Regional Studies

ISSN: 0034-3404 (Print) 1360-0591 (Online) Journal homepage: http://www.tandfonline.com/loi/cres20

Distance-based agglomeration externalities and neighbouring firms' characteristics

Giulio Cainelli & Roberto Ganau

To cite this article: Giulio Cainelli & Roberto Ganau (2018) Distance-based agglomeration externalities and neighbouring firms' characteristics, Regional Studies, 52:7, 922-933, DOI: 10.1080/00343404.2017.1360482

To link to this article: https://doi.org/10.1080/00343404.2017.1360482

REGIONAL STUDIES, 2018
VOL. 52, NO. 7, 922–933
https://doi.org/10.1080/00343404.2017.1360482

Distance-based agglomeration externalities and neighbouring firms' characteristics

Giulio Cainelli[a] and Roberto Ganau[b]

ABSTRACT

This paper tests the hypothesis that firms with different characteristics can differ in their capability to produce local externalities by investigating the relationship between firm-specific distance-based weighted agglomeration measures and firms' short-run productivity growth in the Italian manufacturing industry. The results suggest that positive localization economies increase with distance when neighbouring firms' characteristics are accounted for. Diversification-type forces have negative effects on productivity growth at short distances, while there are positive effects at longer distances regardless of the weighting scheme considered. Moreover, the negative effect of inter-industry externalities seems to persist over distance when neighbouring firms' characteristics are accounted for.

KEYWORDS
distance-based agglomeration; neighbourhood characteristics; total factor productivity; manufacturing; Italy

JEL C3, D24, R12
HISTORY Received 8 June 2016; in revised form 3 July 2017

INTRODUCTION

The spatial agglomeration of economic activities is a remarkable feature of the economic geography (EG) of many countries, regions and local systems (Porter, 1990). Silicon Valley (Saxenian, 1994), the carpet manufacturing industry in Dalton, Georgia (Krugman, 1991), and the Italian industrial districts (Becattini, 1990) are well-known examples of a general and complex phenomenon.

Since Marshall's (1920) seminal contribution, investigations into the determinants and main features of geographically agglomerated areas have proliferated in the fields of economics and EG, and have identified three different mechanisms which may induce firms to co-localize: the availability of skilled labour (labour market pooling), the access to specialized suppliers (shared inputs), and the spread of interfirm knowledge spillovers (Glaeser, Kallal, Scheinkman, & Schleifer, 1992; Henderson, Kuncoro, & Turner, 1995). Theoretical and empirical contributions suggest that firms located in an agglomerated area benefit from these local externalities, which contribute to reductions in production costs. Several studies have attempted to understand whether agglomeration forces (in particular, localization and diversification economies) play a role in explaining firms' economic performance, in particular measured as total factor productivity (TFP).

This literature generally assumes that agglomerated firms have similar characteristics (e.g., in terms of size, productivity and technology). The hypothesis of firm homogeneity (Alcácer & Chung, 2007) implies that all the firms located in a given geographical area contribute in the same way, and with the same intensity, to the production of agglomeration externalities. For example, Shaver and Flyer (2000) underline that the theoretical economic models proposed by Romer (1986), David and Rosenbloom (1990) and Rauch (1993) make this assumption (or assume that firms are 'small' relative to the industry output), thus assuming also that firms do not have differential effects on local externalities.

Some recent EG contributions have shown that the co-existence of firms with different characteristics is a widespread phenomenon within clusters, industrial districts and local productive systems (Wang, 2015). Within agglomerated areas, firms can be small or large, technological leaders or laggards, more or less productive. If firms differ in some characteristics, they should also differ in their capability to produce externalities. In other words, a firm cannot be seen only as a 'receiver' of local externalities as suggested by many theoretical economic models, but also

CONTACT
[a] giulio.cainelli@unipd.it
Department of Economics and Management 'Marco Fanno', University of Padua, Padua, Italy.
[b] **(Corresponding author)** roberto.ganau@unipd.it
Department of Economics and Management 'Marco Fanno', University of Padua, Padua, Italy.
Supplemental data for this article can be accessed http://dx.doi.org/10.1080/00343404.2017.1360482.

as a potential 'source' of these local effects. In fact, as underlined by Alcácer and Chung (2007, p. 761) 'firms are neither equally equipped to receive knowledge nor homogeneously willing to serve as sources of spillovers'.

The aim of this paper is to contribute to this debate. It analyses a large sample of Italian manufacturing firms to investigate empirically the relationship between agglomeration (localization- and diversification-type) externalities and firms' short-run productivity growth by explicitly testing the hypothesis that neighbouring firms' characteristics influence the generating process of local externalities. This hypothesis is tested using firm-specific distance-based weighted agglomeration measures constructed to account for the size and TFP of the neighbouring firms. This allows one to capture the role played by firms' characteristics in generating externalities. The idea is that spatial agglomeration forces may depend not only on the number of co-localized firms (i.e., critical mass effect) but also on their characteristics since firms may contribute differently to the production of local externalities depending on their characteristics.

This paper makes another contribution to this literature by relaxing the modifiable areal unit problem (MAUP), which refers to the arbitrary choice of the spatial partition used to analyse geographically based phenomena (Arbia, 2001).[1] Many studies on the agglomeration–firm productivity relationship use predefined geographical units of analysis to capture agglomeration forces. However, local labour markets (LLMs) and administrative (e.g., NUTS-2 or NUTS-3) regions do not necessarily coincide with real economic areas. This paper tackles this issue by using distance-based agglomeration measures computed within continuous and non-overlapping distance bands defined around each firm in the sample.

The paper is organized as follows. The next section discusses the related literature. The third section presents the data and the methodology adopted. The fourth section reports and discusses the empirical results. The fifth section concludes.

RELATED LITERATURE

Agglomeration and firm productivity

The literature on agglomeration economies identifies two main forms of local externalities arising from the geographical concentration of economic activities, i.e., localization and diversification externalities. Localization externalities arise from the spatial concentration of firms operating in the same industry, and their relevance dates back to Marshall's (1920) contribution on the industrial district concept. The key idea is that firms located close to other firms operating in the same industry benefit from reduced transportation costs, emergence of external-scale economies, availability of specialized workers and suppliers, and diffusion of intra-industry knowledge and technological spillovers which reduce production costs, thus fostering efficiency and economic growth (Glaeser et al., 1992; Duranton & Puga, 2004). Conversely, diversification externalities arise from the geographical concentration of firms operating in different industries. The main advantages derived from location in a highly diversified environment are related to the availability of inputs from suppliers operating at different stages in the production chain, and cross-fertilization among existing ideas and technologies favoured by the variety of the local economic structure (Jacobs, 1969).

Studies of the role played by these types of agglomeration economies on productivity and firms' TFP growth have become especially relevant in the last 15 years (e.g., Cainelli, Iacobucci, & Ganau, 2016; Cainelli & Lupi, 2010; Cingano & Schivardi, 2004; De Lucio, Herce, & Goilcolea, 2002; Henderson, 2003; Marrocu, Paci, & Usai, 2013; Martin, Mayer, & Mayneris, 2011). However, similar to investigations of the impact of agglomeration forces on employment growth (e.g., Cainelli & Leoncini, 1999; Mameli, Faggian, & McCann, 2014; Paci & Usai, 2008), the empirical results from this research strand are rather puzzling. For example, De Lucio et al. (2002) find a positive effect of industrial variety and a 'U'-shaped effect of localization externalities on labour productivity at province level in Spain. Henderson (2003) finds positive effects of localization economies on productivity at plant level in US high-tech industries, but not in machinery industries, and finds little evidence of diversification economies. Cingano and Schivardi (2004) find positive effects of localization, but negligible effects of diversification externalities on TFP growth at the LLM level in Italy. They also find negative effects of localization and positive effects of diversification externalities on employment growth. Martin et al. (2011) find that French firms' productivity benefits from localization, but not from diversification economies. However, the benefits from industrial clustering (although highly significant from a statistical point of view) are quite modest in magnitude. Fazio and Maltese (2015) find that the effect of agglomeration forces on Italian small- and medium-sized firms' performance varies depending on whether the level or growth of TFP is considered: TFP levels are influenced mostly by localization externalities, while TFP growth is higher in the presence of diversification (and competition) externalities. Finally, Cainelli et al. (2016) find that the effect of localization externalities is stronger than the effect of diversification externalities on Italian manufacturing firms' TFP.

Neighbouring firms' characteristics and agglomeration externalities

A limitation of this literature concerns the hypothesis of firm homogeneity. Previous works assume (explicitly or implicitly) that firms located in an agglomerated area have similar characteristics, e.g., in terms of size, productivity and technology. It is not surprising that 'firms operating in industrial districts have been traditionally modelled as undifferentiated and characterized by low variance in their strategies and business models' (Munari, Sobrero, & Malipiero, 2012, p. 430). Since 'firms not only capture benefits from agglomeration economies, but they also contribute to agglomeration economies' (Shaver & Flyer, 2000, p. 1175), the homogeneity hypothesis implies that all the firms located in a given geographical area contribute in the same way, and with the same intensity, to the production of local externalities.

This implication sounds unrealistic. As suggested by some EG papers, firms differ not only across countries and industries but also across regions and local systems within the same country (Almeida & Kogut, 1999; Saxenian, 1994). This means that firms with different characteristics can differ in their ability to identify, absorb, utilize and generate new knowledge and information (Alcácer & Chung, 2014). This is true both when firms act as 'receivers' of local externalities, and when they act as potential 'sources' of local externalities. Therefore, firms with different characteristics may influence in different ways the production of externalities (Alcácer & Chung, 2007). For example, firms with a higher technological endowment may generate more externalities (e.g., local knowledge spillovers) than firms with a lower technological endowment. Similarly, firms employing highly educated workers may generate more local externalities than firms employing less educated workers. It follows that this form of 'firm heterogeneity in technological capabilities' (Wang, 2015) may determine different contributions by firms to the production of externalities.

This strand of EG studies suggests that the agglomeration phenomenon cannot be seen only as a mechanism of 'appropriation' of local externalities; it is also a mechanism for their generation. Hence, neighbouring firms' characteristics should be accounted for to capture this second dimension of the agglomerative process. In this respect, 'new' weighted agglomeration measures should capture better the agglomeration phenomenon as a mechanism generating local externalities if it is true that firms' characteristics influence the way these units contribute to the externality-generation process.

The spatial dimension of agglomeration

Another weakness of this literature concerns the choice of the geographical unit used to analyse the agglomeration–firms' productivity relationship. According to the MAUP, the discretionary choice of the spatial unit may introduce statistical biases related to the level of aggregation and the geographical scale (Arbia, 1989). The majority of these studies captures agglomeration forces through predefined geographical units (e.g., LLMs, NUTS regions) which vary in size and shape, and are characterized by arbitrary predefined boundaries (independently of the criteria adopted for their definition). In particular, these geographical units do not necessarily coincide with real economic areas.

Therefore, the MAUP may help explaining the different results in the literature. In fact, these differences may be due to the different geographical units considered, besides differences in agglomeration measures (Beaudry & Schiffauerova, 2009; Burger, Van Oort, & Van der Knaap, 2010; Rosenthal & Strange, 2003). For example, Briant, Combes, and Lafourcade (2010) compare different French zoning systems to analyse the magnitude of the distortion arising from the MAUP in the context of spatial concentration, agglomeration economies and trade determinants. They show that the size component matters, while the shape component matters less. Similarly, Burger et al. (2010) find that the effect of agglomeration forces on employment growth in the Netherlands significantly varies over scales considering different geographical levels of aggregation (municipality, district and region).

It follows that the geographical scale at which agglomeration phenomena are analysed is a critical issue since agglomeration forces may produce different effects at different spatial scales (Olsen, 2002). Moreover, their effects are likely to attenuate over space (Rosenthal & Strange, 2008). This is likely when distinguishing between localization and diversification externalities, as well as between market- and knowledge-based externalities within each type of agglomeration force (Martin, 1999).

Arbia (2001) suggests relaxing the MAUP using micro-geographical data, thus moving the analysis from the meso- to the micro-geographical level.[2] The idea is to consider the firm as the spatial unit of analysis and to treat the space as continuous in order to avoid the use of predefined spatial areas. Therefore, the sample of firms is treated as a spatial points pattern where each unit represents an individual point identified by its geographical coordinates.

Moving from the contribution by Sorenson and Audia (2000), who capture local density effects in the US footwear industry through cumulative pairwise distances among plants, the use of micro-geographical data in the analysis of spatial agglomeration forces has increased. Two main micro-geographical distance-based methods have been explored in the analysis of agglomeration forces. One method employs distance-decay functions, where cumulated geocoded firms are discounted by geographical or travel-time distances (e.g., see Duschl, Schimke, Brenner, and Luxen, 2014, and Duschl, Scholl, Brenner, Luxen, and Raschke, 2015, for the German case). The other approach employs distance bands and consists in counting geocoded firms within (or at) specific radii, so that distance-based agglomeration measures are computed within distance bands identified around each firm in the sample to evaluate the space component of the agglomeration phenomenon.

Focusing on the latter approach, i.e., the one adopted in this paper, only few contributions have tried to overcome the MAUP modelling agglomeration forces through distance bands. Among these works, Rosenthal and Strange (2003) provide an industry-level analysis of localization and diversification externalities computing employment-based agglomeration variables through a set of concentric rings defined around the centroid of each ZIP code in the United States up to 15 miles. Baldwin, Beckstead, Brown, and Rigby (2008) analyse the effects of Marshall's (1920) three sources of agglomeration economies on Canadian plants' productivity, finding that the agglomeration of plants in the same industry has positive effects only within 10 km from the reference firm. Eriksson (2011) computes plant-specific measures defined within distance bands of radii 0.5, 5 and 50 km to analyse the impact of spillovers and knowledge flows on Swedish plants' productivity growth. In this work, the author finds that the density of economic activities has positive effects within short distances, intra-industry spillovers have greater impacts at increased distances, while the presence of different but related industries matters at short distances. Deltas, De Silva, and McComb (2015) capture localization externalities in the Texas software industry through employment- and firm-based agglomeration variables computed within concentric rings of radius up to 25 miles, which approximates a county-level analysis. Cainelli and Lupi (2010) and Gabriele, Giuliani, Corsino, and Espa (2013) represent the only contributions adopting such an approach for the Italian case. Cainelli and Lupi (2010) analyse a sample of about 23,000 Italian manufacturing firms observed over the period 1998–2001, and find that localization effects are positive within 2 km, but decreasing over distance. On the contrary, diversification effects are negative up to 10 km, but positive between 10 and 30 km. Gabriele et al. (2013) analyse a sample of about 8300 Italian manufacturing firms observed over the period 1996–2004 and find that small-sized firms' growth is fostered by localization externalities, while medium- and large-sized firms benefit more from diversification externalities. However, they do not find evidence of spatial nonlinearities of agglomeration forces.[3]

Although these contributions relax the MAUP by adopting a distance bands-based approach, they also introduce an arbitrary element concerning the definition of the distance bands within which agglomeration forces are captured. This work contributes to this issue by performing a hierarchical cluster analysis in order to limit the arbitrary component in the identification process of the distance bands. Specifically, distance bands are identified in three steps: first, a maximum distance threshold value is chosen; second, a set of distance bands is defined within it, and density measures are computed within each band; and finally, a hierarchical cluster analysis is performed on the arbitrary predefined distance bands to reduce redundancy and statistically identify a reduced number of bands which may be meaningful to capture agglomeration forces.

DATA AND METHODOLOGY

The dataset

The empirical analysis employs an Italian firm-level balance sheet dataset covering the period 2003–12, which is drawn from the AIDA databank (Bureau Van Dijk). The investigation is conducted using three different, but nested, samples in order to maximize the sample size at each of the three steps of the empirical exercise. The original sample was cleaned to remove firms with missing or inconsistent data on value added, tangible assets, labour costs and intermediate inputs. Firms reporting a value added-to-turnover ratio < 0 and > 1, and firms observed for fewer than seven consecutive years during the period 2003–12 were excluded. This led to an unbalanced panel of 69,933 firms observed over the period 2003–12, which is used to estimate firms' TFP.

Agglomeration measures are constructed using sample rather than census data.[4] To this aim, the sample was further cleaned removing firms without information on the exact address. In fact, it is necessary to know the exact geographical location of each firm in order to compute distance-based agglomeration measures. Firms with missing data for number of employees were also excluded in order to construct size-weighted agglomeration variables (e.g., Duranton & Overman, 2005). The year 2009 was selected to construct the agglomeration variables since it presents the largest number of valid observations, i.e., firms reporting data on geographical coordinates, employment and estimated TFP. This second cleaning procedure led to a sample of 41,574 firms observed in 2009, which is used to construct the agglomeration measures.

Starting from this last sample, a third cleaning procedure was performed to construct the final dataset to be used in the empirical analysis of the relationship between agglomeration forces and firms' productivity growth. Firms with missing or inconsistent data on net income and annual depreciation for 2009, and firms with missing data for year of establishment were excluded. This led to a final sample of 28,597 firms observed over the period 2009–12.

Measuring spatial agglomeration

Distance-based agglomeration measures are computed within continuous and non-overlapping distance bands identified around each firm in a sample of 41,574 units observed in 2009. Then, specific characteristics (i.e., size and estimated TFP) of the neighbouring firms located within each distance band are used to construct weighted agglomeration measures to test whether neighbouring firms' characteristics matter in the externality-generation process.

Intra-industry (localization-type) and inter-industry (diversification-type) externalities are captured through absolute density measures, which are computed within three continuous and non-overlapping distance bands of radius r (d_r) up to a maximum distance of 30 km, i.e., $0 \leq d_5 \leq 5$, $5 < d_{10} \leq 15$ and $15 < d_{15} \leq 30$.[5]

Comparison of the surfaces of the three distance bands (Table 1) with the average surfaces of the usually employed spatial units of analysis for the Italian case (Table 2) suggests that the three distance bands provide a relatively good partition of the continuous space in order to capture potential geographical nonlinearities of the agglomeration

Table 1. Surface covered by spatial bands.

Distance band	Radius (km)	Surface area (km^2)
$0 \leq d_5 \leq 5$	5	79
$5 < d_{10} \leq 15$	10 [15]	628 [707]
$15 < d_{15} \leq 30$	15 [30]	2121 [2827]

Notes: Cumulative values are shown in brackets.

Table 2. Average surface of predefined geographical units.

Geographical unit	Average surface area (km^2)
8177 Municipalities	37
611 Local labour markets	494
110 Provinces	2739

forces, which could not be captured using standard (predefined) spatial partitions. In fact, the (cumulative) areas of the three distance bands encompass the average areas of municipalities, LLMs and provinces. Distances exceeding 30 km are not considered because they are not particularly relevant for the Italian case (e.g., Cainelli & Lupi, 2010). Since we are specifically interested in the agglomeration effects generated by nearby firms, a distance of 30 km represents a proper threshold as it identifies an area corresponding approximately to an Italian province; conversely, larger distances would imply the analysis of the agglomeration effects on firms located across different (administrative) regions. Moreover, previous contributions find negligible agglomeration effects beyond similar threshold distances (e.g., Baldwin et al., 2008; Deltas et al., 2015).

The three distance bands are identified through a hierarchical cluster analysis performed on estimated densities computed within a larger set of seven distance bands. The detailed explanation of the statistical approach adopted is presented in Appendix A in the supplemental data online. The motivation driving the use of a hierarchical cluster analysis is twofold. First, it allows one to reduce redundancy and identify a reduced number of distance bands that may be meaningful to capture spatial agglomeration forces, e.g., if the estimated densities computed within two (or more) distance bands present high correlation. Second, although the arbitrary definition of the maximum distance value of 30 km and the seven distance bands identified within it, the use of a hierarchical cluster analysis to identify distances at which agglomeration phenomena may matter represents a step forward with respect to previous contributions with analyses based exclusively on arbitrary distances.

Two main types of agglomeration variables are constructed: unweighted and weighted. Unweighted agglomeration variables represent the baseline measures since they are built on the firm homogeneity hypothesis (explicitly or implicitly) assumed in previous contributions. In fact, they are defined considering the number of neighbouring firms located within a certain distance, without accounting for their specific characteristics. In contrast, weighted agglomeration variables are constructed accounting for neighbouring firms' characteristics, i.e., relaxing the firm homogeneity hypothesis. It follows that weighted intra- and inter-industry agglomeration variables allow to test whether firm-specific characteristics influence the way firms located within an agglomerated area contribute to the production of local externalities.

Two firm-specific characteristics are considered as weights: size, defined in terms of employment, and TFP. Employment-based indexes have been proposed in the literature to proxy for localization (e.g., specialization indexes) and diversification (e.g., Herfindahl–Hirschman indexes) externalities (e.g., Glaeser et al., 1992). These measures are generally constructed considering employment of an industry-area pair with respect to the national dimension or total employment in an industry or area. Contributions using employment-based agglomeration indicators 'implicitly' consider the role of neighbouring firms' characteristics in the agglomerative space, although they neither make assumptions with respect to the externality-generation process nor compare agglomeration indicators constructed with and without the employment dimension. Moreover, employment-based variables computed within spatial units of different sizes do not allow one to capture the role of firm employment in the generating process of externalities because this is likely to be influenced by the size of the local system: a larger area is likely to host a higher number of firms so the overall number of employees in a given area may depend on its surface. However, the use of areas characterized by the same surface may facilitate the comparison between standard and employment-based agglomeration indicators, under the assumption that firms are homogeneously distributed over space, as well as the identification of the role ascribable to co-localized firms' employment size in the production of externalities. Employment-based measures are proposed by Rosenthal and Strange (2003), Gabriele et al. (2013) and Deltas et al. (2015) in the context of distance-based agglomeration indexes. However, their analyses neither assume a role for neighbouring firms' characteristics nor explicitly compare the results of unweighted and weighted variables.

The second weighting component is firm TFP, which is estimated as the residual of a Cobb–Douglas production function using Wooldridge's (2009) approach. Appendix B in the supplemental data online discusses the estimation approach, and reports the estimated inputs' elasticities. TFP-weighted agglomeration measures may represent a better proxy to capture the role played by neighbouring firms' characteristics in the generation of local externalities than size-weighted variables, and this paper is the first attempt to account for this dimension when constructing agglomeration variables. A firm's TFP is correlated with its technological endowment and, consequently, its (potential) ability to produce externalities. On the contrary, size-weighted variables can be only a rough proxy for the human capital endowment of co-localized firms if the skill dimension of employment cannot be identified.

Agglomeration variables are constructed for each firm in the sample and within each distance band as follows:

$$\hat{D}_{x_i^s}(d_r) = \frac{e(x_i^s)\left[\sum_{\substack{j=1 \\ j \neq i}}^{N} 1(\|x_i^s - x_j^g\| \in d_r)w(x_j^g)\right]}{A_{x_i^s}(d_r)} \tag{1}$$

where d_r denotes the distance band with radius r defined in kilometres, such that $0 \leq d_5 \leq 5$, $5 < d_{10} \leq 15$ and $15 < d_{15} \leq 30$; the denominator is the (net) area of the distance band (i.e., the circle) centred in the reference firm i belonging to the two-digit industry s, which is denoted by x_i^s as a spatial point identified by its geographical coordinates; the numerator is the sum of all the neighbouring firms j belonging to the two-digit industry g (denoted by x_j^g as spatial points) and located within a certain distance band, with $s = g$ in the intra-industry case and $s \neq g$ in the inter-industry case; the term $||x_i^s - x_j^g||$ denotes the Euclidean distance between the reference firm i and each neighbouring firm j; $1(\cdot)$ is an indicator function; the term $w(x_j^g)$ denotes the weighting scheme capturing neighbouring firms' characteristics within each distance band, such that $w(x_j^g) = 1$ in the unweighted case, $w(x_j^g) = \text{size}_j^g$ in the size-weighted case (where size_j^g denotes employment) and $w(x_j^g) = \text{tfp}_j^g$ in the TFP-weighted case (where tfp_j^g denotes the estimated TFP of a firm in level); and the term $e(x_i^s)$ denotes Ripley's (1977) edge correction, which is defined as follows:

$$e(x_i^s) = \frac{2\pi r}{\text{length}[c(x_i^s, r) \cap W]} \tag{2}$$

where the numerator defines the circumference of the circle with radius r; and the denominator is the length of the overlap between the circle c centred in x_i^s with radius r and the window W defining the study region (i.e., Italy). This correction term allows one to account for edge effects which may influence the agglomeration phenomenon around firms located close to the boundaries of the study region. In fact, those firms may be surrounded by fewer neighbours with respect to other firms located at longer distances from the study region's boundaries.[6]

Therefore, two forms of agglomeration externalities are captured through unweighted and size- and TFP-weighted agglomeration variables: intra-industry externalities arising from the spatial concentration of firms operating in the same industry as the reference firm (i.e., localization-type forces), and inter-industry externalities arising from the spatial concentration of firms operating in industries different from the industry of the reference firm (i.e., diversification-type forces).[7]

The comparison of the estimated coefficients of unweighted and weighted agglomeration variables allows one to evaluate both the 'critical-mass effect' of agglomeration forces and which type of co-localized firms contributes more to the externality-generation process. As both intra- and inter-industry agglomeration variables are normalized with respect to the (net) area of each distance band, independently of the weighting scheme, the estimated coefficients of the unweighted and weighted variables can be directly compared. While unweighted variables capture the number of neighbouring firms per km^2, i.e., the density of economic activity, size- and TFP-weighted variables capture the number of employees and the level of productivity, respectively, per km^2. Therefore, given a certain number of neighbouring firms within a certain distance band, the comparison of the estimated coefficients for the different weighting schemes allows one to understand whether a firm's productivity growth benefits more from externalities generated by small versus large and lowly versus highly productive neighbouring firms.

The growth equation

The analysis is based on the estimation of a productivity growth equation specified as follows:

$$\Delta\text{TFP}_{is} = \alpha + \sum_{k=1}^{K} \beta_k X_{is}^k + \sum_{d=1}^{D} \delta_{1d}\text{INTRA}_{is}^d + \sum_{d=1}^{D} \delta_{2d}\text{INTER}_{is}^d + \boldsymbol{\gamma}_c + \boldsymbol{v}_m + \varepsilon_{is} \tag{3}$$

where $\Delta\text{TFP}_{is} = \text{TFP}_{is}^{2012} - \text{TFP}_{is}^{2009}$ denotes the productivity growth of firm i operating in the two-digit industry s over the period 2009–12, where TFP_{is}^{2009} and TFP_{is}^{2012} denote the estimated TFP (in logarithmic form); the vector X_{is}^k of log-transformed firm-specific control variables includes the beginning-of-the-period TFP (TFP_{is}), the number of employees at the beginning of the growth period (SIZE_{is}), the difference between 2009 and the year of a firm's set up (AGE_{is}), the ratio between acquired services and total acquired inputs in 2009 to proxy for services outsourcing (OUTSOURCING_{is}) and the cash flow defined as net income plus annual depreciation over tangible assets at the beginning of the growth period (CASH_{is}); the terms INTRA_{is}^d and INTER_{is}^d capture the log-transformed variables for, respectively, intra- and inter-industry agglomeration externalities computed within the three distance bands; the term $\boldsymbol{\gamma}_c$ refers to a set of industrial category dummy variables; the term $\boldsymbol{v}_m$ refers to a set of macro-geographical dummy variables defined at the NUTS-1 level; and ε_{is} denotes the error term. Appendix C in the supplemental data online reports and discusses descriptive statistics and correlation matrices of firm-level and agglomeration variables (see Tables C1–C7). Table C8, also online, compares the final sample with the population of Italian manufacturing firms, suggesting a good representativeness in terms of size and geographical distribution. Finally, Table C9, again online, suggests a good industrial representativeness of the final sample with respect to the larger sample used to estimate firm-level TFP.

The identification strategy

The ordinary least squares (OLS) estimation of equation (3) is likely to be affected by sample selection since productivity growth is observed only for the subsample of firms surviving over the growth period. Therefore, a two-step sample-selection model à la Heckman (1979) is estimated to account for firm exit over the period 2009–12. A first-stage reduced-form selection equation is estimated by maximum likelihood specifying the dependent variable as a dummy (SURVIVAL_{is}), which equals 1 if the firm observed at the beginning of the growth period is observed also at the end of the growth period, and 0 otherwise. The selection equation is identified including on its right-hand side the explanatory variables entering equation (3) plus an exclusion restriction (TURBULENCE_s) capturing the average entry/exit rate over the period 2006–08, and defined at the two-digit industry level. A high value of this variable is likely to be associated with a low (current) probability of firm survival, without necessarily being associated with the surviving firms' economic performance.

Having estimated the selection equation on the whole sample of firms based on a Probit model, the computed inverse Mills ratio (λ) is added to the right-hand side of the productivity growth equation to correct for the sample selection bias. Thus, the augmented version of equation (3) is estimated via OLS for the subsample of firms surviving during the period 2009–12 (Wooldridge, 2010).

EMPIRICAL RESULTS

Table 3 reports the results of the selection and productivity growth equations including unweighted and weighted agglomeration variables. The exclusion restriction identifying the selection equations shows negative and significant coefficients, suggesting that firms' survival probability is negatively affected by the level of industry-specific turbulence, while the estimated inverse Mills ratio shows positive and significant coefficients. Therefore, the results suggest the need to correct for sample selection.

The focus here is on the results of the productivity growth equation. The results of the unweighted agglomeration variables confirm the findings of Cainelli and Lupi (2010). What emerges is a positive effect of intra-industry externalities within 15 km, which is decreasing in distance, while the short-distance negative effect of inter-industry externalities (up to 5 km) becomes positive at longer distances (15–30 km). This evidence has two interpretations. First, what emerges is a sort of substitution effect between localization and diversification forces: firms' TFP growth benefits from industry similarity at short distances, while it benefits from industry diversification at longer distances. Second, the negative and significant effect of inter-industry forces at short distances highlights the emergence of diversification diseconomies, which can be due to congestion costs and a higher demand for local resources by neighbouring firms operating in the other local industries. Negative congestion costs seem to prevail on the positive effects associated with inter-industry knowledge spillovers at short distances, thus producing negative effects on a firm's performance. This evidence confirms the results of Cainelli and Lupi (2010), as well as those proposed by Cainelli, Fracasso, and Vittucci Marzetti (2015) showing that diversification externalities have positive effects on productivity only beyond a certain threshold level of diversity.

The impact of spatial agglomeration significantly changes if firm-specific characteristics are accounted for. The positive effect of intra-industry externalities increases with distance, and this pattern is particularly strong when the size of neighbouring firms is considered, rather than their TFP. This means that positive externalities related to localization forces tend to be higher the bigger and more productive are the neighbouring firms operating in the same industry. In other words, this result suggests that intra-industry externalities do not attenuate over distance when neighbouring firms' characteristics are considered.

The negative effect of inter-industry externalities seems to be reinforced as it persists over distance (up to 15 km) when neighbouring firms' characteristics are accounted for. Then, the effect becomes positive between 15 and 30 km. As already mentioned, congestion costs associated with the localization of firms in other industries seem to prevail on inter-industry knowledge spillovers at short distances, thus generating negative effects on firms' TFP growth. Interestingly, these negative effects tend to be reinforced the bigger and more productive are the neighbouring firms operating in the other industries. In addition, the substitution effect characterizing intra- and inter-industry externalities seems to attenuate at longer distances if neighbouring firms' characteristics are accounted for when capturing agglomeration forces.

Overall, the results suggest that co-localized firms participate in the generating process of both positive externalities and agglomeration diseconomies, and that their contribution primarily depends on their specific characteristics. Bigger and more productive firms seem to contribute more to the production of positive agglomeration externalities in the context of localization-type forces, as well as to the generation of agglomeration diseconomies in the case of diversification-type externalities. The nonlinear effects of agglomeration forces depend on the spatial distance considered, thus confirming that the MAUP must be taken into account in this kind of analysis.

The estimated coefficients of the firm-level controls have the same signs and significance levels in all the TFP growth specifications reported in Table 3. The coefficients of the beginning-of-the-period TFP variable are negative, suggesting a β-convergence effect. A firm's TFP growth is positively affected by its size, age and level of services outsourced. A positive productivity-to-cash flow sensitivity emerges, meaning that firms' productivity growth tends to be affected by credit rationing (i.e., firm growth is pushed by internally generated resources).

Appendix D in the supplemental data online presents two exercises performed to test the robustness of the results. The first test concerns the potential endogeneity of the agglomeration variables, which could emerge due to spatial sorting, and consists in replicating the unweighted case on the subsample of firms located in a specific point in space at least 10 years before the reference year of the agglomeration variables. The second test aims at verifying the validity of the proposed density measures, and consists in replicating the unweighted case using alternative specialization and diversification variables as proposed by Cingano and Schivardi (2004). Overall, the results of the two tests confirm the main results (see Tables D1 and D2 in Appendix D in the supplemental data online).

CONCLUSIONS

A new stream of the EG literature has emphasized the role played by neighbouring firms' characteristics in spatially agglomerated areas. The main idea of these studies is simple: firms with different characteristics generally coexist within clusters, industrial districts and local

Table 3. Agglomeration externalities and productivity growth.

Model	(1)		(2)		(3)	
Weighting scheme	**Unweighted**		**Size weighted**		**TFP weighted**	
Dependent variable	**SURVIVAL$_{is}$**	**ΔTFP$_{is}$**	**SURVIVAL$_{is}$**	**ΔTFP$_{is}$**	**SURVIVAL$_{is}$**	**ΔTFP$_{is}$**
TFP$_{is}$	0.249***	–0.355***	0.249***	–0.346***	0.249***	–0.342***
	(0.017)	(0.020)	(0.017)	(0.021)	(0.017)	(0.021)
SIZE$_{is}$	0.057***	0.143***	0.056***	0.143***	0.056***	0.145***
	(0.009)	(0.006)	(0.009)	(0.006)	(0.009)	(0.006)
AGE$_{is}$	0.098***	0.025***	0.098***	0.030***	0.098***	0.030***
	(0.015)	(0.010)	(0.015)	(0.010)	(0.015)	(0.010)
OUTSOURCING$_{is}$	0.098***	0.073***	0.099***	0.079***	0.098***	0.078***
	(0.015)	(0.011)	(0.015)	(0.011)	(0.015)	(0.011)
CASH$_{is}$	0.018***	0.027***	0.018***	0.028***	0.018***	0.028***
	(0.004)	(0.002)	(0.004)	(0.002)	(0.004)	(0.002)
INTRA$_{is}^{0-5}$	0.005	0.021***	–0.004	0.009***	0.002	0.021***
	(0.012)	(0.005)	(0.007)	(0.003)	(0.011)	(0.004)
INTRA$_{is}^{5-15}$	0.010	0.018***	0.009	0.014***	0.014	0.022***
	(0.016)	(0.006)	(0.010)	(0.004)	(0.014)	(0.006)
INTRA$_{is}^{15-30}$	0.026*	0.005	0.021*	0.019***	0.033**	0.023***
	(0.015)	(0.006)	(0.012)	(0.005)	(0.015)	(0.006)
INTER$_{is}^{0-5}$	–0.056***	–0.052***	–0.041***	–0.037***	–0.054***	–0.052***
	(0.014)	(0.007)	(0.009)	(0.005)	(0.012)	(0.006)
INTER$_{is}^{5-15}$	–0.002	–0.008	–0.015	–0.017***	–0.010	–0.016**
	(0.021)	(0.008)	(0.014)	(0.005)	(0.018)	(0.007)
INTER$_{is}^{15-30}$	0.049**	0.054***	0.051***	0.040***	0.034*	0.033***
	(0.020)	(0.009)	(0.016)	(0.007)	(0.019)	(0.008)
TURBULENCE$_{s}$	–1.321**	…	–1.185*	…	–1.153*	…
	(0.606)		(0.610)		(0.610)	
λ	…	1.630***	…	1.717***	…	1.753***
		(0.204)		(0.206)		(0.207)

(*Continued*)

Table 3. Continued.

Model	**(1)**		**(2)**		**(3)**	
Weighting scheme	**Unweighted**		**Size weighted**		**TFP weighted**	
Dependent variable	**SURVIVAL$_{is}$**	**ΔTFP$_{is}$**	**SURVIVAL$_{is}$**	**ΔTFP$_{is}$**	**SURVIVAL$_{is}$**	**ΔTFP$_{is}$**
Observations	28,597	22,239	28,597	22,239	28,597	22,239
Censored observations	…	6358	…	6358	…	6358
Pseudo-R^2	0.056	…	0.056	…	0.056	…
Log-likelihood	–14,307.72	…	–14,303.20	…	–14,303.18	…
Wald χ^2	1623.96***	…	1617.72***	…	1629.92***	…
Adjusted R^2	…	0.254	…	0.255	…	0.255
F-statistic	…	79.45***	…	78.17***	…	79.85***
Mean VIF	2.80	7.98	2.37	7.62	2.59	8.04

Notes: All specifications include a constant term, industrial category and NUTS-1 dummy variables. Bootstrapped-robust standard errors are shown in parentheses. λ denotes the inverse Mills ratio parameter from first-step selection equation. TFP, total factor productivity; VIF, variance inflation factor.
*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

production systems. If firms differ in terms of size, productivity, technological and human capital endowment, they should also differ in their capability to produce externalities. In fact, firms cannot be seen only as 'receivers' of local externalities as suggested by many theoretical economic models, but also as potential 'sources' of these local effects.

This paper is an attempt to contribute to this debate. This hypothesis is tested by accounting for the role played by firm-specific characteristics (size and TFP) in the generation of local externalities. Moreover, agglomeration externalities are captured through firm-specific distance-based variables computed within continuous and non-overlapping bands, thus relaxing the MAUP.

The empirical results suggest that neighbouring firms' characteristics matter for the generation of externalities, both in the context of intra- and inter-industry agglomeration forces, thus confirming recent EG insights based on the analysis of case studies (Wang, 2015). In particular, we find that the positive effect of localization economies increases with distance when specific characteristics of the neighbouring firms (operating in the same industry as the reference firm) are accounted for. It emerges that unweighted and weighted diversification-type forces have a negative and significant effect on firms' TFP growth at short distances (up to 5 km), but a positive effect at longer distances (15–30 km) regardless of the weighting scheme considered. Moreover, the negative effect of inter-industry externalities seems to persist over distance (up to 15 km) when neighbouring firms' characteristics are accounted for. Therefore, the results support the theoretical intuition of some EG studies (e.g., Alcácer & Chung, 2007, 2014; Wang, 2015): firms with different characteristics contribute differently to the production of both positive local externalities and agglomeration diseconomies.

The results proposed in this paper underline several limitations characterizing the empirical analysis of spatial agglomeration forces, i.e., the assumption of homogeneous firms and the use of predefined spatial partitions. However, this study has two main weaknesses that should be addressed in further research. First, agglomeration variables are computed using sample rather than census data, with the consequence that only a (selected) subsample of the population of Italian manufacturing firms is included in the analysis. Second, the size and TFP of neighbouring firms are rough proxies for a firm's capabilities to produce agglomeration externalities. Alternative firm-specific characteristics (e.g., research and development (R&D), innovativeness, education of employees) should be considered to capture the role of firm diversity in generating local externalities.

ACKNOWLEDGEMENTS

The authors thank the editor, Thomas Kemeny, and two anonymous reviewers for numerous suggestions that helped to improve this paper. The authors also acknowledge Eric Marcon (AgroParisTech) and Steve Gibbons (LSE) for valuable advice about the computation of the distance-based agglomeration measures, as well as the participants at the 55th Congress of the European Regional Science Association, Lisbon, Portugal, August 2015, for useful comments on a previous version of the paper. The usual disclaimers apply.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

NOTES

1. The MAUP has been widely investigated by statisticians and quantitative geographers (e.g., Arbia, 1989; Amrhein, 1995; Wong & Amrhein, 1996).
2. The literature proposes alternative solutions to mitigate the MAUP. Some contributions suggest controlling for extra-region spillovers through spatially lagged agglomeration variables computed within administrative or labour market regions (e.g., Burger et al., 2010; Van Oort, 2007). Others propose a multilevel approach to enable simultaneous modelling at micro- and macro-levels of analysis (e.g., Van Oort, Burger, Knoben, & Raspe, 2012).
3. The reviewed works extend to the measurement of agglomeration economies the micro-geographical approaches used to identify the geographical concentration of economic activities through spatial statistics (probability or cumulative density functions) which use pair distances between firms to evaluate at which geographical scale a particular industry shows a clustering pattern (e.g., Arbia & Espa, 1996; Duranton & Overman, 2005; Marcon & Puech, 2010; Scholl & Brenner, 2016).
4. The main reason for this choice is lack of georeferenced micro-data at the census level. We are conscious that this represents a limit of our analysis. Cainelli and Lupi (2010) and Gabriele et al. (2013) adopt a similar approach for the Italian case. Similarly, Martin et al. (2011) construct agglomeration measures using sample data drawn from the French Annual Business Survey, a dataset covering all firms with more than 20 employees. The advantage of our dataset is that it also covers firms with fewer than 20 employees, which represent the majority in our sample. In our opinion, this degree of coverage is particularly important in the context of spatial agglomeration.
5. Distance-based agglomeration measures are firm specific, them being centred at each single firm. The centroid of each firm corresponds to the geographical coordinates obtained from the firm's exact address. This means that each firm, for which the exact address was available, has been geocoded. Although this approach provides great precision in the identification of a firm's neighbourhood, it may also present disadvantages if geographical coordinates refer to the headquarter rather than to (productive or commercial) plants of a multi-plant firm. As the AIDA databank provides information on the headquarters, this issue may represent a drawback of this

study. However, the nature and characteristics of the Italian industrial structure limit this drawback, it being driven by many small firms such that the headquarter tends to coincide with the principal (or unique) operative plant. For instance, according to the 2009 ASIA Archive provided by the Istituto Nazionale di Statistica (ISTAT), multi-plant firms represent only 10.94% of active firms operating in the manufacturing industries covered in this analysis.

6. The agglomeration variables in equation (1) are computed using the R Project for Statistical Computing. Original coding is based on the 'dbmss' R package developed by Marcon, Lang, Traissac, and Puech (2012).

7. Rosenthal and Strange (2003), Gabriele et al. (2013) and Deltas et al. (2015), among others, capture diversification externalities simply considering industries different from the reference one.

ORCiD

Giulio Cainelli http://orcid.org/0000-0003-0902-1847

REFERENCES

Alcácer, J., & Chung, W. (2007). Location strategies and knowledge spillovers. *Management Science*, *53*, 760–776. doi:10.1287/mnsc.1060.0637

Alcácer, J., & Chung, W. (2014). Location strategies for agglomeration economies. *Strategic Management Journal*, *35*, 1749–1761. doi:10.1002/smj.2186

Almeida, P., & Kogut, B. (1999). Localization of knowledge and the mobility of engineers in regional networks. *Management Science*, *45*, 905–917. doi:10.1287/mnsc.45.7.905

Amrhein, C. (1995). Searching for the elusive aggregation effect: Evidence from statistical simulations. *Environment and Planning A*, *27*, 105–119. doi:10.1068/a270105

Arbia, G. (1989). *Spatial data configuration in statistical analysis of regional economic and related problems*. Dordrecht: Kluwer.

Arbia, G. (2001). Modelling the geography of economic activities on a continuous space. *Papers in Regional Science*, *80*, 411–424. doi:10.1007/PL00013646

Arbia, G., & Espa, G. (1996). *Statistica economica territoriale*. Padua: CEDAM.

Baldwin, J. R., Beckstead, D., Brown, W. M., & Rigby, D. L. (2008). Agglomeration and the geography of localization economies in Canada. *Regional Studies*, *42*, 117–132. doi:10.1080/00343400701543199

Beaudry, C., & Schiffauerova, A. (2009). Who's right, Marshall or Jacobs? The localization versus urbanization debate. *Research Policy*, *38*, 318–337. doi:10.1016/j.respol.2008.11.010

Becattini, G. (1990). The Marshallian industrial district as a socio-economic notion. In F. Pyke, G. Becattini, & W. Sengenberger (Eds.), *Industrial districts and inter-firm cooperation in Italy* (pp. 37–51). Geneva: International Institute for Labour Studies (ILO).

Briant, A., Combes, P.-P., & Lafourcade, M. (2010). Dots to boxes: Do the size and shape of spatial units jeopardize economic geography estimations? *Journal of Urban Economics*, *67*, 287–302. doi:10.1016/j.jue.2009.09.014

Burger, M. J., Van Oort, F. G., & Van der Knaap, B. (2010). A treatise on the scale-dependence of agglomeration externalities and the MAUP. *Scienze Regionali*, *9*, 19–40.

Cainelli, G., Fracasso, A., & Vittucci Marzetti, G. (2015). Spatial agglomeration and productivity in Italy: A panel smooth transition regression approach. *Papers in Regional Science*, *94*, S39–S67. doi:10.1111/pirs.12103

Cainelli, G., Iacobucci, D., & Ganau, R. (2016). Do geographic concentration and vertically related variety foster firm productivity? Micro-evidence from Italy. *Growth and Change*, *47*, 197–217. doi:10.1111/grow.12112

Cainelli, G., & Leoncini, R. (1999). Externalities and long-term local industrial development. Some empirical evidence from Italy. *Revue d'Économie Industrielle*, *90*, 25–39. doi:10.3406/rei.1999.1762

Cainelli, G., & Lupi, C. (2010). Does spatial proximity matter? Micro-evidence from Italy. In N. De Liso, & R. Leoncini (Eds.), *Internationalization, technological change and the theory of the firm* (pp. 163–186). London: Routledge.

Cingano, F., & Schivardi, F. (2004). Identifying the sources of local productivity growth. *Journal of the European Economic Association*, *2*, 720–742. doi:10.1162/1542476041423322

David, P., & Rosenbloom, J. (1990). Marshallian factor market externalities and the dynamics of industrial localization. *Journal of Urban Economics*, *28*, 349–370. doi:10.1016/0094-1190(90)90033-J

De Lucio, J. J., Herce, J. A., & Goilcolea, A. (2002). The effects of externalities on productivity growth in Spanish industry. *Regional Science and Urban Economics*, *32*, 241–258. doi:10.1016/S0166-0462(01)00081-3

Deltas, G., De Silva, D., & McComb, R. P. (2015). *Industrial agglomeration and spatial persistence of employment in software publishing* (Working Paper No. 85393182). Lancaster: Management School, Department of Economics, Lancaster University.

Duranton, G., & Overman, H. G. (2005). Testing for localization using micro-geographic data. *Review of Economic Studies*, *72*, 1077–1106. doi:10.1111/0034-6527.00362

Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. In J. V. Henderson & J.-F. Thisse (Eds.), *Handbook of regional and urban economics, Vol. 4* (pp. 2064–2117). Amsterdam: Elsevier.

Duschl, M., Schimke, A., Brenner, T., & Luxen, D. (2014). Firm growth and the spatial impact of geolocated external factors. *Journal of Economics and Statistics (Jahrbücher fuer Nationalökonomie und Statistik)*, *234*, 234–256.

Duschl, M., Scholl, T., Brenner, T., Luxen, D., & Raschke, F. (2015). Industry-specific firm growth and agglomeration. *Regional Studies*, *49*, 1822–1839. doi:10.1080/00343404.2013.861059

Eriksson, R. H. (2011). Localized spillovers and knowledge flows: How does proximity influence the performance of plants? *Economic Geography*, *87*, 127–152. doi:10.1111/j.1944-8287.2011.01112.x

Fazio, G., & Maltese, E. (2015). Agglomeration externalities and the productivity of Italian firms. *Growth and Change*, *46*, 354–378. doi:10.1111/grow.12093

Gabriele, R., Giuliani, D., Corsino, M., & Espa, G. (2013). *Reassessing the spatial determinants of the growth of Italian SMEs* (Discussion Paper No. 2013/06). Trento: Department of Economics and Management (DEM), University of Trento.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Schleifer, A. (1992). Growth in cities. *Journal of Political Economy*, *100*, 1126–1152. doi:10.1086/261856

Heckman, J. (1979). Sample selection bias as a specification error. *Econometrica*, *47*, 153–161. doi:10.2307/1912352

Henderson, V. J. (2003). Marshall's scale economies. *Journal of Urban Economics*, *53*, 1–28. doi:10.1016/S0094-1190(02)00505-3

Henderson, V. J., Kuncoro, A., & Turner, M. (1995). Industrial development in cities. *Journal of Political Economy*, *103*, 1067–1090. doi:10.1086/262013

Jacobs, J. (1969). *The economy of cities*. New York: Vintage.
Krugman, P. (1991). *Geography and trade*. Cambridge, MA: MIT Press.
Mameli, F., Faggian, A., & McCann, P. (2014). Estimation of local employment growth: Do sectoral aggregation and industry definition matter? *Regional Studies*, *48*, 1813–1828. doi:10.1080/00343404.2012.756578
Marcon, E., Lang, G., Traissac, S., & Puech, F. (2012). *dbmss: Distance-based measures of spatial structures*. Retrieved from http://cran.r-project.org/web/packages/dbmss/
Marcon, E., & Puech, F. (2010). Measures of the geographic concentration of industries: Improving distance-based methods. *Journal of Economic Geography*, *10*, 745–762. doi:10.1093/jeg/lbp056
Marrocu, E., Paci, R., & Usai, S. (2013). Productivity growth in the old and new Europe: The role of agglomeration externalities. *Journal of Regional Science*, *53*, 418–442. doi:10.1111/jors.12000
Marshall, A. (1920). *Principles of economics*. London: Macmillan.
Martin, F., Mayer, T., & Mayneris, F. (2011). Spatial concentration and plant-level productivity in France. *Journal of Urban Economics*, *69*, 182–195. doi:10.1016/j.jue.2010.09.002
Martin, R. (1999). Critical survey. The new 'geographical turn' in economics: Some critical reflections. *Cambridge Journal of Economics*, *23*, 65–91. doi:10.1093/cje/23.1.65
Munari, F., Sobrero, M., & Malipiero, A. (2012). Absorptive capacity and localized spillovers: Focal firms as technological gatekeepers in industrial districts. *Industrial and Corporate Change*, *21*, 429–462. doi:10.1093/icc/dtr053
Olsen, J. (2002). On the units of geographical economics. *Geoforum*, *33*, 153–164. doi:10.1016/S0016-7185(01)00034-3
Paci, R., & Usai, S. (2008). Agglomeration economies, spatial dependence and local industry growth. *Revue d'Économie Industrielle*, *123*, 87–109. doi:10.4000/rei.3917
Porter, M. E. (1990). *The competitive advantage of nations*. New York: Free Press.
Rauch, J. E. (1993). Does history matter only when it matters little? The case of city–industry location. *Quarterly Journal of Economics*, *108*, 843–867. doi:10.2307/2118410
Ripley, B. D. (1977). Modelling spatial patterns. *Journal of the Royal Statistical Society B*, *39*, 172–212.
Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy*, *94*, 1002–1037. doi:10.1086/261420
Rosenthal, S. S., & Strange, W. C. (2003). Geography, industrial organization, and agglomeration. *Review of Economics and Statistics*, *85*, 377–393. doi:10.1162/003465303765299882
Rosenthal, S. S., & Strange, W. C. (2008). The attenuation of human capital spillovers. *Journal of Urban Economics*, *64*, 373–389. doi:10.1016/j.jue.2008.02.006
Saxenian, A. (1994). *Regional advantage. Culture and competition in Silicon Valley and Route 128*. Cambridge, MA: Harvard University Press.
Scholl, T., & Brenner, T. (2016). Detecting spatial clustering using a firm-level cluster index. *Regional Studies*, *50*, 1054–1068. doi:10.1080/00343404.2014.958456
Shaver, M. J., & Flyer, F. (2000). Agglomeration economies, firm heterogeneity, and foreign direct investment in the United States. *Strategic Management Journal*, *21*, 1175–1193.
Sorenson, O., & Audia, P. G. (2000). The social structure of entrepreneurial activity: Geographic concentration of footwear production in the United States, 1940–1989. *American Journal of Sociology*, *106*, 424–462. doi:10.1086/316962
Van Oort, F. G. (2007). Spatial and sectoral composition effects of agglomeration economies in the Netherlands. *Papers in Regional Science*, *86*, 5–30. doi:10.1111/j.1435-5957.2006.00088.x
Van Oort, F. G., Burger, M. J., Knoben, J., & Raspe, O. (2012). Multilevel approaches and the firm-agglomeration ambiguity in economic growth studies. *Journal of Economic Surveys*, *26*, 468–491. doi:10.1111/j.1467-6419.2012.00723.x
Wang, C. C. (2015). Geography of knowledge sourcing, search breadth and depth patterns, and innovative performance: A firm heterogeneity perspective. *Environment and Planning A*, *47*, 744–761. doi:10.1068/a140024p
Wong, D., & Amrhein, C. G. (1996). Research on the MAUP: Old wine in a new bottle or real breakthrough? *Geographical Systems*, *3*, 73–76.
Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters*, *104*, 112–114. doi:10.1016/j.econlet.2009.04.026
Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.