SPATIAL AGGLOMERATION, INNOVATION AND FIRM SURVIVAL FOR ITALIAN MANUFACTURING FIRMS

Arnab Bhattacharjee
Economics, Heriot-Watt University and National Institute of Economic and Social Research, UK

Ornella Maietta
Deceased, formerly CSEF and University of Naples Federico II, Italy

Fernanda Mazzotta
Department of Economics and Statistics and CELPE, University of Salerno, Italy

NIESR Discussion Paper No. 546
23 February 2023

About the National Institute of Economic and Social Research

The National Institute of Economic and Social Research is Britain's longest established independent research institute, founded in 1938. The vision of our founders was to carry out research to improve understanding of the economic and social forces that affect people's lives, and the ways in which policy can bring about change. Over eighty years later, this remains central to NIESR's ethos. We continue to apply our expertise in both quantitative and qualitative methods and our understanding of economic and social issues to current debates and to influence policy. The Institute is independent of all party-political interests.

National Institute of Economic and Social Research
2 Dean Trench St
London SW1P 3HE
T: +44 (0)20 7222 7665
E: enquiries@niesr.ac.uk
www.niesr.ac.uk
Registered charity no. 306083

This paper was first published in February 2023

Competing Interest Statement: The views expressed in this paper are based on research and are not attributed to the organizations to which the researchers are affiliated. There are no conflicts of interest. The usual disclaimer applies.

Spatial Agglomeration, Innovation and Firm Survival for Italian Manufacturing Firms‡

Arnab Bhattacharjee[1], Ornella Maietta[2] and Fernanda Mazzotta[3]

Abstract

Innovativeness of a firm improves not only its own survival chances but can also generate externalities on its neighboring firms. We empirically examine the role of agglomeration economies in how innovativeness affects firm survival in Southern Italy, using spatial weights to model spillovers. Spatial Durbin probit model estimates confirm that innovation is a determinant of firm survival not only for firms that are themselves innovative but also ones located close to other innovative firms. Definition of spatial scale and weight plays an important role. Spillover benefits are enhanced by agglomeration economies, but only at a very local scale.

Classification: L20, O3, D22, C21, C41

Keywords: Firm survival, Spatial models, Innovation, Spillovers, Southern Italian SMEs

‡ The paper is dedicated to the fond and loving memory of Ornella Maietta, who died while the first draft of the paper was being prepared. The paper has benefited from comments by participants at the 60th Annual Conference of the Italian Economic Association (Società Italiana di Economia) in 2019. Together, the authors are grateful to Franco Giordano and Gianluigi Coppola for help with understanding the survey design and Taps Maiti for discussions on subsampling and superpopulation models in survey sampling, as well as Roberto Basile, Sean Holly and Cristina Tealdi for many helpful comments and suggestions. The usual disclaimer applies.

[1] Economics, Heriot-Watt University and National Institute of Economic and Social Research, UK. [2] Deceased, formerly CSEF and University of Naples Federico II, Italy. [3] Corresponding author, Department of Economics and Statistics and CELPE, University of Salerno, Italy. Email: mazzotta@unisa.it.

Introduction

We study, empirically, the relationship between innovation and firm survival. While only a few studies explore the relationship between innovation, output and survival of firms (Cefis & Marsili, 2006; Cauchie & Vaillant, 2016), there is general consensus that knowledge is a critical input and a primary source of value generating cumulative effects (Grant, 1996). Innovation depends on a firm's absorptive capacity to understand and transform flows of external knowledge (Farace & Mazzotta, 2015). Geographical proximity is an important channel to spread knowledge, enabling firms to exchange tacit information. Moreover, through repeated interactions, spatial agglomeration can also generate new knowledge spillovers and learning-by-interacting among co-located firms. Many studies debate the impact of agglomeration economies on measures of productivity (labor productivity and TFP), innovativeness, real wages and employment growth (Beaudry & Schiffauerova, 2009; Ferragina & Nunziante, 2018). Importantly, Destefanis (2001) emphasizes the relevance of institutional and technological environment in shaping the lower total factor productivity of Southern Italian firms.

Our contribution is to empirically examine the role of agglomeration economies and spatial spillovers in innovativeness determining firm survival in Southern Italy. We confirm that survival chances increase if firms are innovative. However, spatial agglomeration effects are nuanced and depend both on the specific spatial context and also on spatial scale or granularity. If geographically proximate firms survive, or if closer firms are innovative, this may improve survival chances because of the local business environment and because of external knowledge. Forces of competition can mitigate, and even overturn, such spillover effects, but this can also depend on how close the neighborhood under consideration is. Moreover, a firm's survival chances can increase with higher labor productivity resulting from educational level and training (internal knowledge), as well as proximity to knowledge (university, Science and Technology Park (PST) and/or research laboratory of the Ministry of Agriculture (MIPAA). But this is true only when the knowledge is relevant to their business and if the supply of knowledge is not heavily constrained (Borowiecki, 2015).

New economic geography (NEG) theoretically underpins current understanding as to why local agglomeration occurs, including knowledge spillovers and innovation externalities, specialization and diversification economies, and reduced transportation costs (Fujita, Krugman, & Venables, 1999). However, discourse about the effects and sources of geographic concentration go back to Marshall (1890), highlighting knowledge spillovers which relate to flows of knowledge between workers located in the same vicinity. A central idea is that local aggregation of firms belonging to analogous industries favor the exchange of knowledge between these firms, which in turn drive greater innovation and higher growth rates; but competition and adverse selection can mitigate against agglomeration effects (Rosenthal & Strange, 2003).Less investigated is the role of agglomeration economies with respect to industrial demography, specifically firm entry and exit. Firm entry has received some attention. The set-up of new establishments or start-ups has been analyzed, for example, for new entrepreneurial activities (Rosenthal and Strange, 2003) and multinational enterprises (Mariotti, Piscitello, & Elia, 2010). However, studies on the exit of the firms from the market are relatively sparse, with only a few exceptions (Piacentino, Aronica, Giuliani, Mazzitelli, & Cracolici, 2021). Rather, the literature on firm exit and survival has focused on the role of firm- and sector-specific factors (Caves, 1998; Disney, Haskel, & Heden, 2003; Mata & Portugal, 1994), including innovative activity (Audretsch & Mahmood, 1995) and macroeconomic conditions (Bhattacharjee, Higson, Holly, & Kattuman, 2009), while geographic and spillover aspects have been less studied.

Most studies of the effect of agglomeration on firm exit have been conducted at the region, province or city level. However, firm-level studies are useful in segregating the effect of location characteristics and location choice effects from agglomeration economies. Ferragina and Mazzotta (2015) analyzed the exit of firms and agglomeration economies using Italian firm-level data, using a multilevel approach that allows explicit clustering into homogeneous geographical areas to estimate the spatial variability of both exit and its determinants. However, conventional multilevel models are based on the strong assumption that location random effects are independent. This assumption is untenable in firm dynamics because spatial concentration and clustering would imply correlations between random effects located in close geographic proximity, which then leads to inefficiency and even biased estimates (Bivand & Piras, 2015); we explicitly modelling these effects using spatial econometric models. Based on these estimated models, we ask the following research question: “To what extent do innovativeness and proximity to innovative firms, and thereby agglomeration, enhance survival of SMEs?”

To answer this question, first we need to choose between a menu of several alternate spatial models – Spatial Autoregressive (SAR), Spatial Error model (SEM) and Spatial Durbin model (SDM) – as to which of these might be more appropriate in analyzing the spatial effect of innovation on the probability of survival. We make this model choice based on the likelihood principle, that is choosing the model with the highest likelihood (or posterior score), potentially accounting for model complexity (Aikaike Indormation Criterion, AIC), as the model best supported by the data. In our case, this turns out to be the SDM (LeSage & Pace, 2009; Elhorst, 2010), where both the spatially lagged dependent variable and the spatial lag of a single key independent variable (innovation) are included in the specification. This emphasizes agglomerative clustering of survival chances and learning from neighbors, together with knowledge spillovers.

We use data from the Permanent Observatory on Firms in Salerno province (OPIS) survey. Surveyed firms are small and medium-sized enterprises (SMEs) based in Salerno, a large province in southern Italy with more than one million inhabitants. In 2001, there were 20 local labor systems (LLSs) in the province, out of a total of 49 in the Campania region. Salerno had two industrial districts within its territory, one in the food sector and the other in the chemical sector, while firms in other sectors (particularly ceramics, tourism, and metals) were distributed throughout the province without significant concentration in specific districts. Salerno is representative of a coastal southern Italian province, generally characterized by a higher presence of SMEs. Our micro-spatial analysis may shed light on the determinants of the survival of firms located in sub-regions within similar contextual and cultural factors affecting entrepreneurship (Soo, 2018).

Beyond the context of Salerno, the analysis of survival and innovation in the traditional SME sector is particularly important in Italy, where 95% of firms are the so-called Made-in-Italy firms[1] and 94% have fewer than 10 employees.[2] Hence, our findings provide important implications for industrial and regional policy aimed at enhancing innovativeness at the local level and contributing to improving the absorptive capability and ultimately survival of Italian SMEs. This is a unique contribution in the literature assessing the innovation-survival relationship at the firm level controlling for knowledge spillovers.

The rest of the paper is organized as follows. In Section 2, we review the literature on firm survival and highlight the stylized facts uncovered in previous studies. In Section 3, we present our firm-level dataset, econometric model and methodology, followed in Section 4 by our empirical findings and their discussion. Finally, Section 5 collects conclusions. An online Appendix reports descriptive statistics and details of the survey.

2. Innovation, Proximity and Firm Survival: Theoretical and Empirical Literature

2.1. Innovation and Survival

Survival is a key measure of firm performance; see, for example, Audretsch and Mahmood (1995), Caves (1998), Disney et al. (2003) and Bhattacharjee et al. (2009). Firms that are able to successfully innovate establish competitive advantage in the market which promotes survival (Wagner, 1990).

---

[1] ICE, the Italian Agency for international trade, which promotes internationalisation of Italian firms.
[2] One of the highest proportions in the whole of the European Union (European Commission, SBA Fact Sheet Italy, https://www.hbaproject.org/wp-content/uploads/Italy-2016-SBA-Fact-Sheet.pdf).

Such successful firms may ultimately exit much later either through acquisition or voluntary liquidation.

However, innovation is risky and can have either positive or negative effect on a firm's survival prospects. Radical innovations entail fundamental uncertainty and may increase the probability of firm exit, particularly in highly uncertain environments subject to institutional or policy changes (Hyytinen, Pajarinen, & Rouvinen, 2015). A positive role of innovations, variously defined, on survival is confirmed by some studies (Cefis & Marsili, 2006; Cauchie & Vaillant, 2016), while others have more ambiguous findings (Wagner, 1990). Audretsch (1995) underlines that innovative industries have higher neo-natal exit rates, but for firms surviving beyond the first few years, survival is higher in innovative industries. This leads us to our first hypothesis:

Hypothesis 1. Firm survival chances increase if firms are innovative.

2.2. Agglomeration Economies

Innovation is costly, and this cost may be too high for some firms. Hence they may prefer to imitate rather than introduce their own effective innovations. Thus, diffusion of knowledge and the "geography of innovation" becomes relevant, particularly localized spillovers from R&D spending (Audretsch & Feldman, 1996). Then, the private technology of each firm can become public knowledge and spill over to neighboring firms increasing their productivity. Rosenthal and Strange (2003) consider the relevance of matching, input sharing, and knowledge spillovers for manufacturing firms at various levels of geographic disaggregation. Other studies have found that knowledge spillovers tend to decay rapidly with distance (Audretsch & Feldman, 1996). Geographic concentration generates dynamic processes of knowledge transfer (diffusion and synergies), knowledge creation, learning and innovation. As a result, the cluster becomes a center of accumulated competence through a range of related industries and business interactions.

Related literature on agglomeration economies is extensive, seminal and influential (Marshall, 1890; Jacobs, 1969; Audretsch & Feldman, 1996; Porter, 1998; Fujita et al., 1999; Rosenthal & Strange, 2003). It highlights positive effects of technology transfers and competitive forces, leading to increased competition, reallocation of resources towards more productive firms, and productivity improvements of incumbent firms. Two main types of externalities are identified: diversification economies and localization (or specialization) economies. Diversification economies (Jacobs, 1969) highlights that local knowledge spillovers across different industries promote innovation and growth (Beaudry & Schiffauerova, 2009). This reflects external economies passed on to enterprises through the large-scale operation of agglomeration economies, independent of the industry structure. By contrast, localization economies arise from industry specialization available to local firms within the same sector (Marshall-Arrow-Romer or MAR externalities) and through intra-industry transmission of knowledge as firms learn from other firms in the same industry (Porter, 1998). This explains the development of industrial districts (ID). Thus, relatively densely populated areas are more likely to house universities, research laboratories and other knowledge generating facilities. The post-war Italian economic development literature emphasized the so-called "district effects" quantifying Marshallian advantages, as opposed to the "urban effects" arising from Jacobs (1969) type externalities. Empirical evidence is mixed; Di Giacinto et al. (2014) found stable productivity advantages for firms located in urban areas but only weak advantages traditionally associated with Italian industrial districts.

The theory of agglomeration economies also argues that positive knowledge spillovers are more likely to occur if firms are located in the same area, as geographical proximity promotes the diffusion of ideas and technology due to the concentration of consumers and suppliers, worker mobility, and informal contacts (Greenstone, Hornbeck, & Moretti, 2010). Technology transfers (intra and inter industry knowledge spillovers) may arise from horizontal linkages (imitation, collaboration among firms, concentration of customers and supplier workers mobility, as well as informal contacts) and via vertical linkages (along the supply chain). This suggests policy initiatives to strengthen collaborative ties among key innovation system actors. For Italian firms, Basile and Pittiglio (2017) consider spatial externalities measured at the local labor system (LLS) level through localization/specialization, diversification and population density/urbanization. They found that related variety, which contributes to the generation and diffusion of new knowledge, has a positive effect on firm survival in manufacturing sectors, while unrelated variety, which may work as a portfolio strategy, plays a positive role in services sectors. Localization economies positively influence firm survival only in services sectors. Recently Piacentino et al. (2021) studied new accommodation firms in Sicily to explore whether agglomeration economies play an important role in survival. Following the literature on agglomeration economies (Arbia, Espa, & Giuliani, 2015) they used the localization, urbanization and relatedness index and also control for the distance from the coast. They find that new firms in the accommodation industry suffer considerably from a 'congestion effect', as measured by the spatial concentration of firms in the same industry. They found a sort of barrier to entry in the presence of a spatial concentration of firms specialized in the accommodation industry.

Besides MAR externalities, diversity and concentration are alternate sources of knowledge spillovers. Indeed, inter-industry rather than intra-industry knowledge spillovers (Jacobs, 1969) constitute an important mechanism for economic growth. Then, sectoral diversity triggers innovation and a more diverse economy promotes Jacobs' externalities. Porter (1998) suggests intra-industry concentration and competition provide the best incentives to exchange knowledge and innovate. Empirical research is divided as to whether it is concentration, or diversity, or competition which matters most for growth and innovation (Beaudry & Schiffauerova, 2009). Overall, geographical proximity facilitates interactions, knowledge exchange and face-to-face contacts, thus leading to higher productivity and innovation capacity (Audretsch & Feldman, 1996; Jaffe, 1989). However, proximity by itself would not generate cooperation or knowledge spillovers. For effective spillovers, we need cognitive proximity, that is, similarity in how firms and innovators understand and interpret information, so that firms and scientists can communicate and understand each other effectively by sharing a common vocabulary and framework (Boschma, 2005). Marshallian externalities are also important, both for manufacturing industries and increasingly services and research itself (Fujita et al., 1999).

There is persuasive evidence that proximity to a research institution and the size of its research enterprise affect the probabilities of technology start-ups and innovation (Woodward, Figueiredo, & Guimarães, 2006). However, empirical support for the link between geographic proximity to knowledge centers and firm survival is uncertain. De Silva and McComb (2012) argue that, if local barriers to entry are low, positive spillovers will attract entry up to the point where private benefits from the shared, unpriced inputs are competed away. Thus, one would expect to see higher local levels of entry but exit rates that are similar to those of the broader industry. An alternative interpretation is that when localized technology spillovers are present and start-ups are thus facilitated in the same product space or in a very close substitute, there will be more intense competitive pressure for local firms to rapidly commercialize the given R&D and be the first to market. This may result in less investment in proof-of-concept and thoughtful market strategy which bears higher risk (Kor, 2006). Overall, the presence of more and closer substitutes would effectively increase the exit risk of the local firms that share the unpriced input. Hence our third hypothesis is:

Hypothesis 2. Firms' survival chances can increase or decrease if geographically proximate firms survive and if closer firms are innovative, depending on the relative effects of external knowledge and congestion (competition). Effectively, spillover benefits from other neighboring innovative firms are moderated by the negative effects of competition between neighboring firms.

Competitive advantage stems from knowledge creation, accumulation, and application. A substantial part of this knowledge resides in academia and higher education institutions or non-university public research organizations, such as the Italian Parco Scientifico e Tecnologico (Science and Technology Parks, PSTs). Firms may either be unaware of the economic value of this knowledge or unwilling to use it because they wish to protect their established product portfolio. Also, higher education and research institutes may either have no incentive to commercialize their knowledge or permitted to do so by their status as non-profit organizations. Then, it is crucial to connect firms and research institutions. Hence, regional science highlights knowledge spillovers as sensitive to geographic distance and concentrated in spatial proximity to the respective source (Acs, Audretsch, & Feldman, 1992; Boschma, 2005). Thus, innovative firms may choose to locate close to academic institutions. Among others, Jaffe (1989), Acs et al. (1992), and Fischer and Varga (2003) confirm that contributions of university research and technical societies indirectly enhance the effectiveness of firm R&D. Universities provide geographically specific access to resources such as libraries, faculty, and a ready pool of graduates at all levels. Research laboratories and institutions conduct basic research, creating and diffusing knowledge. This new knowledge spills over most readily into the locality and should result in localized private sector innovation. Moreover, universities increasingly facilitate faculty start-ups and aim to enhance access to university resources to support regional entrepreneurs.

On the other hand, inefficient firms that manage to enter strategic locations or partnerships can create adverse selection particularly if there are supply constraints (Rosenthal & Strange, 2003; Borowiecki, 2015; Piacentino et al., 2021). From these considerations, our third hypothesis is:

Hypothesis 3. The relationship between firm survival chances and proximity to a strategic university department, or a Science and Technology Park (PST) or research laboratory of the Ministry of Agriculture (MIPAA) can be mixed. True, these facilities tend to be spatially concentrated and beneficial for innovativeness of SMEs. But there is also high location sorting and severe competition particularly for start-up businesses which can adversely affect survival.

Innovation arises from a firm's ability to manage and learn from both external and internal sources of knowledge but few studies analyze the link with tacit knowledge (embedded in employees or organization) and how this can influence firm's innovation and survival; notable exceptions are Chen, Jiao & Zhao (2016) on Chinese biotechnology industry and Hyytinen et al. (2015) on Finnish start ups, while Ortiz-Villajos and Satoca (2018) study the value of prior experience of the founder for UK firms. Furthermore, Lokshin, Belderbos, & Carree (2007) examined the impact of internal and external R&D on labor productivity in a panel of Dutch manufacturing firms and found complementarity between internal and external R&D, with a positive effect of external R&D evident only in case of sufficient internal R&D. Hence, our fourth and final hypothesis is:

Hypothesis 4. Firms' survival chances increase with higher worker productivity resulting from educational level and training (internal knowledge).

2.4. Spatial Econometric Studies

The above analyses have prompted further recent studies on agglomeration economies for Italian firms adopting spatial methodologies both at regional and at firm level. For example, Antonelli, Patrucco, and Quatraro (2011) applied spatial econometric methods to model innovation spillovers at the regional level. This follows the approach of Anselin, Varga, and Acs (1997) in applying spatial econometric techniques to innovation models; see also Autant-Bernard, LeSage & Parent (2007). Lamieri and Sangalli (2013) estimated the impact of patents on total factor productivity (TFP) of Italian manufacturing firms using a spatial autoregressive (SAR) model. Also adopting a SAR specification, Cardamone (2017) show that firm productivity is affected by the productivity of nearby firms and that the indirect effect of innovation is stronger than the direct effect. Productivity spillovers at industry level also matter (Carboni, 2013), with evidence that in their R&D decisions firms benefit from spillovers originating from closely related industries.

As discussed, several contributions provided evidence on the positive role of R&D activities at the firm level; see, for example Aiello and Cardamone (2008). However, in order to adequately evaluate the effect of R&D on productivity, productivity spillovers should also be taken into account. Such productivity spillovers could arise because of face-to-face contacts, labor mobility and R&D cooperation between firms (Baltagi, Feng, & Kao, 2012). Both Baltagi et al. (2012) and Lamieri and Sangalli (2013) show that productivity spillovers matter. Moreover, a positive non-linear relationship between R&D investment or product innovation and the probability of firm survival was found by Fontana and Nesta (2009). Ferragina and Mazzotta (2015) take the multidimensional spatial structure into account analyzing how local economies differently affect firms with different levels of global activities. They find that domestic firms not involved in FDI do not benefit sufficiently from the social capital that spills over from industrial districts, while foreign multinationals have higher survival rates. Ferragina and Nunziante (2018) use spatial autoregressive and spatial error models to study Italian manufacturing firms, finding strong spatial productivity spillovers.

However, specific analyses of the link between innovation, agglomeration and survival of firms is lacking, which makes our study unique. Cardamone (2017) used a SAR model to study the role of R&D on firm productivity, but not survival, finding that R&D significantly affects Italian firm productivity and productivity spillovers across firms matter. Moreover, productivity is found to be positively affected by intrasectoral R&D spillovers, while intersectoral R&D spillovers do not have a significant effect. At the regional level, a number of studies have employed spatial econometric tools to take productivity spillovers into account when evaluating the effect of innovative efforts; see, for example, LeSage and Fischer (2008) and Antonelli et al. (2011).

3. Data, Models and Methodology

In this section, we discuss our data and context, empirical models and econometric methodology. All elements are carefully tuned to the central purpose of our empirical work: to verify the effects of innovative behavior and spillovers upon firm survival. Importantly, we focus not only on the direct impacts of own-firm innovative behavior but on indirect effects from neighboring firms.

3.1. Data and Context

The data used in this study are derived from the OPIS database, which contains the results of a survey of 462 manufacturing firms from the province of Salerno, which is a NUTS3 area located in the Campania region in Italy (where NUTS stands for Nomenclatura delle Unità Territoriali Statistiche). The survey sample is statistically representative of Salerno's economic system at the territorial and sectoral levels (Coppola, Farace, Giordano, & Mazzotta, 1999). The survey was conducted by face-to-face interviews in 1998 and 1999 and it provides useful firm-level information, such as innovation, the number of employees, their education level, their training, and their involvement in firm management, as well as the firm's legal form, the industry sector, the source of start-up capital (whether it is the entrepreneur's own capital or comes from family finance, banks, or subsidies), and product markets (local, national, or international). Moreover, there was a follow-up survey for the period from 1999 to 2013 by checking the register of all active firms in the local area (Camera di Commercio) to find out whether each of the original surveyed enterprises was still operating or had ceased to exist. Thus, for each of the firms in the final sample (456 firms for which we have information on the key variables used in our model) we have verified whether or not they survived until 2013 and for those which did not survive we collate from the register the exact date they either ceased activity or went bankrupt; see Appendix for further details of the survey.

For each innovative firm, the survey records their three main innovation-specific partners.[3] The most common partnerships are: suppliers of equipment and plants for product and process innovations; and consultants/commercial labs for organizational innovation. The questionnaire also asks about technological knowledge from the University of Salerno,[4] and the most important public research institution in the province that the firm was interested in for its future innovation strategies. Over the period 2004-2010, the departments of the University of Salerno most actively involved in

[3] Question: "Who were the principal partners with whom you implemented the innovation (max. three)?"
[4] Question: "According to your innovation strategy for the future, what field of knowledge available in the University of Salerno do you consider of particular interest?". Responses: chemistry, engineering, computer science, business, etc.

collaborations with industry were chemistry, computer science, and engineering (ANVUR, 2013).[5] Two municipalities in Salerno host research laboratories of the Ministry of Agriculture (MIPAAF) and one hosts a science and technology park (PST);[6] knowledge spillovers from these centers are captured by a dummy variable equal to one if the municipality where the firm is located hosts the PST together with another dummy variable equal to one if the municipality hosts one of the MIPAAF labs. Appendix Table A1 reports descriptive statistics. Survival of innovative firms is significantly higher than non-innovative firms (Appendix Table A2).

Simple random sampling was not used, but rather survey sampling methods. Each sampled firm has a sampling weight which indicates how many "similar” firms in the population it represents. This has implications for our empirical analysis, to which we return later.

3.2. Econometric Models

For analyzing the survival of firms taking into account agglomeration effects, we model probability of exit allowing for spatial or network dependence. In line with the literature (LeSage & Pace, 2009; Elhorst, 2010), spatial spillover effects are modeled using the standard device in spatial econometrics and statistics – the local spatial average or spatial lags (Anselin, 1988; Anselin et al., 1997; LeSage & Pace, 2009). Construction of spatial lags requires specifying a $n \times n$ (spatial) weights matrix, whose empirical specification in our estimation is discussed later. For each index firm, spatial weights determine which other firms are considered its neighbors. Then, we model sample data on exits for firms at specific locations in space using spatial binary-choice regression models to specify probabilities of a binary outcome (in our case, non-exit or being active in 2013).

Beyond the effect of own explanatory variables (X), the spatial econometrics literature distinguishes between mainly three different types of interaction effects. For any index spatial unit (in our case, a firm): (a) endogenous interaction effects arise from a change in the dependent variable (y^*) in the neighborhood (that is, among neighboring firms); (b) exogenous interaction effects capture the impact of independent variables (Z) for neighboring firms; and (c) interaction effects among the error terms capture spillovers of shocks. A linear spatial regression model (General Nesting Spatial model - GNS) with all these three types of interaction effects takes the following form (LeSage and Pace, 2009):

---

[5] There were more patent activities in the chemistry department (11 patents out a total of 21 for the University of Salerno) and more contract research in the engineering and computer science departments, whereas spin-off creation was equally frequent in the chemistry and engineering departments (two out a total of six for the University of Salerno).
[6] The Istituto Sperimentale per l’Orticoltura (Experimental Institute for horticulture) is located in Pontecagnano Faiano, the Istituto Sperimentale per il Tabacco (Experimental Institute for the cultivation and transformation of tobacco) in Scafati and the Science and Technology Park in Salerno.

$$y^* = \rho W y + X\beta + WZ\gamma + \mu$$
[1]

$$\mu = \lambda W\mu + \varepsilon$$

$$\varepsilon \sim N(0, \sigma^2 I_n)$$

where y^* represents an $n \times 1$ vector of the dependent variable across the n observations, X an $n \times k$ matrix of independent variables (including the intercept), Z an $n \times l$ matrix of independent variables through which there are spatial Durbin effects (Z can be the same as X, or a subset, or can even include additional independent variables), I_n is the $n \times n$ identity matrix, β is the $k \times 1$ vector of regression coefficients, and γ is the $l \times 1$ vector of spatial Durbin coefficients. Further, ρ and λ are structural scalar spatial spillover parameters, each restricted to the range [−1, 1]; ρ, λ and γ are estimated together with α and β.

Here, ρ is called the spatial autoregressive (or lag) parameter and captures endogenous spatial dependence created by spatial spillovers through the effect of changes in neighboring firms' dependent variable. By contrast, γ is the spatial Durbin vector that leads to spatial dependence through the effect of k determinants for neighboring firms through proximity, while λ is the spatial error (autoregressive) parameter which models spillovers of unmodeled shocks across the firms. The reduced form of model [1] is:

$$y^* = (I - \rho W)^{-1} X\beta + (I - \rho W)^{-1} WZ\gamma + (I - \rho W)^{-1}(I - \lambda W)^{-1}\varepsilon.$$
[2]

The spatial weights matrix W contains information on the spatial relationships between observations. Wy^* denote the spatial lag of the dependent variable, which is in effect a weighted combination of dependent variables for neighboring firms. WZ is an exogenous interaction effect among one or more independent variables (in our application, innovation) capturing the idea that a neighboring firm's innovativeness can have a potential impact on survival of the index firm. By contrast, $W\mu$ is the interaction effect among the disturbance terms for different firms (Elhorst, 2010; p. 11-12). Typically, one needs to make a choice between which of the spatial effects is appropriate in a given context and accordingly some of the spatial parameters (ρ, γ and λ) are set to zero.[7]

In our case, the dependent variable y is binary, an exit dummy (1/0). Then, the spatial lag Wy is not interpretable as the dependent variable for the neighboring firms because it is no longer binary. One

[7] For example, the spatial autoregressive (SAR) model has $\rho \neq 0$ but sets $\gamma = \lambda = 0$. By contrast, the SLX model has $\gamma \neq 0$ but $\rho = \lambda = 0$ and the spatial error model with $\lambda \neq 0, \rho = \gamma = 0$. One can have combination of spatial effects as well. For example, the spatial Durbin model sets $\lambda = 0$, but allows ρ and γ to have non-zero effects.

needs to use a binary regression model like probit, applied to the dummy variable $y = \mathrm{I}\ (y^* > 0)$, where $\mathrm{I}(.)$ is the indicator function. Then, the spatial lag can be applied to the underlying continuous variable y^*. Following Elhorst, Heijnen, Samarina, & Jacobs (2017), this leads into a spatial probit model in the same form as [1-2], except that y is observed not the latent variable y^*; for identification purposes, $\varepsilon \sim N(0, I_n)$.

3.3. Variables

Our key covariates (X) for the firm survival models are: an innovator dummy (1/0); dummy variables for technological knowledge relevant to future innovation strategies accessible to the firms through proximity to the University of Salerno and technology parks; and the entrepreneur's human capital as captured by the owner's level of education, the owner's age as a proxy for experience, and dummy variables for the owner's previous job (employee, student or unemployed, self-employed individual, or an entrepreneur in another firm); and finally employee characteristics (training and involvement in management).

Other firm and market characteristics, such as firm age and size, single-product innovation, scale economies and capital intensity are also potential determinants of failure for new businesses (Caves, 1998). The relationship with age may not be linear (Mata & Portugal, 1994) – the probability of exit is initially low, increases to a certain point and decreases afterwards (Bhattacharjee et al., 2009). Exit risk can also be higher in mature firms because of structural inertia constraining the ability to respond to changes in business environment or market conditions. Even in stable environments, the accumulation of rules and routines can over time decrease firm efficiency and decrease survival chances. Furthermore, young firms with higher exit risk can benefit more from innovation to survive into the long term (Cefis & Marsili, 2006).

The probability of survival increases with firm size. Large firms are more likely to have output levels close to their industry minimum efficient scale, and thus are less likely to be vulnerable (Audretsch & Mahmood, 1995). Also, large firms are usually more diversified; this reduces their risk of exit, since adverse conditions in one market can be offset by better conditions in others. Further, in the firm and industry dynamics literature, firm size and age represent the efficiency differences arising from different experience, managerial abilities, production technology and firm organization. Large firms may also find it easier to raise capital, face better tax conditions and be in a better position to recruit qualified workers and more skilled and talented managers. On the other hand, consistent with theories of industry evolution (Audretsch & Mahmood, 1995) and strategic niches (Porter, 1998), firms may remain small because they occupy product niches that are not easily accessible or profitable for large firms. Most empirical studies find that size increases the likelihood of survival in the most technologically advanced industries, but not necessarily in traditional sectors (Caves, 1998).

Together, export intensity, industry structure and technology as well as aggregate macroeconomic conditions and instability can affect firm exits (Caves, 1998; Bhattacharjee et al., 2009).

As control variables for firm survival, we included the above principal factors identified in the literature: firm age, size, and start-up capital. We also controlled for market size (local, national or international); type of products; whether the firm was founded by the previous generation of the entrepreneur; location characteristics (population density in the municipality, industrial district presence, and MIPAAF or technological scientific laboratory presence); and sectors.

3.4. Sampling and spatial weights

Since our data are drawn from the population using a weighted sampling design, standard spatial models are not directly applicable (Benedetti, Suesse, & Piersimoni, 2020).[8] Unfortunately, standard methods, though very generally defined, usually do not take into account the fact that the sample data are drawn from a fixed population. Data obtained from units selected by complex sample designs need to be accounted for. This is because the sampled data need not be representative of the population. Selection criteria need to be explicitly included in the specification of models for survey data and variances of parameter estimates need to be computed in a manner that reflects the impact of the sample design. While this is true of any modeling exercise and is well known in the survey sampling literature (Chambers, Steel, Wang, & Welsh, 2012), the issue is compounded in spatial models and particularly in finite populations because neighboring firms which affect the outcome for an index firm, may not be included in the sample.

Three main approaches have been proposed in the literature to address sample selection in spatial survey data (Benedetti et al., 2020). The first, and somewhat simplistic, approach is to use sampling weights as additional regressors to account for selection probabilities. This acts as a proxy for Mill's ratio for sample selection. As discussed by Benedetti et al. (2020), this is unlikely to be adequate, particularly when the model is nonlinear or limited dependent, as in our case. The second approach, further developed and advanced by Benedetti et al. (2020), is the marginal likelihood approach (Chambers et al., 2012). Unfortunately, this approach is applicable only to maximum likelihood (or likelihood-based) methods. Therefore, we abstract from this approach in our attempt to keep methods more broad-based. The third approach relates to having data on all population units (Benedetti et al., 2020; Besag & Kooperberg, 1995; LeSage & Pace, 2009). This approach is simple to implement, and we apply this to our setting by making a small methodological innovation. We feel

[8] The issue of selection bias in spatial modelling has attracted only limited attention, and specifically for spatial probit models there is no literature.

that this approach is more general and applicable to wider contexts, with more general spatial sampling designs beyond simple random sampling, and to methods beyond maximum likelihood.

In standard non-spatial contexts, considering all population units is a standard approach. This is equivalent to using frequency weighting based on (survey) sampling weights. However, frequency weighting does not work with spatial data, as articulately explained in Benedetti et al. (2020), because this would not adequately account for (spatial) dependence between the units. To address this issue, we expand our sample data to the entire population by making as many copies of each sample data point as this individual data point represents in the population. This way, we construct (pseudo) data for the entire population by using all available information including the survey sampling weights. In our case, the sample data includes $n = 456$ firms, but expanded to the entire population, our inferences are based on $N = 7{,}248$ firms. To emphasize, unlike the first approach, we do not use sampling weights or sample inclusion probabilities as additional regressors to control for selection bias. We use sampling weights to expand our sample to the entire population, and then use the entire population data so generated to estimate our models.

Having compiled our population data, we next need to construct a suitable W matrix of reciprocal influences between firms. Our central measure of W is based on a geographic distance matrix which is a quadratic $N \times N$ matrix with zero diagonal elements. Here, N is the number of firms in our expanded sample to the population of 7,248 firms. The sample of $n = 456$ firms is expanded to the population level using corresponding sector- and municipality- level sampling weights, to a total of $N = 7{,}248$ firms. The generic elements w_{ij} are referred to as "spatial weights", measuring the strength of the relationship between a firm *i* and each neighboring firm *j*. Choice of spatial weights is critical in a study of agglomeration effects, particularly the spatial scale. The focus here is on agglomeration economies where the fine balance between spillover benefits and competition depends on spatial granularity and neighborhood. This is the primary reason why we use distance-band weights that clearly specify the extent of neighborhood for each firm. Further, distance band weights extend easily to our approach of data expansion, without the need to make any assumptions about locations for the non-sampled firms. However, estimates of spatial econometric models can also be sensitive to the choice of spatial weights matrix. To verify robustness with respect to spatial scale, we also estimate our models using several alternate distance thresholds. In addition, we also use inverse-distance weights based on the reciprocal of each distance: $v_{ij} = 1/d_{ij}$.

Given the above approach to expand sample data to the entire population, and the desire to consider spatial scale effects, distance band spatial weights are a natural choice in our case. Hence, we use distance band weights matrices with 10 km and 50 km distances, with binary elements $v_{ij} = 1$ if the distance between firms i and j is below the threshold, and zero otherwise. Finally W is obtained by row-standardizing the spatial weight matrices, that is, $w_{ij} = \frac{v_{ij}}{\sum_j v_{ij}}$. Then our *W* matrix is composed of distance-band weights w_{ij} organized as $W_{N\times N} = \left((w_{ij})\right)_{i,j=1,\cdots,N}$. In order to exclude self-neighbors, the diagonal elements $\mathrm{w_{ij}}$ are set equal to zero. Summary statistics of all the variables in our estimation are reported in Appendix Table A1.

3.5. Estimation

Elhorst et al. (2017) include elaborate discussion of alternate estimation methodology, including their relative merits and demerits. We recognize that there are two main approaches – likelihood-based and moment-based. Within likelihood-based methods, the leading method is maximum likelihood (ML). ML relies strongly on underlying model and distributional assumptions, but perhaps these assumptions are not avoidable. Quasi-ML (QML) and Bayesian methods relax some of these assumptions in specific ways. QML only uses information on selected moments which in turn is related somewhat to moment-based methods. Bayesian methods incorporate prior information; strong (informative) prior information can mitigate against potentially strong model assumptions but the priors themselves are not testable. Elhorst et al. (2017) favor ML, however acknowledging that Bayesian methods are more popular, arguing this may be because of computation intensity and availability of MATLAB programs. We remain relatively agnostic about estimation methods and associated philosophy but choose a likelihood-based Bayesian methodology as our leading method to exploit high-dimensional sparse matrix computation methods and open-source software available with the "*spatialprobit*" R package of Wilhelm and de Matos (2013).[9]

Beyond likelihood-based methods, we are also cognizant of errors in variables problems that potentially arise as data on our outcome (exit) and covariates are expanded to the population level. A similar problem also emerge with geographic distance weights if location data have errors (Arbia et al., 2015). Anselin and Lozano-Gracia (2008) address these problems by a 2SLS approach using latitude, longitude and nonlinear transformations as instruments. Likewise, we use a GMM methodology based on extended set of instruments and implemented using the *spprobit* routine in the "McSpatial" R package (McMillen, 2013); see also McMillen (1992) and Klier and McMillen (2008).

Our findings are robust to these alternate modeling approaches. Indeed, a major reason to consider entire population data in our case, rather than a marginal likelihood approach, was to ensure that any

[9] We also attempted to estimate our models by ML, but faced prohibitively intensive computations given our large population (sample) size of 7,248 firms and a relatively flat log-likelihood surface.

methodology applicable to population data is adequate in our context. This allows us to be agnostic about alternate methods and verify robustness of our empirical findings.

3.6. Model Selection

The three spillover parameters (ρ, γ and λ) in our models [1-2] have different interpretations. We consider spatial Durbin effects only on the innovation dummy variable. Then, the spatial Durbin parameter γ on innovation reflects the (indirect) influence of a neighboring firm's innovativeness upon the index firm's survival, whereas a corresponding coefficient in β would capture direct impacts of own innovation upon survival. A clear distinction between these direct and indirect effects is necessary for adequate understanding of spillovers and agglomeration economies. By contrast, the spatial lag parameter, ρ, has the structural interpretation of capturing the effect of a neighboring firm's exit (or otherwise, survival) on the index firm's exit; this structural mechanism generates endogenous spatial dependence and renders inference particularly challenging. Finally, the spatial error autoregressive parameter, λ, captures the dependence in the error process, whereby shocks affecting a neighboring firm (for example, demand and supply shocks) can have a spillover effect on the index firm. Each spillover parameter can in principle generate positive or negative externalities. If the agglomeration effects dominate competition effects, there are positive externalities, and negative if otherwise. In this way, a clear focus on structural interpretations allows us to understand separately the different channels through which spillover effects influence survival outcomes in a firm and can be modelled using spatial regression models.

Unfortunately, modelling all three spatial spillover parameters together using finite real data leads to serious identification and implementation issues. One view, expressed strongly in LeSage and Pace (2009)is that the General Nesting Spatial model (GNS) as the most general spatial regression model since it includes all types of interaction effects. But one major problem is that the parameters of this GNS model are only weakly identified, which then favors the spatial Durbin model as a general starting point for discussion of spatial regression model estimation. One can then consider either the SAR (or spatial lag) model which only includes the spatially lagged dependent variable, or the SLX which only considers spatial spillovers through X, as an alternative, but retain the SDM as the richer and preferred structural model for spillovers.

Here we take another view, in the rich tradition of specification testing (Anselin, 1988), to use the likelihood principle and choose the model that is best supported by the data. Then, following Elhorst (2010), we compare the log-likelihoods for the candidate models (SEM, SAR and SDM), but in a Bayesian setting. Thus, we estimate all the models using flat priors, so that the posterior log-scores can be interpreted approximately as the log-likelihood of the data. Secondly, again based on log-scores, we implement an approximate LR test to compare the three models and accept the SDM as better supported by the data. The same conclusion arises if we base model selection upon the Akaike Information Criteria (AIC); in this case, the SDM model is chosen as the one with the lowest AIC. The fact that the data supports the spatial Durbin model (with a single spatial Durbin parameter relating to innovativeness of the neighboring firms) highlights spillover agglomeration effects of firm behavior (specifically innovation) together with endogenous local clustering of firm performance (or, exit chances).

Based on the above model selection approach, we follow Elhorst et al. (2017) in computing spatial direct and indirect effects for our central model – a spatial Durbin model. Then, for every covariate k other than the spatial Durbin covariate innovation, we have:

$$\begin{pmatrix} \frac{\partial E(y)}{\partial x_{1k}} & \cdots & \frac{\partial E(y)}{\partial x_{Nk}} \end{pmatrix}_{N\times N} = diag\big(\phi(\hat{\eta})\big)[(I-\rho W)^{-1} I_N \beta_k], \quad [3]$$

and for the spatial Durbin covariate z (innovation dummy):

$$\begin{pmatrix} \frac{\partial E(y)}{\partial z} & \cdots & \frac{\partial E(y)}{\partial z} \end{pmatrix}_{N\times N} = diag\big(\phi(\hat{\eta})\big)[(I-\rho W)^{-1}(I_N \beta_z + W\gamma)], \quad [4]$$

where $\hat{\eta}$ denotes the vector of predicted values of y, and $\phi(.)$ is the standard normal pdf. Note that the above matrices are of order $N \times N$ where in our case N is population (expanded sample) size. Following LeSage and Pace (2009) and Elhorst et al. (2017), we define the *direct effect* as the average diagonal element of the full matrix expression on the right-hand side of equations [3-4], the *indirect effect* as the average row or column sums of the off-diagonal elements of the matrix expression, and the *total effect* as the sum of the two.

We compute direct and indirect effects, as well as Bayesian MCMC estimates, using the R package “spatialprobit” (Wilhelm & de Matos, 2013). Corresponding GMM estimates are obtained using the “McSpatial” R package (McMillen, 2013).[10] This renders our computations replicable using the open source R software.

[10] The distances were computed using the R package spdep, a collection of functions to create spatial weights matrix (Bivand & Piras, 2015) in particular we used the function “dnearneigh”, which identifies neighbors of region points by Great Circle distance in kilometers between lower (zero) and upper band (5, 10, 20 and 50 km). For all the tests (LR, Geary, Moran and join count test for spatial association), we use routines in R.

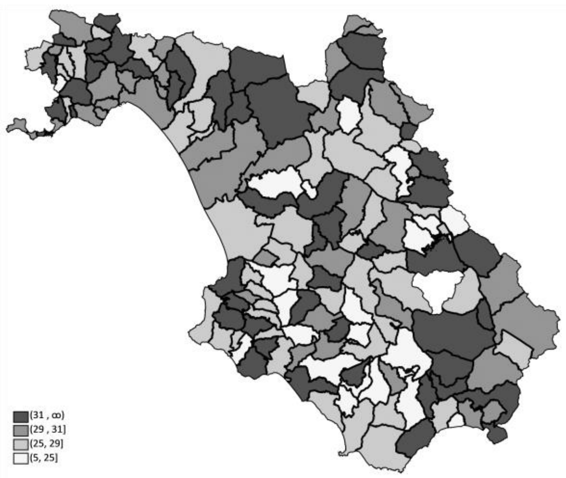
4. Results

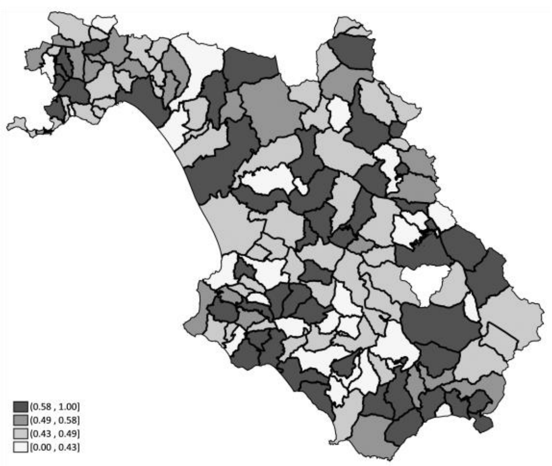
Now, we turn to reporting our empirical findings. We start with preliminary exploratory analyses of the data, followed by empirical estimates of spatial probit models of firm survival together with estimation of spatial direct, indirect and total marginal effects.

Figure 1 shows maps of Salerno province (and its constituent municipalities) together with incidence, by municipality, of three key firm-level variables: longevity of surviving firms, innovativeness and survival (over the period 1990 to 2013). Patterns of concentration are visible, as well as co-incidence across the variables. However, it is also clear from spatial patterns that the relationship between innovativeness and survival is complex. Firms in the municipality of Salerno (on the western coast towards the north) have higher innovativeness but lower survival, potentially because of supply constraints and competition effects mitigating against agglomeration benefits, or potentially adverse selection into central locations (Rosenthal & Strange, 2003). However, in the municipalities of Eboli (south of Salerno), Acerno (east of Salerno) and Sanza (far south and interior), high innovativeness is collocated with better survival. This underlines the need for adequate empirical analysis at a more granular spatial scale.

Figure 1: Maps of Salerno showing concentration of firms by key variables

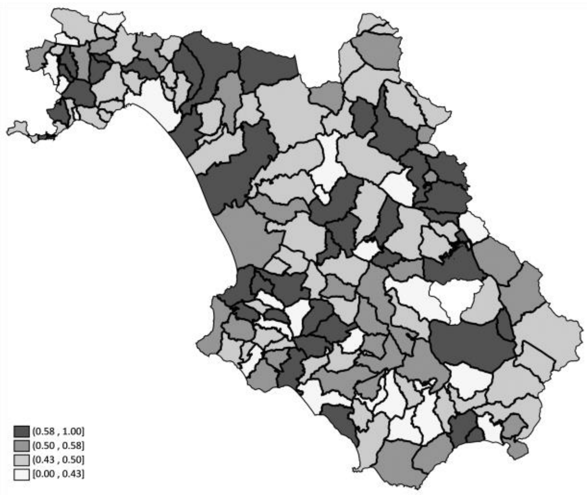
(a) Longevity of survived firms (years) (b) Innovative firms (frequency)





(c) Surviving firms, 1990 to 2013 (frequency) (d) *Comunes* (Municipalities) of Salerno

[Inset: Salerno within Italy]



Hence, we verify spatial concentration using suitable measures of spatial dependence. For continuous variables (firms' duration and owner's age) and ordinal variable (employee involvement in firm management), we use Moran's I and Geary's c measures of spatial autocorrelation based on a 10 km distance band spatial weights matrix (Table 1). Only sampled data units ($n = 456$) are included in this estimation because the expanded sample would artificially inflate positive autocorrelation. For binary variables, choice of autocorrelation measure is critical. Join count statistics for spatial association are more popular for limited dependent variables, but they are not appropriate for spatial point process data as in our case. One can convert the spatial firm locations into regions by Delaunay triangulations and apply contiguity weights, or otherwise use distance band spatial weights. For selected binary variables, we compute both join count statistics and Moran's I using 10 km distance band weights (Table 1). While neither is ideal in our setting, the purpose of this exercise is largely descriptive, and this approach is closely aligned with our estimation strategy.

We cannot reject the null hypothesis of zero spatial autocorrelation for firm longevity (duration) and exit, but there is good degree of spatial concentration in many of our covariates. In particular, the Moran's I statistic reflects statistically significant spatial autocorrelation in our innovation dummy variable and owner's university education. Together, employee training and involvement in firm management is highly concentrated. As expected, spatial concentration is prominent in several sectors, importantly food and related, textiles/leather, non-metallic minerals (ceramics and bricks), and wood/metal products. The complex spatial patterns particularly for exit and innovation highlight specific challenges that our estimated models need to unpick.

Table 1. Moran's I and Geary's c Global Test for spatial autocorrelation (selected variables)

Continuous variables	**Moran's I (p-value)**	**Geary's c (p-value)**
Duration (in years)	-0.1646 (0.565)	-0.1453 (0.558)
Owner's age	-0.4421 (0.671)	-0.6474 (0.741)
Employee involvement in firm mgmt. (0=no, …, 3=high)	**3.7160 (1.0E-04)**	**3.3888 (4.0E-04)**
Categorical variables	**Moran's I (p-value)**	**Join count statistic (p-value)**
Exit dummy		
(exit = 0; alive)	-0.0438 (0.517)	-0.3929 (0.653)
(exit = 1; liquidated)	0.0529 (0.479)	0.4744 (0.318)
Innovation dummy (Innovator = 1)	**2.0602 (0.020)**	1.1989 (0.115)
Owner education – (University degree = 1)	**1.8913 (0.029)**	-0.6241 (0.734)
Employee involvement in firm mgmt. (3 = high)	**2.5758 (0.005)**	**2.6787 (0.004)**
Employee training dummy (Training = 1)	**3.5922 (1.6E-04)**	**3.0258 (0.001)**
Local knowledge dummy (econ./ag.econ./business)	**1.9894 (0.023)**	**1.9621 (0.025)**
Firm size dummy (20 ≤ workers < 50)	0.1057 (0.458)	0.0067 (0.497)
Science/Tech. park dummy (PST in municipality = 1)	**32.40 (< 2.2E-16)**	**31.03 (< 2.2E-16)**
Research lab dummy (MIPAAF lab in municipality = 1)	**8.2919 (< 2.2E-16)**	**7.4157 (6.0E-14)**
Sector dummies		
(Food, drinks, tobacco)	**3.4517 (2.8E-04)**	**3.2682 (5.4E-04)**
(Textiles/leather)	**2.6585 (0.004)**	**2.6835 (0.004)**
(Wood/metal products)	**2.0975 (0.018)**	**1.9086 (0.028)**
(Paper, printing/publishing)	0.4597 (0.323)	0.3956 (0.346)
(Chemicals and Rubber)	0.9710 (0.166)	0.7722 (0.220)
(Non-metallic minerals)	**2.6001 (0.005)**	**2.6419 (0.004)**
(Mechanical prod.)	1.2366 (0.108)	1.1709 (0.121)

Tests are based on $n = 456$ rather than the expanded sample $N = 7{,}248$. This is because sample expansion induces spatial autocorrelation because we have no variation in location.

Spatial autocorrelation in innovativeness reflects that it can promote agglomeration effects in survival in the spatial Durbin model, subject to countervailing competition and congestion effects. Similarly, concentration in owner's education, previous experience and employee training suggests potential agglomeration effects through diffusion of information among neighboring firms. Such effects may be contingent upon business competence and access to scarce resources (such as finance and knowledge). Spatial concentration by sector also suggests externalities of the MAR or Jacobs types.

Against this backdrop, we put our baseline hypotheses and insights from descriptive analysis to test by formal estimation of spatial regression models. We show three probit models: a non-spatial probit, a probit SAR model and a probit SDM (with only a single variable with spatial Durbin effects – innovativeness). Together, we also estimate average direct, indirect and total marginal effects. These effects are computed as discussed before, averaged over the entire population of $N = 7{,}248$ firm observations (expanded from a sample of $n = 456$ firms) providing a summary measure of the impact arising from changes in the i-th observation of each key explanatory variable. For example, if an index firm makes an innovation, the estimated direct effect is the average impact on the exit probabilities on the same firm, averaged across the entire population. This takes into account feedback loops that run through the inter-firm network as modeled by our spatial weights matrix W. Similar computations and interpretation apply to total and indirect effects as well. The average indirect effect quantifies the impact, upon an average firm, of all other firms making an innovation.

Table 2a reports the estimates of the nonspatial probit and probit SDM models, with 10 km and 50 km distance band spatial weights matrices. Also reported are log-likelihoods (to be precise, posterior log-scores) and Akaike Information Criterion (AIC) for alternate models SAR and SEM.[11] Clearly, among all alternate candidate models considered, the spatial Durbin model (SDM) dominates, and the SDM model with 10 km distance band weights is the best. The, Table 2b reports average direct, indirect and total marginal effects for the two SDM models, for 10 km and 50 km distance bands.

[11] The Akaike Information Criterion is defined as $AIC = 2k - 2L$, where k denotes the number of parameters in the model and L the log likelihood. It balances model complexity against model fit and is also related to the likelihood ratio test for nested models.

Table 2a Non-spatial probit, Probit and Probit SDM spatial models of firm exit (Distance Band Weights with Threshold Distance at 50km and 10 km)

Variables	**Non-spatial Probit**		**Probit Spatial Durbin (PSDM) (50km)**		**Probit Spatial Durbin (PSDM) (10km)**	
	Coefficient	z-value	Coefficient	z-value	Coefficient	z-value
Spatial Durbin lag – Innovator dummy			-3.080	**-3.96**	-0.075	-0.56
Innovation dummy (Innovator = 1)	-0.255	**-7.20**	-0.255	**-7.16**	-0.226	**-6.24**
Firm size – Omitted category (workers < 10)						
(10 ≤ workers < 20)	-0.113	**-2.22**	-0.118	**-2.35**	-0.123	**-2.48**
(20 ≤ workers < 50)	-0.076	-1.30	-0.085	-1.52	-0.058	-1.03
(workers ≥ 50)	0.640	**7.88**	0.646	**7.86**	0.565	**7.00**
Dummy – firm founded by current owner	-0.104	**-2.81**	-0.099	**-2.70**	-0.120	**-3.32**
Owner education – Omitted (Univ degree)						
(Low education)	0.081	**2.13**	0.070	1.89	0.062	1.66
(Higher secondary)	-0.366	**-6.47**	-0.374	**-6.84**	-0.334	**-6.13**
Owner's age	-0.020	**-4.53**	-0.020	**-4.63**	-0.020	**-4.67**
Owner's age - squared	3.1E-04	**5.59**	2.97E-04	**5.65**	3.1E-04	**5.94**
Owner's previous job – Omitted (Unemployed)						
(salaried employee)	-0.022	-0.52	-0.015	-0.35	-0.022	-0.50
(self-employed)	0.052	0.69	0.050	0.70	0.055	0.78
(entrepreneur)	0.073	1.68	0.063	1.38	0.086	1.92
(homemaker)	0.083	0.56	0.082	0.51	0.039	0.24
Vertical chain – Omitted (Final goods or both)						
(intermediate goods)	-0.092	-1.92	-0.102	**-2.12**	-0.051	-1.06
Market – Omitted (International market)						
(Local market)	0.444	**4.89**	0.443	**4.81**	0.518	**5.68**
(National market)	0.955	**8.74**	0.937	**8.34**	0.955	**8.59**
Employee involvement in firm mgmt. (0=no, …, 3=high)	0.096	**6.53**	0.094	**6.26**	0.094	**6.39**
Employee training dummy (Training = 1)	-0.636	**-16.61**	-0.629	**-16.00**	-0.586	**-15.01**
Science/Tech. park dummy (PST in municipality = 1)	0.042	0.99	0.017	0.38	0.056	1.31
Research lab dummy (MIPAAF lab in municipality = 1)	1.021	**11.96**	1.056	**12.28**	0.830	**10.23**
Bank financing dummy (bank financing = 1)	0.283	**3.74**	0.266	**3.55**	0.223	**2.98**
Local knowledge dummy – Omitted (None/others)						
(econ./ag.econ./business)	-0.062	-1.20	-0.067	-1.30	3.0E-04	-0.01
(chem./comp.sc./engg.)	0.127	**2.08**	0.130	**2.05**	0.121	1.92
Sector dummies – Omitted (Mechanical prod.)						
(Food, drinks, tobacco)	-0.222	**-3.54**	-0.226	**-3.61**	-0.219	**-3.57**
(Textiles/leather)	0.282	**4.71**	0.296	**4.79**	0.231	**3.72**
(Wood/metal products)	0.255	**4.10**	0.258	**4.21**	0.242	**3.99**
(Paper, printing/publishing)	0.157	**2.50**	0.153	**2.39**	0.153	**2.41**
(Chemicals and Rubber)	0.181	1.89	0.202	**2.26**	0.156	1.75
(Non-metallic minerals)	0.199	**3.31**	0.201	**3.20**	0.263	**4.23**
Intercept	-0.091	-0.60	1.571	**3.49**	-0.150	-0.89
Spatial lag			-0.385	-1.18	0.743	**17.13**
Log likelihood (Posterior log-score)	-4560.561		-4553.852		-4605.271	
AIC (Akaike Information Criterion)	9181.12		9171.70		9274.54	
Alternate models - Log likelihood, AIC						
- Spatial Autoregressive Model			-4556.1852	9174.37	-4608.252	9278.51
- Spatial Error Model			-4589.4874	9242.97	-4611.855	9287.71

A key finding is that the estimated spatial autoregressive coefficient ρ, which reflects the strength of endogenous spatial dependence, is positive, large (0.74) and statistically significant at the 10 km distance band but not statistically significant at the 50 km threshold. This reflects agglomeration economies at a very local spatial scale but stronger spatial competition effects at the wider scale. If a neighboring firm exits, it reflects higher propensity for immediately neighboring firms to also exit, but has marginal effect on firms in the wider neighborhood as they are perhaps able to capture a greater market. This provides an intuitive but nuanced understanding of the nature of agglomeration economies in the survival of firms. This also highlights the importance of spatial scale in empirical analyses of firm dynamics – an issue that has not been discussed much in the literature.

Our central, and a priori expected, finding is that innovation substantially improves survival chances. An innovative firm's exit probability is estimated to be significantly lower across all the three model specifications. However, innovation in neighboring firms also plays an important role; here, too, interpretation is very nuanced and depends on the spatial scale of analysis. This agglomeration effect is ignored in the non-spatial model, and the SAR probit model that only allows spillovers in survival chances. Our estimated SDM at the 50 km threshold reflects very strong positive spillover externalities of innovativeness, but the same is not true at the finer scale 10 km threshold. Innovative activity of neighboring firms in a wider neighborhood of 50 km is beneficial through external knowledge spillovers, but any spillover gains from innovativeness of immediately neighboring firms are somewhat dissipated through competition and congestion effects. Nevertheless, at the 10 km distance band, agglomeration effects ensure that both the direct effects and indirect effects are strong and statistically significant. However, at the wider neighborhood of 50 km, indirect effects are relatively small. The prominence of competition and congestion effects reflect strong adverse selection and supply constraints in resources (Rosenthal & Strange, 2003; Borowiecki, 2015).

Similar negative indirect effects are also observed for proximity to research lab (MIPAAF) and access to financing. MIPAAF is specialized in agriculture, but the regional high-tech specialization in Salerno is mainly in the ceramics and chemicals sectors. Hence, this evidence could reflect reallocation of resources towards the food sector to detrimental effect on survival. On the other hand, training of employees has positive effects and generates positive externalities, potentially because of mobility of skilled labor. However, employee training has the expected positive significant effect on survival.

Table 2b SDM probit spatial autoregressive models: Total, Direct and Indirect Marginal effects on the probability of exit (selected variables), z-values in parentheses Distance Band Weights with Threshold Distance at 50km and 10 km

Variables	Non-spatial probit		Spatial probit Durbin model: marginal effects & z-value (50km)						Spatial probit Durbin model: marginal effects & z-value (10km)					
			Total effects		Direct Effects		Indirect effects		Total effects		Direct effects		Indirect effects	
	marginal effect	z-value	marginal effect	z-value	marginal effect	z-value	marginal effect	z-value	marginal effect	z-value	marginal effect	z-value	marginal effect	z-value
Innovation dummy (Innovator = 1)	-0.101	**-7.36**	-0.070	-1.93	-0.092	**-4.36**	0.021	0.64	-0.315	**-4.68**	-0.079	**-6.21**	-0.235	**-3.81**
Firm size – Omitted category (workers < 10)														
(10 ≤ workers < 20)	-0.045	**-2.23**	-0.033	-1.18	-0.042	-1.41	0.010	0.54	-0.172	**-2.29**	-0.043	**-2.49**	-0.129	**-2.15**
(20 ≤ workers < 50)	-0.030	-1.30	-0.024	-0.82	-0.031	-0.90	0.007	0.46	-0.082	-1.03	-0.020	-1.04	-0.061	-1.01
(workers ≥ 50)	0.240	**8.36**	0.178	**1.98**	0.232	**4.78**	-0.054	-0.65	0.787	**4.72**	0.198	**7.31**	0.588	**3.84**
Dummy – firm founded by current owner	-0.041	**-2.86**	-0.027	-1.36	-0.036	-1.62	0.008	0.57	-0.168	**-2.90**	-0.042	**-3.30**	-0.125	**-2.70**
Owner education – Omitted (Univ degree)														
(Low education)	0.032	**2.17**	0.019	0.99	0.025	1.12	-0.006	-0.51	0.086	1.60	0.022	1.60	0.064	1.56
(Higher secondary)	-0.144	**-6.37**	-0.103	**-2.03**	-0.134	**-4.09**	0.031	0.63	-0.465	**-4.48**	-0.117	**-6.21**	-0.348	**-3.77**
Owner's age	-0.008	**-4.16**	-0.005	-1.69	-0.007	**-3.09**	0.002	0.60	-0.028	**-3.72**	-0.007	**-5.19**	-0.021	**-3.11**
Owner's age - squared	1.2E-04	**5.16**	8.2E-05	**-3.37**	1.1E-04	**-2.75**	-2.5E-05	-0.88	4.4E-04	**1.99**	1.1E-04	**-4.48**	3.3E-04	**-5.71**
Owner's previous job – Omitted (Unemployed)														
(salaried employee)	-0.009	-0.53	-0.004	-0.21	-0.005	-0.20	0.001	0.14	-0.031	-0.49	-0.008	-0.49	-0.023	-0.49
(self-employed)	0.021	0.67	0.014	0.40	0.018	0.43	-0.004	-0.30	0.076	0.77	0.019	0.81	0.057	0.74
(entrepreneur)	0.029	1.69	0.017	0.75	0.023	0.81	-0.005	-0.46	0.120	1.82	0.030	1.88	0.090	1.76
(homemaker)	0.033	0.60	0.023	0.30	0.030	0.30	-0.007	-0.20	0.054	0.25	0.014	0.24	0.041	0.24
Vertical chain – Omitted (Final goods or both)														
(intermediate goods)	-0.037	-1.90	-0.028	-1.04	-0.037	-1.26	0.008	0.53	-0.071	-1.02	-0.018	-1.05	-0.053	-1.01
Market – Omitted (International market)														
(Local market)	0.173	**5.50**	0.122	1.76	0.159	**2.79**	-0.037	-0.63	0.722	**4.11**	0.182	**5.67**	0.540	**3.55**
(National market)	0.336	**9.47**	0.259	**2.08**	0.336	**5.03**	-0.078	-0.66	1.331	**4.93**	0.335	**8.71**	0.996	**4.06**
Employee involvement in firm mgmt. (0=no, …, 3=high)	0.038	**6.56**	0.026	**1.96**	0.034	**3.75**	-0.008	-0.64	0.130	**4.71**	0.033	**6.27**	0.097	**3.91**
Employee training dummy (Training = 1)	-0.248	**-16.81**	-0.173	**-2.10**	-0.226	**-10.33**	0.052	0.66	-0.816	**-5.93**	-0.206	**-16.10**	-0.610	**-4.46**
Science/Tech. park dummy (PST in municipality = 1)	0.017	0.98	0.005	0.22	0.006	0.23	-0.001	-0.16	0.078	1.26	0.020	1.35	0.058	1.25
Research lab dummy (MIPAAF lab in municipality = 1)	0.351	**11.44**	0.291	**2.13**	0.379	**7.60**	-0.088	-0.65	1.155	**5.57**	0.291	**10.45**	0.864	**4.27**
Bank financing dummy (bank financing = 1)	0.111	3.91	0.073	1.58	0.095	**2.13**	-0.022	-0.61	0.309	**2.80**	0.078	**3.01**	0.230	**2.61**
Local knowledge dummy – Omitted (None/others)														
(econ./ag.econ./business)	-2.5E-02	-1.20	-0.019	-0.73	-0.024	-0.78	0.005	0.40	8.4E-05	0.00	-1.4E-04	-0.01	2.2E-04	0.00
(chem./comp.sc./engg.)	0.050	**2.15**	0.036	1.04	0.047	1.23	-0.011	-0.54	0.170	1.81	0.043	1.92	0.127	1.76
Sector dummies – Omitted (Mechanical prod.)														
(Food, drinks, tobacco)	-0.088	**-3.63**	-0.062	-1.58	-0.081	**-2.17**	0.019	0.59	-0.305	**-3.16**	-0.077	**-3.58**	-0.228	**-2.87**
(Textiles/leather)	0.112	**4.69**	0.082	1.80	0.106	**2.83**	-0.024	-0.62	0.322	**3.13**	0.081	**3.61**	0.241	**2.82**
(Wood/metal products)	0.101	**4.17**	0.071	1.67	0.092	**2.55**	-0.021	-0.60	0.338	**3.34**	0.085	**4.09**	0.253	**3.05**
(Paper, printing/publishing)	0.062	**2.51**	0.042	1.22	0.055	1.42	-0.013	-0.53	0.214	**2.20**	0.054	**2.46**	0.160	**2.05**
(Chemicals and Rubber)	0.072	1.75	0.056	1.12	0.073	1.33	-0.016	-0.53	0.219	1.69	0.055	1.76	0.164	1.60
(Non-metallic minerals)	0.079	**3.33**	0.055	1.40	0.072	1.92	-0.017	-0.58	0.367	**3.40**	0.092	**4.23**	0.275	**3.10**

Our central findings provide good insights into the industrial organization of SMEs in Salerno, a highly innovative Southern Italian coastal province. Three main features emerge. First, the industry is highly competitive and there are negative externalities due to congestion and supply constraints (Rosenthal & Strange, 2003; Borowiecki, 2015). This is reflected not only in the negative effects of research lab and access to finance, but also in the mobility of skilled labor. Second, innovation promotes survival not only of the innovative firms themselves but can also bring spillover benefits. Hence policies to promote innovation are highly desired, together with training opportunities for employees. Third, and however, placement of an agricultural laboratory is detrimental to survival, potentially because it reallocates resources away from sectors of high-value local specialization. This emphasizes that placement of research facilities must not be mindless and need to be underpinned by careful regional industrial planning. Importantly, the findings underscore a nuanced understanding of the nature of competition and agglomeration economies, as well as industrial and education policy. In the context of spatial econometric models, it also highlights the importance of choosing an appropriate spatial scale of analysis.

Robustness of our findings is verified in several ways. First, we estimate a suite of three models and compare the findings. The comparison is underpinned by careful model selection and the choice of spatial Durbin model as our preferred model in itself provides insights into the nature of agglomeration economies at play. Second, while our central results are based on a distance-band binary weights matrix, we also check the validity of our findings against an alternate construction of spatial weights matrix – an inverse distance spatial weights matrix. Third, and finally, we also benchmark our empirical analysis against alternate estimates by GMM (Table 3), which largely provide similar results. We also estimate a SLX model using parametric (Weibull) and semiparametric (Cox) survival models to robustly study the spillover effects of innovation. These results are confirmatory of our main findings and we do not report these in the paper for reasons of brevity.

Table 3. GMM estimates of Probit spatial Durbin models of firm exit (Distance Band Weights with Threshold Distance at 50km and 10 km)

Variables	**Probit Spatial Durbin (PSDM) (50km)**		**Probit Spatial Durbin (PSDM) (10km)**	
	Coefficient	z-value	Coefficient	z-value
Spatial Durbin lag – Innovator dummy	-2.244	**-2.40**	-0.272	**-2.00**
Innovation dummy (Innovator = 1)	-0.245	**-7.05**	-0.218	**-6.23**
Firm size – Omitted category (workers < 10)				
(10 ≤ workers < 20)	-0.104	**-2.04**	-0.088	-1.74
(20 ≤ workers < 50)	-0.079	-1.36	-0.043	-0.75
(workers ≥ 50)	0.625	**8.10**	0.568	**7.61**
Dummy – firm founded by current owner	-0.093	**-2.57**	-0.102	**-2.88**
Owner education – Omitted (Univ degree)				
(Low education)	0.063	1.69	0.089	**2.46**
(Higher secondary)	-0.358	**-6.17**	-0.334	**-5.94**
Owner's age	-0.020	**-4.07**	-0.020	**-4.26**
Owner's age - squared	0.000	**4.89**	0.000	**5.26**
Owner's previous job – Omitted (Unemployed)				
(salaried employee)	-0.017	-0.39	-0.005	-0.12
(self-employed)	0.043	0.56	0.034	0.44
(entrepreneur)	0.073	1.69	0.125	**3.00**
(homemaker)	0.091	0.66	0.095	0.70
Vertical chain – Omitted (Final goods or both)				
(intermediate goods)	-0.100	**-2.06**	-0.102	**-2.11**
Market – Omitted (International market)				
(Local market)	0.419	**5.18**	0.426	**5.37**
(National market)	0.890	**8.82**	0.942	**9.48**
Employee involvement in firm mgmt. (0=no, …, 3=high)	0.093	**6.28**	0.077	**5.61**
Employee training dummy (Training = 1)	-0.615	**-16.13**	-0.581	**-15.00**
Science/Tech. park dummy (PST in municipality = 1)	0.018	0.41	0.070	1.82
Research lab dummy (MIPAAF lab in municipality = 1)	1.037	**11.36**	0.919	**11.18**
Bank financing dummy (bank financing = 1)	0.255	**3.53**	0.290	**4.10**
Local knowledge dummy – Omitted (None/others)				
(econ./ag.econ./business)	-0.072	-1.38	-0.094	-1.84
(chem./comp.sc./engg.)	0.139	**2.35**	0.084	1.42
Sector dummies – Omitted (Mechanical prod.)				
(Food, drinks, tobacco)	-0.201	**-3.25**	-0.182	**-3.01**
(Textiles/leather)	0.294	**4.88**	0.262	**4.35**
(Wood/metal products)	0.255	**4.17**	0.218	**3.65**
(Paper, printing/publishing)	0.156	**2.50**	0.138	**2.21**
(Chemicals and Rubber)	0.196	1.89	0.261	**2.54**
(Non-metallic minerals)	0.207	**3.47**	0.215	**3.65**
Intercept	1.139	**2.16**	0.039	0.24
Spatial lag	-0.277	-0.66	0.474	**6.41**

5. Conclusions

The South of Italy is characterized by low productivity and higher relative labor costs. A possible reason is that while overall in Italy, innovative activities are far from the level reached in the more industrialized countries, in the South innovative activities are even lower. In fact, the South is particularly characterized by small traditional manufacturing sectors such as food, textile, and ceramics with traditional, low skilled and labor-intensive technologies. Then exploring the pattern to improve the innovativeness and also verifying the link between innovativeness and firm's survival can be of great importance for policy makers in designing and implementing appropriate industrial, education and development policies. Previous empirical literature has emphasized the relevance of knowledge flows and provides evidence consistent with the hypothesis of positive knowledge externalities.

In this context, the spatial dimension stressed in some of these studies (Fischer & Varga, 2003) is of primary interest for regional policy makers. Then, in this paper, we take account of spatial autocorrelation and spatial heterogeneity not only within an area but also between firms operating within a spatial context. Specifically, we consider the reciprocal influences between firms based on their geographical distance and apply Spatial Autoregressive model (SAR) and the spatial Durbin models (SDM) to data on firm survival and its determinants, most importantly innovative behavior. We do so by using spatial probit models that are appropriate to this context. There is a single previous study applying the spatial probit model to analyze whether innovation strategies are related to a spatially defined cohort of nearby firms' strategy choices (Autant-Bernard et al., 2007), but no other study considering firm innovation and survival within a geographical context. Both our estimates and model selection approach provide exciting new insights on the nature of competition and agglomeration economies. We calculate the direct and the indirect effect of each variables mediated by an index firm's innovativeness as well as those of the neighboring firms. The central hypothesis is that the innovativeness and survival of the firms is driven by the diffusion of internal and external knowledge (which come from other firms in close geographical proximity and also the presence of nearby university research centers).

The principal finding is that a firm's own innovativeness (internal to the firm) and the external innovativeness of neighboring firms positively affect the survival of firms, if these firms are in the immediate neighborhood. However, this positive effect is counterbalanced by a negative indirect effects of research lab and access to finance, particularly if the index firm is not innovative itself. This competition is manifested mainly through the increase in productivity of workers, as indicated in the higher indirect effect of training for workers of neighboring firms. A policy implication therefore may be that closer firms can benefit from the innovation of each other. If they have an autonomous capacity to absorb the external knowledge stored within the innovative idea, then the innovation externalities come only if the firms has the absorptive capability to understand and transform flows of external knowledge. This observation provides important insights not only on the nature of agglomeration economies but also industrial policy. Also, we find that policy for creation of knowledge is context specific and must be carefully tuned to the needs of local industry; this suggests the importance of careful spatial and regional planning.

Our work includes some methodological innovations particularly in the application of probit spatial regression models and model selection in this context. Future work will extend such methods also to spatial duration models, and in this context, also to the examination of latent inter-firm knowledge spillover networks.

References

Acs, Z. J., Audretsch, D. B. & Feldman, M. P. (1992). Real Effects of Academic Research: Comment. *American Economic Review*, *82*(1), 363–367.

Aiello, F. & Cardamone, P. (2008). R&D spillovers and firms' performance in Italy. *Empirical Economics*, *34*(1), 143–166.

Anselin, L. (1988). *Spatial Econometrics: Methods and Models* (Vol. 4). Dordrecht: Springer Netherlands.

Anselin, L. & Lozano-Gracia, N. (2008). Errors in variables and spatial effects in hedonic house price models of ambient air quality. *Empirical Economics*, *34*(1), 5–34.

Anselin, L., Varga, A. & Acs, Z. (1997). Local Geographic Spillovers between University Research and High Technology Innovations. *Journal of Urban Economics*, *42*(3), 422–448.

Antonelli, C., Patrucco, P. P. & Quatraro, F. (2011). Productivity Growth and Pecuniary Knowledge Externalities: An Empirical Analysis of Agglomeration Economies in European Regions. *Economic Geography*, *87*(1), 23–50.

ANVUR. (2013). Valutazione della Qualità della Ricerca 2004-2010 (VQR 2004-2010) Rapporto finale. ANVUR, Roma.

Arbia, G., Espa, G., & Giuliani, D. (2015). Measurement errors arising when using distances in microeconometric modelling and the individuals' position is geo-masked for confidentiality. *Econometrics* 3(4), 709-718.

Audretsch, D. B. (1995). Innovation, growth and survival. *International Journal of Industrial Organization*, *13*(4), 441–457.

Audretsch, D. B. & Feldman, M. P. (1996). R&D spillovers and the geography of innovation and production. *American Economic Review*, *86*(3).

Audretsch, D. B. & Mahmood, T. (1995). New Firm Survival: New Results Using a Hazard Function. *Review of Economics and Statistics*, *77*(1), 97–103.

Autant-Bernard, C., Lesage, J.P. & Parent, O. (2007). Firm Innovation Strategies: A Spatial Cohort Multinomial Probit Approach. *Annales d'Économie et de Statistique*, (87/88), 63.

Baltagi, B. H., Feng, Q. & Kao, C. (2012). A Lagrange Multiplier test for cross-sectional dependence in a fixed effects panel data model. *Journal of Econometrics*, *170*(1), 164–177.

Basile, R., Pittiglio, R. & Reganati, F. (2017). Do agglomeration externalities affect firm survival?. *Regional Studies*, *51*(4), 548–562.

Beaudry, C. & Schiffauerova, A. (2009). Who's right, Marshall or Jacobs? The localization versus urbanization debate. *Research Policy*, *38*(2), 318–337.

Benedetti, R., Suesse, T. & Piersimoni, F. (2020). Spatial auto-correlation and auto-regressive models estimation from sample survey data. *Biometrical Journal*, *62*(6), 1494–1507.

Besag, J. & Kooperberg, C. (1995). On Conditional and Intrinsic Autoregression. *Biometrika*, *82*(4), 733.

Bhattacharjee, A., Higson, C., Holly, S. & Kattuman, P. (2009). Macroeconomic Instability and Business Exit: Determinants of Failures and Acquisitions of UK Firms. *Economica*, *76*(301), 108–131.

Bivand, R. & Piras, G. (2015). Comparing Implementations of Estimation Methods for Spatial Econometrics. *Journal of Statistical Software*, *63*(18), 1–36.

Borowiecki, K. J. (2015). Agglomeration economies in classical music. *Papers in Regional Science*, *94*(3), 443–468.

Boschma, R. (2005). Proximity and Innovation: A Critical Assessment. *Regional Studies*, *39*(1), 61–74.

Carboni, O. A. (2013). A spatial analysis of R&D: the role of industry proximity. *Economics of Innovation and New Technology*, *22*(8), 820–839.

Cardamone, P. (2017). A Spatial Analysis of the R&D-Productivity Nexus at Firm Level. *Growth and Change*, *48*(3), 313–335.

Cauchie, G. & Vaillant, N. G. (2016). New Firm Survival: Isolating the Role of Founders' Human Capital in Accounting for Firm Longevity. *Journal of Human Capital*, *10*(2), 186–211.

Caves, R. E. (1998). Industrial Organization and New Findings on the Turnover and Mobility of Firms. *Journal of Economic Literature*, *36*(4), 1947–1982.

Cefis, E. & Marsili, O. (2006). Survivor: The role of innovation in firms' survival. *Research Policy*, *35*(5), 626–641.

Chambers, R. L., Steel, D. G., Wang, S. & Welsh, A. (2012). *Maximum Likelihood Estimation for Sample Surveys*. CRC Press.

Chen, J., Jiao, H. & Zhao, X. (2016). A knowledge-based theory of the firm: managing innovation in biotechnology. *Chinese Management Studies*, *10*(1), 41–58.

Coppola, G., Farace, S., Giordano, F. & Mazzotta, F. (1999). Industrial District in the South of Italy. A new databank for the analysis of the Local Labour Market Area LLMA: methods and first results. In G.

Capaldo & M. Raffa (Eds.), *Innovation and Economic Development: the role of Entrepreneurship and Small and Medium Enterprices, Conference proceedings "International Council of Small Business", ESI*. Naples.

Destefanis, S. Pietro. (2001). Differenziali Territoriali di Produttivita' ed Efficienza negli Anni '90: i Livelli e l'Andamento. Discussion Papers No. 59. Celpe -University of Salerno, Salerno.

Di Giacinto, V., Gomellini, M., Micucci, G. & Pagnini, M. (2014). Mapping local productivity advantages in Italy: industrial districts, cities or both?. *Journal of Economic Geography*, *14*(2), 365–394.

Disney, R., Haskel, J. & Heden, Y. (2003). Restructuring and productivity growth in UK manufacturing. *Economic Journal*, *113*(489), 666–694.

Elhorst, J. P. (2010). Applied Spatial Econometrics: Raising the Bar. *Spatial Economic Analysis*, *5*(1), 9–28.

Elhorst, J. P., Heijnen P., Samarina H., & Jacobs J.P.A.M. (2017). State transfers at different moments in time: a spatial probit approach. *Journal of Applied Econometrics* 32(2), 422-439.

Farace, S. & Mazzotta, F. (2015). The effect of human capital and networks on knowledge and innovation in SMEs. *Journal of Innovation Economics*, *16*(1), 39–71.

Ferragina, A. M. & Mazzotta, F. (2015). Agglomeration economies in Italy: impact on heterogeneous firms' exit in a multilevel framework. *Economia e Politica Industriale*, *42*(4).

Ferragina, A. M. & Nunziante, G. (2018). Are Italian firms performances influenced by innovation of domestic and foreign firms nearby in space and sectors?. *Economia e Politica Industriale*, *45*(3), 335–360.

Fischer, M. M. & Varga, A. (2003). Spatial knowledge spillovers and university research: Evidence from Austria. *Annals of Regional Science*, *37*(2), 303–322.

Fontana, R. & Nesta, L. (2009). Product Innovation and Survival in a High-Tech Industry. *Review of Industrial Organization*, *34*(4), 287–306.

Fujita, M., Krugman, P. R. & Venables, A. (1999). *The Spatial Economy: Cities, Regions and International Trade*. MIT Press.

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, *17*(S2), 109–122.

Greenstone, M., Hornbeck, R. & Moretti, E. (2010). Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings. *Journal of Political Economy*, *118*(3), 536–598.

Hyytinen, A., Pajarinen, M. & Rouvinen, P. (2015). Does innovativeness reduce startup survival rates?. *Journal of Business Venturing*, *30*(4), 564–581.

Jacobs, J. (1969). The City: The Economy of the Cities. Random House.

Jaffe, A. B. (1989). Real Effects of Academic Research. *American Economic Review*, *79*(5), 957–970.

Klier, T. & McMillen, D. P. (2008). Clustering of Auto Supplier Plants in the United States. *Journal of Business & Economic Statistics*, *26*(4), 460–471.

Kor, Y. Y. (2006). Direct and interaction effects of top management team and board compositions on R&D investment strategy. *Strategic Management Journal*, *27*(11), 1081–1099.

Lamieri, M. & Sangalli, I. (2013). Produttività e innovazione: un approccio spaziale. *Economia e Politica Industriale*, (2), 67–89.

LeSage, J. P. & Fischer, M. M. (2008). Spatial Growth Regressions: Model Specification, Estimation and Interpretation. *Spatial Economic Analysis*, *3*(3), 275–304.

LeSage, J. P. & Pace, R. K. (2009). *Introduction to Spatial Econometrics*. CRC Press.

Lokshin, B., Belderbos, R. & Carree, M. (2007). The productivity effects of internal and external R&D: Evidence from a dynamic panel data model. Maastricht, NL : UNU-MERIT.

Mariotti, S., Piscitello, L. & Elia, S. (2010). Spatial agglomeration of multinational enterprises: the role of information externalities and knowledge spillovers. *Journal of Economic Geography*, *10*(4), 519–538.

Marshall, A. (1890). *Principles of Economics* (8th ed.). London: Great Minds Series.

Mata, J. & Portugal, P. (1994). Life Duration of New Firms. *Journal of Industrial Economics*, *42*(3), 227.

McMillen, D. (2013). Package '*McSpatial*': Nonparametric spatial data analysis. Version 2.0.

McMillen, D. P. (1992). Probit with Spatial Autocorrelation. *Journal of Regional Science*, *32*(3), 335–348.

Ortiz-Villajos, J. M. & Sotoca, S. (2018). Innovation and business survival: A long-term approach. *Research Policy*, *47*(8), 1418–1436.

Piacentino, D., Aronica, M., Giuliani, D., Mazzitelli, A. & Cracolici, M. F. (2021). The effect of agglomeration economies and geography on the survival of accommodation businesses in Sicily. *Spatial Economic Analysis*, *16*(2), 176–193.

Porter, M. E. (1980). Competitive Strategy: Techniques for Analyzing Industries and Competitors. New York: Free Press.

Rosenthal, S. S. & Strange, W. C. (2003). Geography, Industrial Organization, and Agglomeration. *Review of Economics and Statistics*, *85*(2), 377–393.

Soo, K. T. (2018). Innovation across cities. *Journal of Regional Science*, *58*(2), 295–314.

Wagner, J. (1990). Who exists from German manufacturing industries and why? Evidence from the Hanover firm-panel study. In D. B. Audretsch & A. Thurik (Eds.), *Innovation, Industry Evolution and Employment* (pp. 253-264.). Cambridge, UK: Cambridge University Press.

Wilhelm, S. & de Matos, M.G. (2013). Estimating Spatial Probit Models in R. *The R Journal*, *5*(1), 130-143.

Woodward, D., Figueiredo, O. & Guimarães, P. (2006). Beyond the Silicon Valley: University R&D and high-technology location. *Journal of Urban Economics*, *60*(1), 15–32.

Appendix A

Survey Data Description

Salerno is a large province in the Campania region of southern Italy with more than 1 million inhabitants. In 2001, it had 20 Local Labour Systems (LLS) out of 49 in the Campania region, and since 2011 it has hosted two industrial districts, one each in the food and chemicals sectors.[12] Salerno is also characterized by high SME innovation and industry-university partnerships (Calenda 2017; Morning Future 2018), particularly in the food, ceramics, tourism, metals, and plastic sectors.[13]

The Permanent Observatory on Firms in Salerno province (OPIS) survey collected data on small and medium-sized enterprises (SMEs) based in Salerno in a way that is statistically representative of the population of SMEs in the province. There is extensive information on 462 manufacturing firms – all aspects of business including finance, sector, innovation, human resources, owner endowments, and training, all of which are critical for our work. Particularly relevant to our spatial models, precise geographical location of the firms is recorded, and hence we can account for location characteristics and compute distances and proximity between each pair of firms.

The survey is based on sampling weights, whereby each surveyed firm represents a certain number of SMEs, by municipality and sector. For the purpose of analysis, we replicate (take copies) of each firm according to how many firms it represents in the population, and this provides a working random sample of 1,250 firms. Descriptive statistics, reported in Appendix Table A1, reveal that 48% of firms introduced at least one innovation,[14] whereas 50% survived till April 2013. The type of innovation (process, product and organizational) is also recoded together with the sources of new knowledge. Among innovative firms, 62% survived over the

---

[12] The Italian labor market is divided into LLS., which are sub-regional geographical areas where most of the workers reside, and where firms can find the largest concentration of labor force necessary for jobs. Likewise, production is spatially organized into industrial districts, which are specialized and cover about two-third of national manufacturing employment (Coppola et al. 1999; Staber 2001).

[13] Artisanal ceramics is an ancient traditional and now highly specialised activity in Salerno. There were about 1,400 employees in this sector in 2011 with employment share of 18.5% as against 8.8% for the Campania region as a whole. Together, the food sector had the highest specialization index (2.6) and it was the third largest industry by employee share, occupying about 10,000 units (24.3% of the entire manufacturing sector). The high specialization index indicates the significantly higher importance of the sector in Salerno as compared to Italy, where it employed only 9.2% of the workers. This, together with the weight of exports (about 50% of the total), confirms the importance of the food sector in the province. It has 5.5 employees per local unit, just slightly less than the mechanical/machinery industry for which the relative weight is lower than the national average (Amendola et al. 2013).

[14] Innovation is recorded as response to the question: "Did your enterprise introduce any (product, process or organizational) innovation?" Subsequent questions record when the innovation was introduced, and we only consider innovations within 10 years prior to the survey.

period 1999-2013, while among non-innovative firms, only 40% survived (Table A2); this reflects a very large 22 percentage points difference.

Table A1. Descriptive statistics of the variables (Number of observations/firms = 456 and 7248)

	Sample $n = 456$		Population $N = 7248$	
Variables	Mean	Std. Dev.	Mean	Std. Dev.
Innovation dummy (Innovator = 1)	0.471	0.500	0.528	0.499
Firm size – Omitted category (workers < 10)	0.772	0.420	0.707	0.455
Size Dummy (10 ≤ workers < 20)	0.116	0.320	0.132	0.339
Size Dummy (20 ≤ workers < 50)	0.075	0.264	0.113	0.316
Size Dummy (workers ≥ 50)	0.036	0.187	0.033	0.179
Dummy – firm founded by current owner	0.683	0.466	0.705	0.456
Owner education– Omitted category (Low education=1)	0.422	0.494	0.389	0.487
Owner education dummy (Higher secondary)	0.441	0.497	0.466	0.499
Owner education dummy (univ degree)	0.137	0.344	0.146	0.353
Owner's age (years)	43.440	12.246	43.115	12.787
Owner's age – squared	2036.661	1060.008	2022.389	1049.580
Owner's previous job – Omitted category (Unemployed)	0.272	0.446	0.299	0.458
Previous job dummy (salaried employee)	0.385	0.487	0.361	0.480
Previous job dummy (self-employed)	0.053	0.224	0.054	0.226
Previous job dummy (entrepreneur)	0.279	0.449	0.274	0.446
Previous job dummy (homemaker)	0.011	0.102	0.012	0.107

Vertical chain – Omitted category (Final goods)	0.776	0.417	0.732	0.443
Vertical chain – Omitted category (Both: final and intermediate)	0.117	0.322	0.132	0.339
Vertical chain dummy (Only intermediate goods)	0.107	0.309	0.136	0.343
Market – Omitted category (International market)	0.015	0.124	0.039	0.193
Market dummy (Local market)	0.927	0.260	0.900	0.300
Market dummy (National market)	0.057	0.232	0.061	0.240
Employee involvement in firm mgmt. (0=no, …, 3=high)	1.013	1.138	1.096	1.138
Employee training dummy (Training = 1)	0.310	0.463	0.312	0.463
Science/Tech. park dummy (PST in municipality = 1)	0.136	0.344	0.184	0.387
Research lab dummy (MIPAAF lab in municipality = 1)	0.053	0.224	0.044	0.206
Bank financing dummy (bank financing = 1)	0.042	0.200	0.049	0.216
Local knowledge dummy (econ/ag.econ/business)	0.106	0.308	0.128	0.334
Local knowledge – (chem/comp.sc./engg)	0.088	0.283	0.085	0.279
Sector dummies – Omitted category (Mechanical prod.)	0.226	0.419	0.119	0.324
Sector dummy (Food, drinks, tobacco)	0.226	0.419	0.156	0.363
Sector dummy (Textiles/leather)	0.122	0.328	0.233	0.423
Sector dummy (Wood/metal products)	0.265	0.442	0.144	0.351
Sector dummy (Paper, printing/publishing)	0.055	0.228	0.145	0.353
Sector dummy (Chemicals and Rubber)	0.031	0.172	0.037	0.188

Sector dummy (Non-metallic minerals)	0.075	0.264	0.166	0.372

The survey provides detailed information at the firm level, such as the number of employees, their education level, their training and their involvement in firm management, firm's legal form, industry sector, source of start-up capital (entrepreneur's own or family financing, banks or subsidies) and geographical product market (local,[15] national or international). We classify firm size by the number of workers in 1999:[16] less than 10, 10–19, 20–49 and at least 50. Each firm was assigned to an industry sector based on the Italian Chambers of Commerce two-digit ATECO code. The survey also includes characteristics of the entrepreneur[17] such as age and educational level.

Table A2 Innovative and Surviving Firms from 1999 to 2013 (total no. $\boldsymbol{n} = \mathbf{456}$)

	Per cent
Innovative Firms	48.4
Survived Firms	50.3
Survived Firms | Innovative	61.8
Survived Firms | Not Innovative	39.6

Estimates of alternate models are available, beyond Tables 2a, 2b and 3 in the paper. These include probit SAR, probit SEM and SLX duration models. These are not reported for the sake of brevity but are available with the authors. All computations are implemented using open-source R packages and are fully replicable.

Table A3 Loglikelihood and AIC

	10 km		50 km	
	Loglikelihood	AIC	Loglikelihood	AIC
SDM	-4605.27	9274.542	-4553.85	9171.70

[15] Local market is defined by the province of Salerno, the Campania region or southern Italy.
[16] Some firms have only one worker, which is the owner. For this reason, size is defined by the number of workers less one (the owner).
[17] The manager and the owner are almost always the same person in traditional sector SMEs.

SAR	-4608.25	9278.505	-4556.19	9174.37
SEM	-4611.86	9287.711	-4589.49	9242.97

Additional references

Calenda, C. (2017). Annual report to Parliament on the implementation and impact of legislation in support of innovative startups and SMEs. Italian Minister of Economic Development.

Morning FUTURE (2018). From Apple to Cisco, the tech giants are looking South. https://www.morningfuture.com/en/article/2018/03/19/cisco-apple-technology-south-work-academy-campania-innovation-startup/258/ [Accessed: 25 Janaury 2021].