Empir Econ
DOI 10.1007/s00181-014-0834-6

A micro spatial analysis of firm demography: the case of food stores in the area of Trento (Italy)

Giuseppe Arbia · Patrizia Cella · Giuseppe Espa · Diego Giuliani

Received: 31 October 2012 / Accepted: 24 February 2014

Abstract The potential offered by the increasing availability of micro-geographic data is tremendous and still largely not exploited. In this paper, we propose a new methodology to study the spatial dynamics of firm demography making use of such rich source of information. Exploiting the techniques of stochastic spatial point processes (Diggle, Statistical analysis of spatial point patterns, 2003), it is indeed possible to model the firm birth, growth and death explaining the behavior of the individual economic agent. We consider the spatial distribution of the economic activities as the result of a dynamic process occurring in space and time, and we model the micro-spatial patterns as single realizations of marked space time survival point processes (Rathbun and Cressie, J Am Stat Assoc 89:1164–1174, 1994). Within such a methodological framework firms are created at some random location and some point of time and then operate, grow, and attract (or repulse) the localization of other firms in their neighborhood. The proposed model is fitted to some empirical data on firms' locations sourced from the ASIA harmonised archive of the Italian National Institute of Statistics. The empirical analysis reveals the presence of local competition behavior in firm creation among small retail food stores, thus shedding light on the phenomena of firm demography.

Keywords Firm demography · Spatial point pattern analysis · Economic geography

G. Arbia (✉)
Department of Statistics, Università Cattolica del Sacro Cuore, Rome, Italy
e-mail: giuseppe.arbia@rm.unicatt.it

P. Cella
Italian National Institute of Statistics (ISTAT), Rome, Italy

G. Espa · D. Giuliani
Department of Economics and Management, University of Trento, Trento, Italy

Published online: 19 July 2014

JEL Classification C15 C21 R12

1 Introduction

The current state-of-the-art in the study of firm demography includes a vast variety of contributions and empirical methodologies mainly for data aggregated at a macro or meso territorial levels, in which the typical observations consist of administrative units such as regions, counties or municipalities. In this field, it is well known that the distribution of firms is spatially uneven, and that policies to promote new firm formation should take spatial location and spatial proximity into account due to the presence of neighborhood and spatial contagion effects (see e.g., Storey 1984, 1985). The emphasis of the studies in the fields, however, has changed dramatically in the last decades. From a mere attempt to explain regional variations in new firm formation using mostly unsophisticated statistical tools (Reynolds et al. 1994), more recently studies have shifted towards the use of more sophisticated time series and cross/sectional techniques aiming at clarifying the role of entrepreneurship in the process of new firm formation and of their spatial diffusion (see e.g., Zoltan and Storey 2004; Audretsch and Keilbach 2004; Arauzo et al. 2010).

In this stream of studies, as said, the typical observation is represented by a region (however defined), while comparatively less attention has been devoted to the development of a systematic approach to the analysis of micro-territorial data where the observations are represented by the locations of each single firm. Indeed, this fact has strongly limited the possibility of obtaining robust evidences about new firm formation and firm survival phenomena in two respects. First of all, geographically aggregated data (such as regional firm birth and death rates) are based on an arbitrary definition of the spatial units and, as a consequence, they introduce a statistical bias arising from the discretional characterization of space adopted (the so-called *modifiable areal unit* problem, or MAUP, bias. See Arbia 1989). Secondly, most theoretical models of firm demography speculate on the individual behavior of the economic agents (see Hopenhayn 1992; Krueger 2003; Lazear 2005 among others). This implies that empirical methods based on regional aggregates can provide proper evidence to the theoretical models only under the restrictive, and most of the time unrealistic, assumption of homogeneity of firm's behavior within the regions. In the few remarkable cases where a genuine micro approach was adopted, the results confirm the relevance of neighborhood effects that reveal interesting scenarios for future researches. For instance, Igami (2011) shows that the entry of a large supermarket in one area increases the chance of exit of larger stores in the neighborhood, but increases the probability of survival of smaller incumbents. Analyzing a set of data collected in the food retailing sector, Borraz et al. (2013) show that the entry of a supermarket in the neighborhood of a small store significantly increases the probability of the small store to go out of business in the same year.

Considering the plausible scenario that geocoded information on individual productive plants will become more and more available in the near future, this article is devoted to the proposal of a new approach to the analysis of the spatial dynamics of firm demography based on micro-geographic data. Indeed, exploiting stochastic point processes theory (Diggle 2003), we show that it is possible to uncover some aspects of firm demography and model the links of spatial interactions among the individual economic agents. In this way, by treating the spatial distribution of the economic activities as the result of a dynamic process occurring in space and time, the observed micro-geographic patterns can be modeled as realizations of a marked space time survival point processes (Rathbun and Cressie 1994), where firms are created at some random location and at some point in time and then operate, grow, and attract (or repulse) other firms in their neighborhood. Following the reductionist approach proposed by Rathbun and Cressie (1994) and already exploited in Arbia (2001), we formalize three different model components, namely a birth model, a growth model, and a death/survival model. The estimated model parameters provide useful indications to validate the different paradigmatic economic-theoretical cases like, e.g., the presence of spatial spillover effects, the effects of positive spatial externalities, or hypotheses concerning the spatial inhibition processes among economic agents. As a case study, we apply the proposed methodology to the observed spatial distribution of small retail food stores located in the city of Trento, Italy.

The paper is divided into six sections. In particular, in Sect. 2, we will describe the data that will form the basis of our case study. Section 3 will be devoted to the specification and estimation of a model of birth of new firms, whereas Sects. 4 and 5 will contain, respectively, a discussion of the model of growth and survival of the existing firms. Finally, Sect. 6 contains some conclusions and priorities for further research in the field.

2 Data description

In the past, most of the studies on firm demography were based on data on firm location aggregated at the level of geographical partitions such as regions, counties or countries. The scenario, however, is rapidly changing and the accessibility to micro geographical data related to the single economic agents is becoming more and more common in many applied studies. In particular, when dealing with firm demography, many existing official databases are enriched with detailed information related to the single firm including its geographical coordinates and a set of relevant variables such as production, capital and labor inputs, level of technology, and many others. A good example of such an informative database is represented by the Statistical Register of Active Enterprises (ASIA), managed and updated by the Italian Statistical Institute (ISTAT). At a firm level, this database currently contains, for the period 2004–2009, information about firm code, tax code, business name, sector of activity (according to the Nace classification), number of firm's employees, legal status (according to the current classification), class of sales, firm's birth and (if applicable) termination date. Drawn from such a rich database, the data employed in the case study analyzed in the present paper refer to the geographical location and the number of employees of small retail food stores and big supermarkets/hypermarkets[1], also selling food

[1] The identification of these two kinds of economic activities has been made referring to the OECD/Eurostat classification scheme (NACE Rev 2).

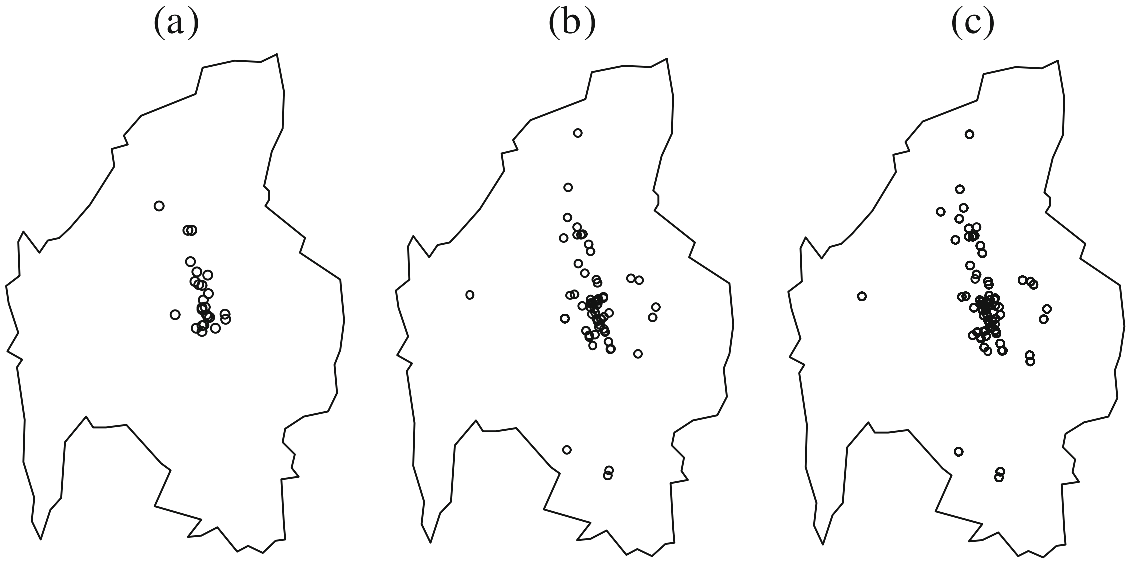

Fig. 1 Spatial point patterns of the small retail food stores used to estimate, respectively, a the firm birth process in the period 2004–2007, b the firm growth process in the period 2004–2007 and c the firm survival process in the period 2004–2007. Municipality of Trento, Italy

products, operating in the town of Trento (North of Italy) in the period 2004–2007. The model consists of three basic components: a model describing the process of birth of new firms, a model describing the process of growth of the existing firms, and a model describing the survival of the existing firms after some years of their creation. We conceive the three models as a way of describing different aspects of the phenomenon. Consequently, (and consistently with Rathbun and Cressie 1994) the three models are estimated separately and not simultaneously with no interaction effects, thus eliminating possible problems of simultaneity that could potentially arise in the estimation methods undermining their validity. More specifically, the case study refers to the food stores in Trento, the birth component model will be fitted to the observed spatial distribution of the 26 small retail food stores born after 2004 and which were still active in 2007. Similarly, the growth component model will be fitted to the observed spatial pattern of the 72 existing small retail food stores, which (born in 2004 or before) were always active in the period from year 2004 and 2007. Finally, the death/survival component process will be fitted to the 229 small retail food stores which were born in 2004 or before and that survived at least until 2007, or ceased to operate in some year during the period 2004-2007. The three point patterns relative, respectively, to birth, growth, and survival, are represented in Fig. 1.

3 A spatial birth process

3.1 The birth model

In our methodological framework, the observed spatial point pattern of new small food stores births is assumed to be the realization of a point process conditional on the locations of existing small food stores and big supermarkets in that moment. Our model aims at testing the hypothesis that big supermarkets inhibit the birth of small food stores in the nearby locations while, due to the presence of positive spatial externalities, the presence of small food stores attracts the location of other small food stores (see Igami 2011; Borraz et al. 2013).

In order to formalize our model, we rely on the spatial point process methodology (Diggle 2003). Within this framework, a spatial point process is considered as a stochastic mechanism that generates patterns of point on a planar map (such as those depicted in Fig. 1). The basic characteristic of a spatial point process is the so-termed intensity function, denoted with the symbol $\lambda(x)$, and formally defined as

$$\lambda(x) = \lim_{dx \to 0} \left\{ \frac{E[N(dx)]}{dx} \right\}$$

where (in accordance with to the notation adopted in Diggle 2003), x is a generic point of coordinates, say, $x = (x_1, x_2)$, dx an infinitesimal portion of space containing point x, dx its surface area, and $N(dx)$ the number of points falling within dx. In this setting, the intensity $\lambda(x)$ represents the expected number of points located within an infinitesimal region centered at the generic point x. Thus, by definition, the higher is $\lambda(x)$, the higher is the expected concentration of points around x (see Arbia et al. 2008).

Following the modeling framework originally proposed in a seminal paper by Rathbun and Cressie (1994) and imported by Arbia (2001) in the field of regional economics, the new firm formation process of the small food stores during period 2004–2007 is modeled as an inhomogeneous Poisson process (see Diggle 2003) with intensity function $\lambda(x)$ driven by the potential interaction effects of the existing small food stores and big supermarkets. The values of $\lambda(x)$ constitute a realization of a random function parametrically specified by the following model:

$$\lambda(x) = \exp \quad \alpha \quad \beta_{ss} n_{ss}(x) \quad \beta_{bs} n_{bs}(x) \quad \beta_{nh} n_{nh}(x) \quad \beta_{af} n_{af}(x) \quad , \tag{1}$$

where β_{ss}, β_{bs}, β_{nh}, and β_{af} are parameters to be estimated. The variables $n_{ss}(x)$ and $n_{bs}(x)$ measure the overall number of employees of the small food stores and, respectively, of big supermarkets, existing from before 2005 which are located around the arbitrary point x.

The two additional variables $n_{nh}(x)$ and $n_{af}(x)$ control for other factors that can affect the spatial intensity of the new small food stores formation process. On one hand, locational choices of firms, and in particular of retail activities, can strongly depend on the potential market demand. As a proxy of the spatial distribution of potential customers in the city of Trento we use the number of households by census tract in 2004, which is the finest level of spatial resolution. In order to properly include this data in the modeling framework of Eq. (1) we constructed a marked point pattern where the points are the centroids of the census tracts and the associated marks represent the number of household per census tract (see Fig. 2). Then we could define the control variable $n_{nh}(x)$, which measures the number of households in 2004 that reside close to the arbitrary point x.

On the other hand, the decision to start a new firm in a particular location is also affected by the spatial characteristics of the territory (such as the urban structure, the presence of useful infrastructures or the environmental, and administrative limits). To

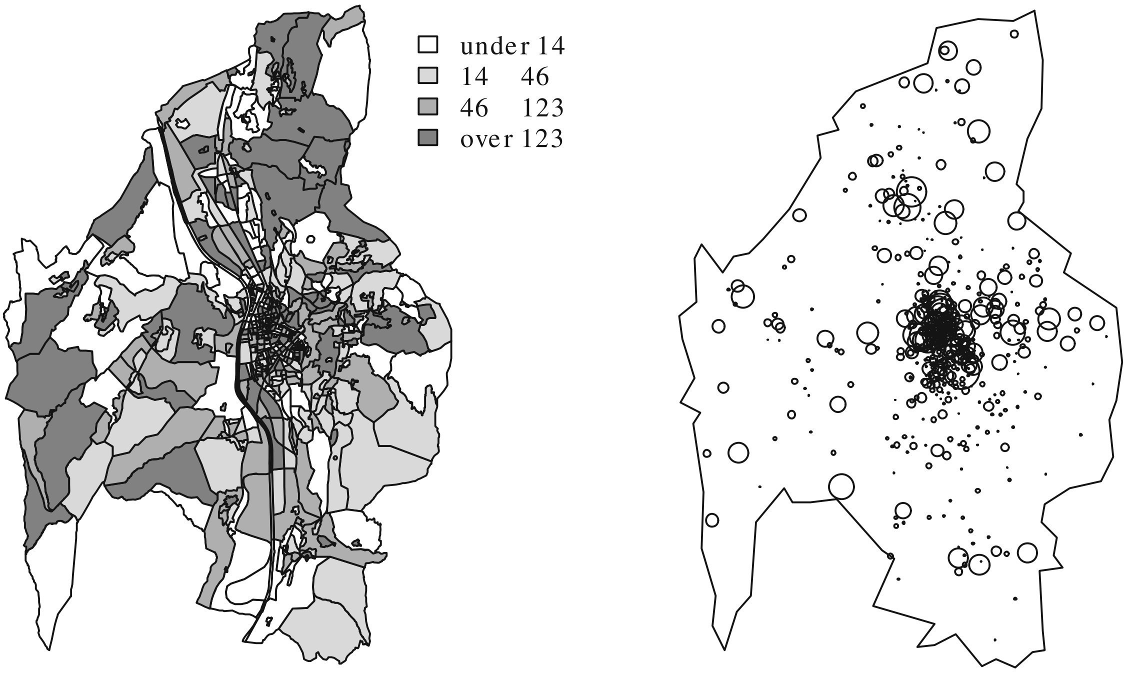

Fig. 2 Number of households by census tract in 2004 in the city of Trento: a as a census tracts map of the quartile distribution and b as a marked point pattern

control for these unidentified sources of spatial heterogeneity, we included variable $n_{af}(x)$, which is constructed as the overall number of employees of all firms of all industries, operating from before 2005, located around the generic point x. The use of this specific control variable is motivated by the assumption that the main unobserved exogenous spatial factors affecting the locational choices of firms are common for all the economic agents.

The model is based on the working assumption that the locations of incumbent economic agents operating before 2005 are exogenously given. As already remarked by Arbia (2001), this hypothesis is consistent with Krugman's idea of "historical initial conditions" (Krugman 1991).

The logarithmic transformation allows to fit the model maximizing the log-pseudolikelihood for $\lambda(x)$ (see Besag 1975) based on the points x constituting the observed point pattern. According to the current state-of-the-art in the spatial statistics literature, the most efficient and versatile method of maximizing the log-pseudolikelihood (thus obtaining unbiased estimates of the parameters) is the technique proposed by Berman and Turner (1992) (For a clear and detailed discussion of the method we refer to Baddeley and Turner 2000). As shown by Strauss and Ikeda (1990), maximum pseudolikelihood is equivalent to maximum likelihood in the case of a Poisson stochastic process. Therefore, it is possible to test the significance of the estimated model parameters using standard formal likelihood ratio criteria based on the χ^2 distribution.

3.2 Empirical results

The maximum pseudo-likelihood estimates of the parameters for the birth process model (1) are reported in Table 1.

Table 1 Estimates for the spatial birth point process of new small food stores

Parameter	Estimate	Standard error	z-test
α	18.900	0.8809	
β_{ss}	0.074	0.0234	**
β_{bs}	0.028	0.0075	***
β_{nh}	0.004	0.0012	***
β_{af}	0.001	0.0003	***

** Significant at 5%;
*** Significant at 1%

The significantly positive value of the estimate of β_{ss} indicates that births of small food stores are positively dependent on locations and sizes of the existing small food stores, while the significantly negative value of the estimate of β_{bs} indicates that they are negatively dependent on locations and sizes of the existing big supermarkets. Therefore, the probability of the birth of new small food stores is higher in the locations characterized by the presence of other existing small food stores in the neighborhood, thus highlighting the presence of positive spatial externalities. On the other hand, such probability is lower in the locations characterized by the near presence of existing big supermarkets, which indicates the presence of negative spatial externalities. The model also reveals the presence of a positive significant relationship between the spatial intensity and the two proxies of market potential and the urban structure (respectively the parameters β_{nh} and β_{af}).

In the model considered, no measure is available to test the goodness of fit playing a role similar to that of R^2 in the OLS standard regression. To this aim, however, it is possible to rely on Monte Carlo simulations. The adequacy of the model to the observed data can indeed be visually assessed by looking at the behavior of the empirical inhomogeneous K-function (Baddeley et al. 2000) of the observed 26 new born small retail food stores point pattern (see Fig. 1a) plotted against the behavior of the inhomogeneous K-function, in terms of confidence bands, derived from 999 simulations of the estimated model.

The inhomogeneous K-function, $K_I(d)$, is a summary measure of the characteristics of a spatial point process. In particular, at each possible distance d, it identifies the level of spatial dependence among points of a point process identified by a non-constant intensity function $\lambda(x)$ (Espa et al. 2013). Baddeley et al. (2000) showed that, for an inhomogeneous Poisson process with a given intensity function $\lambda(x)$, the $K_I(d)$ function is equal to $K_I(d) = \pi d^2$. As a consequence, any observed point pattern, such that its empirical inhomogeneous K-function (that is $\hat{K}_I(d)$), does not deviate significantly from the theoretical inhomogeneous K-function (that is πd^2) for each value of d, can be considered as a realization of that model.

Following Baddeley et al. (2000), $K_I(d)$ can be estimated using the edge-corrected unbiased estimator, $\hat{K}_I(d)$, given by

$$\hat{K}_I(d) = \frac{1}{A} \sum_{i=1}^{n} \sum_{j \neq i} \frac{w_{ij} I\left(d_{ij} \leq d\right)}{\hat{\lambda}(x_i)\hat{\lambda}\left(x_j\right)}$$

where A is the total surface of the study area A, the term d_{ij} is the Euclidean spatial distance between the ith and jth observed points and $I(d_{ij} \leq d)$ represents the indi-

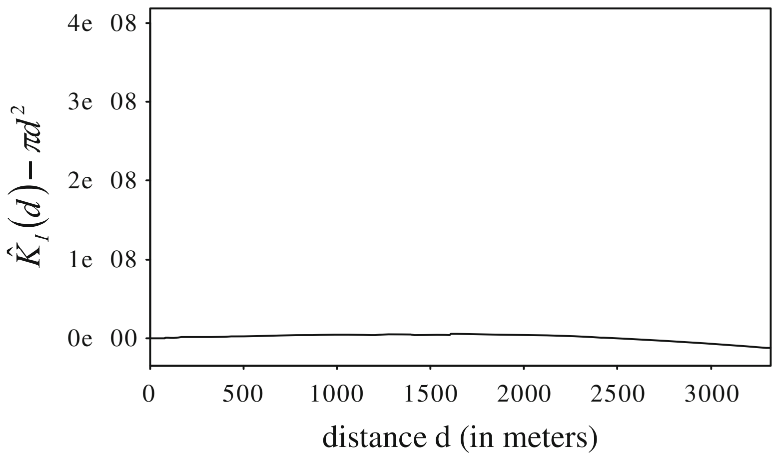

Fig. 3 Behavior of the empirical $\hat{K}_I(d) \quad \pi d^2$ (continuous line) and the corresponding 99.9% confidence bands (dashed lines)

cator function such that $I = 1$ if $d_{ij} \leq d$ and 0 otherwise. Due to the presence of edge effects arising from the arbitrary delimitation of the study area, the adjustment factor w_{ij} is introduced to avoid potential biases in the estimates close to the boundary of A. More specifically, the weight function w_{ij} expresses the reciprocal of the proportion of the surface area of a circle centred on the ith point, passing through the jth point, which lies within A (Boots and Getis 1988).

Baddeley et al. (2000) proposed to estimate the intensity function of the observed pattern using the following non-parametric method of kernel smoothing (see Silverman 1986):

$$\hat{\lambda}_h(x_i) = \sum_{i \neq j} h^{\ 2} k\left(\frac{x_j \quad x_i}{h}\right) \Big/ C_h(x_j),$$

where k( ) is a radially symmetric bivariate probability density function, h represents the bandwidth (the parameter controlling the smoothness of the surface), and $C_h(x_j)$ is an edge-correction factor[2].

The graph reported in Fig. 3 shows the behavior of the empirical $\hat{K}_I(d) \quad \pi d^2$ calculated from the observed point pattern of the new-born small food stores against the upper and lower confidence bands calculated from 999 realizations of the estimated model. The benchmark value representative of a good fit is zero for each distance d. As it can be noted, the empirical function tends to be close to zero and lies entirely within the confidence bands thus indicating that the estimated model describes adequately the spatial birth phenomenon of small food stores.

4 A spatial growth process

4.1 The growth model

The spatial growth dynamic of small food stores is modeled using the observed point pattern of small food stores that were created in 2004 or before and that survived for the

[2] The kind of edge-correction factor Baddeley et al. (2000) used is essentially a slightly modified version of that proposed in Berman and Diggle (1989), $C_h(x_j) = \int_A k_h(x_j \quad u)\, du$.

whole period 2005–2007 (see Fig. 1b). In this context, the growth of a single small food store (proxied by the growth in the number of employees), is assumed to be a function of its stage of development at the beginning of the period and of the competitive (or co-opetitive) influences of the other food stores located in the neighborhood.

Let $x_{ss,i} : i = 1, \ldots, n$ and $x_{bs,i} : i = 1, \ldots, m$ denote, respectively, the spatial coordinates of small food stores and of big supermarkets existing at least from 2004 and that have survived until 2007. Let also $Z04_{ss,i}(Z04_{bs,i})$ and $Z07_{ss,i}(Z07_{bs,i})$ denote the number of employees of the ith small store (big supermarket) in 2004 and 2007, respectively, and let $g_i = Z07_{ss,i}/Z07_{ss,i} Z04_{ss,i}$. $Z04_{ss,i}$ be a proxy of the growth of the ith small store.

Following the framework suggested by Rathbun and Cressie (1994), the growth of the ith small food store can be modeled as follows:

$$g_i = \alpha \quad \beta_Z Z04_{ss,i} \quad \beta_{ss} W_{ss,i} \quad \beta_{bs} W_{bs,i} \quad \varepsilon_i, \tag{2}$$

where

$$W_{ss,i} \equiv \sum_{j=1}^{n} \Phi_{ij} \left(Z04_{ss,j}, d_{ij} = \left\| x_{ss,i} \quad x_{ss,j} \right\| \right)$$

and

$$W_{bs,i} \equiv \sum_{j=1}^{m} \Phi_{ij} \left(Z04_{bs,j}, d_{ij} = \left\| x_{ss,i} \quad x_{bs,j} \right\| \right) \tag{3}$$

represent measures of the level of spatial interaction of the ith small food store with, respectively, the other small food stores and big supermarkets, α, β_Z, β_{ss}, β_{ss} are parameters to be estimated and the ε_i's are independently and normally distributed errors with zero mean and a finite variance. Relying on the idea that the level of spatial interaction of a neighboring economic activity should depend on the distance to that economic activity and on its size (here approximated by the number of employees), the two measures of spatial interaction reported in Eq. (3) are derived by specifying the functional form of $\Phi_{ij}(Z, d)$. Rathbun and Cressie (1994) proposed to chose between the following functional forms: $\Phi_{ij}(Z, d) = 1/d$, $\Phi_{ij}(Z, d) = 1/d^2$, $\Phi_{ij}(Z, d) = Z/d$, $\Phi_{ij}(Z, d) = (Z/d)^2$, and $\Phi_{ij}(Z, d) = Z^2/d$.

Having chosen a proper specification of $\Phi_{ij}(Z, d)$, and hence given Eq. (3), the parameters of the growth model in Eq. (2) can then be estimated using OLS.

4.2 Empirical results

The OLS estimates of the parameters for the growth process model (2) fitted to the data of the small food stores are reported in Table 2. In our analysis, we specified the spatial interaction term as $\Phi_{ij}(Z, d) = Z^2/d$, which is the one maximizing the fit in terms of the coefficient of determination R^2.

As it can be noted, the parameters β_Z and β_{ss} are not significant, thus implying that in our specific case, the growth of small food stores is not affected by their size at the

Table 2 Estimates for the growth birth point process of new small food stores

Estimates	Full model	Restricted model
α	1.1600 *** (0.1145)	1.1190 *** (0.0622)
βZ	–0.0317 (0.0364)	–
β_{ss}	0.0446 (0.0987)	–
β_{bs}	0.0000 *** (0.0000)	0.0000 *** (0.0000)
R^2	0.6811	0.6921

*** Significant at 1%

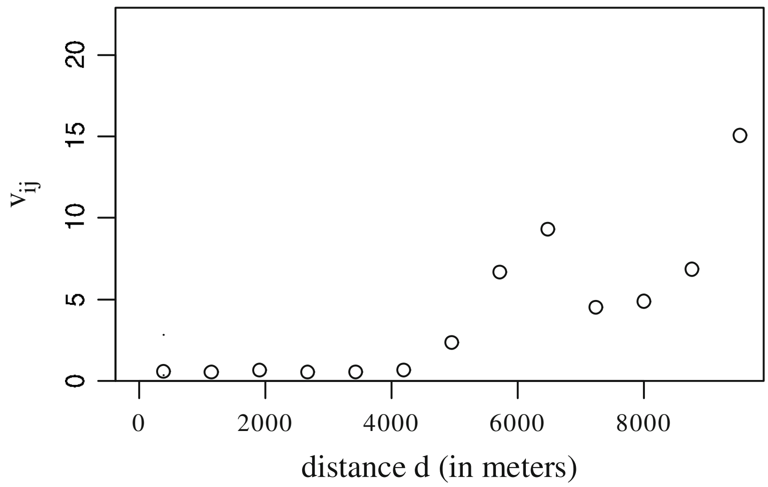

Fig. 4 Monte Carlo envelopes for the variogram of the model (2) studentized residuals (dashed lines) and empirical variogram of residuals (circles)

beginning of the period nor by the closeness to other economic activities of the same typology. On the other hand, the parameter β_{bs} is significant and positive (although very close to zero), thus indicating that spatial interactions with big supermarkets foster firm growth. In other words, growth rates are higher for small food stores located in the proximity of big supermarkets than for those that are far from them. This evidence suggests the presence of some sort of co-opetitive behavior among competitive economic agents. Similar results have been found by Igami (2011) and Borraz et al. (2013).

Model (2) is based on the implicit assumption that growth rates are spatially independent (Rathbun and Cressie 1994). The presence of spatial variability, however, may lead to the violation of this assumption, which in turn could result in spatial autocorrelation of model residuals. The proper diagnostic tool to test whether the residuals of a micro-geographic data model are spatially correlated is the variogram. For the studentized model residuals the empirical variogram, the ordinates are represented by the quantities $v_{ij} = \frac{1}{2}\left(r_i \quad r_j\right)^2$, where r_i and r_j are the studentized residuals corresponding to the small food stores at the locations $x_{ss,i}$ and $x_{ss,j}$, respectively (Diggle and Ribeiro 2007). A plot of v_{ij} against the corresponding distance $d_{ij} = \|x_{ss,i} \quad x_{ss,j}\|$ compared with the envelope of the empirical variograms computed from random permutations of the residuals, holding their locations fixed, allows to detect spatial autocorrelation. Figure 4 shows a variogram envelope obtained from 999 independent random permutations of the studentized residuals, with values averaged within distance bands. Since all the empirical variogram ordinates are within the simulation envelopes, we are eager to conclude that there is no spatial dependence among the model residuals.

5 A spatial death/survival process

5.1 The survival model

The last component of the model is devoted to uncover the death/survival process that together with the other two components contributes to a full understanding of the whole spatial demographic phenomenon. The death/survival process is fitted to the small food stores existing in 2004 and which have survived or ceased to operate during the period 2005–2007 (Fig. 1c). Therefore, the data of interest consist of the spatial point pattern of small food stores observed at the time periods t = 2005, 2006, 2007. According to the methodological framework proposed by Rathbun and Cressie (1994), a death/survival process can be developed as a survival conditional probability model which will be now described. Let $x_{ss,i}(t) : i = 1, \ldots, n$ and $x_{bs,i}(t) : i = 1, \ldots, m$ denote the spatial coordinates of small food stores and big supermarkets, respectively, which survives at time t. Let also $Z_{ss,i,t}$ and $Z_{ss,i,t}$ represent the number of employees of the ith small store and of the ith big supermarket, respectively, at time t. Finally let $M_i(t)$ denote a survivorship indicator variable such that $M_i(t) = 1$ if the ith small food store survives at time t and $M_i(t) = 0$ if the ith small food store ceases its activity at time t.The survival probability of the ith small food store can then be defined as:

$$p_i\,(t;\theta) \equiv P \quad M_i\,(t;\theta) = 1 \; M_i\,(t \quad 1;\theta) = 1$$

where θ is a set of unknown parameters to be estimated. By assuming that the survival at time t is a function of the dimension of the small food store at time $t-1$ and of the competitive or co-opetitive influences of the other neighboring small food stores and big supermarkets also operating at time $t-1$, $p_i(t;\theta)$ is then modeled using the following space–time logistic regression (Rathbun and Cressie 1994):

$$\log \frac{p_i\,(t;\theta)}{1 \quad p_i\,(t;\theta)} = \alpha \quad \beta_Z Z_{ss,i,t \quad 1} \quad \beta_{ss} W_{ss,i,t \quad 1} \quad \beta_{bs} W_{bs,i,t \quad 1} \tag{4}$$

where

$$W_{ss,i,t \quad 1} \equiv \sum_{j=1}^{n} \Phi_{ij} \left(Z_{ss,j,t \quad 1}, d_{ij} = \left\| x_{ss,i}\,(t \quad 1) \quad x_{ss,j}\,(t \quad 1) \right\|\right) \tag{5a}$$

and

$$W_{bs,i,t \quad 1} \equiv \sum_{j=1}^{m} \Phi_{ij} \left(Z_{bs,j,t \quad 1}, d_{ij} = \left\| x_{ss,i}\,(t \quad 1) \quad x_{bs,j}\,(t \quad 1) \right\|\right) \tag{5b}$$

Table 3 Estimates for the death/survival point process of new small food stores

Estimates	Full model	Restricted model
α	1.7220 *** (0.4842)	2.3076 *** (0.2345)
β_Z	0.1811 (0.1648)	–
β_{ss}	– 0.0003 *** (0.0000)	– 0.0003 *** (0.0000)
β_{bs}	2.6880 (5.6140)	–
Log-likelihood	–70.7685	–72.4076

*** Significant at 1%

represent measures of the level of spatial interaction of the ith small food store with, respectively, the other small food stores and big supermarkets, $\Phi_{ij}(Z, d)$ is a known function which can be specified by the same functional forms already proposed for the growth model's Eqs. (2) and (3), and $\theta \equiv (\alpha, \beta_Z, \beta_{ss}, \beta_{bs})'$ is the vector parameters to be estimated.

Chosen a proper specification of $\Phi_{ij}(Z, d)$ and hence given Eqs. (5a–5b), θ for model (4) can be estimated using the standard maximum likelihood (ML) procedure.

5.2 Empirical results

For estimation purposes, first of all, we specified the spatial interaction term (similarly to Sect. 4.2) as $\Phi_{ij}(Z, d) = Z^2/d$, which is the specification that produces the better fit in terms of the value of the log-likelihood. The ML estimates of the parameters for the death/survival model expressed in Equation (4) are shown in Table 3.

The model results show that parameters β_Z and β_{bs} are not significant, thus implying that the survivorship of small food stores is unaffected by their size and by the closeness to big supermarkets. Furthermore, since β_{ss} is significant and negative (although very close to zero), the estimated model indicates that spatial interactions with the other small food stores result in a relatively lower probability to survive. This evidence reinforces the conjecture of a competitive behavior between small food stores.

6 Conclusions

In this paper, we aim at proposing a novel model-based approach to the analysis of the dynamics of firm demography based on micro-geographic data. Following the reductionist modeling framework introduced by Rathbun and Cressie (1994) and already proposed in Arbia (2001), we presented a methodology to estimate a three equations model dealing, respectively, with firms birth, growth, and survival. Consistently with Rathbun and Cressie (1994), the three equations are estimated independently thus eliminating possible problems of simultaneity in the estimation phase. We argue that decomposing firm demography processes into these three sub-processes allows to uncover the relative importance of competitive and co-opetitive spatial interactions in determining the spatial distribution of economic activities.

The case study examined referred to the demography of small retail food stores located in the municipality of Trento in the period 2004–2007. The analysis indicates that space and spatial proximity are relevant factors in determining the phenomena of birth of new firms. On the contrary, the process of growth and survival of the existing firms seems to be unaffected by spatial elements and by the presence of other economic agents in the neighborhood. With particular reference to the data examined, we showed that the creation of small retail food stores tends to occur nearby other small food stores, but far from the big supermarkets/hypermarkets. In contrast, once the small food stores are operative, their tendency to grow in their dimension and to survive in time seems to be unaffected by spatial factors such as the influence of neighboring competitor or co-opetitor firms.

It is important to note that in the present context we could only use information about the spatial distribution of firm entries, exits, and incumbent firms. Therefore, we limited ourselves to providing evidence of the presence of significant spatial interactions among economic agents and their effects on firm demography. On the other hand, in order to uncover the entire locational process of economic agents, it would be necessary to have access to a larger information set on structural variables other than just the mere geographical location of firms, such as, for instance, the characteristics of the local demand, workforce skill and urban structure. For this reason, this purpose is not undertaken in the present study and is left to some future analysis is beyond the scope of this paper.

The present empirical application shows the potential of the proposed methodologies to describe some characteristics of the firm demography belonging to a single economic sector. However, the practical analysis of a single sector is only one of the possible approaches that could be employed. In some future studies, the techniques presented here will be extended to include more comprehensive studies which model the role of spatial proximity in the process of co-agglomeration and in the analysis of the joint locational behavior of the different economic sectors. In this way, the global pattern of firm location, growth, and survival observed at the level of the economy as a whole will be modeled as the outcome of the individual firm choices and their interaction in space and time. Following this more comprehensive multivariate analysis dealing with all industries simultaneously will also imply the fact that the analysis cannot be limited any more to a single region, but it should take into account a whole country, possibly considering the model parameters to change smoothly in space. This approach is conversely not undertaken here where the emphasis is on the regional characteristics of firm formation and where we refer to a single industry. A comprehensive approach to the modeling of the micro behavior of economic agents has been suggested for the first time some decades ago by Durlauf (1989, 1999) who established a parallel between physics and economic analysis. In his pioneering work, he observed that the basic idea underlying statistical mechanics is that the behavior of one atom is influenced by the behavior of other atoms located nearby. In a similar fashion, the hypothesis that forms the basis of all spatial economic studies is that individual or collective choices depend on the decisions taken by other actors located nearby. At the time of the papers by Durlauf, the approach suggested was not susceptible of any empirical validation due to the lack of data, the lack of statistical models, and the reduced computer capabilities. In contrast, nowadays we have the data, the statistical methodologies and the computational power to treat in a rigorous way the immense richness of observational spatial data that are currently available. This paper presents a contribution in this direction.

References

Arbia G (2001) Modeling the geography of economic activities on a continuous space. Pap Reg Sci 80: 411–424
Arbia G (2006) Spatial econometrics: statistical foundations and applications to regional convergence. Springer, Berlin
Arbia G (2011) A lustrum of SEA: recent research trends following the creation of the spatial econometrics association (2007–2011). Spat Econ Anal 6(4):377–395
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
Arauzo JM, Liviano D, Manjón M (2010) Empirical studies in industrial location: an assessment of their methods and results. J Reg Sci 50(3):685–711
Audretsch DB, Fritsch M (1999) The industry component of regional new firm formation processes. Rev Ind Organ 15:239–252
Audretsch DB, Keilbach M (2004) Entrepreneurship capital and economic performance. Reg Stud 38(8):949–959
Baddeley A, Møller J, Waagepetersen R (2000) Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. Stat Ned 54(3):329–350
Baddeley A, Turner R (2000) Practical maximum pseudolikelihood for spatial point patterns. Aust Nz J Stat 42(3):283–322
Bartelsman E, Scarpetta S, Schivardi F (2003) Comparative analysis of firm demographics and survival: micro-level evidence for the OECD countries, OECD economics department working papers 348, OECD Publishing
Berman M, Diggle PJ (1989) Estimating weighted integrals of the second-order intensity of a spatial point process. J R Stat Soc, B 51:81–92
Berman M, Turner TR (1992) Approximating point process likelihoods with GLIM. Appl Stat 41(1):31–38
Besag J (1975) Statistical analysis of non-lattice data. Stat 24(3):179–195
Boots BN, Getis A (1988) Point pattern analysis, sage scientific geography series, vol 8. Sage Publications, London
Borraz F, Dubra J, Ferrés D, Zipitría L (2013) Supermarket entry and the survival of small stores. Rev Ind Organ 44:73–93. doi:10.1007/s11151-013-9379-7
Diggle PJ (2003) Statistical analysis of spatial point patterns, 2nd edn. Edward Arnold, London
Diggle PJ, Ribeiro PJ Jr (2007) Model-based geostatistics. Springer, New York
Durlauf SN (1989) Locally interacting systems, coordination failure and the long run behaviour of aggregate activity, Working paper Dept. of Economics, Stanford
Durlauf SN (1999) How can statistical mechanics contribute to the study of social science? P Natl A Sci 96:10582–10584
Espa G, Giuliani D, Arbia G (2013) Conditional versus unconditional industrial agglomeration: disentangling spatial dependence and spatial heterogeneity in the analysis of ICT firms' distribution in Milan. J Geogr Syst 15(1):31–50
Eurostat-OECD (2007) EUROSTAT - OECD Manual on Business Demography Statistics, OECD, ISBN 978-92-64-04187-5
Hopenhayn H (1992) Entry, exit and firm dynamics in long run equilibrium. Econometrica 60:1127–1150
Igami M (2011) Does big drive out small? entry, exit, and differentiation in the supermarket industry. Rev Ind Organ 38:1–21
Krueger NF Jr (2003) The cognitive psychology of entrepreneurship. In: Acs Z, Audrestsch DB (eds) Handbook of entrepreneurial research. Kluwer Law International, London, pp 105–140
Krueger NF Jr (2003) The cognitive psychology of entrepreneurship. In: Acs Z, Audrestsch DB (eds) Handbook of entrepreneurial research. Kluwer Law International, London, pp 105–140
Lazear EP (2005) Entrepreneurship. J Labor Econ 23(4):649–680

Rathbun SL, Cressie N (1994) A space time survival point process for a longleaf pine forest in Southern Georgia. J Am Stat Assoc 89:1164–1174
Reynolds P, Storey DJ, Westhead P (1994) Cross national comparison of the variation in new firm formation rates. Reg Stud 28:443–456
Silverman BW (1986) Density estimation for statistics and data analysis. Chapman & Hall, London
Storey DJ (1984) Editorial. Reg Stud 18(3):187–188
Storey DJ (1985) Introduction. In: Storey DJ (ed) Small firms and regional economic development: Britain, Ireland and the United States. Cambridge University Press, London, pp 1–5
Strauss D, Ikeda M (1990) Pseudolikelihood estimation for social networks. J Am Stat Assoc 85(409): 204–212
Zoltan JA, Storey DJ (2004) Introduction: entrepreneurship and economic development. Reg Stud 38(8):871–877