CONTRIBUTI DI RICERCA CRENOS

A SPATIAL ANALYSIS OF R&D: THE ROLE OF INDUSTRY PROXIMITY

Oliviero A. Carboni

WORKING PAPERS

2012/04

CUEC

CENTRO RICERCHE ECONOMICHE NORD SUD
(CRENOS)
UNIVERSITÀ DI CAGLIARI
UNIVERSITÀ DI SASSARI

CRENOS was set up in 1993 with the purpose of organising the joint research effort of economists from the two Sardinian universities (Cagliari and Sassari) investigating dualism at the international and regional level. CRENoS' primary aim is to improve knowledge on the economic gap between areas and to provide useful information for policy intervention. Particular attention is paid to the role of institutions, technological progress and diffusion of innovation in the process of convergence or divergence between economic areas. To carry out its research, CRENoS collaborates with research centres and universities at both national and international level. The centre is also active in the field of scientific dissemination, organizing conferences and workshops along with other activities such as seminars and summer schools.
CRENoS creates and manages several databases of various socio-economic variables on Italy and Sardinia. At the local level, CRENoS promotes and participates to projects impacting on the most relevant issues in the Sardinian economy, such as tourism, environment, transports and macroeconomic forecasts.

www.crenos.it
info@crenos.it

CRENOS – CAGLIARI
VIA SAN GIORGIO 12, I-09100 CAGLIARI, ITALIA
TEL. +39-070-6756406; FAX +39-070- 6756402

CRENOS - SASSARI
VIA TORRE TONDA 34, I-07100 SASSARI, ITALIA
TEL. +39-079-2017301; FAX +39-079-2017312

Title: A SPATIAL ANALYSIS OF R&D: THE ROLE OF INDUSTRY PROXIMITY

ISBN: 978 88 84 67 727 3

First Edition: February 2012

Via Is Mirrionis, 1
09123 Cagliari
Tel./Fax 070 291201
www.cuec.it

A Spatial Analysis of R&D: the Role of Industry Proximity

Oliviero A. Carboni
University of Sassari and CRENoS

Abstract

This paper employs individual firm data in order to check the existence of industry-spatial effects alongside other microeconomic determinants of R&D investment. Spatial proximity is defined by a measure of firms' industry distance based on trade intensity between sectors. The spatial model specified here refers to the combined spatial autoregressive model with autoregressive disturbances (SARAR). In modelling the outcome for each location as dependent on a weighted average of the outcomes of other locations, outcomes are determined simultaneously. The results of the spatial two stage least square estimation suggest that in their R&D decision firms benefit from spillovers originating from neighbouring industries.

Keywords: spatial weights, spatial dependence, spatial models, R&D
Jel Classifications: C31, 010, R11, O31

Acknowledgments. I am indebted to David M. Drukker, Maurizio Pisati and Rafal Raciborski for helpful suggestions for the spatial analysis. I also thank Giuseppe Medda and Juan de Dios Tena-Horillo for valuable comments on the weight matrix construction. All the usual disclaims apply.

1. Introduction

Many economists have found that R&D expenditures of individual firms are crucial for the competitiveness ofirms, sectors and for the sustained long-run growth of an economy (Grossman and Helpman, 1990; Romer, 1990). While firms invest in R&D in order to increase their productivity and profits, the privately generated technology of individual firms spills over to other firms and becomes public knowledge. This gives rise to an external effect which increases the productivity of all firms. The basic idea is that a higher level of knowledge, which is the result of the general level of R&D commitment, generates positive externalities and promotes the generation and introduction of new processes or products. Externalities in the form of knowledge spillovers have a potentially important role in shaping the incentives for research and development activities of private firms. All this contributes to foster the economic growth of a country.

While theories of knowledge spillovers were traditionally formulated to explain the concentration of industries in general, geographic distance is not the only relevant factor for R&D spillovers as the trade intensity between two firms might also be important. The rationale is that the amount of research in one sector might be correlated to that in nearby sectors. In this sense, firms benefit from spillovers originating from their own or from close industries. Such benefits depend on the extent to which a sector trades with the other sectors. In such a view spillovers might be particularly important for explaining the clustering of R&D expenditure in specific sectors.

More than most industries, R&D depends on new knowledge and the sectoral concentration of R&D activity is likely to favour the exchange of information among firms about new goods or new ways of production. Since R&D expenditures tend to be relatively concentrated in a number of high-tech sectors, such as computers, electronics, drugs, instruments, and aerospace, spatial spillover analysis may be useful to investigate the mechanism by which R&D expenditures are distributed over industries. In the OECD (1996) study, input-output matrices are employed to calculate R&D flows between sectors. The conclusion from this analysis, which compares distributed R&D with own R&D, is that medium and low-technology industries gain more than high-technology industries. As a corollary, R&D spillovers tend to reduce technological differences and bring the three industry groups closer. This takes one to the question whether the research engagement of one firm might be influenced by the R&D decision of nearby firms or sectors and represents the main issue of the present work.

Following the industry spillover literature, this work examines how a firm's technology level may depend on R&D efforts by the firms belonging to the same sector as well as by other related sectors. This paper looks at the spatial dimension of private R&D activity. Compared to other studies in innovation economics (for instance Audretsch, 1998), this work pays particular attention to the sectoral element in analyzing firms' in R&D investment. The analysis is carried out at micro level and is based on a sample of 1060 Italian firms investing in research and it is applied to a cross-section for the period 2001-2003. Spatial analysis is employed to investigate the existence of industry-spatial effects alongside other microeconomic determinants of R&D investment in order to improve statistical conclusions. Given its gap in terms of R&D investment compared to most other industrialized countries, Italy provides an interesting case study. Moreover, since this kind of an analysis has never been carried out for the Italian case, this paper represents a novelty in this regard.

The consequences of spatial autocorrelation are the same as those associated with serial correlation and heteroscedasticity. However, if the spatial correlation is due to the direct influence of neighbouring sectors, OLS estimation is biased and inefficient (Anselin, 1988). In this work external technology is measured by the weighted sum of R&D capital of other firms in the same or different industries (Griliches, 1979; Bernstein and Nadiri, 1989; Jaffe, 1986; Los and Verspagen, 2000; Aiello and Cardamone, 2008). Spatial proximity is defined by a measure of firms' industry distance based on trade intensity between sectors. The weighting framework is built on a row-standardized spatial link matrix of the bilateral supply shares (Coe and Helpman, 1995) and a symmetric kind of trade intensity is employed to measure the distance between two firms/sectors. The spatial influence corresponds to the weighted sum of the R&D expenditure in each industry j, where the weights are given by the share of sales (over total sales) of industry i to industry j (for $i \neq j$).

The model specified here refers to the combined spatial autoregressive model with autoregressive disturbances (SARAR). In modelling the outcome for each location as dependent on a weighted average of the outcomes of other locations, outcomes are determined simultaneously. Hence, differently from traditional spatial analysis which typically considers only one spatial dependence at time (either lag or error) while the other type of dependence is set equal to zero, in this work spatial autoregressive lag dependence and spatial autoregressive error dependence are modelled simultaneously (Anselin and Florax, 1995). The results of the spatial two stage least square estimation suggest that in their R&D decision firms benefit from spillovers originating from neighbouring industries.

The article is organized as follows. Section 2 provides a description of the literature. Section 3 presents the data set and the variables used. Section 4 describes the theory underlying the spatial regression models. Section 5 explains the construction of the spatial weight matrix. Section 6 presents the general results. Section 7 illustrates the main conclusions.

2. Literature review

There is a general consensus on the fact that significant R&D spillovers between firms or industries occur within and across nations (Griliches, 1992, 1998; Nadiri, 1993). Griliches (1979) emphasizes the role of knowledge spillovers in economic growth. Coe and Helpman (1995) highlight international spillovers of technology through the trade of intermediate goods and show that productivity depends on domestic and on foreign R&D capital stocks. They use the cumulative spending for R&D of a country to measure the domestic stock of knowledge, while the foreign stock of knowledge is calculated as import-weighted sum of cumulated R&D expenditures of the trade partners of the country. Positive empirical evidence at firm level, which has shown how R&D influences growth and the rate of return of innovative investments, has been extensively reviewed in Griliches (1992) and Nadiri (1993).

Since Marshall (1890) several theories have been proposed to explain firms' tendency to cluster spatially. Firms may wish to minimize transport costs by locating close to a natural resource, to their suppliers or markets. Firms may also cluster to share inputs such as specialized workers. Finally, firms may cluster to better capture the knowledge that spills over from other neighbouring firms. Concerning labour markets, Marshall argues the ability of a large market to supply a constant market for skills where employers and workers benefit from the imperfect correlations between market components. In relation to input sharing, firms' proximity in specific industries can facilitate a mutually helpful market of upstream input suppliers. Referring to knowledge spillovers, Marshall argues that spatial concentration results in a situation where the trade activity is an implicit exchange of extra-market information. The potential sharing of inputs and especially of knowledge spillovers is likely to be greatly relevant when choosing a location.

Within this debate a less widespread stream of research has tried to examine the rationale of the spatial distribution of R&D activity. It is common wisdom that proximity matters for information to circulate. Tacit non-rival knowledge (Arrow,1962) can easily spill over and it can be exploited in various economic applications. Hence, proximity to an external source of information may increase or incentivise the effect of spillovers from that source. Employing U.S. manufacturing data, Rosenthal and Strange (2001) and Ellison et al (2007) consider the importance of input sharing, matching, and knowledge spillovers for manufacturing firms at various levels of geographic disaggregation (state, MSA, county, and zip code levels) while several other studies have found that knowledge spillovers tend to vanish rapidly as distance increases (Audretsch and Feldman, 1996; Keller, 2002).

A strand of econometric literature has been concerned with the issue of detecting and measuring R&D spillovers between firms and sectors. The main characteristic associated to spillovers is their non-market nature, which makes their detection cumbersome. R&D spillovers, as well as other forms of knowledge accumulation, such as learning-by-doing, are not considered by the traditional R&D intensity variable.

Despite the considerable theoretical and empirical contributions to this field, conclusions about the extent and effect of R&D spillovers are still not clearly defined and it remains still difficult to reliably quantify the effect of R&D externalities. While R&D spillovers are commonly recognized as generating major externalities for firms undertaking R&D, it is less clear whether external R&D is a complement or a substitute for the firm's own R&D. For example, a positive effect of knowledge externalities on R&D spending can be interpreted as evidence of complementarity between external and internal R&D, but also as being consistent with simple technological competition among firms. Moreover, it may well be that the effects of R&D spillovers might depend on the individual firm's efforts to look for information.

A recent contribution in the estimation of R&D spillovers between firms and sectors concerns the introduction of space. The geography of innovation literature stresses that geography matters in the innovation process because of better and easier interpersonal relationships and contacts. It is assumed that the benefit a firm can derive from other firms' technological efforts is inversely related to its distance from the firm emanating the externality (Wolff and Nadiri, 1993; Keller, 2002). In some studies the borders are determined geographically (Adams and

Jaffe, 1996; Orlando, 2004) while other studies define economic spaces within manufacturing sectors to explore intra-industry spillovers (Bernstein and Nadiri, 1989; Los and Verspagen, 2000).

A general technique to measure spillovers is to proxy Marshallian and Jacobian spillovers directly. The Marshallian criterion of agglomeration externalities is often expressed by a specialization index such as the share of industry output or employment to which a firm belongs to in the region (Harris and Li, 2009) or some form of location index (de Vor and de Groot, 2009). Jacobian measures of diversity can be constructed by simple counts of the number of industries located in a certain region (Harris and Li, 2009), or by similar proxies such as the Krugman specialisation index (de Vor and de Groot, 2009). It is worth highlighting that such spillover variables do not require a specific form of weighting. They assume spatial externalities to be confined to the region in which the firm is settled. To the extent that spillovers are inter-regional, such measures are therefore mis-specified.

Several papers have used economic theory to predict forms of spatial correlation (Funke and Niebuhr, 2005; Bottazzi and Peri, 2003). In measuring embodied inter-industry R&D spillovers Griliches (1979, 1992) clearly separates embodied rent and disembodied knowledge spillovers. Embodied spillovers are mainly linked to transmission of goods and are commonly computed through input–output data of R&D embodied in the intermediate and capital goods. Terleckyj (1980) distinguishes the effect of R&D within industries and the effect of embodied R&D spillovers from other industries in terms of capital and intermediate goods. Scherer (1982) employs the product R&D data to measure the R&D spillovers and constructs an inter-industry flow matrix.

Aiello and Cardamone (2008) generate an R&D spillover variable to investigate firm's productivity in a sample of Italian firms. They weight the (external) R&D capital stock of all firms in the dataset by a variable that reflects firms' technological similarity and geographical proximity. Technological similarity represents technological flows between two firms. Their results show that R&D spillovers positively affect firms' production and that geography matters in determining the role of the external technology. Andersson and Gråsjö (2009), employing a gravity model type approach, create variables that proxy spillovers as a measure of accessibility (and thus the potential for interaction). Spatial spillovers are assumed to be associated only to physical distance, and a particular distance decay function is imposed. Nevertheless, it would be possible to use different weight matrices and different (also combined) types of distance.

3. Data and variables

The data used in this study are taken from the Survey of Manufacturing Firms (SMF), carried out by the Research department of Capitalia Bank (2003). The SMF surveyed a stratified sample of Italian firms with 11 to 500 employees. It also included all manufacturing firms with more than 500 employees. The data was stratified according to the number of employees, the sector, and the geographical location. It used the Census of Italian Firms as a benchmark. The SMF contains questionnaire information about firms' structure and behaviour, and fifteen years of data on their balance sheets (1989-2003). Unfortunately, access to longitudinal data is limited. Since only a small fraction of the observations overlap, only the 2001-2003 survey is used in the empirical application. This clearly prevents the analysis from addressing long-term considerations.

Considering the sectoral composition, the sample under investigation is dominated by firms in basic metal and in textiles, clothing, metals, metallic products, industrial machinery. On the other hand, petroleum oil and coal industries are represented by only few firms. Since not all sectors contain a sufficient number of observations to allow running estimations, some sectors have been grouped according to their technological similarities and finally 14 sectors were obtained (table 1). After data cleaning the final sample contains 1060 observations.

The survey supplies information about the total amount of R&D investment. Within each sector the differences in R&D intensity are substantial, as indicated by the fact that the standard deviation is larger than the mean in most of the 14 sectors. It is straightforward to observe that the average amount of research per worker for the whole sample is 3'735€ and that firms investing in R&D are not uniformly spread across industries. R&D per worker is the lowest for the wood and wood products sector (1'490€) and for the petroleum and coal sector. It is highest in the auto, motor vehicles and other transportation equipment (8817€). The ratio of the largest to smallest sector is 6, thus confirming that the distribution of investment in research is rather heterogeneous across the sample. This represents a starting information for this paper which attempts to ascertain if such heterogeneity has effects on the R&D intensity in individual sectors. Finally, the 14 sectors obtained are used to construct the sectoral links weight matrix employed for the spatial analysis described below.

3.1. Variables description

The variable of interest is the amount of R&D expenditure over the three year period (2001-2003). This is divided by the number of workers to provide a measure of the intensity of R&D ($LogR\&D_{EMP}$). Hence, unlike many other studies, this work considers continuous dependent variables when exploring the relative importance of different factors in R&D decision.

In line with the existing literature, firm size is included as an explanatory variable. This variable is measured as the logarithm of the number of employees and refers to the initial year (*LogEMP*). Size can affect R&D decisions in several ways, such as better organization, easier access to the financial markets, specialization of activities and routines, and investment in complementary activities to R&D. Moreover, since size can help to overcome the fixed cost barrier, it becomes an important factor in determining whether or not the firm invests in R&D. Ever since Schumpeter, this idea has been investigated from a theoretical point of view (Arrow, 1962) and mixed empirical evidence to support it has been found (Cohen and Levintal, 1989; Audretsch, 1991; Breschi, Malerba and Orsenigo, 2000). The squared term of firms' size is included to control for possible non-linear effects (Log^2EMP).

A measure of stable R&D commitment ($LogEMP_{R\&D\text{-}EMP}$), measured as the number of R&D personnel compared to total personnel at the initial period, is also included among the regressors. This supplies a proxy for a firm's engagement in R&D and approximates the firm's human capital intensity. Knowledge is a crucial intangible asset in R&D engagement decision and having its own R&D department is considered a factor that reduces risks (Kleinknecht and van Reijnen, 1992).

Capital intensity, expressed in logarithmic term ($LogKAP_{INT}$), is measured as physical assets per employee, to account for the fact that firms in more capital-intensive productions may have a higher propensity to be committed in R&D projects.

The model checks for the possible role of R&D subsidies by including a dummy variable which indicates whether the firm received R&D subsidies during the three year period ($GRANT_{R\&D}$). Market failures in real and financial markets offer justification for public support, as the return may be not sufficient to justify private investment. The broad consensus on the use of public support is based on the inefficiencies of the market. These create a gap between the private and social return on R&D and may yield to a less than optimal level of research. This is because of incomplete appropriability of research output and externalities deriving from the public good nature of R&D (Nelson, 1959; Arrow, 1962). As a result, public funding tends to have a positive influence on firms' R&D spending (Czarnitzki, D. 2006; Carboni, 2011, 2012). $GRANT_{OTH}$ is a dummy=1 if the firm received other public grants.

An export dummy (*EXPORT*) is included because firms that compete in foreign markets tend to be more innovative than others. Operating in more competitive environments, exporting firms are more inclined to invest in research and to improve R&D strategies. There may also be an indirect effect, deriving from the richer network of customers, suppliers or competitors that exporting firms may have access to, which may make R&D investment more likely. Franco et al (2011) for instance, investigate the effect of trade-related R&D spillovers on Total Factor Productivity and find that the impact on TFP of the available foreign R&D stock is greater than that of the domestic one.

A measure of the financial constraints is also included (*RATION*). Such constraints are in general good at explaining under-investment in technology and in R&D expenditure. This variable provides a proxy of credit market efficiency. The total cost of research may vary across firms due to differences in the availability and cost of financial resources. Arguments such as risks, sunk costs and other forms of market failures are commonly seen as having particularly severe effects in this field (Czarnitzki, D. 2006). A measure of indebtedness is also included in order to control for the potential of the firms to find financial sources. It is expressed as the ratio of debt to banks over average value added ($DEBT_{AVY}$).

Industry dummies are included among the regressors in order to control for potential sectoral systematic differences in the amount of research. These are: traditional 'supplier dominated' ($PAVITT_1$), 'scale-intensive' ($PAVITT_2$), 'specialized equipment suppliers' ($PAVITT_3$) and 'science-based' ($PAVITT_4$). There might be significant cross-sectional differences in technological opportunity, appropriability conditions which may affect innovation behavior and competences of individual establishments. Moreover in some industries fixed costs will be lower than in others. Controls for intercept effects may be desirable in such cases, so that some of these unobservable effects can be captured.

Finally, where useful, variables are divided by labour units so as to reduce collinearity with firm size and log-transformed in order to avoid dimensional effects. Variables are referred to the initial period in order to mitigate possible endogeneity with government grants.

4. Spatial Regression Models

There are two approaches in the literature dealing with spatial dependence: spatial lag model and spatial error model. Spatial lag model (SAR) can be employed when the aim is to investigate the existence and strength of spatial interaction. In spatial lag model, not only Y depends on its characteristics (y_i) but it also depend on the value of its neighbours (y_j). It assumes that the spatially weighted sum of neighbourhood (the spatial lag) enters as an explanatory variable in the specification of housing price formation:

$$y_i = \lambda \sum_{j=1}^{n} w_{ij} y_j + \sum_{p=1}^{k} x_{ip} \beta_p + u_i \tag{1}$$

in matrix notation:

$$\mathbf{Y} = \lambda \mathbf{WY} + \mathbf{X\beta} + \mathbf{u} \tag{2}$$

Hence in the spatial lag model the spatially lagged variable WY is included as an additional regressor. λ is the spatial dependence parameter typically referred to as spatial-autoregressive parameter. W is a $n \times n$ standardized spatial weight matrix (where n is the number of observations). X is an $n \times k$ matrix of observations on k right-hand-side exogenous variables. β is the corresponding $k \times 1$ parameter vector. The spatial weight matrix, W, tells us whether any pair of observations are neighbours. The resulting spatial lag WY can be viewed as a spatial weighted average of observations at neighbouring locations and represent the corresponding scalar parameters typically referred to as spatial-autoregressive parameters. ε are *i.i.d.* disturbances. In this case the spatially lagged regressor is correlated with the error term and OLS estimation turns out to be biased and inconsistent due to the simultaneity bias (Anselin, 1988). The spatial lag model will be estimated using maximum likelihood approach.

In the spatial error model (SARE), spatial dependence is modelled as a spatial autoregressive process in the error term:

$$y_i = \sum_{p=1}^{k} x_{ip}\beta_p + u_i \qquad (3)$$

$$u_i = \rho \sum_{j=1}^{n} m_{ij} u_j + \varepsilon_i \qquad (4)$$

in matrix notation:

$$\mathbf{Y} = \mathbf{X\beta} + \mathbf{u}\ ;\ \mathbf{u} = \rho\mathbf{Mu} + \boldsymbol{\varepsilon} \qquad (5)$$

Where Y is an $n\times1$ vector of observations on the dependent variable, ε are again *i.i.d* disturbances, ρ is the spatial error parameter and M is a $n\times n$ spatial link matrix with zero diagonal elements. Ignoring spatial dependence in the error term does not yield biased least squares estimates though their variance will be, thus resulting in misleading inference (Anselin 1988, 1990).

Spatial models have many similarities to the moving average (MA) model in time series econometrics, in which the error of certain observations may be affected by errors of other observation. In such a case, OLS estimation of spatial error model will be inefficient because it violates the assumption of independence among disturbance term. Hence, the classical estimators for standard errors are biased.

One crucial feature of spatial analysis is that it takes into account the spatial arrangement of the observational units (locations). This spatial arrangement is represented by a spatial weights matrix W whose non-zero off-elements w_{ij} express the presence or absence (binary weights matrix) or the degree (non-binary weights matrix) of potential spatial interaction between each i^{th} and j^{th} possible pair of locations.

Spatial-weighting matrices are employed to compute weighted averages in which more weight is placed on nearby observations than on distant observations (Cliff and Ord, 1981; Haining, 2003) and parameterize Tobler's law of geography "Everything is related to everything else, but near things are more related than distant things" (Tobler, 1970). This issue rises concerns on how to measure the distance or contiguity between the observations at different locations.

There are several different spatial link matrices available to measure the contiguity between locations. Neighbourhood matrices are symmetric and binary $n\times n$ matrices with w_{ij}=1 if two observations are neighbours and w_{ij}=0 if not. These matrices strongly depend on the definitions of the neighbourship. Spatial connectivity matrices are similar to neighbourhood matrices, but they are non-binary. They are symmetric $n \times n$ spatial link matrices, where the elements w_{ij} measure the degree of the closeness. Similar to these connectivity matrices are distance matrices which are again non-binary symmetric $n \times n$ matrices, here the elements w_{ij} measure the distance between locations.

In a contiguity matrix W,

$$w_{ij} = \begin{cases} d_{ij} & \text{if } i \text{ and } j \text{ are neighbors} \\ 0 & \text{otherwise} \end{cases} ; \quad \text{where } d_{ij} \text{ is a weight} \qquad (6)$$

In inverse-distance spatial-weighting matrices, the weights are inversely related to some measure of distance between the locations. In an inverse-distance matrix W, $w_{ij}=1/d_{ij}$ where d_{ij} is the distance between places i and j. Inverse-distance matrices may allow for all spatial objects to affect each other and are usually normalized to limit dependence. Sometimes objects outside a given radius are set to have zero weight. This also limits dependence. To sum up, spatial weight matrices measure the similarities or dissimilarities between spatial locations: the higher the connectivity, the smaller the distance and vice versa.

Spatial influence enters network autocorrelation models through W (the structure matrix). Entry w_{ij} represents the extent to which y_i is dependent on y_j, and thus to what extent actor j influences i. Constructing an *a priori* constructed spatial weights matrix has the big advantage that spatial interactions across "regions" are collapsed into a single (weighted) variable. However, its limit is that it does not directly test which regions interact with each other nor the strength of such interactions (Harris and Moffat, 2011).

The spatial weight matrix is usually standardized. This enables one to interpret the spatial lag term in a spatial model as a mere spatially-weighted average of observed neighbouring values (y_j). For instance, cohesion suggests that actors are influenced by adjacent actors, normalization then decreases the individual strength of influence with the number of influencers. The most common normalization method, which is also used in this work, is the row-normalization.

In a row-normalized matrix, the $(i, j)^{th}$ element of $\tilde{W}$ becomes e $\tilde{w} = w_{ij} / \Sigma r_i$, where Σr_i is the sum of the i^{th} row of $\tilde{W}$. Thus Σr_i denotes the number of actors with whom i has a tie. After row normalization, each row will sum to one and every actor receives the same total amount of influence from all actors. Influence of *i* by *j* decreases with the number of actors influencing *i*. Finally even though the original matrix is symmetric, the weight matrix $\tilde{W}$ deriving from row (or column) normalization is likely to become asymmetric since:

$$\sum_j w_{ij} \neq \sum_i w_{ji} \tag{7}$$

By construction, whatever the type of proximity chosen, the spatial lag WY is an endogenous variable. Hence, in autocorrelation models the specification of W is of crucially important since this is aimed at estimating ρ and λ (the spatial-autoregressive parameters which measure the extent of these interactions) or β (Leenders, 2002). Sometimes these models are used to remove the bias due to the interdependence of units from OLS parameters estimation. The literature supplies several methods. The most common is the maximum likelihood (ML) estimation procedure which explicitly incorporates the weight matrix W.[1]

4.1 The sectoral weight matrix construction

Several physical and social phenomena take place within networks of interdependencies. This type of phenomenon is typically modelled as a network autocorrelation model where parameter estimates and inferences are based upon the specification of a weight matrix, the elements of which indicate the influence pattern in the network. The issue here is that in integrated economies it might well be the case that is not the geographic distance but rather the trade intensity between two firms that becomes relevant for R&D decisions. The assumption in this case, is that sectors can be regarded as regions and, with an appropriate spatial link matrix, a spatial analysis can be done. This raises a question about how to adequately measure the distance between firms in different industries (locations).

For example, to construct industry weights, previous studies have used trade flows statistics at the sectoral level. Bartelsman et al. (1994) and Morrison and Siegel (1999)employ a method that distinguishes between the potential spillover from downstream linkages (demand-driven spillovers) and upstream linkages (supply-driven spillovers). This

---

[1] Further discussions of spatial-weighting matrices and the parameter space for the spatial-autoregressive parameter can be found in Kelejian and Prucha (2010) and Drukker et al. (2011,a).

method is based on the assumption that the more industry *i* acquires from and sells to industry *j*, the more it can be influenced by industry *j* (Audretsch and Feldman,1996; Peri, 2005; Piga and Poyago-Theotoky, 2005).

To take into account spatial effects this work employs a weight matrix that is based on sectoral distance. The assumption is that each firm belonging to manufacturing sector *i* can potentially be affected from facts occurring in sector *j*, and that the magnitude of the influence depends on the intensity of trade flows between the two sectors. To construct a measure of trade intensity between sectors data from the input-output matrix for the Italian manufacturing sectors is used (Istat, 2004; Medda and Piga, 2007). Hence the spatial sector indicator captures the intensity of the potential R&D influence that a firm in industry *i* receives from the R&D performed in all the other *j* industries that supply industry *i*, where weights are proportional to the inter-industry trade flows and are derived from input-output matrix coefficients (Terleckyj, 1974; Wolff and Nadiri, 1993). The spatial influence corresponds to the weighted sum of the R&D expenditure in each industry *j*, where the weights are given by the share of sales (over total sales) of industry *i* to industry *j*, for $i \neq j$ (table 2).[2]

Following Coe and Helpman (1995) the bilateral supply shares are used to construct the row-standardized spatial link matrix. Therefore a symmetric kind of trade intensity is employed to measure the distance between two firms (Gumprecht, 2007). Each element w_{ij} (of the distance matrix W) is defined as the average of the bilateral import-shares between sectors *i* and *j* (Anselin and Bera, 1998) and represents the intensity of effects between these two sectors. The elements of the symmetric spatial connectivity matrix are simply calculated by:

$$w_{ij} = \frac{z_{ij} + z_{ji}}{2} \qquad (8)$$

[2] Some studies measure distance between firms by considering inter-sectorial flows of intermediate goods. Other works employ patents of innovations to construct technology spaces. Adams and Jaffe (1996) and Orlando (2004) employ a measure of geographical distance between firms, while Macdissi and Negassi (2002) model the external technological spillover on the basis of firms' resources devoted to cooperation and capital flows.

where $i \neq j$, and z_{ij} are the bilateral import-shares of sector i from sector j, and by definition $w_{ij}=0$ for $i=j$. The distance between two sectors is simply the inverse connectivity and is used to produce a trade-intensity space:

$$d_{ij} = \frac{1}{w_{ij}}; \text{ where } d_{ij} = 0 \text{ (i.e. the main diagonal takes zero values)}$$

After computing variables and accounting for missing values, this yields a 1060×1060 dimension matrix of weights. It is worth noting from table 2, that most of the trade takes place within the same sector. Nevertheless, each sector does show input exchange with the remaining industries. This information is included in the connectivity matrix which has been created with the *spmat* command in STATA (Drukker et al., 2011,a) and summarized in table 3. It emerges that there are 1'118'839 total links ranging from a minimum value of 1'009 to a maximum value of 1'059 indicating quite a spread (though with different intensities) trade among the 1'060 firms considered.

5. Empirical analysis and results

Following the recent literature on empirical R&D models, the paper proposes a cross-section analysis based on a sample of Italian manufacturing firms. The aim is to investigate the determinants of R&D and check for potential spatial effects. The conjecture, then, is that R&D investment decisions in one industry may be correlated with a cluster in nearby industries. The analysis follows two steps. Firstly, the OLS model is run and tested also for spatial autocorrelation and heteroscedasticity. Then, the combined spatial-autoregressive model with spatial-autoregressive disturbances is estimated.

The R&D equation is:

$$\begin{aligned} LogR\&D_{INT} = {} & LogAVY_{INT}, LogEMP_{R\&D}, LogEMP, LogEMP^2,\ LogK_{INT}, \\ & DEBT_{AVY}, RATION, GRANT_{R\&D}, GRANT_{OTH}, EXP, PAVITT_{1,2,3} \end{aligned} \quad (9)$$

One of the main assumptions for the ordinary least squares regression is the homogeneity of variance of the residuals. Before proceeding further, the Breusch–Pagan testis employed on the residuals of the original linear model. The chi-square value is small, indicating that heteroscedasticity is not likely to be a problem in the sample used (χ^2=0.34 with a p-value =0.56). It is useful to recall that the OLS model does not take into account spatial spillovers among the units.

One first question with spatial analysis is to detect potential spatial dependence among observations. If not, there is no need for using special models or methods in the analysis. The most common global test of spatial autocorrelation is based on a statistic developed by Moran (1950). This statistics compares the value of the observed variable at any location with the value of the same variable at neighbouring locations. The Moran coefficient is given by:

$$I = \frac{\sum_i \sum_j w_{ij}\left(z_i z_j\right)}{\sum_i z_i^2} \tag{10}$$

Where w_{ij} denotes the elements of the spatial weights matrix, $z_{ij}=y_i-\mu$ the variable of interest centred on the sample mean μ. Under the null hypothesis of no global spatial autocorrelation, the expected value is:

$$E(I) = \frac{1}{N-1} \tag{11}$$

Hence, Moran's *I* is used here to analyse the spatial association of the R&D investment intensity at the level of the establishment. This coefficient is of fairly simple computation and interpretation. The Moran coefficient is zero in the case of no spatial autocorrelation irrespective of the analysed variable or spatial system (Hordijk, 1974). If Moran's *I* is larger than its expected value, then the overall distribution of variable under observation can be seen as characterized by positive spatial autocorrelation, meaning that the value of R&D investment intensity at each location *i* tends to be similar to the values taken on by the same variable at spatially contiguous locations (Pisati, 2001).

Table 4 depicts the results of the Moran's *I* test. The value of this statistic is 0.0205 while its mean is -0.0009 so positive spatial autocorrelation is detected with a highly robust significance (*p-value*=0.0000) both, with normal approximation and randomization assumptions.

Beside the Moran's *I* test the Lagrange multiplier (LM) test, and a robust Lagrange multiplier test (robust LM) are performed both for the spatial lag model and for the spatial error model. RLM-error test corrects for the presence of local spatial lag dependence, assuming λ=0. Likewise, the RLM-lag assumes ρ=0. LM tests are distributed χ^2. The Moran test supplies reliable results for alternative forms of ignored spatial dependence, whereas the LM tests supply indications about the kind of spatial dependence (Anselin and Florax, 1995; Anselin and Bera 1998). It is worth underlying that these tests explicitly incorporate the weight distance matrix W discussed above.

This result is confirmed when this statistic is derived from the OLS estimation (2.69, *p-value*=0.007). The null hypothesis is safely rejected according to the Moran's *I* test (*p-value*=0.007). The results for spatial error show no evidence of spatial error dependence both in the LM-error (*p-value*=0.189) and the RLM-error (*p-value*=0.534). The RLM- error and the LM-lag statistics are respectively 4.41 (*p-value*=0.036) and 3.06 (*p-value*=0.080) suggesting that spatial lag dependence is likely to be an issue in this specification. It is important to highlight that the Moran's *I* test is a global statistic, meaning that it accounts for spatial autocorrelation for all the units but it does not supply information about the contribution of each single unit. Local measure of spatial correlation are more indicated to account for this drawback.

Since spatial autocorrelation is detected, and given the absence of heteroscedasticity, the model is then re-estimated incorporating a correction for both spatial error and spatial lag. For this purpose it is employed the spreg-gs2sls (spatial two stage least square estimation) routine available in STATA, developed by Drukker et al (2011,b). This represents a combined spatial-autoregressive model with spatial-autoregressive disturbances (SARAR model in the terminology of Anselin and Florax, 1995). This model also allows for the disturbances to be generated by a spatial-autoregressive process. In modelling the outcome for each observation as related to a weighted average of the outcomes of other units, this model determines the outcomes simultaneously (Drukker et al., 2010):[3]

$$y_i = \lambda \sum_{j=1}^{n} w_{ij} y_j + \sum_{p=1}^{k} x_{ip} \beta_p + u_i \tag{12}$$

$$u_i = \rho \sum_{j=1}^{n} m_{ij} u_j + \varepsilon_i \tag{13}$$

[3] For a discussion of the estimation theory for the implemented GS2SLS estimator see Kelejian and Prucha (1998, 1999, 2010), Arraiz et al. (2010) and Drukker et al. (2010).

In matrix notation:

$$\mathbf{Y} = \lambda\mathbf{WY} + \mathbf{X\beta} + \mathbf{u}\,; \quad \mathbf{u} = \rho\mathbf{Mu} + \boldsymbol{\varepsilon} \tag{14}$$

The spatial-weighting matrices W and M are known and non-stochastic and are part of the model definition. In this application it is considered the case where W=M (i.e. spatial lag and spatial error are modelled on the same weight matrix).

When ρ=0 and λ≠0, the model in equations (12-14) reduces to the spatial-autoregressive model(SAR). When ρ≠0 and λ=0 the model becomes the spatial-autoregressive error model (SARE). For ρ = 0 and λ= 0 the model is simply a linear regression (LR) model with exogenous variables. Finally, for $\rho \neq 0$ and $\lambda \neq 0$, we have the spatial lag model with a spatial autoregressive disturbance (SARAR). Typically in the SAR and SARE models only one test for one type of dependence is carried out while the other type is considered zero $(H_0 : \rho = 0$ and $\lambda = 0$ and vice versa). The SARAR model allows to check the spatial-autoregressive lag and spatial autoregressive disturbance simultaneously and it is employed to carry out the empirical analysis.

Table 5 reports the results for the OLS and the regression that corrects for spatial dependence. Although the results for the two models are quite similar, the spatial two stage least square estimation shows clear evidence of sectoral R&D spillovers. The null hypothesis of zero spatial error (λ=0) as well as the null hypothesis of zero spatial lag (ρ=0) can be safely rejected. The parameter λ is positive and strongly significant, indicating spatial-autoregressive dependence in R&D intensity. This simple means that firm's investment in research in a given sector is affected by investment in research of the neighbouring sectors. The parameter ρ is positive and significant at 5% spatial-autoregressive dependence in the error term. In other words, an exogenous change to one sector is very likely to cause changes in neighbouring sectors. These results support the hypothesis of interdependence among sectors in this sample of 1060 Italian firms.

It is worth underlying that the two spatial parameters have ranges larger than (-1, 1). This is a consequence of the row normalization procedure where each row of the spatial-weighting matrix is rescaled by a different scalar to be within (-1,1) so the optimizer might take them (Kelejian and Prucha, 2010). Moreover, the routine does not check or constrain the spatial parameters to be within (-1,1) so the optimizer might take them outside of the desired range.

Table 5 also shows the results of the spatial analysis. As expected, the researchers intensity measure affects positively the amount of investment in research. The level of added value per worker, which poxies firms' efficiency in this work, also has a strongly significant positive effect on R&D. This is likely to be due to the very characteristics of the research sector where higher efficiency is needed than what required in other industries and where skills of workers are particularly crucial.

The two firm size variables, *LogEMP* and its square value *Log^2EMP* are strongly significant and suggest a non-linear effect of firm's size on research. Larger firms are more likely to have the threshold size and technical capability to enter R&D project than small firms do.

The analysis shows that obtaining a public R&D subsidy has a positive and statistically significant impact on a firm's R&D engagement. Also, being export oriented has a positive and significant impact on the amount of research expenditure. The capital intensity variable does not offer a clear cut off. While the OLS estimate show a negative though weak influence on the decision to invest on R&D, the spatial analysis does not suggest statistical evidence. Being indebted, being credit rationed and receiving other forms of public support are found to exert no statistically significant influence on R&D decision.

Finally, industry dummies suggest differences among sectors in R&D investment this is mostly due to their different technological trajectories. Traditional 'supplier dominated', 'scale-intensive' and 'specialized equipment suppliers' industries are different from 'science based' firms. Their coefficients are, in fact, negative and strongly significant. Such firms are likely to rely more on innovative strategies based on the acquisition of innovation embodied in capital goods developed by external suppliers. By contrast the more high-tech ($PAVITT_4$) category tend to have a stronger propensity to invest in research. Differences in the amount of R&D among industries are confirmed by the test of joint significance of the industry dummy variables.

6. Conclusion

This paper attempts to shed some light onto the spatial dimension of R&D investment. The analysis is applied to a sample of 1060 manufacturing firms in Italy, a country that has been little investigated from this perspective so far. Following the industry spillover literature, this work tries to examine whether a firm's spending in research may depend on R&D efforts by other firms belonging to the same sector as well as to other related sectors. The rationale is that R&D spatial spillover effects are unobserved and may affect firms in a given industry through intra-industry (or inter-industry) sales and some kind of spatial location.

The spatial model specified in the empirical analysis refers to the combined model of spatial autoregressive model with autoregressive disturbances (SARAR). In modelling the outcome for each location as dependent on a weighted average of the outcomes of other locations, such procedure determines outcomes simultaneously. Within the econometrics literature, the method employed here has never been applied to R&D studies so far. In this respect, this paper adds a new contribution to the empirical analysis.

Spatial proximity is defined by a measure of firms' industry distance based on trade intensity between sectors. A symmetric kind of trade intensity is employed to measure the distance between two firms/sectors. The spatial influence corresponds to the weighted sum of the R&D expenditure in each industry, where the weights are given by the share of sales (over total sales) among industries. Hence, the potential spillover flow considered here consists of intangible effects that take place by means that are somehow linked to the amount of market transactions.

The results of the spatial two stage least square estimation suggest that in their R&D decision firms in the sample benefit from spillovers originating from neighbouring industries. This indicates sectoral complementarities among firms' R&D decision. Such benefits depend upon the amount of trade among sectors. The higher level of knowledge resulting from the general level of R&D commitment promotes the generation and introduction of new processes or products. Externalities in the form of knowledge spillovers have a potentially important role in shaping the incentives for research and development activities of private firms. Moreover, given the trade relation among sectors, the cross-spread of spillovers is likely to favour a more equal distribution of knowledge over sectors given that firms from the more dynamic technological sectors possibly pull the whole system.

In terms of policy design the results of this analysis provide interesting indications. The fact that the propensity to engage in research projects increases with enterprise size, has a special meaning in Italy where small and medium firms typically constitute the majority of the productive system. Policies and incentives aiming at joining forces or promoting research joint ventures, would have positive effects on the overall level of R&D. Public support specifically aimed at research activity plays an important role in increasing firm's research commitment. This is in line with theoretical considerations on market failures in real and financial markets. In this respect government funding can mitigate financial constraints and have positive effects on the borrowing capacity of firms, particularly when one considers the inefficiencies inherent in innovation. All this assumes particular relevance in a country like Italy that traditionally lags behind in terms of R&D spending if compared to the other industrialized countries. Public financial support not specifically oriented at R&D does not have influence on the amount of research, reinforcing the need to design specific R&D schemes in order to encourage technological projects.

References

Adams, J.D. & Jaffe, A.B. (1996) Bounding the Effects of R&D: An Investigation Using Matched Establishment-Firm Data, *Rand Journal of Economics*, 27(4), 700-721.

Aiello, F. & Cardamone, P. (2008) R&D Spillovers and Firms' Performance in Italy: Evidence from a Flexible Production Function, *Empirical Economics*, 34(1), 143-166.

Andersson, M. & Gråsjö, U. (2009) Spatial dependence and the representation of space in empirical models, *The Annals of Regional Science*, 43(1), 159-180.

Anselin, L. (1988) Spatial Econometrics: Methods and Models. Boston: Kluwer Academic Publishers.

Anselin, L. & Florax R. (1995) Small Sample Properties of Tests for Spatial Dependence in Regression Models: Some Further Results, in Luc Anselin and Raymond Florax (eds.), New Directions in Spatial Econometrics. New York: Springer, 75-95.

Anselin, L. (1990) Some robust approaches to testing and estimation in spatial econometrics, *Regional Science and Urban Economics*, 20, 141-163.

Anselin, L. & Bera, A.K. (1998) Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics, in: Giles, D., Ullah, A. (Eds.), Handbook of Applied Economic Statistics, Marcel Dekker, New York, 237-289.

Arraiz, I., Drukker D. M., Kelejian H. H., & Prucha I. R. (2010) A spatial Cliff-Ord-type model with heteroskedastic innovations: Small and large sample results, *Journal of Regional Science*, 50(2): 592-614.

Arrow, K. (1962), Economic welfare and the allocation of resources for inventions, in Nelson R. (ed.), The Rate and Direction of Inventive Activity, Princeton University Press, Princeton, NJ.

Audretsch, D.B. (1991) New firm survival and the technological regime, *Review of Economics and Statistics*, 73, 441-450.

Audretsch, D.B. (1998) Agglomeration and the Location of Innovative Activity*, Oxford Review of Economic Policy*, 14(2), 18-29.

Audretsch, D.B. & Feldman M. (1996): R&D Spillovers and the Geography of Innovation and Production, *American Economic Review*, 86(3), 630-640.

Bartelsman, E.J., Caballero R.J., & Lyons, R.K. (1994) Customer-and Supplier-Driven Externalities, *American Economic Review*, 84(4), 1075-1084.

Bernstein, J.I. & Nadiri, M.I. (1989) Research and Development and intra-industry spillovers: an empirical application of dynamic duality, *Review of Economic Studies*, 56, 249-269.

Bottazzi, L. & Peri, G. (2003) Innovation and spillovers in regions: Evidence from 23 European patent data, *European Economic Review*, 47, 687-710.

Breschi, S., Malerba, F. & Orsenigo, L. (2000) Technological regimes and Schumpeterian patterns of innovation, *Economic Journal*, 110, 388-410.

Capitalia (2003), Indagine sulle imprese manifatturiere. Rapporto sull'industria italiana e sulla politica industriale.

Carboni, O.A. (2011) R&D subsidies and private R&D expenditures: evidence from Italian manufacturing data, *International Review of Applied Economics*, 25(4), 414-439.

Carboni, O.A. (2012) An Empirical Investigation of the Determinants of R&D Cooperation: an Application of the Inverse Hyperbolic Sine Transformation, *Research in Economics*, (forthcoming).

Cliff, A. & Ord J.K. (1981), Spatial processes: Models and applications (Pion, London).

Coe, D.T. & Helpman, E. (1995) International R&D spillovers, *European Economic Review*, 39, 859-887.

Cohen, W.M. & Levinthal, D.A. (1989) Innovation and learning: the two faces of R&D, *Economic Journal*, 99, 569-596.

Czarnitzki, D. (2006) Research and Development in Small and Medium-Sized Enterprises: The Role of Financial Constraints and Public Funding, *Scottish Journal of Political Economy*, 53, 335-357.

Drukker, D.M., Peng, H., Prucha I.R. & Raciborski, R. (2011,a) Creating and managing spatial-weighting matrices using the spmat command, Technical report, Stata.

Drukker, D.M., Prucha I. R. & Raciborski R. (2011,b) A command for estimating spatial-autoregressive models with spatial-autoregressive disturbances and additional endogenous variables. Technical report, Stata.

Drukker, D.M., Egger, P. & Prucha I.R. (2010) On two-step estimation of a spatial autoregressive model with autoregressive autoregressive disturbances and endogenous regressors. Technical report, Department of Economics, University of Maryland.

Ellison, G., Glaeser, E.L. & Kerr, W.R. (2007). What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns. Harvard Business School: Working Papers: 07-064.

Franco, C., Montresora, S. & Marzetti, G.V. (2011) On indirect trade-related R&D spillovers: The "Average Propagation Length" of foreign R&D, *Structural Change and Economic Dynamics*, 22, 227-237.

Funke, M. & Niebuhr, A. (2005) Regional Geographic R&D Spillovers and Economic Growth Evidence from West Germany, *Regional Studies*, 39, 143-154.

Grossman, G.M. & Helpman, E. (1990) Innovation and growth in the global economy, Cambridge, Mass.: MIT Press.

Gumprecht, D. (2007) R&D Spillovers: A Non-Spatial and a Spatial Examination, *Austrian Journal Of Statistics*, 36(1), 3-15.

Griliches, Z. (1979) Issues in assessing the contribution of R&D to productivity growth, *Bell Journal of Economics,* 10, 92-116.

Griliches, Z.(1992) The search for R&D spillovers, *Scandinavian Journal of Economics*, 94 (Suppl.), 29-47.

Griliches, Z.(1998) R&D and Productivity: The Econometric Evidence. University of Chicago Press, Chicago.

Harris, R., Moffat, J. & Kravtsova, V. (2011) In Search of 'W', *Spatial Economic Analysis*, 6(3), 249-270.
Haining, R.P. (2003). Spatial Data Analysis. Cambridge: Cambridge University Press.
Harris, R.I.D., & Li, Q.C. (2009) Exporting, R&D and Absorptive Capacity in UK Establishments. *Oxford Economic Papers*, 61(1), 74-103.
Hordijk, L. (1974) Spatial Correlation in the Disturbances of Spatial Correlation in the Disturbances of a linear Interregional Model, *Regional and Urban Economics*, 4,117-140.
ISTAT (2004). Il nuovo sistema input-output, Roma, Istat.
Kelejian, H.H. & Prucha I.R. (2010) Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances, *Journal of Econometrics*, 140(1), 53-130.
Keller, W. (2002) Trade and the Transmission of Technology, *Journal of Economic Growth*, 7, 5-24.
Keller, W. (2002) Geographic Localization of International Technology Diffusion, *American Economic Review*, 92, 120-42.
Leenders, R.Th.A.J. (2002) Modelling social influence through network autocorrelation: constructing the weight matrix, *Social Networks*, 24, 21-47.
Los, B. & Verspagen, B. (2000), R&D Spillovers and Productivity: Evidence from U.S. Manufacturing Microdata", *Empirical Economics*, 25, I, 127-48.
Macdissi, C. & Negassi, S. (2002) International R&D spillovers: an empirical study, *Economics of Innovation and New Technologies*, 11(2), 77-91.
Marshall, A. (1890) Principles of Economics (Revised Edition ed.). London: Macmillan.
Medda, G. & Piga C.A. (2007)Technological Spillovers and Productivity in Italian Manufacturing Firms, Loughborough University, WP 2007-16.
Morrison, C.J. & Siegel D.S. (1999)Scale Economies and Industry Agglomeration Externalities: A Dynamic cost Function Approach, *American Economic Review*, 89(1), 272-290.
Nadiri, I.M.(1993) Innovations and technological spillovers, NBER Working Paper, 4423.
Nelson, R. R. (1959) The simple economics of basic scientific research, *Journal of Political Economy*, 67, 297-306.

OECD (1996) Technology and Industrial Performance. Technology Diffusion, Productivity, Employment and Skills, International Competitiveness. OECD, Paris.

Orlando, M. J. (2004) Measuring Spillovers from Industrial R&D: On the Importance of Geographic and Technological Proximity, *The RAND Journal of Economics*, 35(4), 777-786

Moran, P. (1950) A test for serial independence of residuals, *Biometrika*, (37), 178-181.

Peri, G. (2005) Determinants of Knowledge Flows and their Effect on Innovation, *Review of Economics and Statistics*, 87(2), 308-322.

Piga, C.A., & Poyago-Theotoky, J. (2005) Endogenous R&D Spillovers and Locational Choice, *Regional Science and Urban Economics*, 35,127-139.

Pisati, M. (2001) Tools for spatial data analysis, Stata Technical Bulletin, STB-60, March.

Romer, P.M. (1990) Endogenous Technological Change, *Journal of Political Economy*, XCVIII: S71-S103.

Rosenthal, S.S. & Strange, W.C. (2008) The attenuation of human capital spillovers, *Journal of Urban Economics*, 64(2), 373-389.

Scherer, F.M. (1982) Inter-industrial technology flows and productivity growth, *The Review of Economics and Statistics*,64, 627-634.

Terleckyj, N.E. (1974) Effects of R&D on the Productivity Growth of Industries: An Exploratory Study, Washington, D.C.: National Planning Association.

Terleckyj, N.E. (1980) Direct and indirect effects of industries research and development on the productivity growth of industries. In: Kendrick, W., Vaccara, B. (Eds.), New Development in Productivity Measurement. National Bureau of Economic Research, New York.

Tobler, W.R. (1970) A computer movie simulating urban growth in the Detroit region, *Economic Geography*, 46, 234-240.

de Vor, F., & de Groot, H.L.F. (2009) Agglomeration externalities and localized employment growth: the performance of industrial sites in Amsterdam, *The Annals of Regional Science*, 44, 409-431.

Wolff, E.N. & Nadiri, M.I. (1993) Spillover Effects, Linkage Structure, and Research and Development, *Structural Change and Economic Dynamics*, 4, 315-331.

TABLE 1: Descriptive statistics: sectors

Manufacturing sector (Nace 2-digit)	Number of firms	R&D intensity (€, industry mean)	s.d.
15, 16: Food, tobacco	78	2436.437	2731.817
17, 18:Textiles, Clothing	129	3836.038	4464.507
19: Shoes, leather	37	3142.146	4767.628
20: Wood and wood products (no furniture)	13	1490.916	1529.66
21, 22: Paper, printing and publishing	32	1992.448	2570.239
23: Petroleum, coal	6	1723.056	1674.074
24: Chemicals	79	5075.802	6942.087
25:Rubber, plastics	65	3130.119	3096.451
26: Non metallic minerals	50	1848.734	1795.395
27, 28: Metals, metallic products	124	3041.212	4383.834
29: Industrial machinery	238	3624.304	4294.113
30, 31, 32, 33: Professional instruments, electric and electronic equipment, radio, TV and telecommunications, Optical, jewelry, measurement equipments	113	6335.567	6732.95
34, 35: Auto and moto vehicles, other transportation equipment	30	8817.395	24074.56
36: Misc.: furniture, musical instruments, toys	66	2234.709	2727.957
Total	1060	3735.036	6165.128

TABLE 2: bilateral import-shares between sector *i* and sector *j*

Sectors	**15, 16**	**17, 18**	**19**	**20**	**21, 22**	**23**	**24**	**25**	**26**	**27, 28**	**29**	**30, 31, 32,33**	**34, 35**	**36**
15, 16	40.8	0.3	3	0.7	4.55	1.35	1.85	2.8	2.65	1.05	1.05	0.85	0.6	0.5
17, 18	0.3	83.9	5.55	0.45	1.4	0.5	3.55	1.85	0.1	0.7	0.85	1	0.55	3.85
19	3	5.55	76.9	0.2	1.15	0.15	0.8	0.75	0	1.15	0.45	0.5	0.4	3.2
20	0.7	0.45	0.2	34.7	0.55	0.15	0.6	0.75	1.1	1.9	0.65	0.65	0.8	18.25
21, 22	4.55	1.4	1.15	0.55	71	0.15	6.6	2.5	2.05	2.95	2.8	2.45	1	1.4
23	1.35	0.5	0.15	0.15	0.15	6.4	4.53	0.25	0.95	3.2	0.45	0.45	0.25	0.15
24	1.85	3.55	0.8	0.6	6.6	4.53	28.9	9.45	3.5	2.7	1.45	1.95	0.7	1.3
25	2.8	1.85	0.75	0.75	2.5	0.25	9.45	12.1	1.05	2.9	5.55	5.1	5.45	1.65
26	2.65	0.1	0	1.1	2.05	0.95	3.5	1.05	18.1	4.25	1.3	2.35	1	1.35
27, 28	1.05	0.7	1.15	1.9	2.95	3.2	2.7	2.9	4.25	63.2	24.6	10.8	11.1	6.2
29	1.05	0.85	0.45	0.65	2.8	0.45	1.45	5.55	1.3	24.6	47.5	16.8	5.7	2.05
30, 31, 32,33	0.85	1	0.5	0.65	2.45	0.45	1.95	5.1	2.35	10.8	16.8	28	11.65	2.9
34, 35	0.6	0.55	0.4	0.8	1	0.25	0.7	5.45	1	11.1	5.7	11.65	83.8	1.55
36	0.5	3.85	3.2	18.25	1.4	0.15	1.3	1.65	1.35	6.2	2.05	2.9	1.55	34.9

TABLE 3: Summary of spatial-weighting matrix

Dimensions		1060 x 1060
Values:		
	Min	0
	min>0	8.27e-06
	Mean	.0009
	Max	.0184
Links:		
	Total	1’118’839
	min	1’009
	mean	1’055.508
	Max	1’059

TABLE 4: Tests for Spatial Autocorrelation

Moran's *I* Statistics: Lag spatial		
Tests	Statistic	P-Value
	Normal Approximation	Randomization Assumptions
Moran's *I*	0.0205	0.0205
Mean	-0.0009	-0.0009
Stddev	0.0023	0.0023
P-value*	0.0000	0.0000
Tests for Spatial Autocorrelation in the OLS Model residuals		
Moran's *I*	2.69	0.007
LM Error	1.73	0.189
Robust LM Error	0.39	0.534
LM Lag	4.41	0.036
Robust LM Lag	3.06	0.080

* Two-tailed test

TABLE 5: Regression results

Dependent variable: R&D intensity # 1060	OLS		Spatial autoregressive model: SARAR (GS2SLS estimates)	
	Coef.	SE	Coef.	SE
$LogEMP_{R\&D\text{-}EMPL}$	0.22***	(0.02)	0.22***	(0.02)
$LogAVY_{INT}$	0.35***	(0.08)	0.35***	(0.09)
LogEMP	-0.83***	(0.19)	-0.82***	(0.02)
Log^2EMP	0.07***	(0.02)	0.07***	(0.02)
$LogKAP_{INT}$	-0.06*	(0.03)	-0.04	(0.03)
$DEBT_{AVY}$	0.04	(0.05)	0.05	(0.04)
RATION	0.02	(0.13)	0.04	(0.13)
$GRANT_{R\&D}$	0.48***	(0.07)	0.46***	(0.06)
$GRANT_{OTHER}$	-0.01	(0.08)	0.03	(0.08)
EXPORT	0.24**	(0.11)	0.22**	(0.10)
$PAVITT_1$	-1.03***	(0.13)	-0.80***	(0.12)
$PAVITT_2$	-1.18***	(0.15)	-0.99***	(0.14)
$PAVITT_3$	-0.86***	(0.13)	-0.83***	(0.12)
Cons	9.78***	(0.53)	0.22	(0.17)
Lambda			1.23***	(.18)
Rho			-2.74**	(1.42)
Test on joint significance of industry dummies	χ^2 (3) = 56.25***			

*, ** and *** indicate significance at the 10%, 5% and 1%, respectively

Ultimi Contributi di Ricerca CRENoS

I Paper sono disponibili in: http://www.crenos.it

- **12/03** *Juan Gabriel Brida, Nicolàs Garrido, Francesco Mureddu,* "Club performance dynamics at Italian regional level"
- **12/02** *Emanuela Marrocu, Raffaele Paci,* "Regional Development and Creativity"
- **12/01** *Bianca Biagi, Maria Giovanna Brandano, Claudio Detotto,* "The effect of tourism on crime in Italy: a dynamic panel approach"
- **11/23** *Rinaldo Brau, Anna Maria Pinna,* Movements of People for Movements of Goods?"
- **11/22** *Giorgio Garau, Giovanni Mandras, Lucia Schirru,* "A Statistical Information System supporting Environmental Policies"
- **11/21** *Emanuela Marrocu, Raffaele Paci, Stefano Usai,* "The complementary effects of proximity dimensions on knowledge spillovers"
- **11/20** *Giuseppe Pulina,* "Tax Evasion and Presumptive Taxation Methods. A Case Study in Italy: Sector Studies"
- **11/19** *Juan Gabriel Brida, Giacomo Del Chiappa, Marta Meleddu, Manuela Pulina,* "The perceptions of an island community towards cruise tourism: A factor analysis"
- **11/18** *Fabio Cerina, Francesco Mureddu,* "Structural Change and Growth in a NEG model"
- **11/17** *Juan Gabriel Brida, Manuela Deidda, Nicolàs Garrido, Manuela Pulina* "Exploring the dynamics of the efficiency in the Italian hospitality sector. A regional case study"
- **11/16** *Juan Gabriel Brida, Nicolàs Garrido, Francesco Mureddu,* "Italian economic dualism and convergence clubs at regional level"
- **11/15** *Adriana Di Liberto, Marco Sideri,* "Past dominations, current institutions and Italian regional economic performance"
- **11/14** *Juan Gabriel Brida, Marta Meleddu, Manuela Pulina*, "Museum visitors can be regarded as specific cultural tourists? A length of stay analysis"
- **11/13** *Edoardo Otranto,* "Classification of Volatility in Presence of Changes in Model Parameters"
- **11/12** *Alessandro Fiori Maccioni,* "The risk neutral valuation paradox"
- **11/11** *Leonardo Becchetti, Vittorio Pelligra,* "Don't Be Ashamed to Say You Didn't Get Much: Redistributive Effects of Information Disclosure in Donations and Inequity-Aversion in Charitable Giving"
- **11/10** *Alessandra Colombelli, Marta Foddi, Raffaele Paci,* "The knowledge regions in the enlarged Europe"
- **11/09** *Emanuela Marrocu, Raffaele Paci, Stefano Usai,* "Proximity, Networks and Knowledge Production in Europe"
- **11/08** *Vittorio Pelligra,* "Empathy, Guilt-Aversion and Patterns of Reciprocity"
- **11/07** *Claudio Detotto, Edoardo Otranto,* "Cycles in Crime and Economy Revised"
- **11/06** *Juan Gabriel Brida, Marta Meleddu, Manuela Pulina,* "The attractiveness of a modern and contemporary art museum"

Finito di stampare nel mese di Aprile 2012
Presso **Copy…Right! studio grafico & stampa digitale**
Via Turritana 3/B – Tel. 079.200395 – Fax 079.4360444
07100 Sassari

www.crenos.it

ISBN 978-88-84-67-727-3