



Detecting space–time agglomeration processes over the Great Recession using firm-level micro-geographic data

Giulio Cainelli¹ · Roberto Ganau^{1,2} · Yuting Jiang³

Received: 24 January 2020 / Accepted: 30 June 2020
© The Author(s) 2020

Abstract

We analyze the spatio-temporal agglomeration dynamics that occurred in the Italian manufacturing industry during the recent period of the Great Recession. To study this phenomenon, we employ three different statistical methods—namely, Ellison and Glaeser’s index of industrial geographic concentration, the spatial K -function, and the space–time K -function—, and rely on a large sample of geo-referenced, single-plant manufacturing firms observed over the period 2007–2012. First, we demonstrate that different statistical techniques can lead to (very) different results. Second, we find that most Italian manufacturing sectors experienced spatial dispersion processes during the period of the Great Recession. Finally, we show that space–time dispersion processes occurred at small spatial distances and short time horizon, although we do not detect statistically significant space–time interactions.

Keywords Agglomeration dynamics · Spatial methods · Manufacturing industry · Italy

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10109-020-00332-4>) contains supplementary material, which is available to authorized users.

✉ Roberto Ganau
roberto.ganau@unipd.it; r.ganau1@lse.ac.uk

✉ Yuting Jiang
47-218@gdudf.edu.cn

Giulio Cainelli
giulio.cainelli@unipd.it

¹ Department of Economics and Management “Marco Fanno”, University of Padova, Via del Santo 33, 35123 Padua, Italy

² Department of Geography and Environment, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK

³ Department of Financial Mathematics and Statistics, Guangdong University of Finance, Yingfu Lu 527, Longdong, Guangzhou, Guangdong Province, China

1 Introduction

The spatial agglomeration of economic activities is a key feature of the economic geography of many countries, regions, and clusters (Porter 1990). Well-known examples of such a phenomenon are Silicon Valley and Route 128 (Saxenian 1994), the carpet manufacturing industry in Dalton (Krugman 1991), and the Italian industrial districts (Brusco 1982; Becattini 1990; Cainelli and Zoboli 2004).

The most recent literature on spatial agglomeration has focused on two main research topics. The first one refers to the empirical determinants of the geographic concentration of production activities, and tries to identify the reasons why some geographic areas are more agglomerated than others. According to the Marshallian tradition (Marshall 1920), these factors are generally identified in knowledge spillovers, input sharing, and labor market pooling (Strange and Rosenthal 2001). The second literature stream refers to the effects of spatial agglomeration, and, precisely, of different types of local externalities, on firms' economic performance. Thus, it aims at understanding whether location within an agglomerated area generates positive returns on the economic performance of firms (Henderson 2003; Martin et al. 2011), and, consequently, on the economic dynamism and growth of territories as a whole (Glaeser et al. 1992; Henderson et al. 1995; Combes 2000).

A commonality characterizing most of these studies is that spatial agglomeration is treated as a static—i.e., time-invariant—phenomenon. To the best of our knowledge, only few contributions have investigated empirically the temporal dynamics of spatial agglomeration processes, that is whether and how the spatial agglomeration of firms changes over time (e.g., Arbia et al. 2010; Kang 2010; Arbia et al. 2014), and how these changes influence its firm-level economic returns (e.g., Martin et al. 2011; Cainelli et al. 2016). Indeed, the temporal dynamics is a key dimension of spatial agglomeration (Ellison and Glaeser 1997), which is a complex process changing over both space and time. In fact, some industries experience processes of geographic clustering/concentration, while others spread over space. Moreover, geographic concentration/dispersion processes can accelerate during some years (time clustering), while they can reduce during others (time dispersion). Spatial agglomeration may accelerate over time during the initial stages of the product life-cycle of a local industry/cluster, when a rapid increase in the number of new firms is generally observed (Klepper and Graddy 1990; Klepper 1996; Kang 2010); in years of 'sudden changes' associated with historical accidents, technological revolutions or discoveries, structural transformations (e.g., industrialization), external shocks (e.g., natural disasters, economic or financial crises); and, finally, during some specific periods of the business cycle (Arbia et al. 2010; Kang 2010; Henderson et al. 2018). On the contrary, spatial agglomeration may decelerate in years characterized by mature stages of the product life-cycle of a local industry/cluster, or during periods without significant technological advances or structural changes. In all these cases, the underlying mechanism is the same, i.e., the entry (exit) of firms into (from)

the local industry/cluster that may change the agglomeration structure over both space and time. In other words, both the rate at which firms' birth and death take place and their spatial location can have effects not only on the geographic concentration of economic activities, but also on the intensity of clustering processes (Devereux et al. 2004). This implies that spatial agglomeration cannot be investigated as a static phenomenon, given its intrinsic dynamic nature, such that the temporal dimension of agglomeration—the so-called time clustering—cannot be ignored. Moreover, it should be investigated at the sectoral level, since manufacturing sectors, which may be characterized by different levels of spatial agglomeration, do not behave homogeneously over both space and time.

The aim of this paper is to investigate the agglomeration dynamics that has characterized the Italian manufacturing industry during the period of the Great Recession. We employ a large sample of geo-referenced, single-plant manufacturing firms observed over the period 2007–2012, and located in the Italian continental territory, to analyze the spatial and temporal dimensions of agglomeration processes at the two-digit sector level, as well as their potential interactions. The choice of focusing on the period 2007–2012 is justified by the fact that the effects of the Great Recession have been particularly relevant for the Italian economy during these years. Indeed, during this time period, firm demography—birth and death rates—changed significantly in many sectors, thus affecting their agglomeration structures.

We investigate agglomeration over time using three different statistical approaches. First, we compute the Ellison and Glaeser's (1997) index of industrial geographic concentration for the years 2007 and 2012 at the two-digit sector level by adopting three different spatial units of analysis: the region, corresponding to the level 2 of the *Nomenclature des Unités Territoriales Statistiques* (NUTS) adopted by the European Union (EU); the province, corresponding to the NUTS-3 level; and the Local Labor Market (LLM), corresponding to a functional area defined according to economic—i.e., commuting patterns of workers—rather than administrative criteria.

As it is well-known, the Ellison and Glaeser's (1997) index suffers from the so-called Modifiable Areal Unit Problem (MAUP) that refers to the discretionary choice of the spatial unit used to analyze geographic-based phenomena (Arbia 1989; Amrhein 1995; Wong and Amrhein 1996). In fact, the use of pre-defined geographic units can introduce statistical biases in this kind of analysis. As suggested by the literature, a possible solution to the MAUP consists in relying on micro-geographic data, and adopting statistical methods that treat the space as a continuum (Arbia 2001).

Based on these insights, the second approach we adopt to study agglomeration over time consists in estimating Ripley's (1976) spatial K -function that exploits geographic information on the location (latitude and longitude coordinates) of firms. The spatial K -function is a statistical tool employed in the context of spatial points pattern analysis, and, precisely, it is a distance-based method that estimates the expected number of point events realizing within a certain distance from any randomly sampled point event. In our particular context, point events are represented by firms, and the spatial K -function allows us to evaluate whether sector-specific concentration/dispersion patterns took place—with respect to the benchmark case of complete spatial randomness (CSR) in the location choice of firms—in a given year. Therefore, this approach allows us to evaluate the geographic scale at which a

sector shows a clustering/dispersion pattern, if any, in a ‘single moment of time,’ but without imposing a priori spatial boundaries to the area under investigation (Marcon and Puech 2003; Duranton and Overman 2005; Arbia et al. 2008; Marcon and Puech 2010; Kosfeld et al. 2011; Albert et al. 2012; Scholl and Brenner 2016). Given the ‘static’ nature of the spatial K -function, the temporal dynamics of agglomeration processes can be analyzed in this context simply by comparing the values of the estimated K -functions in different years. In other words, these univariate functions are used in a dynamic fashion by analyzing them separately, i.e., year by year, and then by assessing both whether sector-specific clustering/dispersion patterns took place in the pre-crisis year 2007 and in the last observational year 2012, and how they potentially changed in the year 2012 with respect to the year 2007. This method, however, does not allow us to describe the dynamics of agglomeration patterns taking place over time, namely the agglomeration processes occurring between the years 2007 and 2012. This can be a problem, since ‘different space–time processes can lead to resulting spatial patterns which look the same’ (Arbia et al. 2010, p. 312).

The third approach is aimed precisely at overcoming this limitation, and consists in estimating the space–time K -function (Diggle et al. 1995) that can be used to analyze simultaneously the spatial and temporal dimensions of agglomeration processes, as well as the potential existence of space–time interactions. Thus, the space–time K -function represents an extension of the purely spatial K -function with respect to the analysis of the temporal dimension of spatial points patterns. Specifically, it estimates the expected number of point events per unitary area in the region under investigation and per unit of time that fall at a spatial distance and at a time interval equal to or lower than a certain radius and a certain time interval, respectively, from a reference point (French et al. 2005). Thus, the space–time K -function allows us to identify not only the existence of concentration/dispersion processes in different years, but also the spatial and temporal threshold values at which these concentration/dispersion processes occur over both space and time—for example, a rapid acceleration over time in the process of spatial clustering of a sector. Therefore, this third statistical approach provides us with a more comprehensive tool to analyze the dynamics of agglomeration processes, as it allows us to evaluate whether and how sector-specific agglomerative structures have exhibited different degrees of clustering/dispersion at different spatial scales along the temporal dimension.

The rest of the paper is organized as follows. The second Section describes the dataset used in the empirical analysis. The third Section presents the statistical approaches employed. The fourth Section presents and discusses the empirical evidence on the agglomeration dynamics. The fifth Section concludes the work.

2 The dataset

The analysis employs a large sample of single-plant manufacturing joint-stock companies located in the Italian continental territory. The firm-level data are drawn from the *AIDA* database (*Bureau Van Dijk*) that provides personal information and balance sheet data for Italian firms.¹

The original sample of 230,198 firms was first cleaned by removing multi-plant firms.² The choice of focusing on single-plant firms is driven by the fact that the *AIDA* database provides the exact address of headquarters only. Therefore, the focus on single-plant firms allows us to study the agglomeration dynamics of firms by considering the exact location where the economic activity takes place. Second, firms without information on the exact address were removed, given the necessity to identify the pair of geographic coordinates for each individual observational unit. Third, firms located in the two main islands of Sicily and Sardinia, and into smaller islands, were removed. This choice is driven by the necessity of considering a continuous land space in order to evaluate firms' concentration/dispersion processes through micro-geographic, distance-based statistical approaches. The cleaning procedure left us with a final sample of 149,135 manufacturing firms observed over the period 2007–2012, corresponding to an unbalanced panel dataset of 614,220 observations.

The period of observation starts with the year 2007 that is generally regarded as a pre-crisis year, and ends with the year 2012 that corresponds to the first year the Italian economy entered a second wave of downturn after the recovery peak reached in 2011—see Figure A1 (Supplementary material 1).

The final sample includes firms operating in the two-digit sectors 10–33 of the NACE Rev. 2 classification of economic activities adopted by the EU, except for the two-digit sector '12—Manufacture of tobacco products'.³ Table A1 (Supplementary material 1) reports the sample distribution by two-digit sector, while Table A2 (Supplementary material 1) reports the sample distribution by sector and year, as well as the percentage change in the number of firms between the years 2007 and 2012.

¹ The choice of using the *AIDA* database is somehow compulsory. In fact, alternative firm-level databases providing information on the exact address of firms are not available in Italy for privacy reasons. As it is known, the greatest limitation of the *AIDA* database is that it considers only joint-stock companies, thus excluding partnerships. A potential consequence of this limitation is that many micro- and small-sized firms, which represent in Italy a large share of some sectors, could be ruled out. As a consequence, the representativeness of micro- and small-sized firms could not be good. However, this is not the case. In fact, the coverage of the *AIDA* database in terms of micro- and small-sized firms is generally sufficiently high. This is the reason why firm-level datasets drawn from the *AIDA* database have been used in many empirical studies on the determinants or the effects of spatial agglomeration in Italy (e.g., Cainelli and Lupi 2010; Cainelli and Ganau 2018).

² Single-plant firms in the *AIDA* database have been identified using information derived from the *ASIA Archive* provided by the Italian National Institute of Statistics (Istat) that collects selected information on the entire population of firms and local units operating in Italy.

³ The two-digit sector '12—Manufacture of tobacco products' has been excluded a priori from the analysis due to the very little number of firms operating in it, as well as their peculiar spatial distribution. According to the 2011 Industry and Services Census (Istat), there were only eight active local units operating in Italy, of which one in the Northern part of the country (Piedmont), six in the Central part of the country (two in Tuscany, one in Umbria, one in the Marches, two in Latium), and one in the Southern part of the country (Campania).

Overall, the number of observational units has increased by 3.82% from the pre-crisis year 2007 to the year 2012. Most two-digit sectors have observed an increase in the number of firms, ranging from 0.17%, concerning sector ‘32—Other manufacturing’, to 37.45%, concerning sector ‘33—Repair and installation of machinery and equipment’. On the contrary, eight out of the 23 sectors analyzed have shown a reduction in the number of firms, from -0.37% , concerning sector ‘18—Printing and reproduction of recorded media’, to -7.88% , concerning sector ‘14—Manufacture of wearing apparel’.

Table A3 (Supplementary material 1) complements the previous insights by providing the sample distribution by sector and NUTS-1 geographic area. Overall, 60.3% of sample firms is located in Northern regions, 22.2% is located in Central Italy, while only 17.5% is located in the South. As Table A4 (Supplementary material 1) shows, about all two-digit sectors have recorded a reduction in the average firm size—defined as average number of employees per firm—between the years 2007 and 2012, ranging from -0.27% , concerning sector ‘24—Manufacture of basic metals’, to -20.41% , concerning sector ‘23—Manufacture of other nonmetallic mineral products’. Only two sectors have shown an increase in the average firm size, namely sectors ‘19—Manufacture of coke and refined petroleum products’ (24.34%), and ‘21—Manufacture of basic pharmaceutical products and pharmaceutical preparations’ (5.76%).

Interesting insights emerge looking at firms’ survival rate—see Table A5 (Supplementary material 1), which reports the rate of firms observed in year $t = 2007, \dots, 2011$ and survived over the subsequent periods $t + n$, with $n = 1, \dots, 5$.⁴ First, the survival rate has decreased over time from $t + 1$ to $t + n$. Second, the one-year survival rate for firms observed in the year 2007 is higher than the corresponding value for firms observed in the subsequent years. Third, the one-year survival rate for firms observed in 2010 is slightly higher than the corresponding values for firms observed in the years 2009 and 2011. This last evidence could probably depend on the fact that the Italian economy reached a recovery peak in 2011 before entering a new phase of downturn in 2012.

Figure A2 (Supplementary material 1) maps the spatial distribution of firms by two-digit manufacturing sector in the years 2007 and 2012. Finally, Tables A6–A10 (Supplementary material 1) highlight that the *AIDA* sample provides a good representativeness of the Italian population of active local units included in the *ASIA Archive*.

3 Statistical modeling

3.1 The index of industrial concentration

The first step of our empirical analysis of agglomeration over time is based on Ellison and Glaeser’s (1997) index of industrial geographic concentration. This index

⁴ The survival rate over the period $t, t + n$ is defined as the share of firms observed at time t and survived at time $t + n$ over firms observed at time t .

allows for the cross-sector comparison of the degree of geographic concentration, and, specifically, is employed to identify the clustering behavior in terms of spatial concentration/dispersion of two-digit manufacturing sectors in the years 2007 and 2012, as well as to evaluate how agglomeration patterns potentially changed between the two years.

We have considered three different spatial units to calculate the sector-specific concentration index. First, two administrative geographies defined at the NUTS levels 2 and 3, and corresponding to 18 regions and 90 provinces, respectively.⁵ Second, the LLM that consists of 559 functional areas.⁶ Formally, the index for the two-digit sector $s = 1, \dots, 23$ at time $t = 2007, 2012$ is computed as follows (Ellison and Glaeser 1997):

$$\hat{\gamma}_{st} = \left\{ \sum_{g=1}^G (p_{sgt} - p_{gt})^2 - \left[\left(1 - \sum_{g=1}^G p_{gt}^2 \right) \sum_{k=1}^K z_{skt}^2 \right] \right\} / \left[\left(1 - \sum_{g=1}^G p_{gt}^2 \right) \left(1 - \sum_{k=1}^K z_{skt}^2 \right) \right] \tag{1}$$

where p_{sgt} denotes the employment share in sector s and geography $g = 1, \dots, G$ at time t ; p_{gt} denotes the employment share in geography g at time t ; and z_{skt} denotes the employment share of firm k in sector s at time t .

This index provides a simple and easy-to-understand measure to explore sector-specific clustering behaviors. Given the benchmark case of CSR in the location choice of firms for $E(\hat{\gamma}_s) = 0$, then positive values of the index provide evidence of concentration, while negative values provide evidence of firms locating more diffusely than expected.

However, similarly to the many other region-based indexes of industrial geographic concentration—e.g., the Gini index used by Krugman (1991), or that proposed by Maurel and Sédillot (1999), it suffers from a main shortcoming that is the MAUP. The MAUP refers to the discretionary choice of the spatial partition used to analyze geographic-based phenomena. In the context of the analysis of spatial agglomeration and clustering dynamics, the MAUP emerges because neither administrative regions nor (functional) LLMs can necessarily coincide with the real economic areas where firms’ location processes take place. Consequently, the use of spatial units that differ in shape and size, and are characterized by pre-defined geographic boundaries, can introduce statistical biases related to both the level of aggregation and the geographic scale (Arbia 1989).

⁵ The NUTS-2 insular regions of Sicily and Sardinia, and the NUTS-3 regions located within these two main islands, have been excluded a priori from the analysis. During most of the period investigated, Italy was split into 107 NUTS-3 regions, nine (eight) of which located in Sicily (Sardinia). Considering the abovementioned exclusion criteria, the final number of NUTS-3 regions covered is equal to 90 spatial units.

⁶ Italian LLMs are defined according to the classification adopted by Istat in the 2001 Industry and Services Census that, identified 686 geographic units. The exclusion of all insular municipalities has led to drop 77 LLMs located in Sicily, 45 LLMs located in Sardinia, two LLMs located in the Ischia island (Campania), and two LLMs located in the Elba island (Tuscany). In addition, the cleaning procedure of the firm-level data has led to drop firms located in the municipality of San Marcello Pistoiese (Tuscany), such that the final number of LLMs covered is equal to 559 spatial units.

3.2 The spatial K -function

Following Arbia (2001), a possible solution to relax the MAUP consists in relying on micro-geographic data. This means that spatial agglomeration and clustering dynamics are analyzed at the level of the individual firms, rather than considering spatially-aggregated patterns of industries. The literature suggests to rely on a class of spatial points statistics, namely the K -functions—originally proposed by Ripley (1976), and widely employed in ecology, molecular biology, and epidemiology—, in order to evaluate the geographic scale at which an industry shows a clustering pattern, if any (Marcon and Puech 2003; Duranton and Overman 2005; Arbia et al. 2008; Marcon and Puech 2010; Kosfeld et al. 2011; Albert et al. 2012; Scholl and Brenner 2016). This type of spatial statistical approach is based on the idea of using the firm as the spatial unit of analysis, and of treating the space as continuous, rather than using pre-defined spatial areas.⁷

Therefore, the second step of the analysis is performed by employing Ripley's (1976) spatial K -function to evaluate whether sector-specific clustering patterns took place in—and, potentially, how they changed between—the years 2007 and 2012. Ripley's (1976) K -function is a distance-based method that measures the spatial concentration/dispersion of point events—in this case, firms—by counting the number of neighboring points j occurring within a circle of radius r centered at each reference point i , with $j \neq i$, and then by comparing the observed pattern with the one that would be expected in a CSR situation. Formally, the estimate of the K -function for the two-digit sector s at time $t = 2007, 2012$ can be defined as follows:

$$\hat{K}_{st}(r) = \frac{1}{\hat{\lambda}_{st} N_{st}} \sum_i \sum_{j \neq i} \frac{I(d_{x_{st}^i, x_{st}^j})}{w_{x_{st}^i, x_{st}^j}} \quad (2)$$

where N_{st} denotes the total number of firms operating in sector s at time t and located in the area of the study region (W), with $\hat{\lambda}_{st} = N_{st}/\text{Area}_W$ denoting its estimated density; $d_{x_{st}^i, x_{st}^j}$ denotes the distance in kilometers between each pair of firms i and j operating in sector s and observed at time t —denoted by x_{st}^i and x_{st}^j as spatial points identified by their geographic coordinates; $I(\cdot)$ denotes an indicator function taking value of 1 if $d_{x_{st}^i, x_{st}^j} \leq r$, that is whether the distance between a pair of firms i and j ($d_{x_{st}^i, x_{st}^j}$) is lower than or equal to the radius r , and value of 0 otherwise; $w_{x_{st}^i, x_{st}^j}$ denotes the edge correction parameter that defines the length of the overlap between the circle with radius r centered in the reference firm x_{st}^i and passing through the firm x_{st}^j which lies within the study region (Ripley 1977). The edge correction term avoids biased estimates of $\hat{K}_{st}(r)$ which may occur in proximity to the boundaries of the study region, where increases in r are not accompanied by increases in the number

⁷ Economists and economic geographers have relied on micro-geographic data to analyze also agglomeration-related externalities on the performance of firms (e.g., Sorenson and Audia 2000; Rosenthal and Strange 2003; Baldwin et al. 2008; Cainelli and Lupi 2010; Eriksson 2011; Duschl et al. 2014; Duschl et al. 2015; Cainelli and Ganau 2018).

of firms—indeed, the number of firms can be lower in proximity to rather than at longer distances from the study region’s boundaries, and there are no firms outside the study region.

Having estimated the sector- and time-specific K -functions, it is possible to test for the location pattern characterizing the observed firms against the CSR hypothesis. Under the null hypothesis of CSR, then $\hat{K}_{st}(r) = \pi r^2$, such that it is possible to compute the difference between the empirical value of the K -function for the observed points pattern—namely, $\hat{K}_{st}(r)$ —and the theoretical value under the CSR hypothesis. Following Albert et al. (2012), this difference—defined as M -function—can be formalized as follows:

$$\hat{M}_{st}^{CSR}(r) = \hat{K}_{st}(r) - \pi r^2 \tag{3}$$

such that if $\hat{K}_{st}(r) > \pi r^2$, then the observed firms in sector s at time t will show a clustering pattern at a certain distance r , because the observed density of firms is greater than the theoretical one; if $\hat{K}_{st}(r) < \pi r^2$, then the observed firms will show a dispersion pattern; while if $\hat{K}_{st}(r) = \pi r^2$, then the observed firms will show a random distribution in the space.

Although the spatial K -function represents a valuable statistical tool for testing purposes, as it allows us to evaluate the extent of concentration/dispersion of firms with respect to a CSR scenario, it suffers from a main shortcoming related to its interpretation. In fact, the K -function requires to compare any value to πr^2 , but area values are difficult to interpret directly (Marcon and Puech 2003; Kosfeld et al. 2011). A solution to this problem comes from Besag (1977), who proposes the so-called L -function as a normalization of the K -function with respect to a zero-benchmark case. Formally, the spatial L -function can be defined as follows:

$$\hat{L}_{st}(r) = \sqrt{\frac{\hat{K}_{st}(r)}{\pi}} - r \tag{4}$$

such that if $\hat{L}_{st}(r) = 0$, then the observed firms will show a random distribution in space; if $\hat{L}_{st}(r) > 0$, then the observed firms will show a clustering pattern; while if $\hat{L}_{st}(r) < 0$, then the observed firms will show a dispersion pattern.

It is worth noting, however, that we rely on the L -function simply to exploit its advantages in terms of graphical interpretation with respect to the K -function, such that it is used only to provide a better graphical comparison between the empirical values of the estimated K -functions for the years 2007 and 2012 by two-digit manufacturing sector in order to evaluate whether changes in the degree of concentration/dispersion occurred between the two years.

3.3 The space–time K -function

Although the use of micro-geographic data in the context of spatial points pattern analyses allows us to relax MAUP-related biases, a second issue emerges with respect to the temporal dimension characterizing clustering/dispersion processes

of firms. In fact, the spatial agglomeration of economic activities is the result of a dynamic process that evolves over both space and time.

As these processes occur over time, (sector-specific) agglomerative structures may potentially exhibit different degrees of clustering at different spatial scales along the temporal dimension. The contributions by Getis (1964) and Getis and Boots (1978) are among the first ones to highlight that time matters, besides space, in the analysis of spatial events. They suggest that it is the temporal evolutionary perspective that can help understanding the (observed) resulting spatial structure. Indeed, any spatial structure evolves over time, and, in particular, similar spatial patterns can be the result of very different space–time processes (Arbia et al. 2010). This line of reasoning fits perfectly the analysis of the spatial agglomeration of economic activities. As firm-level demographic and localization phenomena present both a spatial and a temporal feature, industrial clustering/dispersion processes are likely to evolve along both dimensions, and, as a consequence, the observed spatial agglomeration of firms could be the result of their interaction.

The joint analysis of the spatial and temporal processes underlying agglomerative structures has received little attention in the empirical literature. As underlined by Arbia et al. (2010, pp. 311–312), ‘[t]ime series methods have generally disregarded the spatial dimension while spatial clustering models have been essentially static and have only analysed the outcome of the dynamic adjustments as it is observed in one single moment of time.’ To the best of our knowledge, only few contributions have analyzed industrial clustering processes by dealing with the MAUP, and by accounting simultaneously for the spatial and temporal dimensions, as well as for their potential interaction. In particular, Arbia et al. (2010) focus on the long-run localization process of firms located in Rome (Italy), and operating in the information and communication technology sector. Kang (2010) studies relevant clusters of manufacturing and services sectors in the Columbus Metropolitan Statistical Area (USA). Finally, Arbia et al. (2014) consider the entry and exit dynamics of pharmaceutical and medical devices firms located in the Veneto region (Italy) over the period 2004–2009. These three contributions limit their investigation to a particular industrial sector, or a particular sub-national territory—either a city or a region—of a country. On the contrary, we try to provide a more general overview of spatial agglomeration processes by considering the entire manufacturing industry, and by looking at the whole Italian (continental) territory.

Following these previous contributions, the third statistical approach consists in estimating the space–time K -function (Diggle et al. 1995) in order to evaluate the existence of space–time concentration/dispersion processes, as well as to identify the threshold values at which these behavioral patterns take place over the spatial and temporal horizons. Specifically, the space–time K -function allows us to identify simultaneously at which spatial distance and point in time concentration/dispersion processes occur, i.e., it allows us to identify the geographic scale at which concentration/dispersion process take place over the temporal dimension. In addition, this approach allows us to test for the potential existence of space–time interactions, i.e., to test whether increases/decreases in the incidence of point events are spatially and temporally localized.

The space–time K -function measures the expected number of points per unitary area in the study region and per unit of time falling at a spatial distance and at a time interval equal to or lower than a radius r and a time interval t , respectively, from a reference point (French et al. 2005). An unbiased edge-corrected estimator of the space–time K -function for the two-digit sector s over the observational period $t = 1, \dots, T$ —namely, from the year 2007 to the year 2012—can be defined as follows (Diggle et al. 1995):

$$\hat{K}_s(r, t) = \frac{1}{\hat{\lambda}_s^{RT} N_s} \sum_i \sum_{j \neq i} \frac{I_r(d_{x_s^i, x_s^j}) I_t(t_{x_s^i, x_s^j})}{w_{x_s^i, x_s^j} v_{x_s^i, x_s^j}} \tag{5}$$

where N_s denotes the total number of firms operating in the two-digit sector s and located in the area of the study region (W), with $\hat{\lambda}_s^{RT} = N_s / (\text{Area}_W T)$ denoting the spatial and temporal joint intensity of the points process; the terms $d_{x_s^i, x_s^j}$ and $t_{x_s^i, x_s^j}$ denote the spatial distance and the temporal distance, respectively, between each pair of firms i and j operating in sector s ; the terms $I_r(\cdot)$ and $I_t(\cdot)$ denote indicator functions taking value of 1 if $d_{x_s^i, x_s^j} \leq r$ and $t_{x_s^i, x_s^j} \leq t$, respectively, and value of 0 otherwise; the terms $w_{x_s^i, x_s^j}$ and $v_{x_s^i, x_s^j}$ denote edge correction parameters, with $w_{x_s^i, x_s^j}$ defining the length of the overlap between the circle with radius r centered in the reference firm x_s^i and passing through the firm x_s^j which lies within the study region, and $v_{x_s^i, x_s^j}$ defining the time segment of length t centered at the firm x_s^i that lies within the total observed duration time between $t = 1$ and $t = T$.

Similar estimators can be derived for the purely spatial and temporal processes defining the space–time K -function, namely $\hat{K}_s^R(r)$ and $\hat{K}_s^T(t)$:

$$\hat{K}_s^R(r) = \frac{1}{\hat{\lambda}_s^R N_s} \sum_i \sum_{j \neq i} \frac{I_r(d_{x_s^i, x_s^j})}{w_{x_s^i, x_s^j}} \tag{6}$$

$$\hat{K}_s^T(t) = \frac{1}{\hat{\lambda}_s^T N_s} \sum_i \sum_{j \neq i} \frac{I_t(t_{x_s^i, x_s^j})}{v_{x_s^i, x_s^j}} \tag{7}$$

where $\hat{\lambda}_s^R = N / \text{Area}_W$ denotes the spatial intensity, and captures the number of points per unit area of the study region; $\hat{\lambda}_s^T = N / T$ denotes the temporal intensity, and captures the number of points per unit time; and all other terms are defined as for Eq. (5).

Having estimated the sector-specific space–time K -functions, it is possible to test for the independence of the spatial and temporal processes under the hypothesis that $\hat{K}_s(r, t) = \hat{K}_s^R(r) \hat{K}_s^T(t)$ in the absence of space–time interaction. The baseline test statistic takes the following functional form (Gatrell et al. 1996):

$$\hat{D}_s(r, t) = \hat{K}_s(r, t) - \hat{K}_s^R(r) \hat{K}_s^T(t) \tag{8}$$

that is proportional to the increased number of points within distance r and time t compared to a process with the same spatial and temporal structures, but no space–time interaction. If $\hat{D}_s(r, t) > 0$, then we can assume that space–time clustering takes place, while if $\hat{D}_s(r, t) < 0$, then we can assume that space–time dispersion takes place (Kang 2010). Furthermore, visual inspection through a three-dimension plot of $\hat{D}_s(r, t)$ against the spatial and temporal dimensions allows us to uncover the scale and nature of the dependence.

An alternative, and more intuitive, transformation of the functional defined in Eq. (8) has been proposed by Diggle et al. (1995) and French et al. (2005) to consider relative quantities rather than absolute numbers. The so-called ‘Diggle function’ allows for a perspective plot of the $\hat{D}_s(r, t)$ surface through the following functional form:

$$\hat{D}_s^0(r, t) = \frac{\hat{D}_s(r, t)}{\hat{K}_s^R(r)\hat{K}_s^T(t)} \quad (9)$$

that is proportional to the relative increase in the number of points within distance r and time t compared to a process with the same spatial and temporal structures, but no space–time interaction. Similarly to the previous case, the function $\hat{D}_s^0(r, t)$ can be visually inspected through a three-dimension plot against the spatial and temporal dimensions in order to evaluate the existence of space–time interaction in the observed points process.

Furthermore, statistical inference on the empirical values of $\hat{D}_s(r, t)$ can be drawn through a Monte Carlo approach. Since the exact distribution of $\hat{D}_s(r, t)$ is unknown, its variance cannot be evaluated theoretically and no statistical test can be adopted. In order to deal with this issue, Diggle et al. (1995) suggest to obtain m simulated spatio-temporal points patterns under the null hypothesis of no space–time interaction in order to compute m different estimates of $\hat{D}_s(r, t)$, such that the variance of the m estimates, namely $\widehat{\text{Var}}_s(r, t)$, can be used as an estimator of the variance of $\hat{D}_s(r, t)$. Having retrieved the estimated variance of $\hat{D}_s(r, t)$, it is possible to compute the ‘standardized residuals’ as follows:

$$\hat{R}_s(r, t) = \frac{\hat{D}_s(r, t)}{\sqrt{\widehat{\text{Var}}_s(r, t)}} \quad (10)$$

which gives a measure of space–time interaction representing the excess number of points of $\hat{K}_s(r, t)$ relative to $\hat{K}_s^R(r)\hat{K}_s^T(t)$. The advantage of plotting the ‘standardized residuals’ against $\hat{K}_s^R(r)\hat{K}_s^T(t)$ is that a two-dimension plot is easier to visualize, even though the corresponding spatial and temporal scales are not explicit. Under the hypothesis of no space–time interaction, then $E[\hat{R}_s(r, t)] = 0$ and $\text{Var}[\hat{R}_s(r, t)] = 1$, and one would expect to observe approximately 95% of the values lying in the interval $[-2, +2]$ (French et al. 2005; Arbia et al. 2010). Thus, substantial values of $\hat{R}_s(r, t)$ lying outside the interval $[-2, +2]$ indicate the presence of space–time

Table 1 Ellison and Glaeser’s (1997) index of geographic concentration in 2007 and 2012

| NACE Rev. 2 | NUTS-2 Region | | | NUTS-3 Region | | | LLM | | |
|----------------|------------------|--------|------------------------------------|------------------|--------|------------------------------------|------------------|--------|------------------------------------|
| | $\hat{\gamma}_s$ | | $\Delta\hat{\gamma}_s^{2007-2012}$ | $\hat{\gamma}_s$ | | $\Delta\hat{\gamma}_s^{2007-2012}$ | $\hat{\gamma}_s$ | | $\Delta\hat{\gamma}_s^{2007-2012}$ |
| | 2007 | 2012 | | 2007 | 2012 | | 2007 | 2012 | |
| 10 | 0.0233 | 0.0232 | -0.10 | 0.0094 | 0.0076 | -18.78 | 0.0041 | 0.0032 | -20.83 |
| 11 | 0.0367 | 0.0367 | 0.22 | 0.0170 | 0.0178 | 5.18 | 0.0055 | 0.0057 | 3.09 |
| 13 | 0.0617 | 0.0637 | 3.26 | 0.0400 | 0.0401 | 0.13 | 0.0389 | 0.0390 | 0.23 |
| 14 | 0.0205 | 0.0241 | 17.56 | 0.0099 | 0.0098 | -0.71 | 0.0071 | 0.0066 | -6.58 |
| 15 | 0.1287 | 0.1275 | -0.93 | 0.0472 | 0.0452 | -4.15 | 0.0371 | 0.0347 | -6.52 |
| 16 | 0.0137 | 0.0143 | 4.20 | 0.0100 | 0.0089 | -11.28 | 0.0050 | 0.0045 | -10.38 |
| 17 | 0.0027 | 0.0015 | -45.57 | 0.0039 | 0.0029 | -24.36 | 0.0027 | 0.0025 | -7.62 |
| 18 | 0.0061 | 0.0026 | -56.64 | 0.0092 | 0.0060 | -34.30 | 0.0092 | 0.0061 | -33.68 |
| 19 | 0.0281 | 0.0482 | 71.53 | 0.0073 | 0.0075 | 2.85 | -0.0005 | 0.0036 | 0.41 |
| 20 | 0.0319 | 0.0296 | -7.22 | 0.0188 | 0.0213 | 12.98 | 0.0162 | 0.0195 | 20.47 |
| 21 | 0.0857 | 0.0978 | 14.07 | 0.0748 | 0.0764 | 2.20 | 0.0851 | 0.0898 | 5.52 |
| 22 | 0.0078 | 0.0089 | 14.48 | 0.0036 | 0.0033 | -9.16 | 0.0027 | 0.0028 | 3.17 |
| 23 | 0.0209 | 0.0179 | -14.33 | 0.0094 | 0.0069 | -27.25 | 0.0075 | 0.0061 | -18.55 |
| 24 | 0.0344 | 0.0330 | -4.27 | 0.0153 | 0.0177 | 15.80 | 0.0062 | 0.0074 | 20.98 |
| 25 | 0.0025 | 0.0027 | 5.88 | 0.0021 | 0.0021 | -2.21 | 0.0012 | 0.0011 | -6.31 |
| 26 | 0.0166 | 0.0142 | -14.42 | 0.0239 | 0.0176 | -26.25 | 0.0225 | 0.0166 | -26.16 |
| 27 | 0.0090 | 0.0053 | -41.21 | 0.0050 | 0.0036 | -26.94 | 0.0046 | 0.0026 | -44.69 |
| 28 | 0.0137 | 0.0137 | -0.29 | 0.0032 | 0.0033 | 1.86 | 0.0020 | 0.0020 | 0.89 |
| 29 | 0.0465 | 0.0378 | -18.82 | 0.0244 | 0.0140 | -42.61 | 0.0147 | 0.0116 | -20.88 |
| 30 | 0.0329 | 0.0367 | 11.66 | 0.0168 | 0.0233 | 38.58 | 0.0114 | 0.0157 | 37.93 |
| 31 | 0.0640 | 0.0552 | -13.73 | 0.0310 | 0.0308 | -0.76 | 0.0219 | 0.0218 | -0.40 |
| 32 | 0.0263 | 0.0219 | -16.63 | 0.0204 | 0.0191 | -6.36 | 0.0172 | 0.0161 | -6.24 |
| 33 | 0.0092 | 0.0110 | 19.16 | 0.0058 | 0.0053 | -8.48 | 0.0040 | 0.0033 | -16.27 |

Authors’ elaboration on *AIDA* data. The reference for the calculation of the Ellison and Glaeser’s (1997) index is the *AIDA* sample of firms. Variations of the sector-specific indexes between the years 2007 and 2012 are defined in percentage terms

interaction characterizing the observed points pattern, and this space–time structure can be visually interpreted from the plots of $\hat{D}_s(r, t)$ or $\hat{D}_s^0(r, t)$.

However, the interpretation of the ‘standardized residuals’ plot could be misleading in the case of highly dependent residuals. Therefore, a final Monte Carlo-based test can be performed in order to assess the existence of space–time interaction. This test consists in comparing the observed sum of the functionals $\hat{D}_s(r, t)$ over all values of r and t with the empirical distribution of the m sums of the corresponding simulated estimates of $\hat{D}_s(r, t)$ over all values of r and t , such that there is evidence of overall space–time interaction if the observed sum shows a particularly high-ranked position relative to the simulated sums (Arbia et al. 2010).

4 Empirical evidence

4.1 Detecting spatial concentration/dispersion through Ellison and Glaeser's (1997) index

Table 1 reports the calculated sector-specific Ellison and Glaeser's (1997) indexes of geographic concentration for the years 2007 and 2012, as well as their percentage change, considering the three different geographic scales previously discussed. The first interesting insight concerns the high level of heterogeneity characterizing the degree of geographic concentration in the Italian manufacturing industry. In fact, the 2007 and 2012 indexes present high levels of cross-sector variation both within the same spatial unit of analysis, as well as across the three geographies considered. Indeed, the comparison of the sector-specific indexes among the three spatial units suggests clearly that MAUP-related biases are likely to affect the evaluation of the extent of industrial geographic concentration. This is a second relevant insight emerging from our analysis, which also offers a key justification for the use of micro-geographic data to investigate agglomeration-related phenomena, i.e., for analyzing the agglomeration dynamics at the firm level rather than considering spatially-aggregated industrial patterns.

Following the classification proposed by Ellison and Glaeser (1997), who identify sectors as 'lowly' concentrated if $\hat{\gamma}_{st} < 0.02$, 'moderately' concentrated if $0.02 \leq \hat{\gamma}_{st} \leq 0.05$, and 'highly' concentrated if $\hat{\gamma}_{st} > 0.05$, it emerges how individual sectors exhibit a different degree of concentration with respect to the different spatial units. For example, the Italian manufacturing industry in 2007 consisted of four 'highly' concentrated sectors—namely, '13—Manufacture of textiles', '15—Manufacture of leather and related products', '21—Manufacture of basic pharmaceutical products and pharmaceutical preparations', '31—Manufacture of furniture'—when considering NUTS-2 regions, but only sector '21—Manufacture of basic pharmaceutical products and pharmaceutical preparations' presented a value of the concentration index higher than 0.05 when considering the NUTS-3 and LLM level geographies—see also Figure A3 (Supplementary material 1).⁸

In addition, the calculation based on different geographic scales leads to different sector-specific temporal dynamics of concentration. As an example, sector '14—Manufacture of wearing apparel' exhibits an increase in the degree of geographic concentration between the years 2007 and 2012 when considering the NUTS-2 geographic level, while a reduction is observed when the NUTS-3 and LLM geographic levels are considered. The opposite dynamics characterizes, for example, sector '20—Manufacture of chemicals and chemical products'—see also Figure A4 (Supplementary material 1). In other words, the use of different geographic

⁸ It is not surprising that these manufacturing sectors result to be the most geographically concentrated at the NUTS-3 and LLM level, since the main Italian industrial districts/clusters operate in these production activities. In fact, the textile districts of Prato and Carpi, the shoes districts of Fusignano and San Mauro Pascoli, and the biomedical/pharmaceutical sector of the metropolitan area of Milan are specialized in these productions (Brusco 1982; Becattini 1990; Cainelli and Zoboli 2004).

units—regions, provinces, and LLMs—may lead to different results in terms of spatial concentration/dispersion patterns. This means that these findings—and, generally speaking, analyses of agglomeration patterns based on arbitrarily pre-defined spatial units—must be interpreted carefully, since they can lead to a serious misunderstanding of the underlying economic phenomena.

More generally, this means that the choice of the geographic boundaries may have relevant consequences in many empirical studies focusing on the determinants or the effects of spatial agglomeration. It is known that some studies conducted their empirical analysis at the NUTS-3 or LLM level (for example, the majority of Italian studies), while others adopted as unit of analysis the NUTS-2 region (for example, the majority of European studies). According to our findings, the choice of the geographic unit of analysis does not seem to be neutral, thus leading to results and conclusions which can be very different. This is something that has never been emphasized sufficiently.

4.2 Dealing with the MAUP: evidence based on the spatial K -function

The analysis based on Ellison and Glaeser's (1997) index has highlighted not only the existence of heterogeneity in the spatial agglomeration dynamics of firms operating in different sectors, but also that region-based measures of clustering suffer from MAUP-related biases that may lead to (very) imprecise conclusions. Drawing on this last insight, in this section we present and discuss the results obtained by employing the spatial K -function.⁹ Table 2 summarizes the key insights from the analysis, while Figure A5 (Supplementary material 1) reports the plots of the estimated values of the spatial K - and M -functions by two-digit sector for the years 2007 and 2012.

As summarized in Table 2, the empirical evidence suggests that all Italian manufacturing sectors were characterized by a spatial concentration pattern occurring at all distances from zero to about 250 km in both years 2007 and 2012, as $\hat{K}_s(r) > \pi r^2$ for all sectors in both years. Thus, the evidence suggests that the location of Italian manufacturing firms is not 'random,' but rather seems to follow a geographic clustering pattern. In other words, the Italian manufacturing firms (within each sector) tend to be spatially concentrated rather than spatially dispersed or randomly located in the space. This result could appear surprising, as one may expect that only few 'traditional' manufacturing sectors—such as textiles, leather, footwear, furniture—would be characterized by spatial concentration patterns, since the most important Italian industrial districts are specialized in these production activities (Brusco 1982; Becattini 1990; Cainelli and Zoboli 2004). On the contrary, and according to our findings based on the spatial K -function, the non-random spatially concentrated distribution of manufacturing firms seems to be a general empirical 'regularity'.

However, despite this 'common' feature, the detected geographic patterns can differ significantly in intensity and shape across sectors. A deeper analysis of Table 2

⁹ The analysis based on the spatial K -function has been performed using the 'spatstat' library available for the R software.

Table 2 Summary results of the estimated spatial K - and M -functions for the years 2007 and 2012

| Functional | $\hat{K}_s(r)$ | | $\hat{K}_s(r) > \pi r^2$ | Concentration/dispersion | $\hat{K}_s(r) > \pi r^2$ | \hat{K}_s^{2007} versus \hat{K}_s^{2012} |
|--------------------|--------------------------------|------|--------------------------------|--------------------------|---|--|
| | 2007 | 2012 | | | | |
| Year | Concentration/dispersion | | | | | |
| Pattern | Concentration/dispersion | | | | | |
| <i>NACE Rev. 2</i> | | | | | | |
| 10 | Concentration at all distances | Yes | Concentration at all distances | Yes | $\hat{K}_s^{2007} < \hat{K}_s^{2012}$ up to 40 km, then $\hat{K}_s^{2007} > \hat{K}_s^{2012}$ | |
| 11 | Concentration at all distances | Yes | Concentration at all distances | Yes | $\hat{K}_s^{2007} > \hat{K}_s^{2012}$ at all distances | |
| 13 | Concentration at all distances | Yes | Concentration at all distances | Yes | $\hat{K}_s^{2007} > \hat{K}_s^{2012}$ at all distances, but in [190, 200] and [220, 230] km | |
| 14 | Concentration at all distances | Yes | Concentration at all distances | Yes | $\hat{K}_s^{2007} > \hat{K}_s^{2012}$ at all distances | |
| 15 | Concentration at all distances | Yes | Concentration at all distances | Yes | $\hat{K}_s^{2007} < \hat{K}_s^{2012}$ at all distances | |
| 16 | Concentration at all distances | Yes | Concentration at all distances | Yes | $\hat{K}_s^{2007} > \hat{K}_s^{2012}$ at all distances | |
| 17 | Concentration at all distances | Yes | Concentration at all distances | Yes | $\hat{K}_s^{2007} > \hat{K}_s^{2012}$ up to 40 km, then $\hat{K}_s^{2007} < \hat{K}_s^{2012}$ | |
| 18 | Concentration at all distances | Yes | Concentration at all distances | Yes | $\hat{K}_s^{2007} > \hat{K}_s^{2012}$ at all distances | |
| 19 | Concentration at all distances | Yes | Concentration at all distances | Yes | $\hat{K}_s^{2007} > \hat{K}_s^{2012}$ up to 70 km and in [130, 190] km; $\hat{K}_s^{2007} < \hat{K}_s^{2012}$ in [70, 130] and [190–240] km | |
| 20 | Concentration at all distances | Yes | Concentration at all distances | Yes | $\hat{K}_s^{2007} > \hat{K}_s^{2012}$ at all distances | |
| 21 | Concentration at all distances | Yes | Concentration at all distances | Yes | $\hat{K}_s^{2007} > \hat{K}_s^{2012}$ at all distances | |
| 22 | Concentration at all distances | Yes | Concentration at all distances | Yes | $\hat{K}_s^{2007} > \hat{K}_s^{2012}$ at all distances | |
| 23 | Concentration at all distances | Yes | Concentration at all distances | Yes | $\hat{K}_s^{2007} > \hat{K}_s^{2012}$ at all distances | |
| 24 | Concentration at all distances | Yes | Concentration at all distances | Yes | $\hat{K}_s^{2007} < \hat{K}_s^{2012}$ in [0, 20], [40, 80], [100, 110] km; $\hat{K}_s^{2007} > \hat{K}_s^{2012}$ in [20, 40], [80, 100], from 110 km | |
| 25 | Concentration at all distances | Yes | Concentration at all distances | Yes | $\hat{K}_s^{2007} > \hat{K}_s^{2012}$ at all distances | |
| 26 | Concentration at all distances | Yes | Concentration at all distances | Yes | $\hat{K}_s^{2007} > \hat{K}_s^{2012}$ at all distances | |
| 27 | Concentration at all distances | Yes | Concentration at all distances | Yes | $\hat{K}_s^{2007} > \hat{K}_s^{2012}$ at all distances | |
| 28 | Concentration at all distances | Yes | Concentration at all distances | Yes | $\hat{K}_s^{2007} > \hat{K}_s^{2012}$ at all distances | |

Table 2 (continued)

| Functional | $\hat{K}_s(r)$ | | \hat{K}_s^{2007} versus \hat{K}_s^{2012} | |
|------------|--------------------------------|--------------------------|--|--------------------------|
| | 2007 | 2012 | $\hat{K}_s(r) > \pi r^2$ | $\hat{K}_s(r) > \pi r^2$ |
| Year | | | | |
| Pattern | Concentration/dispersion | Concentration/dispersion | Concentration/dispersion | Concentration/dispersion |
| 29 | Concentration at all distances | Yes | Concentration at all distances | Yes |
| 30 | Concentration at all distances | Yes | Concentration at all distances | Yes |
| 31 | Concentration at all distances | Yes | Concentration at all distances | Yes |
| 32 | Concentration at all distances | Yes | Concentration at all distances | Yes |
| 33 | Concentration at all distances | Yes | Concentration at all distances | Yes |

Authors' elaboration on *AIDA* data. Sector-specific comparison of the estimated empirical K values between the years 2007 and 2012 is evaluated considering sector- and year-specific mean values of the estimated empirical K values averaged over distance bands of 10 km in the interval [0, 250]. Plots of the estimated spatial K - and M -functions are reported by two-digit manufacturing sector and year in Figure A5 (Supplementary material 1)

highlights some interesting insights. In fact, the comparison of the estimated empirical values of K in the years 2007 and 2012 suggests that 14 out of 23 two-digit sectors experienced a reduction in the degree of spatial concentration between the years 2007 and 2012 at all distances. This means that these sectors, although characterized by significant spatial concentration in both years, have experienced a slight process of spatial dispersion in the year 2012 with respect to the year 2007. The opposite dynamics has characterized sector ‘15—Manufacture of leather and related products’ that experienced an increase in the degree of spatial concentration at all distances in the year 2012 with respect to the year 2007. The remaining sectors present different dynamics: some of them show an increase (decrease) in the degree of spatial concentration between the years 2007 and 2012 at short distances, while a decrease (increase) at longer distances; other sectors exhibit more complicated dynamics characterized by alternating increases and reductions in the degree of spatial concentration between the years 2007 and 2012 over different distance intervals.¹⁰

Further interesting evidences can be derived from the analysis of the spatial L -function, which comes simply as a normalization of the spatial K -function, but has the advantage of providing a clearer graphical representation of the intensity/strength of concentration at each specific distance. Thus, the L -function allows us to identify and compare the intensity of concentration that has characterized each two-digit sector at different distances in the years 2007 and 2012. Figure A6 (Supplementary material 1) reports the plots of the sector-specific spatial L -functions estimated for the years 2007 and 2012. Two stylized facts emerge from the plots. First, it is confirmed that the majority of the Italian manufacturing sectors experienced a process of spatial dispersion during the years of the Great Recession, i.e., the distance-specific intensity of spatial concentration in the year 2012 has been lower than in the year 2007 for most two-digit sectors. For some sectors, this process of spatial dispersion has been concentrated in the first 50 km, thus affecting the geographic area with the highest firm density of the sector—in some cases, the main ‘pole’ of the industrial district/cluster. For other sectors, this process occurred at longer distances. The first pattern could be explained by a higher death rate of firms located within the ‘core’ of the industrial cluster, while the second pattern could be the result of a higher mortality of firms located in the ‘periphery’ of the cluster. Despite most sectors experienced a process of spatial dispersion, we find also evidence of few two-digit sectors that have been characterized by a general increase in the intensity of spatial concentration in the year 2012 with respect to the year 2007. This increase in the strength of spatial concentration has interested sector ‘15—Manufacture of leather and related products’, sector ‘17—Manufacture of paper and paper products’ (from about 40 km) and, to a lower extent, sector ‘30—Manufacture of other transport equipment’, while an alternating dynamics has characterized sector ‘19—Manufacture of coke and refined petroleum products’.

¹⁰ The sector-specific comparison of the estimated empirical K values between the years 2007 and 2012 is evaluated considering sector- and year-specific mean values of the estimated empirical K values averaged over distance bands of 10 km in the interval [0, 250].

The second stylized fact emerging from the analysis of the spatial L -function is that the intensity of concentration changes at different distances. For some ‘traditional’ sectors, such as textile, leather and furniture, the intensity of agglomeration tends to increase very rapidly, reaching high values within the first 10 to 20 km. This result is consistent with the literature on spatial agglomeration/clustering, which shows that most localization activities occur at distances below 10–20 km, and then tend to decline. This is the typical ‘spatial’ organization of many industrial districts/clusters, where a high number of small firms is co-located within a bounded geographic area. We find a similar pattern also for a high-tech sector such as the pharmaceutical one. In this case, the L -function rises up to a distance of 20 km, and then declines. In other words, also in this case the spatial structure of the sector is typical of an industrial cluster.

Finally, additional insights emerge also from Table A11 (Supplementary material 1), which reports the sector-specific minimum and maximum values of the estimated spatial L -functions with the corresponding distances for the years 2007 and 2012. According to Table A11, the maximum value of this function for the pharmaceutical sector is reached at a distance of about 45 km. In this sense, the pharmaceutical sector is not only the most spatially concentrated manufacturing sector in Italy, but also that characterized by the productive activity where the concentration grows more rapidly at short distances. This is not surprising, since these high-tech firms take advantage of knowledge spillovers whose intensity tends to increase within bounded geographic spaces. A similar result emerges also for the textile industry, where the maximum value of the L -function is reached at about 57 km.

More generally, each sector shows its own geographic pattern, whose intensity at short distances can change significantly. For the ‘traditional’ sectors and the pharmaceutical one, agglomeration effects at short distances may be particularly strong. Thus, it is likely that firms operating in these sectors benefit from intra-industry knowledge spillovers. For other sectors, the geographic clustering process takes place over longer distances that include one or more neighbor regions. It is likely that firms operating in sectors characterized by this pattern of geographic concentration exploit other types of locational advantages, which produce their effects at larger geographic scales. As recently suggested by the literature, some of these localization advantages can be identified with inter-industry knowledge spillovers or input–output (vertical) linkages which generally exert their effects at long distances (e.g., Cainelli and Ganau 2018).

As already mentioned, a relevant limit of this kind of analysis is that it does not allow us to describe the dynamics of agglomeration occurring between two different ‘moments of time,’ i.e., the so-called agglomeration over time. This can be a problem, since, at the sectoral level, the ‘same’ agglomeration pattern can be the result of quite different space–time dynamics. Understanding these processes is therefore quite important in order to identify the (potential) effects produced on these patterns by factors such as the stage of the product life-cycle of a local industry/cluster, historical accidents, technological revolutions, and external shocks (e.g., natural disasters and economic crises).

4.3 The space–time agglomeration dynamics

Agglomeration tends to evolve not only over space, but also over time. Indeed, firms' agglomeration dynamics is not a static phenomenon, such that its analysis requires to evaluate its spatial and temporal dimensions simultaneously. Moving from these premises, and drawing on the previous findings, this sub-section presents the results concerning the agglomeration dynamics obtained by employing the space–time K -function approach, and using the firm-level data for each observational period between the year 2007 and the year 2012.¹¹

Table 3 summarizes the key insights emerging from the analysis, while Figure A7 (Supplementary material 1) reports the plots of the estimated values of the $\hat{D}_s^0(r, t)$ functional, the $\hat{D}_s^1(r, t)$ functional, the 'standardized residuals' $\hat{R}_s(r, t)$ versus $\hat{K}_s^R(r)\hat{K}_s^T(t)$, and the empirical frequency distribution of the sum of the differences between the space–time K -function and the product of the separate space and time K -functions in 99 simulations resulted from the Monte Carlo test.

First, based on the visual inspection of the estimated $\hat{D}_s^0(r, t)$ functionals, we find that, overall, all two-digit manufacturing sectors recorded a space–time dispersion process that occurred with a higher intensity at relatively small distance and short time horizon.¹² This evidence suggests that the Italian manufacturing industry has been characterized by a general dispersion process of firms that occurred simultaneously in space and time, but at a small spatial scale and during the first years of the period analyzed. These evidences have two interesting implications. First, while it is generally recognized that agglomeration processes take place at short spatial distances (Cainelli and Ganau 2018), it is less clear that these dynamics may accelerate during short periods of time. A prejudice underlying the literature is that agglomeration is characterized by strong path-dependence. This is the reason why most studies assume agglomeration as a time-invariant phenomenon. Our analysis shows that this is not true, at least during a period of rapid changes such as the Great Recession. Unfortunately, we do not have a counterfactual. In fact, it is not unlikely that changes in agglomeration patterns could have occurred at a slower rate 'before' or 'after' the Great Recession. However, in a 'specific moment of time'—in our case, associated with the dramatic 2008–2009 economic and financial crisis—an acceleration of spatial dispersion in many manufacturing sectors can be observed. Similar processes might also occur during the initial stages of the product life-cycle of a local industry/cluster, characterized by a rapid increase in the number of new firms, or in years of 'sudden changes' associated with historical accidents, technological revolutions or discoveries, and structural transformations. The second implication of our

¹¹ The analysis based on the space–time K -function has been performed using the 'splancs' library available for the R software.

¹² It is worth noting that the $\hat{D}_s^0(r, t)$ functional does not measure the intensity of spatial concentration/dispersion, but rather describes the space–time dynamics of the agglomeration process. Therefore, while the spatial K - and L -functions provide a static picture of the spatial structure of each sector in a particular year (in our case, the years 2007 and 2012), the space–time K -function offers a dynamic description of the sector-specific agglomeration processes occurred over a period of time (in our case, between the years 2007 and 2012). Thus, we do interpret our findings in this sense.

Table 3 Summary results of the estimated space–time K -functions over the period 2007–2012

| Functional | $\hat{D}_s^0(r, t)$ | Space–time concentration/dispersion (model) | $\hat{R}_s(r, t)$ versus $\hat{K}_s^R(r)\hat{K}_s^T(t)$ | Monte Carlo test (p value) |
|-----------------------------|---------------------|---|---|-------------------------------|
| Pattern | | | | |
| Space–time interaction test | | | | |
| <i>NACE Rev. 2</i> | | | | |
| 10 | | Space–time dispersion at small distance and short time horizon, with time dispersion decreasing over time | (1) No significant interaction | 0.67 |
| 11 | | Space–time dispersion at small distance and short time horizon, with time dispersion decreasing over time | (1) No significant interaction | 0.46 |
| 13 | | Space–time dispersion at small distance and short time horizon, with time dispersion decreasing over time | (1) No significant interaction | 0.70 |
| 14 | | Space–time dispersion at small distance and short time horizon, with time dispersion decreasing over time | (1) No significant interaction | 0.93 |
| 15 | | Space–time dispersion at small distance over the period 2007–2008, then slight time clustering process | (2) No significant interaction | 0.76 |
| 16 | | Space–time dispersion at small distance and short time horizon, with time dispersion decreasing over time | (1) No significant interaction | 0.48 |
| 17 | | Space–time dispersion at small distance and short time horizon, with time dispersion decreasing over time | (1) No significant interaction | 0.74 |
| 18 | | Space–time dispersion at small distance and short time horizon, with time dispersion decreasing over time | (1) No significant interaction | 0.20 |
| 19 | | Space–time dispersion at small distance and short time horizon, with time dispersion decreasing over time | (1) No significant interaction | 0.59 |
| 20 | | Space–time dispersion at small distance and short time horizon, with time dispersion decreasing over time | (1) No significant interaction | 0.66 |
| 21 | | Space–time dispersion at small distance and short time horizon, with time dispersion decreasing over time | (1) No significant interaction | 0.51 |
| 22 | | Space–time dispersion at small distance and short time horizon, with time dispersion decreasing over time | (1) No significant interaction | 0.47 |
| 23 | | Space–time dispersion at small distance and short time horizon, with time dispersion decreasing over time | (1) No significant interaction | 0.35 |
| 24 | | Space–time dispersion at small distance and short time horizon, with time dispersion decreasing over time | (1) No significant interaction | 0.67 |
| 25 | | Space–time dispersion at small distance and short time horizon, with dispersion increasing over time | (3) No significant interaction | 0.19 |
| 26 | | Space–time dispersion at small distance and short time horizon, with time dispersion decreasing over time | (1) No significant interaction | 0.59 |
| 27 | | Space–time dispersion at small distance and short time horizon, with time dispersion decreasing over time | (1) No significant interaction | 0.64 |
| 28 | | Space–time dispersion at small distance and short time horizon, with time dispersion decreasing over time | (1) No significant interaction | 0.63 |
| 29 | | Space–time dispersion at small distance and short time horizon, with time dispersion decreasing over time | (1) No significant interaction | 0.72 |
| 30 | | Space–time dispersion at small distance over the period 2007–2008, then slight time clustering process | (2) No significant interaction | 0.47 |

Table 3 (continued)

| Functional | $\hat{D}_s^0(r, t)$ | | $\hat{R}_s(r, t)$ versus $\hat{K}_s^R(r) \hat{K}_s^T(t)$ | Monte Carlo test (p value) |
|-----------------------------|---|-----|--|-------------------------------|
| Pattern | Space–time concentration/dispersion (model) | | | |
| Space–time interaction test | | | | |
| 31 | Space–time dispersion at small distance and short time horizon, with time dispersion decreasing over time | (1) | No significant interaction | 0.64 |
| 32 | Space–time dispersion at small distance and short time horizon, with time dispersion decreasing over time | (1) | No significant interaction | 0.62 |
| 33 | Space–time dispersion at small distance and short time horizon, with time dispersion decreasing over time | (1) | No significant interaction | 0.94 |

Authors' elaboration on AIDA data. Plots of the estimated $\hat{D}_s^0(r, t)$ and $\hat{D}_s(r, t)$ functionals, 'standardized residuals,' and empirical frequency distribution of the sum of the differences between the space–time K -function and the product of the separate space and time K -functions resulted from the Monte Carlo tests are reported by two-digit manufacturing sector in Figure A7 (Supplementary material 1)

analysis concerns the impact of these dynamics on the structure and organization of local industries. We do not have direct evidences on these phenomena, but it is likely that a process of spatial dispersion at short distances and concentrated in few years has strongly changed the nature and the intensity of local externalities (the so-called Marshallian effects), inter-firm (horizontal and vertical) relationships, and the role and positioning of leader firms. This is coherent with some recent studies suggesting that Italian industrial clusters experienced a process of radical transformation during the Great Recession (e.g., De Marchi and Grandinetti 2014). Our evidences of spatial dispersion in many manufacturing sectors may be interpreted as a further ‘indirect’ support to this ‘vision’ of the evolution of the Italian industrial districts over the Great Recession.

Second, the analysis allows us to identify three main ‘behavioral’ patterns that have characterized different two-digit sectors. The first pattern (Model 1) concerns all but three sectors—namely, ‘15—Manufacture of leather and related products’, ‘25—Manufacture of fabricated metal products, except machinery and equipment’, and ‘30—Manufacture of other transport equipment’. The space–time dispersion process that characterized the sectors belonging to the first group occurred at small distance (within about 30 km) and short time horizon (2007–2009), and exhibited time dispersion decreasing over time (from 2009 to 2012) in its intensity. In addition, it is worth noting that sectors characterized by this general pattern did not behave exactly in the same way. In other words, even within this group it is possible to detect heterogeneity in the time dispersion pattern. As an example, sector ‘13—Manufacture of textiles’ recorded a slightly different time dispersion process that occurred smother over time with respect to the other sectors belonging to this first group. The second pattern (Model 2) concerns sectors ‘15—Manufacture of leather and related products’ and ‘30—Manufacture of other transport equipment’. These two sectors recorded a deep space–time dispersion that occurred at small distance and over the period 2007–2008, but then recorded a light process of time clustering over the period 2009–2012. It is not surprising that these dynamics were also detected through the spatial K -function approach. However, through the space–time K -function approach, we are also able to describe the time evolution of these processes. The third pattern (Model 3) concerns only sector ‘25—Manufacture of fabricated metal products, except machinery and equipment’, that recorded a space–time dispersion process at small distance (about 25 km) over the period 2007–2008, but time dispersion has been increasing from the year 2007 onwards. Figure 1 reports one example of the estimated $\hat{D}_s^0(r, t)$ functional for each identified ‘behavioral’ pattern.

These three different patterns do not seem to depend on specific characteristics of the sectors in terms of industrial organization, technology, or firm size composition. In other words, we are not able to detect any empirical regularity characterizing these patterns due to the fact that the observed firms belong to both traditional and high-tech industries. This is an interesting result that requires further investigation.

Finally, although the analysis highlights clearly the existence of space–time dispersion processes, inferential testing does not point to the existence of statistically significant space–time interaction—see the last two columns of Table 3, and

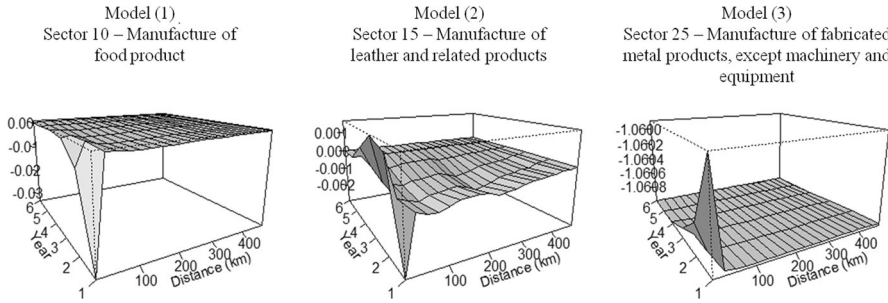


Fig. 1 Examples of 'behavioral' patterns— $\hat{D}_s^0(r, t)$ functional. Note: Estimated $\hat{D}_s^0(r, t)$ functionals for two-digit sectors '10—Manufacture of food products' (Model 1), '15—Manufacture of leather and related products' (Model 2), and '25—Manufacture of fabricated metal products, except machinery and equipment' (Model 3)

Table A12 (Supplementary material 1).¹³ Overall, the absence of space–time interaction could depend on the fact that the time period we analyze is too short, together with the reduced number of replications performed in the Monte Carlo procedure.¹⁴ Both issues will be addressed in future extensions of this work.

5 Conclusions

The spatial and temporal evolution of firms' agglomerates is one of the key phenomena occurring in many industrialized and emerging countries. Despite its relevance, only few studies have investigated the dynamics of agglomeration over both space and time. To the best of our knowledge, this is one of first papers that attempt to investigate empirically these processes by employing a battery of statistical methods, and using a large sample of geo-referenced, single-plant firms covering almost the entire Italian manufacturing industry.

In particular, our contribution to this literature is twofold. The first one is methodological. In fact, we show that different statistical techniques, generally employed in economic geography and regional economics for studying spatial agglomeration phenomena, can lead to different results. In the case of the Ellison and Glaeser's

¹³ As suggested by a Reviewer, the absence of space–time interaction for two-digit sectors belonging to Models (1) and (2) emerges clearly looking at the plots of the estimated $\hat{D}_s^0(r, t)$ functional, as the majority of points lies on the surface representing the 'no interaction' plane. On the contrary, the absence of space–time interaction appears as less obvious from a purely graphical inspection with respect to sector '25—Manufacture of fabricated metal products, except machinery and equipment', i.e., the only sector belonging to Model (3). In this case, the corresponding plot shows large and diffuse deviations over all horizons, but the p -value of the Monte Carlo test (which is equal to 0.19) does not allow us to reject the null hypothesis of no space–time interaction.

¹⁴ The Monte Carlo procedure is based on only 99 replications for reasons related to computational speed. Indeed, the size of our dataset is particularly large when compared to previous studies relying on similar methodologies.

(1997) index, this depends on two main issues, namely the MAUP, and the fact that the temporal dimension characterizing agglomeration processes is not accounted for adequately. The spatial K -function approach (Ripley 1976) allows us to relax MAUP-related issues, but does not help us solving the second problem, i.e., the empirical identification of the agglomeration dynamism. In this context, this can be done only by comparing the values of the estimated spatial K -functions in two separate moments of time. In other words, this statistical method does not allow us to describe the dynamics of agglomeration patterns taking place between different moments of time. The only approach allowing for this ‘temporal’ evaluation is the space–time K -function (Diggle et al. 1995), through which we can analyze simultaneously the spatial and temporal dimensions of agglomeration processes, as well as the potential existence of space–time interactions.

The second contribution of our analysis concerns the empirical identification of the agglomeration dynamics of a sample of Italian manufacturing firms during the period of the Great Recession. In fact, we find 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. Finally, although our analysis does not detect any statistical evidences of space–time interaction, we observe that the process of space–time dispersion has been more intense at small spatial distances and short time horizons. To the best of our knowledge, this is the first paper that describes these types of agglomeration dynamics across different manufacturing sectors.

Our analysis, and, particularly, the identification of space–time processes in agglomeration patterns, have relevant implications in terms of generation and diffusion of local externalities, inter-firm relationships, and level of concentration and hierarchy of local industries. In fact, all these features of a local industry/cluster are likely to change significantly as a consequence of spatial and temporal changes of the agglomeration dynamics. This is an aspect that should be taken into account seriously in the analysis of the empirical determinants of spatial agglomeration, and of its effects on firms’ economic performance and local economic growth.

Of course, our analysis comes with some limitations. First, the period of analysis—i.e., only 6 years—is not long enough to justify the presence of statistically significant space–time interaction. Second, although our intent was to provide a broader picture of the agglomeration dynamics occurred in the Italian manufacturing industry, the use of industrial sectors at the two-digit level could not be completely appropriate for analyzing these phenomena. All these limitations will be addressed in future developments of this research.

Acknowledgements Open access funding provided by Università degli Studi di Padova within the CRUI-CARE Agreement. We are grateful to the Editor-in-Chief, Prof. Andrés Rodríguez-Pose, and three anonymous Reviewers for numerous suggestions and recommendations that helped us to improve the article. We are also grateful to Camilla Lenzi, Claudio Lupi, David C. Mare and Stefano Usai for useful comments, as well as to participants to the XV Workshop c.MET05 (Prato, 2018), the 58th ERSA Congress (Cork, 2018), and the XL AISRe Congress (L’Aquila, 2019). The usual disclaimers apply.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative

Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Albert JM, Casanova MR, Orts V (2012) Spatial location patterns of Spanish manufacturing firms. *Pap Reg Sci* 91(1):107–136
- Amrhein CG (1995) Searching for the elusive aggregation effect: evidence from statistical simulations. *Environ Plan A* 27(1):105–119
- Arbia G (1989) *Spatial data configuration in statistical analysis of regional economic and related problems*. Kluwer, Dordrecht
- Arbia G (2001) Modelling the geography of economic activities on a continuous space. *Pap Reg Sci* 80(4):411–424
- Arbia G, Espa G, Quah D (2008) A class of spatial econometric methods in the empirical analysis of clusters of firms in the space. *Empir Econ* 34:81–103
- Arbia G, Espa G, Giuliani D, Mazzitelli A (2010) Detecting the existence of space-time clustering of firms. *Reg Sci Urban Econ* 40:311–323
- Arbia G, Espa G, Giuliani D, Dickson MM (2014) Spatio-temporal clustering in the pharmaceutical and medical device manufacturing industry: a geographical micro-level analysis. *Reg Sci Urban Econ* 49:298–304
- Baldwin JR, Beckstead D, Brown WM, Rigby DL (2008) Agglomeration and the geography of localization economies in Canada. *Reg Stud* 42(1):117–132
- Becattini G (1990) The Marshallian industrial district as a socioeconomic notion. In: Pyke F, Becattini G, Sengenberger W (eds) *Industrial districts and inter-firm cooperation in Italy*. International Institute for Labour Studies (ILO), Geneva, pp 37–51
- Besag J (1977) Contribution to the discussion of Dr. Ripley's paper. *J R Stat Soc Ser B* 39(2):193–195
- Brusco S (1982) The Emilian model: productive decentralisation and social integration. *Camb J Econ* 6(2):167–184
- Cainelli G, Ganau R (2018) Distance-based agglomeration externalities and neighbouring firms' characteristics. *Reg Stud* 52(7):922–933
- Cainelli G, Lupi C (2010) Does spatial proximity matter? Micro-evidence from Italy. In: De Liso N, Leoncini R (eds) *Internationalization, technological change and the theory of the firm*. Routledge, London, pp 163–186
- Cainelli G, Zoboli R (2004) The evolution of industrial districts. Changing governance, innovation and internationalization of local capitalism in Italy. *Physica-Verlag, Heidelberg*
- Cainelli G, Ganau R, Iacobucci D (2016) Do geographic concentration and vertically-related variety foster firm productivity? Micro-evidence from Italy. *Growth Change* 47(2):197–217
- Combes PP (2000) Economic structure and local growth: france, 1984–1993. *J Urban Econ* 47(3):329–355
- De Marchi V, Grandinetti R (2014) Industrial districts and the collapse of the Marshallian model: looking at the Italian experience. *Compet Change* 18(1):70–87
- Devereux MP, Griffith R, Simpson H (2004) The geographic distribution of production activity in the UK. *Reg Sci Urban Econ* 34(5):533–564
- Diggle PJ, Chetwynd AG, Haggkvist R, Morris SE (1995) Second-order analysis of space time-clustering. *Stat Methods Med Res* 4(2):124–136
- Durantón G, Overman HG (2005) Testing for localization using micro-geographic data. *Rev Econ Stud* 72(4):1077–1106
- Duschl M, Schimke A, Brenner T, Luxen D (2014) Firm growth and the spatial impact of geolocated external factors. *J Econ Stat* 234(2–3):234–256
- Duschl M, Scholl T, Brenner T, Luxen D, Raschke F (2015) Industry-specific firm growth and agglomeration. *Reg Stud* 49(11):1822–1839

- Ellison G, Glaeser EL (1997) Geographic concentration in U.S. manufacturing industries: a dartboard approach. *J Polit Econ* 105(5):889–927
- Eriksson RH (2011) Localized spillovers and knowledge flows: how does proximity influence the performance of plants? *Econ Geogr* 87(2):127–152
- French NP, McCarthy HE, Diggle PJ, Proudman CJ (2005) Clustering of equine grass sickness cases in the United Kingdom: a study considering the effect of position-dependent reporting on the space-time K-function. *Epidemiol Infect* 133(2):343–348
- Gatrell AC, Bailey TC, Diggle PJ, Rowlingson BS (1996) Spatial point pattern analysis and its applications in geographical epidemiology. *Trans Inst Br Geogr* 21(1):256–274
- Getis A (1964) Temporal land-use pattern analysis with the use of nearest neighbor and quadrat methods. *Ann Assoc Am Geogr* 54(3):391–399
- Getis A, Boots B (1978) Models of spatial processes. Cambridge University Press, Cambridge
- Glaeser EL, Kallal HD, Scheinkman JA, Shleifer A (1992) Growth in cities. *J Polit Econ* 100(6):1126–1152
- Henderson JV (2003) Marshall's scale economies. *J Urban Econ* 53(1):1–28
- Henderson V, Kuncoro A, Turner M (1995) Industrial development in cities. *J Polit Econ* 103(5):1067–1090
- Henderson JV, Squires T, Storeygard A, Weil D (2018) The global distribution of economic activity: nature, history, and the role of trade. *Q J Econ* 133(1):357–406
- Kang H (2010) Detecting agglomeration processes using space-time clustering analyses. *Ann Reg Sci* 45(2):291–311
- Klepper S (1996) Entry, exit growth, and innovation over the product life-cycle. *Am Econ Rev* 86(3):562–583
- Klepper S, Graddy E (1990) The evolution of new industries and the determinants of market structure. *RAND J Econ* 21(1):27–44
- Kosfeld R, Eckey H, Lauridsen J (2011) Spatial point pattern analysis and industry concentration. *Ann Reg Sci* 47:311–328
- Krugman PR (1991) Geography and trade. The MIT Press, Cambridge
- Marcon E, Puech F (2003) Evaluating the geographic concentration of industries using distance-based methods. *J Econ Geogr* 3(4):409–428
- Marcon E, Puech F (2010) Measures of the geographic concentration of industries: improving distance-based methods. *J Econ Geogr* 10(5):745–762
- Marshall A (1920) Principles of economics. Palgrave classics in economics. Palgrave, London
- Martin P, Mayer T, Mayneris F (2011) Spatial concentration and plant-level productivity in France. *J Urban Econ* 69(2):182–195
- Maurel F, Sédillot B (1999) A measure of the geographic concentration in French manufacturing industries. *Reg Sci Urban Econ* 29(5):575–604
- Porter ME (1990) The competitive advantage of nations. Free Press, New York
- Ripley BD (1976) The second-order analysis of stationary point processes. *J Appl Probab* 13(2):255–266
- Ripley BD (1977) Modelling spatial patterns. *J R Stat Soc B* 39(2):172–192
- Rosenthal SS, Strange WC (2003) Geography, industrial organization, and agglomeration. *Rev Econ Stat* 85(2):377–393
- Saxenian A (1994) Regional advantage: culture and competition in Silicon Valley and Route 128. Harvard University Press, Cambridge
- Scholl T, Brenner T (2016) Detecting spatial clustering using a firm-level cluster index. *Reg Stud* 50(6):1054–1068
- Sorenson O, Audia PG (2000) The social structure of entrepreneurial activity: geographic concentration of footwear production in the United States, 1940–1989. *Am J Sociol* 106(2):424–462
- Strange WC, Rosenthal SS (2001) The determinants of agglomeration. *J Urban Econ* 50(2):191–229
- Wong D, Amrhein CG (1996) Research on the MAUP: old wine in a new bottle or real breakthrough? *J Geogr Syst Geogr Inf Anal Theory Decis* 3(2):73–76