manufacturing industries that might be the case with *Manufacture of knitted and crocheted apparel (143)*, *Manufacture of basic precious and other non-ferrous metals; reprocessing of nuclear fuels (244)*, *Spinning, weaving and finishing of textiles (131)* and – only for short distance ranges - also *Manufacture of cement, lime and plaster (235)*, *Manufacture of agricultural and forestry machinery (283)*, *Mining and Quarrying n.e.c. (089)* and *Manufacture of refined petroleum products (192)*.
The same confidence in the significance of results for relatively large industries might also be held for some apparently agglomerated service industries and, chiefly, *Camping grounds, recreational vehicle parks and trailer parks (553)*, *Sea and coastal water transport (501)* and – somehow less so – for *Wireless telecommunications activities (612)* and *Other short term accommodation activities (552)*. Indeed, services seem to show much less agglomeration for shorter distance ranges compared to manufacturing: however, if we consider the largest distance range we have computed M for – 30 minutes – we find that also some larger industries as *Activities of call centers (822)*, *Passenger air transport (511)* and *Hotels (551)* are showing some degree of agglomeration.
On the other side of the spectrum, among the most dispersed industries for every distance range – leaving aside a bunch of them that are way too small to provide meaningful results – there seem to be some interesting cases, such as:

- *Manufacture of footwear (152)* and *Manufacture of musical instruments (322),* which are commonly regarded as industries subject to agglomerating forces and are indeed among the most agglomerated nationally (Tidu, 2021);
- *Wired telecommunications activities (611)*, curiously contrasting with *Wireless telecommunications activities (612)* which – as mentioned above – is actually among the most agglomerated;

On the other hand, some industries' dispersion was clearly expected and might be interpreted as a sign that our approximation did not produce misleading results; this was the case with activities commonly spread all over, such as *Postal activities (531)*, *Electric power generation, transmission and distribution (351)*, *Other passenger land transport (493)*, *Monetary intermediation (641)*, *Photographic activities (742)*, *Waste collection (381)*, *Water collection, treatment and supply (360)*, *Electrical, plumbing and other construction installation activities (432)*, *Cleaning activities (812)* and many others.
As already suggested by some of the results mentioned above, correlation with national results (Tidu, 2021) is almost non-existent and the most agglomerated industries in Sardinia are almost unequivocally different than those in Italy as a whole, especially when the smallest and least significant ones are left out. Indeed, a large part of the manufacturing activities included among the most agglomerated industries in Sardinia, barely appear in the middle ranks nationally; chiefly, among them:
- *Manufacture of metal-forming machinery and machine tools* (284), mostly located in Calangianus;
- *Manufacture of tubes, pipes and hollow profiles and of tube or pipe fittings of cast-iron* (242*),* with the bulk of its employees working in one plant in Siniscola*;*

- *Manufacture of basic precious and other non-ferrous metals; reprocessing of nuclear fuels* (244), represented mostly by micro-firms in several municipalities but with a few dominating large plants located in Portoscuso (see figure 1).
On the other hand, those manufacturing activities that are among the most agglomerated industries in Italy (Tidu, 2021) – such as the already cited *Manufacture of musical instruments* (322) and *Manufacture of footwear* (152), but also *Manufacture of clay building material*s (233) and *Manufacture of jewellery, bijouterie and related articles* (321) – are not agglomerated at all in Sardinia (for a visual example, see figure 3). *Logging (022)* and *Extraction of natural gas (062)* curiously appear among the least agglomerated, but the reason is clearly their almost non-existence in Sardinia, with only two plants for each. On the other hand, transport and accommodation activities are far less agglomerated in Sardinia than they are nationally.
When analyzing the same timeframe included in our study, Cainelli, Ganau & Jiang (2020, p. 443) found that "*Italian manufacturing sectors experienced a process of space-time dispersion during the period of the Great Recession, although with slightly different intensity and patterns*"[16]. Indeed, descriptive statistics provided in tables 1 and 2 show an almost imperceptible decrease in the weighted mean of agglomeration results for every distance range. However, a deeper look at changes between 2007 and 2012 results, provided in table 3, shows that the decrease in absolute terms does not have a counterpart when referring to percentage changes, which have actually slightly increased meanwhile.
In 2012, the patterns of the most agglomerated industries do not seem to have changed much, with the same reasons outlined above for 2007 seemingly still valid even five years - and a severe economic crisis - later. Nevertheless, and even leaving aside the smallest industries, M values seem to be generally lower than they were five years earlier: this would be somehow reminiscent of Behrens & Bougna's (2015, p. 48) finding that "*localization is decreasing, i.e., manufacturing industries become less geographically concentrated in Canada*", but it contrasts with findings by Brakman, Garretsen & Zhao (2017) who report increased agglomeration in China between 2002 and 2008.
Sardinia's agglomerating (table 6) and dispersing (table 7) trends are quite reminiscent of national ones, although some differences are remarkable: chiefly, the most agglomerating industry in Sardinia – *Manufacture of agricultural and forestry machinery (283),* which can be visualized in figure 4 – is among the most dispersing nationally, and the same is true not only for other similarly shrinking industries – such as *Construction of utility projects (422)* and *Processing and preserving of meat (101)* – but also for strongly growing ones, as is the case of *Landscape care and maintenance service activities* (813*),* as it is evident from figure 7. Among the most dispersing industries in Sardinia, on the other hand, many are actually agglomerating nationally (Tidu, 2021): the most striking example is *Other accommodation (559),* which comes second both as for the most dispersing in Sardinia and for the most agglomerating in the entire country (see figure 6).
Finally, we were interested in checking and visualizing the relationship between concurrent changes in agglomeration and industry size for industries large enough to produce meaningful

[16] This difference in intensity and patterns is reminiscent of De Dominicis, Arbia & De Groot (2013, p. 5), who observed that <<*whereas manufacturing has been spreading out, service activities have become increasingly clustered*>>.

results (our minimum threshold for an industry was to have at least 15 plants and 100 employees).
As concerns manufacturing (see figure 7), figures 8 and 9 show how two shrinking industries look on a map when their agglomeration index either increases – as is the case of *Treatment and coating of metals; machining (256)* - or decreases – as in *Sawmilling and planing of wood (161)*. It is clear enough how the first industry has lost employees' density where this was low already, while retaining most of its agglomerated hotspots around Cagliari, Sulcis, Sassari, Olbia, Nuoro and Ozieri. On the other hand, the second industry has seen a reduction of its employees in places where there used to be many, thus somehow losing its 2007 agglomerations in Gallura and Medio Campidano, and around Sassari and Isili.
Because of the Great Recession, just a few manufacturing industries increased their size between 2007 and 2012[17]. Therefore, in order to show how a growing industry might look differently depending on whether its agglomeration is increasing or decreasing, we rely to service industries (see figure 10). *Courier activities (532)* is a strongly growing industry (around +50% both in terms of plants and employees) showing decreasing agglomeration for every distance range, as it is immediately apparent when looking at figure 11, where plants are clearly sprouting everywhere regardless of previously settled areas. On the other hand, *Pre-primary education (851)* has been agglomerating while featuring a growth of +26% in employees and +17% in plants. Indeed, in figure 12 it looks like already established hotspots have grown at the expense of other less originally dense municipalities.

[17] Manufacture of basic precious and other non-ferrous metals; reprocessing of nuclear fuels (244) grew in terms of plants (+14,29%) but had a significant drop in terms of employees (-21,82%) and therefore it wouldn't make much sense to use it as a token for "growing industries"; the same is true for Other manufacturing n.e.c. (329). Manufacture of medical and dental instruments and supplies (325) shows growth as concerns both employees and plants, but it is uninteresting in terms of agglomeration change (ranging from -0,93% to +1,85% for different distance ranges); Manufacture of clay building materials (233) grew in terms of plants but went almost unchanged as concerns employees; moreover, the change in agglomeration was not consistent between different ranges, with M(5) showing +146,67% and M(15) featuring a -19,14%. Similar problems occur with Processing and preserving of fish, crustaceans and molluscs (102) and Manufacture of rubber products (221).

5 Conclusion

We use comprehensive data provided by ISTAT – the Italian Institute of Statistics - in order to measure agglomeration for Sardinian industries before and after the Great Recession. Even when accounting for the approximation that we were forced to accept in order to deal with the huge amount of data, we believe our contribution is relevant with respect to both the methodological approach and the accuracy of our results. Indeed, our operationalization suggests an innovative way to exploit an accurate measure - such as Marcon and Puech's M - outside the limited scope of city neighborhoods. This method, thus, allows extending its implementation possibilities to the study of larger geographic regions and even entire countries. This is of utmost importance because it offers an alternative to the passive acceptance of the distortions caused either by the Modifiable Areal Unit Problem or, alternatively, by the absence of a benchmark when relying on more commonly used distance-based methods, such as Duranton & Overman's *Kd*. With micro-geographic data becoming increasingly available (Arbia, 2001), it is crucial trying to exploit their whole potential when researching economics. It now becomes interesting to compare results obtained with our proposed methodology with those obtained with an exact geo-localization of plants, devoid of any sort of approximation: this would certainly confirm the accuracy – and thus the reliability - of our methodology when dealing with larger regions. Indeed, Sardinia was chosen as the target of this study because of the viability of such a comparison, thanks to a demographic and economic size that make the island at the same time manageable and relevant.

Meanwhile, our results already seem plausible and in line with our expectations and with other researchers' findings, both in Italy and abroad, reinforcing Marcon & Puech's (2017, p. 30) proposition that "*cumulative functions are insensitive to errors at smaller scales than the distance they consider: if the uncertainty is a few hectometers, the number of neighbors up to a few kilometers is known with no error except for the more distant ones, which are a small proportion*". Indeed, when scrolling our ranking of the most agglomerated industries (table 6), it is easy to spot those factors that literature traditionally identifies as fundamental in generating agglomeration; and, on the other side of the spectrum as well (table 7), those industries that came out as the most disperse are certainly in line with literature predictions.

Such results are surely interesting in and by themselves, but their relevance grows when they present the opportunity to assess the change that has occurred during such a dramatic event as the Great Recession. Specifically, we believe that some of the most at large considerations of previous literature were confirmed, with agglomeration somehow slightly decreasing (Behrens & Bougna, 2015; Almeida, Neto and Rocha, 2020) during the Great Recession, albeit with the most agglomerated industries – especially manufacturing ones – maintaining a high degree of agglomeration, and sometimes even showing an increase (Behrens & Bougna, 2015). Although the tentative exploration of possible determinants has not yet produced robust enough results to assemble a significant model, our results point to the need of further study and interpretation about how agglomerations behave and react to the crisis. Indeed, there are many accessory ideas that have emerged and that we aim to explore in the future, building up from the results of this study. On the one hand, we aim at replicating the study without approximation in order to assess empirically our methodology's actual value. Moreover, we recognize that the lack of a significance test of results for each industry is a clear limit when trying to distinguish which ones might show a relation between their agglomeration index and other measures, such as size, age and entry/exit rates (Tidu, 2021). However, such a test would require an immense amount of new computations, since we would need to simulate a high number of redistributions of plants along their actual locations. A certain degree of reliability might be offered by weighting each industry by the number of plants that it consists of, in order to give comparatively less importance to results that are much more likely than others to be a consequence of chance alone. However, the possibility to discern which industries actually produce reliable agglomeration results, would make the exploration of determinants much easier, possibly allowing us to identify patterns and to develop a useful model.

References

Aleksandrova, E., Behrens, K., & Kuznetsova, M. (2020). Manufacturing (co) agglomeration in a transition country: Evidence from Russia. *Journal of Regional Science*, *60*(1), 88-128.

Almeida, E. T., Neto, R. M. S., & Rocha, R. M. Manufacturing Location Patterns in Brazil. PhD dissertation in Economics. PIMES, Federal University of Pernambuco – UFPE.

Arbia, G. (2001). Modelling the geography of economic activities on a continuous space. *Papers in Regional Science*, *80*(4), 411-424.

Biagi, B., Dettori, B., Paci, R., & Usai, S. (2021). Economic development in Sardinia: overcoming the insularity gap. *Economic development in Sardinia: overcoming the insularity gap.*, 165-186.

Barlet, M., Briant, A., & Crusson, L. (2013). Location patterns of service industries in France: A distance-based approach. *Regional Science and Urban Economics*, *43*(2), 338-351.

Behrens, K., & Bougna, T. (2015). An anatomy of the geographical concentration of Canadian manufacturing industries. *Regional Science and Urban Economics*, *51*, 47-69.

Biagi, B., Dettori, B., Paci, R., & Usai, S. (2021). Economic development in Sardinia: overcoming the insularity gap. *Economic development in Sardinia: overcoming the insularity gap.*, 165-186.

Bonneu, F., & Thomas-Agnan, C. (2015). Measuring and testing spatial mass concentration with micro-geographic data. *Spatial Economic Analysis*, *10*(3), 289-316.

Brakman, S., Garretsen, H., & Zhao, Z. (2017). Spatial concentration of manufacturing firms in China. *Papers in Regional Science*, *96*, S179-S205.

Cainelli, G., Ganau, R., & Jiang, Y. (2020). Detecting space–time agglomeration processes over the Great Recession using firm-level micro-geographic data. *Journal of Geographical Systems*, *22*(4), 419-445.

Coll‐Martínez, E., Moreno‐Monroy, A. I., & Arauzo‐Carod, J. M. (2019). Agglomeration of creative industries: An intra‐metropolitan analysis for Barcelona. *Papers in Regional Science*, *98*(1), 409-431.

De Dominicis, L., Arbia, G., & De Groot, H. L. (2013). Concentration of manufacturing and service sector activities in Italy: Accounting for spatial dependence and firm size distribution. *Regional Studies*, *47*(3), 405-418.

Duranton, G., & Overman, H. G. (2005). Testing for localization using micro-geographic data. *The Review of Economic Studies*, *72*(4), 1077-1106.

Ellison, G., & Glaeser, E. L. (1997). Geographic concentration in US manufacturing industries: a dartboard approach. *Journal of political economy*, *105*(5), 889-927.

Garretsen, H., & Martin, R. (2010). Rethinking (new) economic geography models: taking geography and history more seriously. *Spatial Economic Analysis*, *5*(2), 127-160.

Gehlke, C. E., & Biehl, K. (1934). Certain effects of grouping upon the size of the correlation coefficient in census tract material. *Journal of the American Statistical Association*, *29*(185A), 169-170.

Gini, C. (1912). Variabilità e mutabilità. *Reprinted in Memorie di metodologica statistica (Ed. Pizetti E)*.

Jensen, P., & Michel, J. (2011). Measuring spatial dispersion: exact results on the variance of random spatial distributions. *The Annals of regional science*, *47*(1), 81-110.

Klier, T., & McMillen, D. P. (2008). Evolving agglomeration in the US auto supplier industry. *Journal of Regional Science*, *48*(1), 245-267.

Koh, H. J., & Riedel, N. (2014). Assessing the localization pattern of German manufacturing and service industries: a distance-based approach. *Regional Studies*, *48*(5), 823-843.

Kolko, J. (2010). Urbanization, agglomeration, and coagglomeration of service industries. In *Agglomeration economics* (pp. 151-180). University of Chicago Press.

Lafourcade, M., & Mion, G. (2007). Concentration, agglomeration and the size of plants. *Regional science and urban economics*, *37*(1), 46-68.

Lang, G., Marcon, E., & Puech, F. (2020). Distance-based measures of spatial concentration: introducing a relative density function. *The Annals of Regional Science*, *64*(2), 243-265.

Marcon, E., & Puech, F. (2010). Measures of the geographic concentration of industries: improving distance-based methods. *Journal of Economic Geography*, *10*(5), 745-762.

Marcon, E., & Puech, F. (2017). A typology of distance-based measures of spatial concentration. *Regional Science and Urban Economics*, *62*, 56-67.

Méndez-Ortega, C., & Arauzo-Carod, J. M. (2019). Locating software, video game, and editing electronics firms: using microgeographic data to study Barcelona. *Journal of Urban Technology*, *26*(3), 81-109.

Moreno‐Monroy, A. I., & Cruz, G. A. G. (2016). Intra‐metropolitan agglomeration of formal and informal manufacturing activity: Evidence from Cali, Colombia. *Tijdschrift voor economische en sociale geografie*, *107*(4), 389-406.

Nakajima, K., Saito, Y. U., & Uesugi, I. (2012). Measuring economic localization: Evidence from Japanese firm-level data. *Journal of the Japanese and International Economies*, *26*(2), 201-220.

Openshaw, S. (1984). Ecological fallacies and the analysis of areal census data. *Environment and planning A*, *16*(1), 17-31.

Openshaw, S., & Taylor, P. (1979). A million or so correlation coefficients: three experiments on the modifiable unit area problem. *Statistical Applications in the Spatial Sciences*, 127-144.

Pagnini, M. (2003). Misura e determinanti dell'agglomerazione spaziale nei comparti industriali in Italia. *Rivista di Politica Economica*, *93*(2), 149-149.

Penttinen, A., Stoyan, D., & Henttonen, H. M. (1992). Marked point processes in forest statistics. *Forest science*, *38*(4), 806-824.

Piacentino, D., Arbia, G., & Espa, G. (2021). Advances in spatial economic data analysis: methods and applications. *Spatial Economic Analysis*, *16*(2), 121-125.

Ripley, B. D. (1976). The second-order analysis of stationary point processes. *Journal of applied probability*, *13*(2), 255-266.

Ripley, B. D. (1977). Modelling spatial patterns. *Journal of the Royal Statistical Society: Series B (Methodological)*, *39*(2), 172-192.

Rosenthal, S. S., & Strange, W. C. (2001). The determinants of agglomeration. *Journal of urban economics*, *50*(2), 191-229.

Sideri, M., & Usai, S. (2016). Economic Outlook: A missed opportunity. In *Surrounded by Water, Landscapes, Seascapes and Cityscapes of Sardinia (A. Cordale and G. Sistu eds.).* Cambridge Scholars Publishing.

Tidu, A. (2021). Economic agglomeration in Italy before and after the Great Recession. *PhD Dissertation.*

Zhang, X., Yao, J., Sila-Nowicka, K., & Song, C. (2021). Geographic concentration of industries in Jiangsu, China: a spatial point pattern analysis using micro-geographic data. *The Annals of Regional Science*, *66*(2), 439-461.

Tables and figures

Table 1 – *Statistics for M results - Sardinia (2007)*

	Obs	Mean	Std. Dev.	Min	Max
M (5 minutes)	224	1,44	2,60	0	358,34
M (10 minutes)	224	1,35	1,78	0	263,24
M (15 minutes)	224	1,26	0,84	0	119,76
M (20 minutes)	224	1,20	0,66	0	101,39
M (30 minutes)	224	1,14	0,39	0	28,77
Employees	224	1659,21	3286,44	2,00	30526,34
Plants	224	540,46	1245,69	2,00	8745,00

Source: Compiled by the authors

Table 2 – *Statistics for M results - Sardinia (2012)*

	Obs	Mean	Std. Dev.	Min	Max
M (5 minutes)	219	1,43	2,31	0	201,05
M (10 minutes)	219	1,34	1,31	0	80,69
M (15 minutes)	219	1,26	0,74	0	42,3
M (20 minutes)	219	1,19	0,58	0	39,49
M (30 minutes)	219	1,14	0,41	0	29,16
Employees	219	1552,30	2816,06	1,20	16941,03
Plants	219	532,47	1177,86	2,00	7336,00

Source: Compiled by the authors

Table 3 – *Descriptive statistics for percentage changes in M in Sardinia between 2007 and 2012*

	Obs	Mean	Std. Dev.	Min	Max
M (5 minutes)	174	2,75	32,82	-100	555,98
M (10 minutes)	174	1,62	29,38	-100	472,68
M (15 minutes)	174	0,66	21,52	-100	202,01
M (20 minutes)	174	0,67	17,98	-100	151,79
M (30 minutes)	174	0,44	14,80	-100	129,25

Source: Compiled by the authors

Table 4 - *Descriptive statistics for percentage changes in M in Manufacturing industries in Sardinia between 2007 and 2012*

	Obs	Mean	Std. Dev.	Min	Max
M (5 minutes)	56	7,47	60,27	-100	555,98
M (10 minutes)	56	8,56	53,24	-100	472,68
M (15 minutes)	56	2,31	34,00	-100	202,01
M (20 minutes)	56	3,33	24,81	-100	151,79
M (30 minutes)	56	3,22	17,56	-100	129,25

Source: Compiled by the authors

Table 5 - *Descriptive statistics for percentage changes in Service industries in Sardinia's M between 2007 and 2012*

	Obs	Mean	Std. Dev.	Min	Max
M (5 minutes)	100	0,48	18,53	-88,61	99,19
M (10 minutes)	100	-1,37	15,86	-89,74	100,81
M (15 minutes)	100	-0,82	15,48	-96,05	93,75
M (20 minutes)	100	-0,89	14,66	-85,06	87,39
M (30 minutes)	100	-0,81	13,32	-83,01	82,20

Source: Compiled by the authors

Table 6 – *20 industries with the largest % increase in agglomeration in Sardinia between 2007 and 2012*

Industry code	Industry description	Employees	Plants	M (5 minutes)	M (10 minutes)	M (15 minutes)	M (20 minutes)	M (30 minutes)	National rank
283	MANUFACTURE OF AGRICULTURAL AND FORESTRY MACHINERY	-68,99	-47,62	555,98	472,68	167,00	151,79	129,25	151
132	WEAVING OF TEXTILES	-67,55	-18,92	159,23	213,70	202,01	67,09	23,10	16
192	MANUFACTURE OF REFINED PETROLEUM PRODUCTS	2,58	14,81	111,93	123,22	115,78	111,79	34,62	3
772	RENTING AND LEASING OF PERSONAL AND HOUSEHOLD GOODS	-27,60	-17,17	99,19	100,81	93,75	87,39	82,20	1
256	TREATMENT AND COATING OF METALS; MACHINING	-40,19	-49,21	168,28	173,10	85,99	15,50	6,67	14
151	TANNING AND DRESSING OF LEATHER; MANUFACTURE OF LUGGAGE; HANDBAGS; SADDLERY AND HARNESS; DRESSING AND DYEING OF FUR	-56,28	-40,00	92,95	84,33	89,88	78,38	61,67	76
152	MANUFACTURE OF FOOTWEAR	-21,63	15,00	130,26	157,35	94,38	-14,81	39,29	29
143	MANUFACTURE OF KNITTED AND CROCHETED APPAREL	-38,88	-23,53	-84,32	113,17	143,62	107,18	113,82	19
324	MANUFACTURE OF GAMES AND TOYS	-75,47	-43,75	489,47	-44,78	-9,70	-35,27	-29,38	66
274	MANUFACTURE OF ELECTRIC LIGHTING EQUIPMENT	-8,68	-26,32	100,88	42,64	126,02	58,62	24,19	32
109	MANUFACTURE OF PREPARED ANIMAL FEEDS	-48,13	-34,29	52,81	171,15	22,08	52,46	26,43	25
803	INVESTIGATION ACTIVITIES	140,78	65,00	40,20	50,58	45,57	44,52	66,95	52
429	CONSTRUCTION OF OTHER CIVIL ENGINGEERING PROJECTS	2,11	-3,55	68,38	81,08	49,55	31,13	15,38	15
422	CONSTRUCTION OF UTILITY PROJECTS	-37,92	-57,89	62,14	109,40	39,85	32,80	-5,47	144
262	MANUFACTURE OF COMPUTERS AND PERIPHERAL EQUIPMENT	-61,62	-52,83	38,17	59,62	50,00	35,66	34,65	4
813	LANDSCAPE CARE AND MAINTENANCE SERVICE ACTIVITIES	170,29	222,73	47,48	40,91	30,43	40,57	33,98	142
351	ELECTRIC POWER GENERATION, TRANSMISSION AND DISTRIBUTION	-6,54	88,54	54,74	47,06	49,32	27,06	3,92	40
101	PROCESSING AND PRESERVING OF MEAT	-17,83	-4,35	33,60	52,08	20,63	38,89	28,13	156
108	MANUFACTURE OF OTHER FOOD PRODUCTS	-14,04	-19,39	36,12	49,82	28,11	33,50	21,92	30
233	MANUFACTURE OF CLAY BUILDING MATERIALS	1,04	27,78	146,67	45,41	-19,14	-8,87	2,46	24

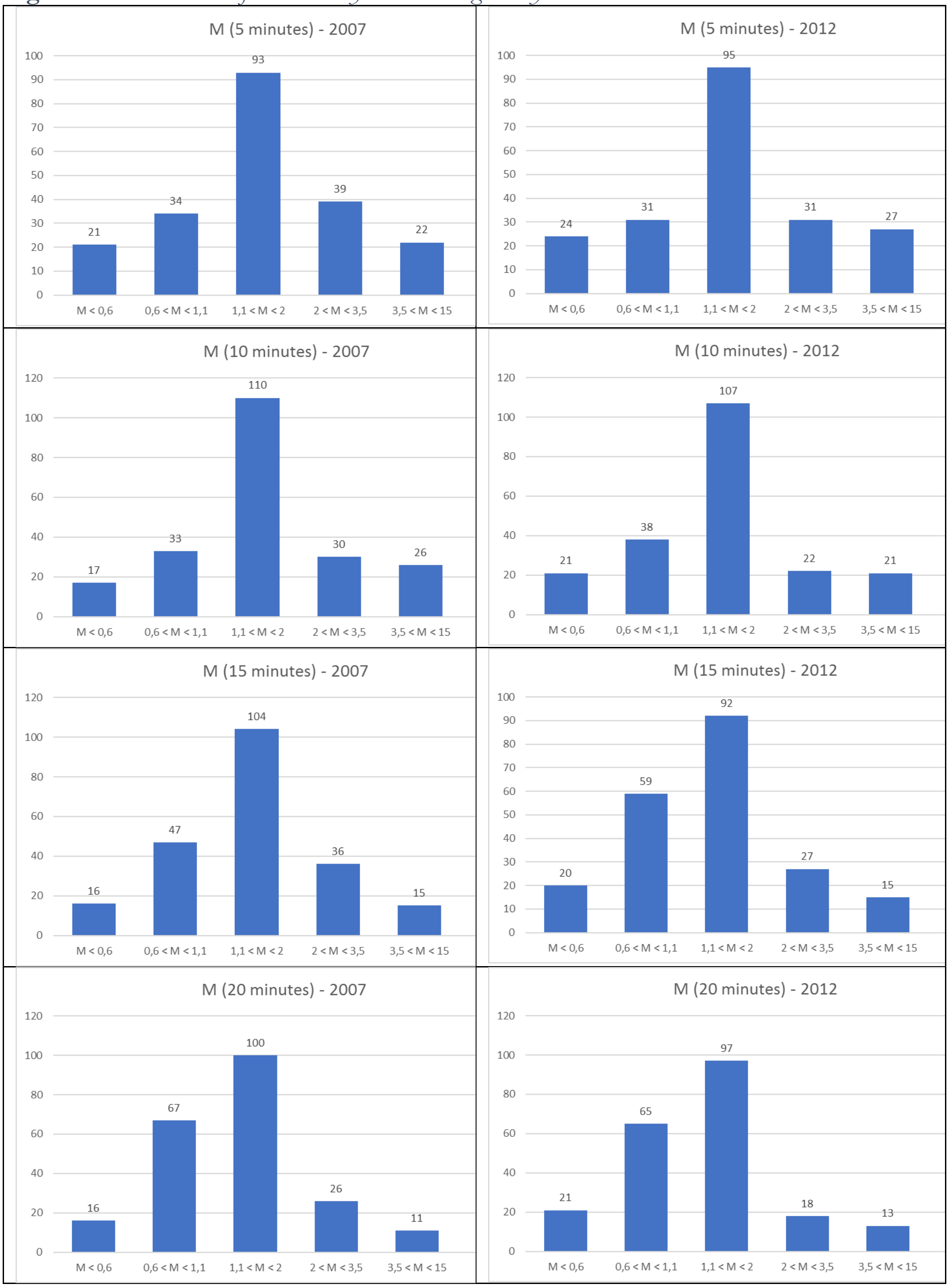
Source: Compiled by the authors

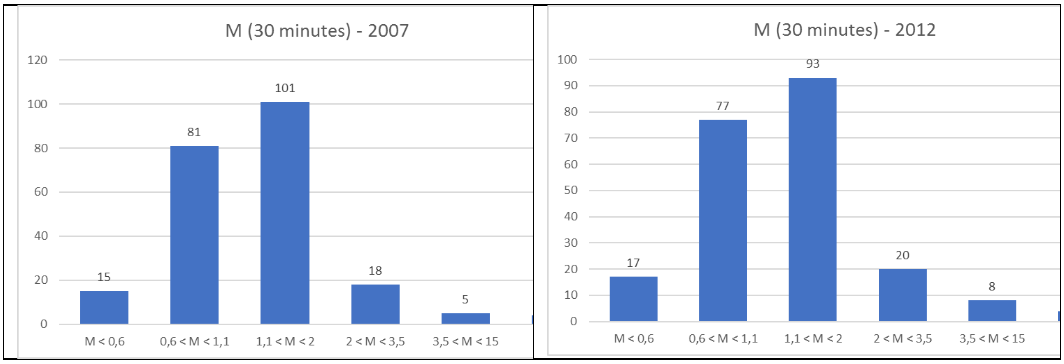
Table 7 – *20 industries with the largest % decrease in agglomeration in Sardinia between 2007 and 2012*

Industry code	Industry description	Employees	Plants	M (5 minutes)	M (10 minutes)	M (15 minutes)	M (20 minutes)	M (30 minutes)	National rank
212	MANUFACTURE OF MEDICINAL CHEMICAL AND BOTANICAL PRODUCTS	-96,23	-81,25	-100,00	-100,00	-100,00	-100,00	-100,00	46
559	OTHER ACCOMODATION	-73,05	-34,78	-88,61	-89,74	-96,05	-85,06	-83,01	2
871	RESIDENTIAL NURSING CARE FACILITIES	-2,32	-51,47	-63,03	-60,38	-77,01	-61,88	-62,50	155
89	MINING AND QUARRYING N.E.C.	-12,18	-14,29	-66,94	-30,28	-60,73	-29,46	-51,60	41
491	PASSENGER RAIL TRANSPORT, INTERURBAN	-75,33	-87,50	57,14	-50,00	-71,56	-78,69	-74,73	99
244	MANUFACTURE OF BASIC PRECIOUS AND OTHER NON-FERROUS METALS; REPROCESSING OF NUCLEAR FUELS	-21,82	14,29	-42,26	-42,29	-45,09	-49,45	-28,74	164
139	MANUFACTURE OF OTHER TEXTILES	-27,39	-28,08	-50,41	-53,63	-49,29	-37,44	-7,09	129
203	MANUFACTURE OF PAINTS, VARNISHES AND SIMILAR COATINGS, PRINTING INK AND MASTICS	-35,07	-34,29	-81,40	-52,00	-28,26	-20,65	-12,37	48
411	PROJECT MANAGEMENT ACTIVITIES RELATED TO CONSTRUCTION	-91,47	-65,66	-17,33	-25,00	-39,74	-55,56	-50,22	91
161	SAWMILLING AND PLANING OF WOOD	-50,52	-46,30	-49,06	-50,00	-37,38	-25,00	-21,62	47
257	MANUFACTURE OF CUTLERY, HAND TOOLS AND GENERAL HARDWARE	-8,59	-10,53	-28,25	-27,52	-43,20	-45,08	-38,70	124
532	COURIER ACTIVITIES	49,36	51,43	-33,93	-39,78	-37,65	-37,65	-30,34	104
553	CAMPING GROUNDS, RECREATIONAL VEHICLE PARKS AND TRAILER PARKS	-25,78	-16,67	-33,17	-28,33	-32,30	-29,45	-25,97	135
106	MANUFACTURE OF GRAIN MILL PRODUCTS, STARCHES AND STARCH PRODUCTS	-15,69	-38,78	-52,25	-29,65	-13,24	-34,34	-13,98	17
822	ACTIVITIES OF CALL CENTERS	6,57	-9,73	-43,05	-30,26	-21,86	-18,55	-22,91	38
799	OTHER RESERVATION SERVICE AND RELATED ACTIVITIES	21,53	33,67	-10,03	-38,18	-27,99	-32,91	-16,77	173
172	MANUFACTURE OF CORRUGATED PAPER AND PAPERBOARD AND OF CONTAINERS OF PAPER AND PAPERBOARD	-14,27	-18,18	-17,11	-13,91	-31,65	-30,08	-32,46	128
332	INSTALLATION OF INDUSTRIAL MACHINERY AND EQUIPMENT	-3,02	4,67	-26,05	-34,88	-20,83	-27,34	-12,93	43
81	QUARRYING OF STONE, SAND AND CLAY	-39,78	-38,46	-35,30	-33,14	-16,03	-22,80	-3,70	96
243	CASTING OF SEMI-FINISHED STEEL PRODUCTS	-57,23	-42,86	0,58	-12,69	-33,09	-41,53	-22,63	147

Source: Compiled by the authors

Figure 1 – *Distribution of M results by distance range and year*



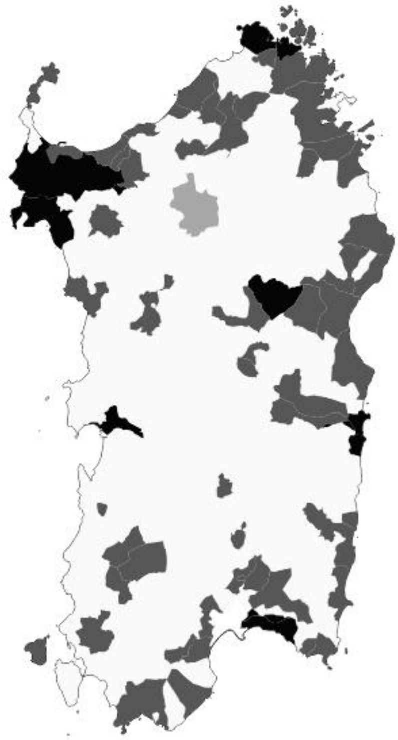


Source: Compiled by the authors

Figure 2 – *Manufacture of basic precious and other non-ferrous metals; reprocessing of nuclear fuels (244) in 2007*

Source: Compiled by the authors

Figure 3 – *Manufacture of jewellery, bijouterie and related articles (321) in 2007*

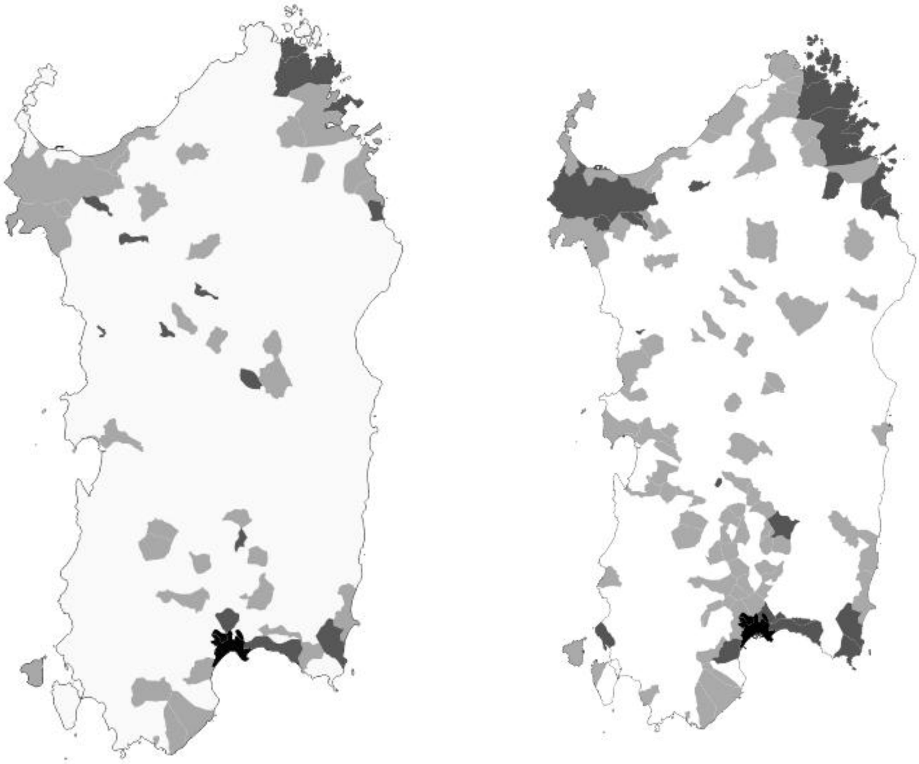


Source: Compiled by the authors

Figure 4 – *Manufacture of agricultural and forestry machines (283) in 2007 and 2012*

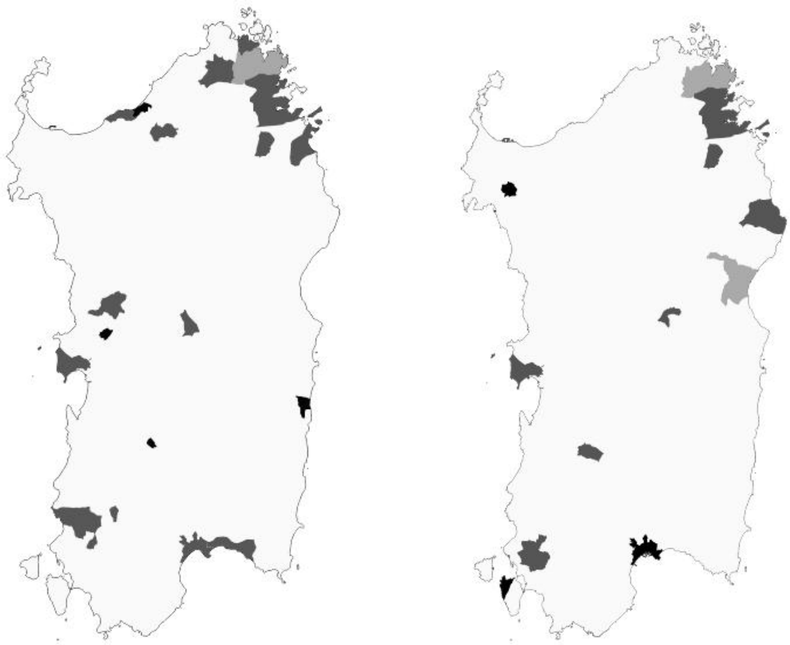
Source: Compiled by the authors

Figure 5 – *Landscape care and mantainance service activities (813) in 2007 and 2012*



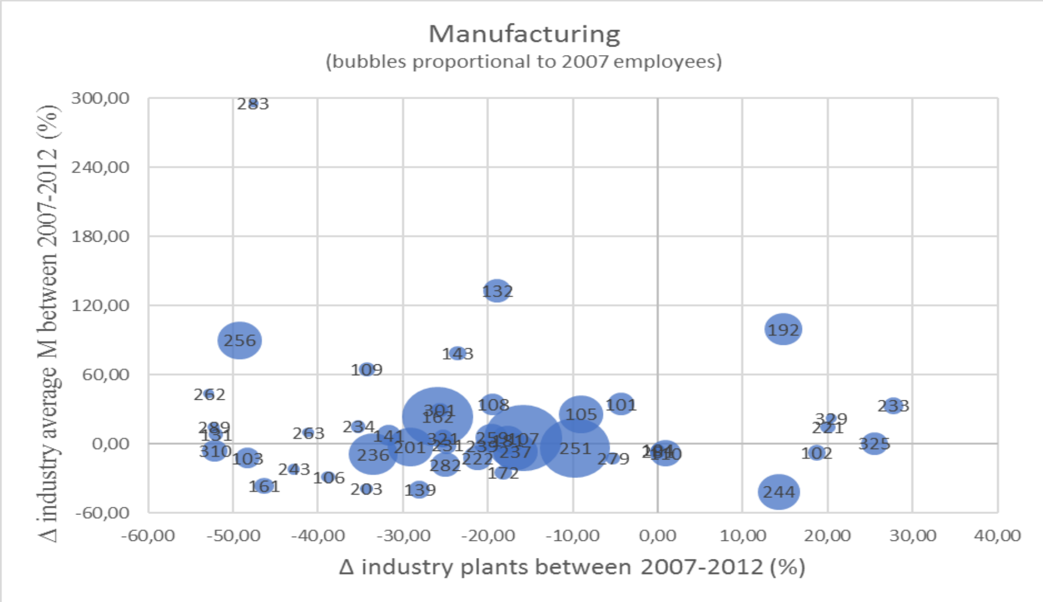
Source: Compiled by the authors

Figure 6 – *Other accomodation (559) in 2007 and 2012*



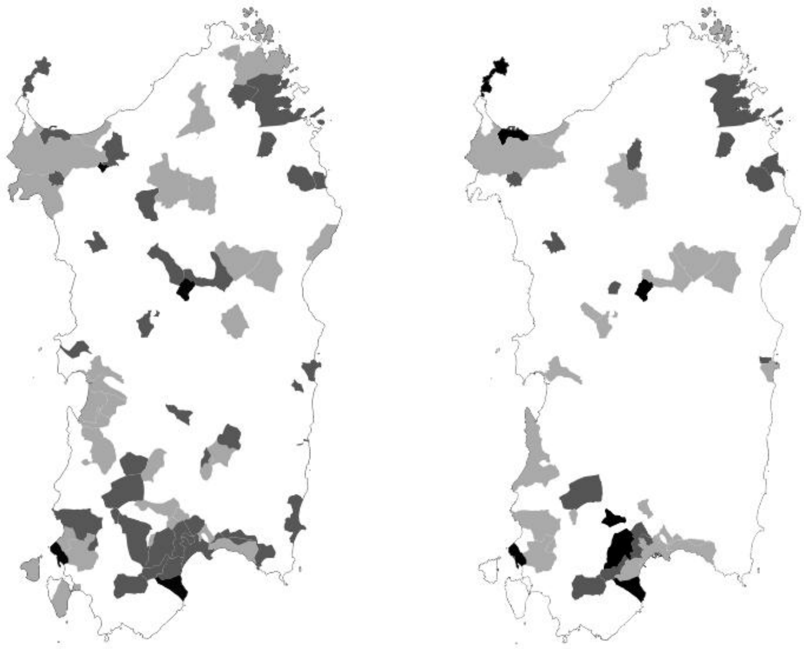
Source: Compiled by the authors

Figure 7 – *Change in agglomeration vs change in industry size for manufacturing industries (2007-2012)*



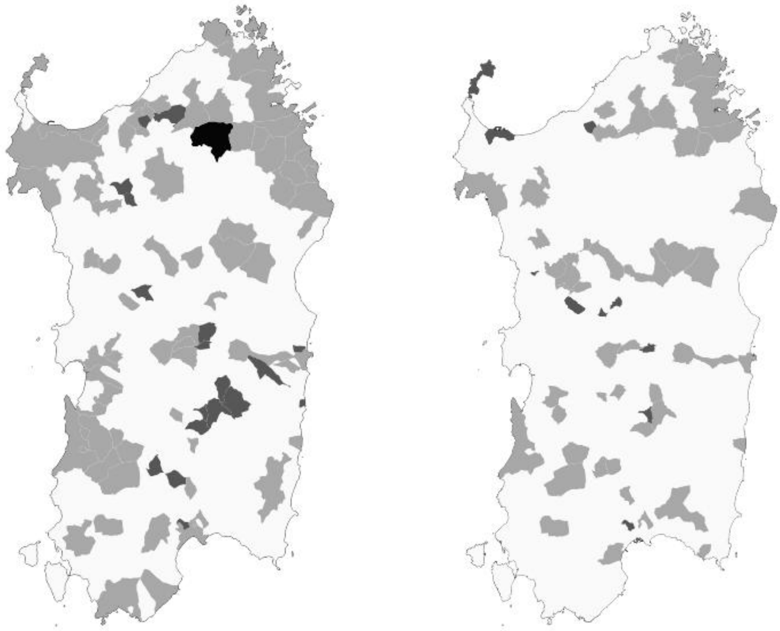
Source: Compiled by the authors

Figure 8 – *Treatment and coating of metals; machining (256) in 2007 and 2012*



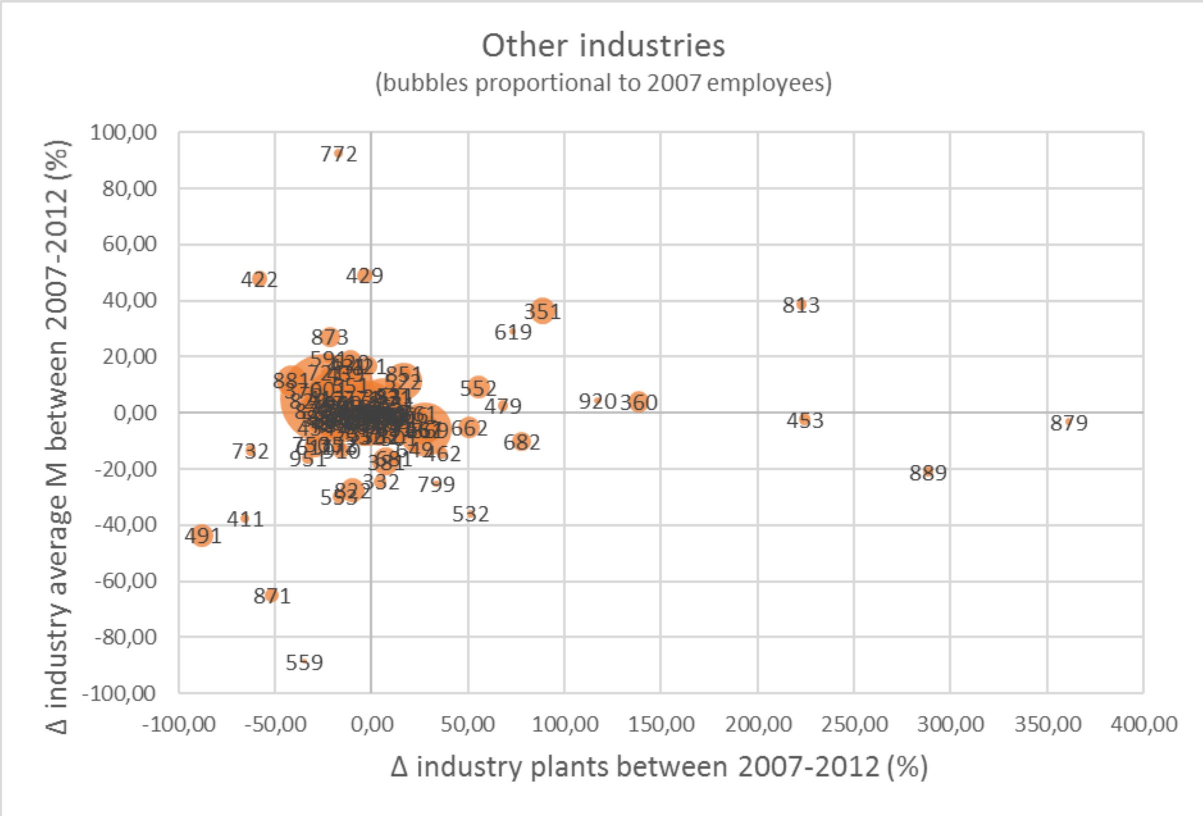
Source: Compiled by the authors

Figure 9 – *Sawmilling and planing of wood (161) in 2007 and 2012*



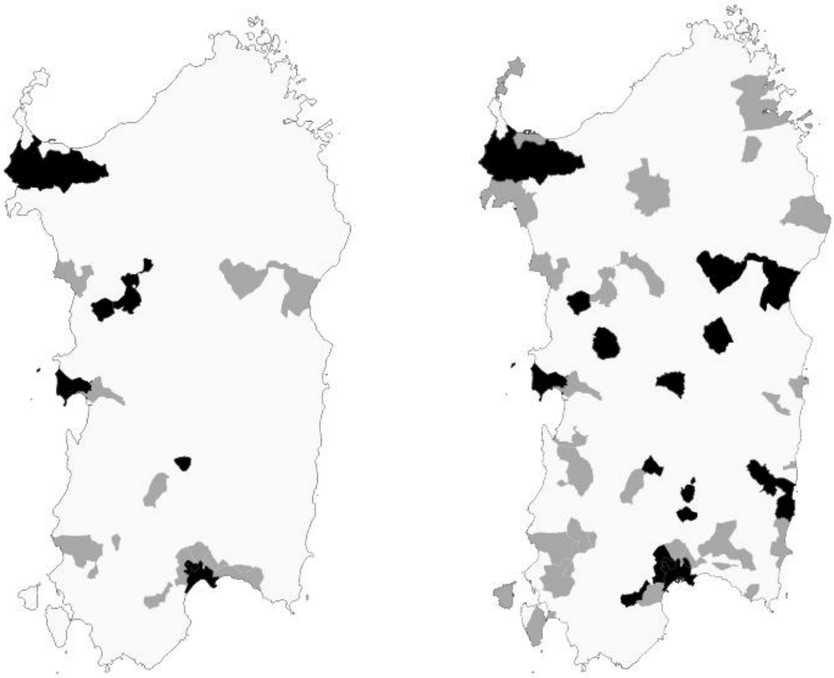
Source: Compiled by the authors

Figure 10 – *Change in agglomeration vs change in industry size for other industries (2007-2012)*



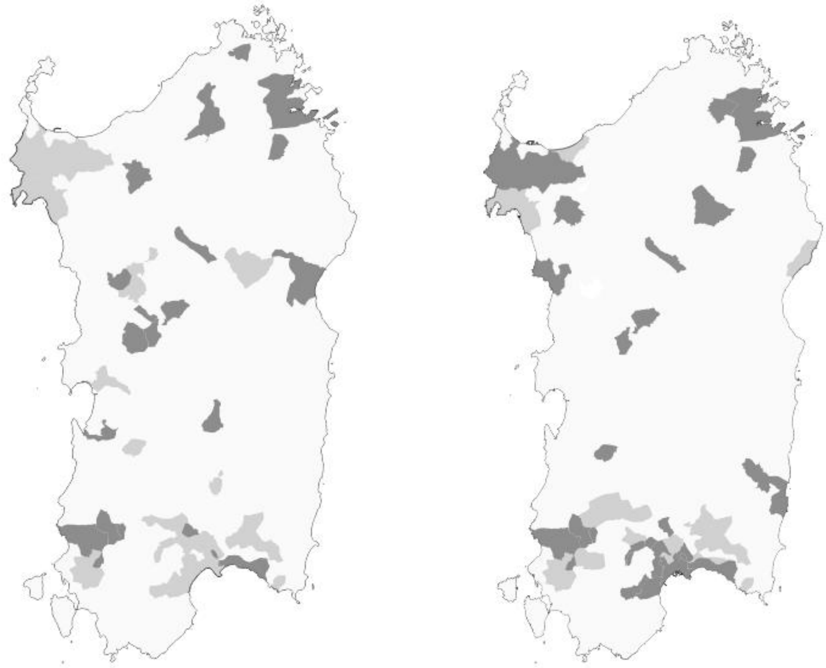
Source: Compiled by the authors

Figure 11 – *Courier activities (532) in 2007 and 2012*



Source: Compiled by the authors

Figure 12 – *Pre-primary education (851) in 2007 and 2012*



Source: Compiled by the authors

Ultimi Contributi di Ricerca CRENoS

I Paper sono disponibili in: http://www.crenos.unica.it

21/09 *Andrea Caria, Fabio Cerina, Marco Nieddu,* "Political Selection and Monetary Incentives in Local Parliamentary Systems"
21/08 *Giorgio Garau, Alessio Tola, Maria Veronica Camerada, Salvatore Lampreu, Silvia Carrus,* "Economic and social polarization dynamics in the EU"
21/07 *Emanuela Marrocu, Raffaele Paci, Stefano Usai,* "Direct and indirect effects of universities on European regional productivity"
21/06 *Pasqualina Arca, Gianfranco Atzeni, Luca Deidda,* "The Signalling Role of Trade Credit on Loan Contracts: Evidence from a Counterfactual Analysis"
21/05 *Grazia Sveva Ascione, Laura Ciucci, Claudio Detotto, Valerio Sterzi,* Do universities look like patent trolls? An Empirical Study of University Patent Infringement Litigation in the United States
21/04 *Michele Battisti, Massimo Del Gatto, Antonio Francesco Gravina, Christopher F. Parmeter,* "Robots versus labor skills: a complementarity / substitutability analysis"
21/03 *William Addessi, Marco Delogu,* "Infrastructure Accumulation in Developing Countries: the Role of the Informal Sector"
21/02 *Luca De Benedictis, Vania Licio, Anna Maria Pinna,* "From the historical Roman road network to modern infrastructure in Italy"
21/01 *Silvia Balia, Rinaldo Brau, Marco G. Nieddu,* "Depowering Risk: Vehicle Power Restriction and Teen Driver Accidents in Italy"
20/08 *Giampiero M. Gallo, Demetrio Lacava, Edoardo Otranto,* "On Classifying the Effects of Policy Announcements on Volatility"
20/07 *Luc Bauwens, Edoardo Otranto,* "Modelling Realized Covariance Matrices: a Class of Hadamard Exponential Models"
20/06 *Demetrio Lacava, Giampiero M. Gallo, Edoardo Otranto,* "Measuring the Effects of Unconventional Policies on Stock Market Volatility"
20/05 *Gianfranco Atzeni, Luca G. Deidda, Marco Delogu, Dimitri Paolini,* "Drop-out decisions in a cohort of Italian university students"
20/04 *Emanuela Marrocu, Raffele Paci, David Rigby, Stefano Usai,* "Smart Specialization Strategy: any relatedness between theory and practice?
20/03 *Giorgio Garau, Stefano Deriu,* "Total Factor Productivity and Relative Prices: the case of Italy"
20/02 *Fabio Cerina, Alessio Moro, Michelle Rendall,* "A Note on Employment and Wage Polarization in the U.S."
20/01 *Elias Carroni, Dimitri Paolini,* "Business models for streaming platforms: content acquisition, advertising and users"
19/16 *Daniela Sonedda*, "Regional variation in apprenticeship and permanent employment rates: which causes?"
19/15 *Daniela Sonedda,* "Regional disparities in the functioning of the labour markets"
19/14 *Bianca Biagi, Barbara Dettori, Raffaele Paci, Stefano Usai,* "Economic development in Sardinia: overcoming the insularity gap"
19/13 *Miguel Casares, Luca Deidda, Jose E. Galdon-Sanchez,* "On financial frictions and firm market power"
19/12 *Massimiliano Bratti, Maurizio Conti, Giovanni Sulis,* "Employment Protection and Firm-provided Training: Quasi-experimental Evidence from a Labour Market Reform"
19/11 *Jessica Goldberg, Mario Macis, Pradeep Chintagunta*, "Incentivized Peer Referrals for Tuberculosis Screening: Evidence from India"
19/10 *Julio J. Elías, Nicola Lacetera, Mario Macis,* "Paying for Kidneys? A Randomized Survey and Choice Experiment"
19/09 *Fabio Cerina, Elisa Dienesch, Alessio Moro, Michelle Rendall,* "Spatial Polarization"

www.crenos.unica.it

ISBN 9788868513795