The analysis of firm demography: an approach based on micro-geographic data

Diego Giuliani[1], Simonetta Cozzi[2] and Giuseppe Espa[3]

Abstract A method is proposed to assess the dynamics of firm demography in the space by using micro-geographic data. The spatial distribution of economic activity is represented by a marked spatial point pattern of firms and their sizes (as measured in terms of number of employees). In such a framework, marked spatial point patterns can be considered as the result of a dynamic economic process that takes place over time and space. Following Rathbun and Cressie (1994) and Arbia (2001), these patterns can then be modelled as realizations of marked space-time survival point processes, where firms are born at some random location and point in time and then operate, grow, and attract or repulse the localization of other entrant firms. To illustrate the method, an attempt to model the spatial birth process of entrant small retail food stores in the city of Trento, Italy, during the period 2005-2007 is pursued.

1 The data

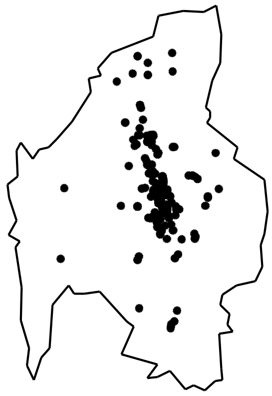
The data are from the Statistical Register of Active Enterprises (ASIA), managed and updated by the Italian Statistical Institute (ISTAT). It consists on the geographical locations and number of employees of small retail food stores and big supermarkets/hypermarkets, also selling food products, operating in the city of Trento area, Italy, during the period 2004-2009. See Figure 1.

Figure 1: Map of small retail food stores and big supermarkets, Trento, 2004-09

1

2

3



2 A spatial birth process

The observed spatial point pattern of births of entrant small food stores is assumed to be the realization of a process conditional on the observed locations of existing small food stores and big supermarkets. On the basis of this assumption, we estimate a model to verify the consistency of the economic stylized fact which postulates that big supermarkets inhibit the birth of small food stores in the nearby locations while, by the effect of positive spatial externalities, the presence of small food stores attracts the location of other small food stores.

Following the modelling framework originally proposed by Rathbun and Cressie (1994) and resumed by Arbia (2001), the spatial birth phenomenon of the new entrant firms during period 2005-2007 is modeled by the means of an inhomogeneous Poisson Cox process (see Diggle, 2003) with random intensity function parametrically specified as

$$\lambda(x) = \exp\{\alpha + \beta_{ss} n_{ss}(x) + \beta_{bs} n_{bs}(x) + \beta_{af} n_{af}(x)\},$$

where $\beta_{ss}, \beta_{bs}, \beta_{af}$ are parameters to be estimated. The intensity function $\lambda(x)$ expresses the expected number of random points located inside an infinitesimal region centered at location *x*. Then, by definition, higher is $\lambda(x)$ higher is, on average, the concentration of firms located around location *x*. Variables $n_{ss}(x)$ and $n_{bs}(x)$ measure, respectively, the overall number of employees of the small food stores and big supermarkets, existing from before 2005, located around the arbitrary point *x*. Variable $n_{af}(x)$ represents a proxy for the unobserved spatial heterogeneity, which is needed in order to control for the effects of unidentified sources of territorial variability. By assuming that the unobserved exogenous factors affecting the location choices of the economic agents (such as environmental limits, the presence of useful infrastructures or the proximity to communication routes) are

substantially the same for all economic sectors, variable $n_{af}(x)$ is constructed as the overall number of employees of all firms of all industries, operating from before 2005, located around the arbitrary point *x*.
The model is based on the working assumption that the economic agents operating before 2005 are fixed at their observed locations. As noted by Arbia (2001), this hypothesis is consistent with Krugman's idea of "historical initial conditions" (Krugman, 1991).

The logarithmic transformation allows to estimate the model parameters by maximizing the log-pseudolikelihood (see Besag, 1975) for $\lambda(x)$ based on the observed locations *x*. In order to maximize the log-pseudolikelihood and then obtain asymptotically unbiased estimates of the β_{ss}, β_{bs}, β_{af} 's parameters we rely on the technique proposed by Berman and Turner (1992). 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 for the significance of the model parameters by using standard formal likelihood ratio criteria based on the χ^2 distribution.

2.1 Results

The maximum pseudolikelihood estimates of the parameters are shown in Table 1.

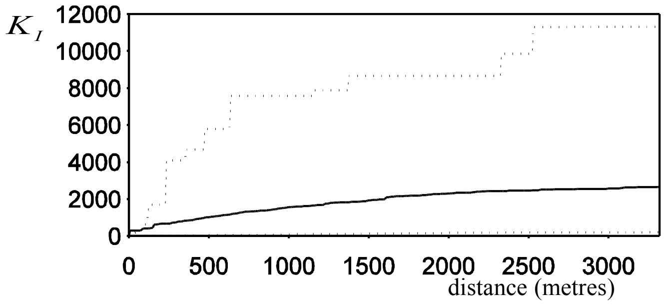
Table 1: Estimates for the Poisson Cox spatial birth process of new small food stores

Parameter	*Estimate*	*Standard Error*	*Z-test*
α	–18.816	0.8614	
β_{ss}	0.075	0.0233	**
β_{bs}	–0.026	0.0073	***
β_{af}	0.001	0.0003	***

The positive value of the estimate of β_{ss} implies that births of small food stores are positively dependent on locations and sizes of the existing small food stores, while the 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 near presence of other existing small food stores, which uncovers the working 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 uncovers the working of negative spatial externalities.
In order to test for goodness of fit of the estimated model we rely on Monte Carlo simulation. The adequacy of model to the observed data is visually represented (Figure 2) by the behavior of the empirical inhomogeneous *K*-function (Baddeley *et al.*, 2000) of the observed points against the behavior of confidence bands derived from 999 simulations of the estimated inhomogeneous Poisson Cox model. The graph

reported in Figure 2 shows that the empirical function lies almost entirely within the confidence bands.

Figure 2: Please use the *caption* style here



References

1. Arbia, G.: Modelling the geography of economic activities on a continuous space. Papers Reg. Sci. 80, 411–424 (2001)
2. Baddeley, A., Møller, J., Waagepetersen, R.: Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. Statistica Nederlandica 54(3), 329–350 (2000)
3. Berman, M., Turner, T.R.: Approximating point process likelihoods with GLIM. Applied Statistics 41(1), 31–38 (1992)
4. Besag, J.: Statistical analysis of non-lattice data. The Statistician 24(3), 179–195 (1975)
5. Diggle, P.J.: Statistical analysis of spatial point patterns, 2nd edition. Edward Arnold, London (2003)
6. Krugman, P: Geography and trade. MIT Press, Cambridge, Massachusetts (1991)
7. Rathbun, S.L., Cressie, N.: A space time survival point process for a longleaf pine forest in Southern Georgia. Journal of the American Statistical Association 89, 1164–1174 (1994)
8. Strauss, D., Ikeda, M. Pseudolikelihood estimation for social networks. Journal of the American Statistical Association 85(409), 204–212 (1990)