MIUR PRIN 2007

Emerging economic regional powers and local systems of production:

new threats or new opportunities?

WP SERIES – N. 07/09

The role of spatial agglomeration in a structural model of innovation, productivity and export: a firm-level analysis

Roberto Antonietti

roberto.antonietti@unipd.it

Department of Economics and Management “Marco Fanno”

University of Padova

Via del Santo 33, 35123 Padova (Italy)

Giulio Cainelli

g.cainelli@ceris.cnr.it

Faculty of Law, University of Bari

Piazza Battisti, Bari (Italy)

and

CERIS-CNR, via Bassini 15, Milan (Italy)

The role of spatial agglomeration in a structural model of innovation, productivity and export: a firm-level analysis

Roberto Antonietti[1] and Giulio Cainelli[2]

Abstract

Using a large sample of Italian manufacturing firms, in this paper we estimate a structural model of research, innovation, productivity and export performance augmented to take account for the role played by spatial agglomeration externalities. This model, which is an 'augmented' version of Crepon*,* Duguet and Mairesse (1998) model, comprises four main equations. The first identifies the factors underlying the intensity of Research and Development (R&D) investments; the second links R&D capital to innovation output; the third focuses on Total Factor Productivity (TFP) as determined by innovation; the fourth relates TFP to export performance. Our estimates show the significant role played by local externalities in these processes. In particular, related variety and urbanization positively affect the creation of new ideas through R&D, while specialization impacts on TFP to complement innovation output. Finally, urbanization economies support TFP in driving firms' export performance.

Keywords: export, innovation, productivity, R&D, spatial agglomeration, related variety

Acknowledgments

The Authors acknowledge with thanks Andreas Stephan and all the participants to the XII Uddevalla Symposium "The geography of innovation and entrepreneurship", Bari, 11-13 June 2009. The Authors acknowledge financial support from the PRIN project "Emerging economic regional powers and local systems of production: new threats or new opportunities?", University of Bologna. All the usual disclaimers apply.

[1] Department of Economics and Management "Marco Fanno", University of Padova, via del Santo 33, 35123 Padova, Italy. Tel: +39 0498271508. E-mail: roberto.antonietti@unipd.it.

[2] Faculty of Law, University of Bari, Piazza Battisti , Bari, Italy and CERIS-CNR, via Bassini 15, Milan, Italy. E-mail: g.cainelli@ceris.cnr.it

1. Introduction

Since the 1990s the literature has greatly emphasised the role of agglomeration, innovation and internationalization as drivers of firms' economic performance. However, while the positive role of innovation in enhancing firm productivity is now acknowledged (Hall and Mairesse 1995; Harhoff 1998; Wakelin 2001; Wang and Tsai 2003; Griffith *et al.*, 2004; Parisi *et al.*, 2006; Cainelli *et al.*, 2006), the effects of spatial agglomeration are still a puzzling question. In fact, this latter relationship has been extensively investigated at both industry and firm levels (De Lucio *et al.*, 2002; Henderson, 2003; Cingano and Schivardi, 2003; Martin *et al.*, 2008), without reaching a clear-cut conclusion.

A recent line of research, linked to the new theory of international trade, show that productivity can affect in turn the export propensity of firms and their internationalization strategies, more generally. These works show that more productive firms tend to have higher export propensities (Wagner, 2005) or, more generally, that firms with different productivity levels – the main source of firm heterogeneity – are engaged in different modes of internationalization. This is because internationalization is characterized by different levels of sunk costs: i.e., firms need to acquire information on foreign market, establish distribution channels, and so on.

In spite of the relevance of these phenomena, the number of empirical studies that model this set of structural relationships within a unitary empirical framework is limited. This paper is an attempt to fill this gap, by identifying the role played by different forms of spatial agglomeration on innovation, productivity, and export.

We model these structural relationships by estimating an 'augmented' version of Crepon, Duguet and Mairesse (CDM hereafter) model, developed to summarize the complex process "that goes from the firm decision to engage in research activities to the use of innovations in its production activities" (Crepon *et al.*, 1998, p.116).

We extend this structural model in two ways, both vertically and horizontally. Vertically, we add to the three equations characterizing the original CDM model - i.e., the "research" equation linking R&D to its main determinants, the "innovation" equation relating research to innovation output, and the "productivity" equation relating innovation output to TFP - a fourth relation that refers to "export" activity. This equation emphasizes the empirical link with firm heterogeneity – being the TFP the main source of – in determining the performance of firms on foreign markets.

In addition, we use this model to empirically test the role played by different forms of agglomeration economies in each stage of the firm process from the decision to engage in R&D to the decision to export. To do this we modify horizontally the model, introducing into each equation a set of measures for local externalities.

Three specific forms of local externalities are considered: (i) localization economies (also known as MAR economies), arising from the spatial concentration of firms belonging to the same industry; (ii) Jacobs's externalities, which are spurred by the variety and diversity of geographically proximate industries, and capture knowledge spillovers from the cross-fertilization of ideas by firms operating in related or unrelated-sectors (Frenken *et al.*, 2007; Boschma and Iammarino, 2009); and (iii) urbanization economies, which mainly involves information spillovers as local public goods, external to both firms and industry, and which are related to the size of the market and the density of the urban area in which the firm is located (Frenken *et al.*, 2007; Chevassus-Lozza and Galliano, 2008).

Controlling for sample selection and simultaneity, we estimate this recursive four equations system using a large sample of over 700 Italian manufacturing firms. Data, for the period 1998-2003, are drawn by merging information from the VIII and IX Survey on Manufacturing Firms conducted by Unicredit-Capitalia (formerly *Mediocredito Centrale*) with ISTAT 1991 census data on manufacturing industries.

We measure research intensity empirically by real investment in R&D per employee. Since our sample includes firms that do not engage in R&D, we account for possible sample selection bias by using the Heckman (1976, 1979) two-step procedure: in the first stage, estimating firm propensity to invest in R&D, and, in the second stage, estimating the determinants of R&D intensity.

Innovation output is measured through firm propensity to introduce a new product or new process in the three years 2001-2003. We also examine the probability of a new product (product innovation) or a new process (process innovation).

Productivity is measured by both TFP and value added per employee. In order to compute TFP we assume a standard, two-input Cobb-Douglas production function with output measured as deflated value added at each three-digit sector level. To avoid simultaneity, we estimate TFP using the Levinsohn and Petrin (2003) semi-parametric approach, where we use a composite index of materials and services in order to control for un-observables.

Finally, we measure export performance (i) as the propensity to export; (ii) as the share of 2003 export sales; and (iii) as the number of macro-areas in which firms exported in 2003.

The paper is organized as follows. In Section 2 we briefly review the related literature. In Section 3 we describe the dataset (3.1), present the modelling strategy (3.2) and discuss the main empirical results (3.3). Section 4 concludes the work.

2. Related literature

2.1. The role of spatial agglomeration on productivity

This kind of studies originated in the early 1990s when the relationships between spatial agglomeration, knowledge spillovers, and economic growth at the urban level were extensively investigated (Glaeser *et al.*, 1992; Henderson *et al.*, 1995).

Using a cross-section of US cities, Glaeser *et al.* (1992) analyze the impact of three different forms of local knowledge spillovers – Marshall-Arrow-Romer (MAR), Porter and Jacobs externalities – on subsequent urban employment growth. Glaeser *et al.* show that localization economies, or MAR externalities, have a negative impact on urban economic growth, while urbanization (or Jacobs) economies positively affect the subsequent growth of a metropolitan area.

Using a similar empirical framework, Henderson *et al.* (1995) find that localization plays a positive role in mature capital-goods sectors, while productive structure differentiation (variety), which should generate the cross-fertilization of ideas between different industries, has a positive impact only in the case of high-tech industries.

Using French data, Combes (2000) also finds a rather negative impact of specialization on employment growth in both industry and service sectors.

Finally, Forni and Paba (2002), using information on a cross section of 995 Italian LLS for the period 1971-1991 find that in most cases specialization and variety positively affect growth, but that the effect of variety is different for each industry. They note also that, consistent with Marshall (1920), in order to capture the spillover-generating process a size effect needs to be added to the specialization effect.

Glaeser *et al.*'s (1992) approach has been replicated in the contexts of different countries in order to provide further evidence on these issues. However, the various results obtained from empirical research in this field are controversial and currently there is not a unique model that explains the link between economic performance, as measured by employment growth, and the structure of the local economy. In particular, some studies referring to the Italian case, find that specialization has a negative impact on local growth, while diversity plays a positive role (see, among others, Cainelli and Leoncini, 1999; Cunat and Peri, 2001; Usai and Paci, 2003; Paci and Usai, 2006; Mameli *et al.*, 2007).

This empirical literature has been extended by several studies that analyze the impact of measures of agglomeration economies on both employment growth (as in the original body of literature referred to), and productivity or firms' TFP growth (De Lucio *et al.*, 2002; Henderson, 2003; Cingano and Schivardi, 2003; Martin *et al.*, 2008). However, the findings from this new strand of empirical research are also rather puzzling.

For example, De Lucio *et al.* (2002) investigate the relationship between labour productivity and spatial agglomeration at the level of the 50 Spanish provinces for the period 1978-1992, and find that variety plays a role in labour productivity growth, and also find a U-shaped effect for specialization. According to their results, low levels of specialization reduce productivity growth, and high levels foster it.

Cingano and Schivardi (2003), on the other hand, use firm-level based TFP indicators to show that specialization, calculated at the level of the 784 Italian LLSs, has a positive impact on firm productivity growth, but that variety has no significant effect. But, taking local employment growth as the dependent variable, they show that the specialization effect is reversed and becomes negative, while variety has a significant and positive impact on employment growth, thus confirming Glaeser *et al.*'s results.

Henderson (2003), using the US Census Bureau Longitudinal Research Database (LRD) shows that localization economies have strong positive effects on productivity at plant level in high-tech, but not in mechanical industries; he also finds little evidence of urbanization economies.

Martin *et al.* (2008), using French individual firm data from 1996 to 2004, find no significant effect of spatial agglomeration on firm productivity. More precisely, they find that French firms benefit from localization, but not from urbanization economies, but that benefits from industrial clustering – even though highly significant from a statistical point of view – are quite modest in terms of magnitude.

Finally, Boschma and Iammarino (2009) estimate the impact of agglomeration economies on regional economic growth, using export and import data for Italian provinces for the period 1995-2003. They find that forms of related-variety affect regional growth. Their main result is that local systems endowed with sectors that are complementary in terms of shared competences (i.e. with related-variety) perform better.[3]

[3] There is yet another body of work that focuses on the positive effects of spatial agglomeration on productivity. This includes studies of Marshallian industrial districts. Starting with Becattini's "re-discovery" in the late 1970s of this analytical category of Marshallian thought (Becattini 1989) and continued in later works (Langlois and Robertson 1995; Best 1990; Boschma and Lambooy 2002; Boschma and ter Wal 2007; Brusco 1982; Brusco *et al.* 1996; Gordon and McCann 2001, 2005; Iammarino and McCann 2006), the industrial district achieved wide notoriety as a specific type of industrial organization, within which long-term informal links among firms generate agglomeration economies, and, more generally, Marshallian externalities which take the place of scale economies internal to individual firms. Alongside this theoretical revisiting of the Marshallian district concept, a new body of empirical

2.2. The role of productivity and agglomeration on export

Most empirical studies show that more productive firms tend to have a higher propensity for export (Bernard and Jensen,, 1995; Wagner, 2005). These studies find that exporters benefit from larger performance *premia* relative to non-exporting firms.

Two different hypotheses on this relationship have been proposed in the literature (Castellani *et al.*, 2009). The first assumes that the presence of sunk costs, such as transport costs or expenses related to establishing and developing distribution channels, induces self-selection of the more productive firms (Roberts and Tybout, 1997; Bernard and Jensen, 1999). The second hypothesis underlines that firms can become more efficient after they start exporting through learning or scale economies effects (Clerides *et al.*, 1998).

The former relationship has given rise to new theoretical models that try to account for these phenomena (Bernard and Jensen, 2004; Melitz, 2003). The basic idea of these models is that firms with different levels of productivity – as supposed to be the main source of firm heterogeneity – will be engaged in different modes of internalization characterized by different sunk costs. These models assume that "servicing a foreign market entails an entry (sunk) cost, due to the fact that, for example, firms need to acquire information on the foreign market, establish distribution channels and find the appropriate suppliers of goods and services" (Castellani and Zanfei, 2006, p. 4). In this sense, these "new international trade" theories are able to provide a theoretical explanation of the link between firm productivity and export performance (Bernard and Jensen, 1995 and 1999; Melitz, 2003).

Finally, most studies also acknowledge the role played by spatial agglomeration on exports (Bagella and Becchetti, 1998; Bagella *et al.*, 1998; Becchetti and Rossi, 2000; Cainelli and Zoboli, 2004). For example, Bagella *et al.* (2000) show, using a sample composed of over 3,800 manufacturing firms drawn from the Mediocredito Centrale database for the period 1989-1991, that spatial agglomeration, captured by localization within the boundary of an industrial district, increases average export intensity by 4 percentage points. This evidence is tested using econometric methods, which confirm the positive effects of geographical agglomeration on export intensity.[4]

3. The empirical investigation

3.1. The dataset

The dataset we use in our empirical investigation consists of a sample of Italian manufacturing firms drawn from the VIII and IX waves of the Survey on Manufacturing Firms (*Indagine sulle Imprese Manifatturiere*) provided by Unicredit-Capitalia (formerly *Mediocredito Centrale*), which covers the period 1998-2003. Interviews have been

---

literature emerged. These works attempt to establish the presence of a "district effect"; that is, they try to identify empirically the agglomerative benefits that firms derive from membership in one of these productive structures. Following Signorini's (1994) contribution, research in this field (see, e.g., Bagella and Becchetti 2000; Fabiani *et al.* 1999) show that firms in industrial districts do indeed benefit from agglomeration advantages. These results were derived from different econometric specifications and different data sets. However, the results on the positive effects of agglomeration on firm performance are unanimous.

[4] This relationship is further confirmed by more recent evidence showing that export intensity in a sample of Italian manufacturing firms tends to be higher for firms in industrial clusters compared to "isolated" firms (Banca Intesa, 2008).

conducted respectively in 2001 and 2004 for the two surveys with all firms with more than 500 employees and over a representative sample with more than 11 and less than 500 employees, stratified by geographical area, industry and employment size. The master datasets, one referring to 1998-2000 and the other to 2001-2003, gather information on 4.680 and 4.289 firms respectively.

We cleaned the data for missing observations, inconsistencies and outliers, and thus achieved a balanced sample of 715 firms, whose distributions are presented Tables 1, 2 and 3.

Table 1 – Sample distribution by size, area and Pavitt's sectors

Size	**1998-2000**	**2001-2003**	**1998-2003**
11-20	39.94	22.15	32.45
21-50	37.14	29.54	35.66
51-250	16.15	36.93	23.92
251-500	3.87	5.27	4.34
> 500	2.91	6.11	3.64
Area			
North West	37.54	35.91	41.54
North East	27.44	30.12	29.51
Centre	20.62	17.65	18.04
South	14.40	16.32	10.91
Pavitt			
Supplier dominated	52.22	50.71	48.81
Scale intensive	18.14	17.42	17.34
Specialized suppliers	24.34	26.92	29.79
Science based	5.30	3.99	4.06
N. obs.	**4.680**	**4.289**	**715**

Table 2 – Sample distribution by regions

Region	**1998-2000**	**2001-2003**	**1998-2003**
Abruzzo	2.37	3.69	2.24
Basilicata	0.28	0.61	0.42
Calabria	0.41	0.54	0.14
Campania	4.51	4.53	2.80
Emilia-Romagna	12.09	12.93	15.24
Friuli Venezia-Giulia	2.74	3.24	2.94
Lazio	2.52	2.66	2.80
Liguria	0.98	1.03	0.56
Lombardia	26.99	25.30	31.47
Marche	4.64	4.18	4.48
Molise	0.26	0.42	0.28
Piemonte	9.19	9.24	8.95
Puglia	3.46	3.36	1.82
Sardegna	0.88	1.03	0.98
Sicilia	2.24	2.17	2.24
Toscana	11.90	9.22	9.51
Trentino Alto Adige	1.07	1.40	1.12
Umbria	1.56	1.52	1.26
Valle d'Aosta	0.13	0.16	0.00
Veneto	11.79	12.79	10.77

Table 3 – Sample distribution by industry

Industry	**%**
Food, beverages and tobacco	8.11
Textile	12.31
Leather	4.06
Wood	3.64
Paper, publishing, printing	5.59
Coke, petroleum, nuclear fuel	0.70
Chemicals	4.06
Rubber, plastics	6.85
Non-metallic mineral products	5.87
Metal products	15.24
Machinery, equipment	16,36
Electrical and optical equipment	9.09
Transport equipment	1.82
Other industries, furniture	6.29
Total	**100.00**

The variables measuring spatial agglomeration come from the Census of Industry and Services conducted by ISTAT in 1991. We draw particularly on census information on employment and population density at the Administrative Provinces level.

Tables 4, 5 and 6 present some descriptive evidence on our main variables. Table 4 shows that exporting firms, on average, are larger, more productive, more likely to introduce new products and/or processes, more specialized and locate din denser urban areas compared to domestic firms.

Table 4 – Mean characteristics of exporting firms (2003)

	Yes	**No**
- Product or process innovation (%)	78.65	21.35
- Product innovation	82.59	17.41
- Process innovation	74.92	25.08
- Intensity (mean share of innovative sales)	12.04	7.78
Productivity 2003	Mean value	Mean value
- Y/L	4.096	3.885
- TFP	4.685	4.475
Agglomeration (natural logarithms, 1991)*	Mean value	Mean value
- Specialization	0.187	0.113
- Related Variety	1.753	1.710
- Unrelated Variety	2.931	2.931
- Density	0.788	0.580
Employment size	109	73

Notes: * for a description of spatial agglomeration variables see Section 3.2.1.

Table 5 shows that firms introducing new products and/or new processes tend to be more willing to export, more productive, more specialized, located in areas characterized by a higher related variety, lower unrelated variety and lower urban density and, finally, of bigger size

Table 5 – Mean characteristics of innovative firms (2003)

	Yes	No
Export propensity (%)	66.60	33.40
- Intensity (mean share of export sales)	31.5	20.9
Productivity	Mean value	Mean value
- Y/L	4.040	4.028
- TFP	4.646	4.591
Agglomeration (natural logarithms, 1991)*	Mean value	Mean value
- Specialization	0.183	0.141
- Related Variety	1.753	1.721
- Unrelated Variety	2.904	2.952
- Density	0.713	0.739
Employment size	126	58

Notes: * for a description of spatial agglomeration variables see Section 3.2.1.

Table 6 presents details on Italian firms' average productivity for the period 2001-2003. It can be seen that the most efficient firms are more present in foreign markets, are more innovative and more concentrated in specialized clusters. Moreover, the level of productivity increases with firm size, and decreases with higher level of industrial density.

Table 6 – Productivity by export, innovation, agglomeration and size

	Ln Y/L 2003	Ln TFP 2003
Export	4.096	4.685
Domestic	3.885	4.475
Innovative	4.040	4.646
Non innovative	4.028	4.591
- Product innovation	4.057	4.614
- No product innovation	4.018	4.632
- Process innovation	4.038	4.671
- No process innovation	4.033	4.591
Specialization (above average)	4.039	4.711
Specialization (below average)	4.031	4.538
Related variety (above average)	4.060	4.531
Related variety (below average)	4.014	4.701
Unrelated variety (above average)	4.045	4.618
Unrelated variety (below average)	4.022	4.632
Density (above average)	3.968	4.558
Density (below average)	4.045	4.634
Small firms (11-50)	3.973	4.448
Medium (51-250)	4.104	4.843
Large (> 250)	4.354	5.469

3.2. The modelling strategy

To model the structural relationships between R&D, innovation, productivity and export performance, we specify and estimate a four equation recursive system which constitutes a modified version of CDM model.

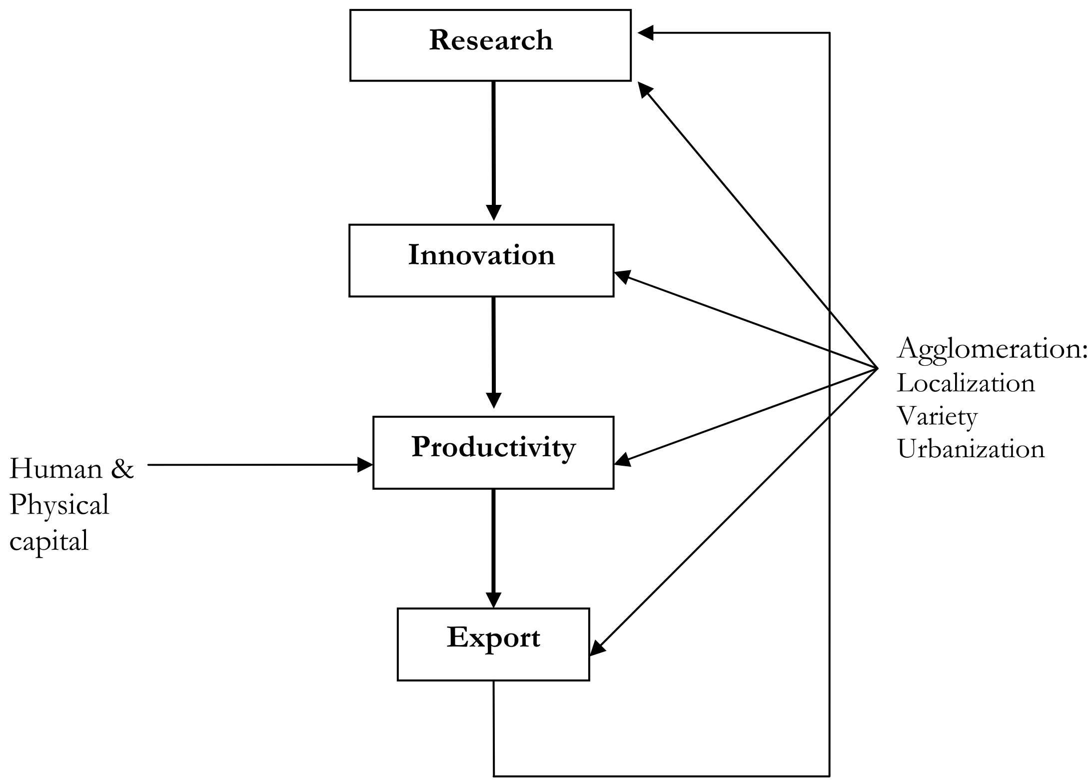
As already mentioned we extend this model in two ways. First, we introduce a new equation, referring to export, which captures the empirical links between firms' productivity and export performance. Second, we empirically test for the effect of different forms of agglomeration economies in each stage, introducing measures of local knowledge spillovers, like: (i) localization economies, arising from the spatial concentration of firms in the same industry; (ii) Jacobs' externalities, spurred by the variety and diversity of geographically proximate industries, and capturing knowledge spillovers from the cross-fertilization of ideas from firms operating in related or unrelated sectors (Frenken *et al.*, 2007; Boschma and Iammarino, 2009); and (iii) urbanization economies, which measure mainly information spillovers as local public goods, external to both firm and industry, and related to the size of the market and the density of the urban area in which the firm is located (Frenken *et al.*, 2007; Chevassus-Lozza and Galliano, 2008).

Note that although we cannot apply panel data techniques, we avoid the problem of simultaneity by measuring all the explanatory variables in each equation three years before the period of the dependent variable.

Our model is depicted in Figure 1. It consists of four equations, one for research, one for innovation, one for total factor productivity, and one for export behaviour. The idea is that research capital is an input that affects the creation of new products/processes, which, in turn, affects firm's TFP level together with human and physical capital. Finally, TFP positively affects export behaviour by reducing the fixed costs associated with access to foreign markets.

All the above relations are assumed to be affected also by spatial agglomeration externalities, stemming from geographical concentration, (related and unrelated) variety and urbanization.

Figure 1 – The structure of the model



3.2.1. The research equation

The first equation in the model concerns firms' research intensity. Since our dataset includes firms that are not involved in research activities, we rely on a two-stage Heckman (1976, 1979) selection model in which we first account for the fact that the firm invested in R&D in 2003, and then we estimate the determinants of R&D intensity. Operationally, we assume that there is a latent variable r^* for firm i given by:

[1] $r_i{*} = x_{0i}\beta_0 + \varepsilon_{0i}$

where x_{0i} is the set of explanatory variables, β_0 is the corresponding vector of the coefficients, ε_0 is the error term and r^* is the latent variable that reflects the criterion for the decision to engage in R&D: for example, the corresponding expected present value of profits. In this case, we observe that the firm invests in R&D if the expected profits are positive or larger than the industry-specific threshold (as in our case since we include a constant term and a set of industry dummies in every equation). In our sample of 715 firms, 49.4% are engaged in R&D compared to 50.6% of firms that are not.
Research intensity r is given by equation [2]:

[2] $r_i = x_{1i}\beta_1 + \varepsilon_{1i}$

where r is the 2003 real observed R&D investment per employee (in natural logarithms), x_{1i} is the vector of the explanatory variables, β_1 is the corresponding vector of coefficients and ε_1 is the error term. In estimating the Heckman two-stage model, we assume that the joint distribution of ε_0 and ε_1 is normal, with zero mean and constant variance-covariance matrix. In order to provide more robust estimates, we include an exclusion restriction in the first-stage selection equation.

Among the explanatory variables, we include:[5] (i) firm market share, computed as firm i sales over average sector sales; (ii) a set of 13 industry-specific dummies; (iii) age (measured up to year 1998); (iv) previous investment in R&D (measured up to year 2000); (v) a dummy variable capturing firm engagement in export in 2000[6]; (vi) a dummy variable equal to 1 if the firm benefited from tax incentives in the period 1998-2000 (for research and technological innovation); (vii) an exclusion restriction based on whether the firm was part of a R&D consortium in the period 1998-2000. This last variable is thought to be correlated with the probability of engagement in R&D, but not necessarily with R&D intensity.

Finally, following the literature on spatial agglomeration economies (Glaeser *et al.*, 1992; Cainelli and Leoncini, 1999; Cingano and Schivardi, 2003; Frenken *et al.*, 2007; Boschma and Iammarino, 2009) we include in this and the following equations: (i) a variable measuring local specialization; (ii) two variables measuring Jacobs externalities through related and unrelated variety; (iii) a variable measuring urbanization economies at Province level.

[5] Unfortunately, our dataset does not include direct information on demand-pull or technology-push variables. Therefore, we deviate slightly from the CDM model, although we do include firm size and sectoral specialization.

[6] In the absence of specific variable, we include market share and previous export activity in order to capture demand-pull factors (at national and international levels), while previous R&D expenditure is used as a proxy for science-push factors and technological opportunities. We use market share as a proxy for firm size.

The first variable captures localization economies, or MAR externalities, i.e. the idea that specialized locations may benefit from within-industry knowledge spillovers from spatial concentrations of firms in the same industry. In line with the literature, we measure MAR externalities using a specialization index. The specialization index of sector *s* located in province *p* is computed as follows:

$$Spec_{s,p} = \frac{\frac{l_{s,p}}{l_p}}{\frac{l_{s,IT}}{l_{IT}}}$$

where $l_{s,p}$ denotes employment in industry *s* in Province *p*, l_p is total manufacturing employment in Province *k*, $l_{s,IT}$ is employment in industry *s* in Italy and finally l_{IT} is total manufacturing employment in Italy.

The second set of variables measures the positive externalities induced by the diversity of local economic activities outside sector *s* as a result of the cross-fertilization of ideas (Jacobs externalities) from related and unrelated sectors. In the former case, following Frenken *et al.*, (2007), we restrict attention to sectors located in province *p* but related to sector *s*. This allows us to calculate an index of *related variety* given by the 5-digit employment share within a 2-digit sector, and thus measuring the level of entropy within each 2-digit sector:[7]

$$\text{Relvar} = \sum_{g=1}^{G} P_g \left(\frac{1}{H_g} \right)$$

where H_g is the inverse of an Herfindahl concentration index calculated at the 5-digit level within each 2-digit level, and P_g is the share of the 2-digit sector *g*.

To measure unrelated variety, we use a standard measure of product diversification in each of the *N* available Italian provinces based simply on the inverse of an Herfindahl concentration index calculated at the 3-digit level. The value of this indicator increases, the more diversified is the productive structure of a Province. This variable is computed as follows:

$$\text{Variety} = \frac{1}{\sum_{i=1}^{N} p_i^2}$$

where p_i denotes the employment share of the 3-digit sector *i*.

Finally, the third variable measures urbanization economies by population density (at each Province level), computed as the number of local units per km^2 of the Province area:

$$\text{Density} = \frac{LU_P}{Area_P}$$

[7] To avoid problems of simultaneity with respect to both the dependent and the other explanatory variables, we calculate our agglomeration variables based on 1991 Census data.

This variable captures knowledge spillovers stemming from the concentration of many different economic activities irrespective of sectoral composition. Moreover, this variable is related to the size of the agglomeration and the importance of collective equipment, facilities, and information as well as to the size of the local (labour and final goods) market.

3.2.2. The innovation equation

The second equation in our extended model is an innovation production function which estimates the drivers of innovation output. Innovation output is measured first as the propensity to create a new product or a new process (*Innovation*). Since we deal with a binary variable, we estimate a Probit model of the following type:

$$[3]\ \Pr(Innovation_i = 1 \mid X = x_i) = \Phi(\alpha_R \hat{r} + x'_{2i}\beta_2 + \varepsilon_{2i})$$

where *Innovation* is the variable underlying the choice to innovate, x_2 is a set of explanatory variables, β_2 is the related vector of coefficients and ε_2 is the error term, distributed according to a standard normal distribution with zero mean and unit variance.

The set of explanatory variables includes absorptive capacity factors and knowledge inputs (Cohen and Levin, 1989). Among the former, we include firm market share, industry dummies, investment intensity (computed as the sum of 1998-2000 real investments in new machinery and equipment, in logs) and spatial agglomeration variables. Including spatial agglomeration in equation [3] allows us to test whether the effect of these variables on innovation output passes completely through the intensity of research or whether there are constant returns to agglomeration.

Knowledge inputs include previous innovation activity (given by a dummy equal to 1 if the firm created a new product or process in the period 1998-2000) and the predicted value of the research equation. This accounts for persistence of innovation activity and innovation inputs, i.e. R&D intensity which, in turn, is explained by variables measured three years earlier.

In addition, we define innovation output in terms of the probability of a new product or a new process being introduced in the three years 2001-2003,[8] with the aim of disentangling the role of local externalities on different types of innovation.

We estimate a bivariate Probit model in which we assume the presence of an unobserved propensity latent variable I^* proportional to the unobserved level of expected profits from each type of innovation (*j* = product, process).

This latent profit is given by:

$$[4a]\ I_j{}^* = \alpha_r \hat{r}_j + x_{3j}\beta_{3j} + \varepsilon_{3j}$$

which can be mapped to an observable binary discrete variable I_j indicating whether a firms introduces (or not) one of the two types of innovations:

[4b] $I_j = 1$ if $I_j^* > 0$ and $I_j = 0$ if $I_j^* \leq 0$ (j = product, process).

The choices to create a new product or a new process are jointly modelled as a system of correlated equations, i.e. a system of equations with correlated error terms,

[8] Unfortunately we do not have information on the number of new products or processes developed.

which are supposed to be jointly distributed according to a multivariate normal distribution with variance-covariance matrix Σ given by:

$$[4c]\ \Sigma = \begin{pmatrix} 1 & \rho_{12} \\ \rho_{12} & 1 \end{pmatrix}$$

in which ρ_{12} represents the correlation coefficient between the two error terms.[9]

3.2.3. The productivity equation

After estimating R&D and innovation, we define our productivity equation by estimating a Cobb-Douglas production function with physical and human capital, firm's network organization strategies, local externalities and innovation output as the main determinants:

$$[5]\ y_i = \alpha_I I_i^* + x_{4i}\beta_4 + \varepsilon_{4i}$$

where y_i is the productivity variable, measured by TFP estimated through the semi-parametric method provided by Levinsohn and Petrin (2003); x_{4i} is the set of explanatory variables, β_4 is the related vector of coefficients, and ε_4 is the error term on which we do not make any *a priori* assumptions.

Among x_{4i} we include: (i) firm size; (ii) physical capital per employee (as the net book value of physical capital adjusted for inflation); (iii) human capital (i.e. the share of managers and executives); (iv) firm organization, as given by two dummies capturing membership to a business group or a consortium in 1998-2000; and, again, (v) spatial agglomeration. Coefficient α_I now represents the marginal effect of innovation output on total factor productivity.

3.2.4. The export equation

Finally, we estimate the equation for the export behaviour of Italian manufacturing firms in year 2003. This is measured in three ways: (i) first, we simply estimate a Probit model for the propensity to export in 2001-2003; (ii) then, we estimate a fractional logit model for the share of 2003 export sales; (iiv) finally, we estimate an ordered Probit model for the number of macro-areas into which firms exported goods in 2003.

The propensity to export is estimates according to equation [6]:

$$[6]\ \Pr(export_i = 1 \mid X = x_i) = \Phi(\alpha_Y \hat{y}_i + x_{5i}\beta_5 + \varepsilon_{5i})$$

where x_{5i} is a set of explanatory variables, β_5 is the corresponding vector of coefficients and ε_5 is the error term, which we assume to have the usual statistical properties. The coefficient α_y reflects the impact of (predicted) productivity (TFP) on the decision to export in 2003.

[9] A correlation coefficient equal to zero implies the error terms are distributed according to the standard normal distribution, so that a univariate Probit specification can be used for both estimations. However, since our two dependent variables (product and process innovation) are significantly correlated (with a coefficient equal to -0.4), we face two non-zero off-diagonal elements of Σ, which means that we need a multivariate probit specification in order to account for correlations across the disturbances of the two latent equations, which, in turn, embodies unobserved characteristics for the same firm.

Relying on the same set of explanatory variables, we next focus on export intensity, computed as the proportion of export sales over total sales. Since this variable is represented by a fractional response bounded between 0 and 1 (including boundaries), we estimate a non-linear fractional Probit model (Papke and Wooldridge, 1996)[10]. In this way, the coefficient α_y represents the marginal impact of predicted productivity on the profitability of export.

Finally, we also provide a non-monetary measure of export intensity. Since export sales may reflect not only the real competitiveness of firms, but also global demand conditions (Malmberg *et al.*, 2000), we measure export intensity by counting the number of macro areas[11] that the firm has served through its exporting activity. In this respect, we define an ordered variable that equals 0 for purely domestic firms, 1, 2 and 3 for firms exporting in one, two and three macro regions respectively, and 4 for firms serving more than three areas.

Following the recent theoretical and empirical literature on the determinants of export performance (Bernard and Jensen, 1999, 2004; Ferragina and Quinitieri, 2001; Sterlacchini, 2001; Melitz, 2003), we include in our explanatory variables: (i) market share; (ii) industry dummies; (iii) age; (iv) membership to a business group or consortium as a measure of firm organization; (v) the propensity to invest in ICT during 1998-2000, and (vi) the predicted level of TFP. In this case, we estimate export intensity with an ordered Probit model in which we assume the error term to be normally distributed with zero mean and constant variance.

3.2.5. The full model

The full model is given by a set of five 'core' equations, two for R&D, one for innovation output, one for productivity and one for export.. We summarize these equations as follows.
Research:

[7] $r_i^* = x_{0i}\beta_0 + \varepsilon_{0i}$

[8] $r_i = x_{1i}\beta_1 + \varepsilon_{1i}$,

innovation:

[9] $\Pr(Innovation_i = 1 \mid X = x_i) = \Phi(\alpha_R \hat{r} + x'_{2i}\beta_2 + \varepsilon_{2i})$

productivity (TFP):

[10] Since the dependent variable is left and right censored to 0 and 1 respectively, we also use a Tobit model for estimating the export intensity equation. The estimated coefficients are very similar to the ones emerging from the fractional Probit specification, but the standard errors are higher and, then, less precise. Results are not reported here, but are available on request to Authors. Alternatively, for the estimation of proportions as dependent variables, one could rely on a beta distribution; however, although bounded between 0 and 1, this distribution does not include the two boundary values.

[11] These macro areas are: (1) EU-15 countries (Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Luxemburg, the Netherlands, United Kingdom, Spain and Sweden); (2) countries that joined the EU in 2004 (Cyprus, Estonia, Latvia, Malta, Poland, Cech Republic, Slovakia, Slovenia, Hungary); (3) Russia, Turkey and other European countries; (4) Africa; (5) Asia (excluding China); (6) China; (7) United States, Canada and Mexico; (8) Centre and South America; (9) Australia.

[10] $y_i = \alpha_I I_i * + x_{4i}\beta_4 + \varepsilon_{4i}$,

export:

[11] $\Pr(export_i = 1 \mid X = x_i) = \Phi(\alpha_Y \hat{y}_i + x_{5i}\beta_5 + \varepsilon_{5i})$

Moreover, we also add four equations, two for product and process innovation and two for export intensity (i.e. the share of export sales and the number of macro-regions of export), as robustness checks for our basic equations 9 and 11 respectively.

3.3. The empirical findings

Tables 7 to 10d report the main results of our analysis. Table 7 refers to the first two equations for R&D capital; Tables 8a and 8b refer to the innovation output functions; Table 9 refers to TFP estimations and Tables 10a-10d refer to exporting.

Table 7 shows that the intensity of R&D capital investments increases with both demand-pull and technology-push factors, i.e. with firm's market share and previous investment in R&D. R&D intensity is driven by firm experience and tax incentives but even more by the strong impact of spatial agglomeration: in particular, related variety and urbanization economies play a highly statistically significant and positive role in increasing the intensity of R&D, once the decision to invest in this activity is taken. Unrelated variety, in contrast, decreases R&D intensity.

The positive relationship between R&D and related variety can be viewed from two perspectives. On the one hand, following Frenken *et al.* (2007), R&D may benefit more from local knowledge spillovers from related sectors, i.e., from new ideas derived from different, but complementary, knowledge sources. On the other hand, unrelated variety may be associated with increased knowledge dispersion and it may decrease the expected returns to R&D, thus lowering levels of R&D investment.

The positive effect of urbanization economies may reflect the fact that large agglomerations, or large cities, may host a variety of activities linked to R&D like universities, research laboratories, advanced services, local credit, trade associations and other knowledge-based organizations. According to Frenken *et al.* (2007, p. 687), "the diverse industry mix in an urbanized locality also improves the opportunities to interact, copy, modify, and recombine ideas, practices and technologies across industries giving rise to Jacobs externalities".

Table 7 – Research capital

Variables	$lnR\&D/L_{2003}$ Coeff. (bootstrapped s.e.)	Selection = $R\&D_{2003}$ Coeff. (bootstrapped s.e.)
Age	0.3706*	-0.0148
	(0.1849)	(0.0700)
Market $share_{2000}$	0.4619**	0.2445***
	(0.1558)	(0.0562)
$lnR\&D_{2000}$	0.2745**	0.1903***
	(0.1040)	(0.0181)
$Export_{2000}$	0.7704	0.2659*
	(0.5713)	(0.1345)
Tax $incentives_{1998-2000}$	0.8051*	0.0114
	(0.3648)	(0.1366)
Specialization	0.1346	-0.0207
	(0.2198)	(0.0842)
Related Variety	3.1144**	0.4778
	(1.0385)	(0.3979)
Unrelated Variety	-1.2602**	-0.1234
	(0.4375)	(0.1926)
Urbanization	0.1544*	0.0359
	(0.0626)	(0.0287)
R&D $Consortium_{2000}$	…	4.7160***
		(0.2927)
N. Obs.	715	
Uncensored Obs.	353	
Mills Lambda	3.929***	
Wald χ^2 (21)	154.36 ($p < 0.000$)	

Notes: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. Estimations include 13 industry-specific dummies.

Table 8a shows that the impact of spatial agglomeration externalities is concentrated in the input phases of the innovation process. The propensity to innovate, in line with the Schumpeterian view of innovation, is affected mainly by R&D and firm size, i.e. market share.

When we investigate product and process innovation separately, we find that, while the former is driven more by past experience and R&D investments, and less by size, the latter depends much more on market share and investment in new equipment. This confirms the results in the literature on innovation in Italy (Parisi *et al.*, 2006), where process innovation is shown to favour large firms and is embodied in new machinery and capital goods.

In addition, we find a significant, but rather weak, role for unrelated variety: being located in a highly diversified productive structure increases the probability of a process innovation. If the introduction of a process innovation reflects capital-embodied technological change, then we can reasonably expect that different capital goods may come from firms belonging to different industries. Therefore, although on the one hand unrelated variety may hinder R&D investment, on the other, it may favour process innovation by increasing the pool of firms providing capital goods.

Table 8a – Innovation output

Variables	Innovation Probit	Innovation dy/dx	Product Bivariate probit	Process Bivariate probit
Age	0.0682	0.025	-0.0174	0.0789
	(0.0700)	(0.026)	(0.0708)	(0.0722)
Market share$_{2000}$	0.1212*	0.046*	0.0388	0.1054*
	(0.0514)	(0.019)	(0.0478)	(0.0474)
Investment$_{2000}$	0.0050	0.002	-0.0080	0.0217
	(0.0123)	(0.005)	(0.0129)	(0.0125)
Innovation$_{98\text{-}00}$	0.1394	0.053	0.2807*	0.1853
	(0.1110)	(0.042)	(0.1090)	(0.1086)
Specialization	0.0172	0.007	-0.0692	0.0404
	(0.0699)	(0.026)	(0.0683)	(0.0700)
Related Variety	-0.3056	-0.116	-0.3854	-0.4934
	(0.3284)	(0.124)	(0.3206)	(0.3263)
Unrelated Var.	0.1898	0.072	0.0374	0.3810*
	(0.1632)	(0.062)	(0.1604)	(0.1664)
Urbanization	-0.0018	-0.000	0.0160	-0.0118
	(0.0266)	(0.010)	(0.0257)	(0.0261)
R&D (predicted)	0.1789***	0.068***	0.1894***	0.1222***
	(0.0328)	(0.012)	(0.0316)	(0.0316)
N. Obs.	715	715	715	
PseudoR^2	0.1256			
Pearson χ^2 (693)	709 (p = 0.331)			
Log Likelihood			-848.65	
Rho			0.369***	

Notes: * p < 0.05, ** p < 0.01, *** p < 0.001. Cluster-robust standard errors in parentheses. Estimations include 13 industry-specific dummies. Pearson chi-squared statistics refers to Hosmer and Lemeshow goodness of fit test.

Table 8b – Marginal effects after bivariate probit estimation

Variables	Product Marginal	Product Joint	Process Marginal	Process Joint
Age	-0.007	-0.018	0.031	0.019
Market share$_{2000}$	0.015	-0.010	0.041*	0.015
Investment$_{2000}$	-0.003	-0.006	0.008*	0.006*
Innovation$_{98\text{-}00}$	0.110*	0.030	0.072	-0.009
Specialization	-0.027	-0.023	0.016	0.020
Related Variety	-0.151	0.002	-1.192	-0.038
Unrelated Var.	0.015	-0.060	0.148*	0.072*
Urbanization	0.006	0.006	-0.005	-0.005
R&D (predicted)	0.075***	0.021**	0.047***	-0.006

Another interesting result emerges from the productivity estimates (Table 9). While, as expected, innovation output contributes directly to TFP (as in Crépon *et al.*, 1998), this impact is statistically significant at the 5% level. Complementary to innovation, size and localization externalities strongly contribute to increasing productivity, while variety does not seem to play any relevant role. This result is in line with the literature on the sources of local productivity growth (Cingano and Schivardi, 2003); however, our structural model allows us to identify a positive role also for variety-based externalities, an effect that is not directly captured in reduced form models, since it is filtered by the impact of R&D on innovation, and innovation on productivity.

Table 9 also shows that the impact of innovation on TFP is due to product rather than process innovation. In this respect, we can speculate that, following results from TBLE 8a, related variety indirectly impacts more on productivity than unrelated variety.

Table 9 – Productivity

	(1)	(2)	(3)
Medium	0.3622	0.3731	0.5582**
	(0.2062)	(0.2031)	(0.2024)
Large	1.2787**	1.3460**	1.6135***
	(0.4162)	(0.4106)	(0.4042)
Group2000	0.0016	-0.0141	0.0455
	(0.2248)	(0.2242)	(0.2233)
Consortium2000	0.5565*	0.5643*	0.6031*
	(0.2616)	(0.2598)	(0.2672)
Human Capital2000	0.1161	0.1193	0.1035
	(0.1110)	(0.1111)	(0.1106)
Physical Capital2000	0.2631***	0.2936***	0.2949***
	(0.0791)	(0.0797)	(0.0796)
Specialization	0.3262***	0.3561***	0.3426***
	(0.0942)	(0.0939)	(0.0958)
Related Variety	-0.7293	-0.7856	-0.5309
	(0.4620)	(0.4612)	(0.4570)
Unrelated Variety	0.1303	0.2270	0.1196
	(0.2095)	(0.2120)	(0.2171)
Urbanization	-0.0517	-0.0572	-0.0462
	(0.0348)	(0.0350)	(0.0343)
Innovation	0.3666*	…	…
	(0.1570)		
Product innovation		1.1297**	…
		(0.3872)	
Process innovation			-0.2038
			(0.5516)
N. Obs.	715	715	715
R^2	0.111	0.1150	0.1049
F (11, 714)	6.1965	6.6708	6.0637

Notes: cluster-robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Finally, we test the role of TFP on export performance through a set of three specifications[12]. First, we focus on the probability to export in year 2003. Table 10a shows that younger, more productive and more "urbanized" firms are more willing to enter foreign markets. Therefore, in addition to TFP, we find that firms use other means to cope with the fixed costs associated with internationalization: the density of the urban

[12] As a robustness check of our empirical model we also reversed the causal relationship between export and productivity and we estimate a four-equation model of R&D, innovation output, export propensity and TFP. Due to the nature of our data and since we are using predicted values, it is not possible to run a weak exogeneity test on our export and TFP variables. However, a simple Hosmer and Lemeshow specification test on the propensity to export as a function of innovation output rejects the null hypothesis of good fit of the Probit model (Pearson $\chi^2 = 886.65$, *p*-value = 0.000), whereas in our original specification (see table 10a) the test does not reject it (*p*-value = 0.16). At the moment, we use this information as a preliminary proof for the correctness of our structural model specification.

structure. As before, the local availability of advanced services, qualified skills, infrastructures, universities, information sources and so on may help firms reduce the sunk costs associated with the sale of goods at the international level.

Table 10a – Propensity to export, 2003 (marginal effects at the mean)

Variables	(1)	(2)	(3)
Age	-0.057*	-0.054*	-0.040
	(0.027)	(0.027)	(0.027)
Market share	0.044	0.046	0.087***
	(0.026)	(0.025)	(0.024)
Group	-0.108	-0.108	-0.084
	(0.059)	(0.059)	(0.059)
Consortium	-0.000	0.000	0.079
	(0.073)	(0.072)	(0.066)
ICT	0.030	0.030	0.035
	(0.048)	(0.048)	(0.048)
Specialization	-0.016	-0.016	0.035
	(0.032)	(0.030)	(0.031)
Related Variety	0.222	0.233	0.116
	(0.121)	(0.121)	(0.119)
Unrelated Variety	-0.057	-0.062	-0.030
	(0.059)	(0.059)	(0.059)
Urbanization	0.031**	0.031**	0.024*
	(0.010)	(0.010)	(0.010)
TFP	0.151**	…	…
	(0.057)		
TFP_product	…	0.149**	…
		(0.052)	
TFP_process	…	…	0.006
			(0.051)
N. Obs.	715	715	715
Pseudo R^2	0.1541		
Log Likelihood	-397.3614	-396.7140	-400.8225
Wald χ^2 (22)	114.82 (p<0.000)		

Cluster-robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Estimations include 13 industry-specific dummies.

These results are confirmed when we measure export intensity through the share of export sales (Table 10b) and the number of macro-regions (Tables 10c and 10d). The higher the intensity of export, the higher the complementary role of TFP and urbanization economies.

Table 10b – Export intensity: share of export sales (marginal effects after fractional Probit)

Variables	(1)	(2)	(3)
Age	-0.026	-0.025	-0.019
	(0.014)	(0.014)	(0.020)
Market share	0.032*	0.032*	0.054***
	(0.015)	(0.015)	(0.014)
Group	-0.036	-0.035	-0.021
	(0.028)	(0.028)	(0.028)
Consortium	-0.003	-0.003	0.039
	(0.034)	(0.034)	(0.036)
ICT	0.014	0.014	0.017
	(0.028)	(0.028)	(0.028)
Specialization	-0.000	-0.000	0.024
	(0.018)	(0.018)	(0.018)
Related Variety	0.084	0.088	0.038
	(0.066)	(0.066)	(0.065)
Unrelated Variety	-0.034	-0.036	-0.020
	(0.034)	(0.034)	(0.034)
Urbanization	0.014**	0.014**	0.010*
	(0.005)	(0.005)	(0.005)
TFP	0.059°	…	…
	(0.044)		
TFP_product	…	0.058*	…
		(0.028)	
TFP_process	…	…	-0.012
			(0.028)
N. Obs.	715	715	715
Residual DF	692	692	692
AIC	0.915	0.915	0.917
BIC	-4254.377	-4254.642	-4252.971
PseudoLog Like	-304.17	- 304.038	-304.873

Cluster-robust standard errors in parentheses. ° $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Estimations include 13 industry-specific dummies.

Table 10c – Export intensity: number of macro-regions (ordered probit)

	(1)	(2)	(3)
Age	-0.0142	-0.0087	0.0139
	(0.0587)	(0.0582)	(0.0581)
Market share	0.2000***	0.2007***	0.2794***
	(0.0518)	(0.0497)	(0.0474)
Group	-0.1221	-0.1233	-0.0803
	(0.1195)	(0.1194)	(0.1205)
ICT	0.2027	0.2024	0.2135*
	(0.1065)	(0.1064)	(0.1063)
Specialization	-0.0254	-0.0265	0.0611
	(0.0686)	(0.0671)	(0.0664)
Related Variety	0.4975	0.5215	0.2999
	(0.2705)	(0.2710)	(0.2667)
Unrelated Variety	-0.2599*	-0.2704*	-0.2080
	(0.1319)	(0.1322)	(0.1321)
Urbanization	0.0804***	0.0801***	0.0685**
	(0.0222)	(0.0221)	(0.0222)
TFP	0.2630*	…	…
	(0.1092)		
TFP_product	…	0.2666**	…
		(0.1010)	
TFP_process	…	…	0.0113
			(0.1008)
N. Obs.	715	715	715
Log Likelihood	-1025.35	-1024.76	-1028.24
Pseudo R^2	0.0874	0.0879	0.0848
LR test	66.30 (*p*=0.364)	66.57 (*p*=0.355)	67.12 (*p*=0.338)

Cluster-robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. LR test refers to approximate likelihood-ratio test for equality of coefficients. Estimations include 13 industry-specific dummies.

Table 10d – Marginal effects of selected variables after ordered probit estimation

N. of countries	1	2	3	4
Market share	-0.015**	0.005*	0.018***	0.056***
ICT(*)	-0.013*	0.008	0.019	0.054*
Specialization	0.002	-0.001	-0.002	-0.070
Related Variety	-0.038	0.014	0.045	0.139
Unrelated Variety	0.020	-0.007	-0.002	-0.073*
Urbanization	-0.006*	0.002*	0.007**	0.022**
TFP	-0.020*	0.007*	0.024*	0.073*
TFP_product	-0.020*	0.007*	0.024*	0.074**
TFP_process	-0.001	0.000	0.001	0.003

(*) dy/dx is for discrete change from 0 to 1.

4. Conclusions

Using a large sample of Italian manufacturing firms we estimated a structural model of research, innovation, productivity and export performance augmented to take account of the role played by spatial agglomeration externalities. In so doing, we combine three main bodies of the recent literature: (i) works concerned with the empirical determinants of innovation; (ii) studies of the role played by agglomeration economies on productivity and (iii) research linking productivity, as the main source of firm heterogeneity, to firms' internationalization.

The model used in the paper is an 'augmented' version of the CDM model, developed to summarize the complex process "that goes from the firm decision to engage in research activities to the use of innovations in its production activities" (Crepon *et al.*, 1998, p.116).

It comprises four main equations: the first identifies the factors underlying the intensity of R&D investments; the second links R&D capital to innovation output; the third focuses on TFP as determined by innovation; the fourth relates export performance to TFP. To avoid problems of simultaneity, in each equation we measure our dependent variables in year 2003, while the explanatory variables refer to the period 1998-2000.

Unlike previous studies, we include in our analysis a set of variables measuring local knowledge spillovers, i.e. specialization, related and unrelated variety and urbanization economies.

Our estimates show that agglomeration economies do play a role in shaping the relationship between innovation, productivity and export performance. In particular, we find that related variety and urbanization economies promote R&D and the generation of new ideas, while specialization affects the exploitation of innovation in terms of higher levels of TFP. Finally, urbanization economies do positively affect both R&D and also the propensity to export and the relative export intensity.

The main contribution of this paper is that, unlike other studies, we do not rely on "reduced form" relationships, but estimate a structural model in which we are able to identify the role played by different forms of spatial agglomeration on innovation, productivity and export. In this sense, this approach can provide a better understanding of these relationships, opening the "black box" of the mechanisms underlying firms' competitive advantage.

References

Bagella, M., & L. Becchetti (2000), *The competitive advantage of industrial districts. Theoretical and empirical analysis*. Heidelberg, Germany: Physica-Verlag.

Banca Intesa (2008), *Monitor dei distretti*, Servizio Studi e Ricerche, Trimestrale n. 21, Milan.

Becattini, G. (1989), Sector and/or districts: Some remarks on the conceptual foundations of industrial economics. In *Small Firms and Industrial Districts in Italy*, ed. E. Goodman and J. Bamford, 123–135. London: Routledge.

Becchetti L. & Rossi S.P.S. (2000), The positive effect of industrial district on the export performance of Italian firms, *Review of Industrial Organization,* V.16: 53-68.

Bernard A.B. & Jensen J.B. (1995), Exporters, jobs and wages in US manufacturing: 1976-1987, *Brooking Papers on Economic Activity, Microeconomics*: 67-119.

Bernard A.B. & Jensen J.B: (1999), Exceptional export performance: cause, effect, or both?, *Journal of International Economics*, V. 47:1-25.

Bernard A.B. & Jensen J.B: (2004), Why Firms Export, *Review of Economics and Statistics*, V. 86: 561-569.

Best, M. (1990), *The new competition. Institutions of industrial restructuring*. Cambridge, MA: Harvard University Press.

Boschma, R.A., & J.G. Lambooy (2002), Knowledge, market structure, and economic coordination: Dynamics of industrial districts, *Growth and Change* 33(3): 291–311.

Boschma, R.A., & A. ter Wal (2007), Knowledge networks and innovative performance in an industrial district. The case of a footwear district in the South of Italy. *Industry and Innovation* 14(2): 177–199.

Boschma R. & Iammarino S. (2009), Related Variety, trade linkages and regional growth in Italy, *Economic Geography*, in press.

Brusco, S. (1982), The Emilian model: Productive decentralisation and social integration. *Cambridge Journal of Economics* 6: 167–184.

Brusco, S., G. Cainelli, F. Forni, M. Franchi, A. Malusardi, & R. Righetti (1996), The evolution of industrial districts in Emilia Romagna. In *Local and regional response to global pressure: The case of Italy and its industrial districts*, ed. F. Cossentino, F. Pyke, and W. Sengenberger, 17–36, Geneva, Switzerland: International Labour Office.

Cainelli G. & Leoncini R. (1999), Externalities and long-term industrial development. Some empirical evidence from Italy, *Revue d'Economie Industrielle*, V. 90: 25-40.

Cainelli G., Evangelista R. & Savona M. (2006), Innovation and economic performance in services. A firm-level analysis, *Cambridge Journal of Economics*, 30, 3, 435-458.

Castellani D. & Zanfei A. (2006), Internalisation, innovation and productivity. How do firms differ in Italy, *The World Economy*.

Castellani D., Serti F. & Tomasi C. (2009), Firms in international trade: importers and exporters heterogeneity in the Italian manufacturing industry, *mimeo*.

Combes P.P. (2000), Economic Structure and Local Growth: France, 1984–1993, *Journal of Urban Economics* , V. 47 (3), 329–355.

Chevassous-Lozza E.& Galliano D. (2008), Local spillovers, firm organization and export behaviour:evidence from the French food industry, *Regional Studies*, V. 37(2): 147-158.

Cingano F. & Schivardi F. (2003), The Sources of Local Productivity Growth, *Bank of Italy*, Working Paper N. 474.

Crépon B., Duguet E. & Mairesse J. (1998), Research, Innovation, and Productivity: An Econometric Analysis at the Firm Level, *Economics of Innovation and New Technology*, V. 7(2): 115-158.

Cunat, A., & Peri, G. (2001) Job creation: Geography, technology and infrastructures. *Giornale degli Economisti e Annali di Economia* 61 (1), 23–54.

De Lucio J. J., Herce J.A. & Goilcolea A., 2002. The effects of externalities on productivity growth in Spanish industry, *Regional Science and Urban Economics* 32 (2), 241–258.

Fabiani, S., G. Pellegrini, E. Romagnano, & L.F. Signorini (1999), Efficiency and localisation: The Italian case. In *Micro- and macrodata of firms, statistical analysis and international* *comparison* (Contributions to Statistics), ed. S. Biffignardi, 457–485. Heidelberg, Germany: Physica-Verlag.

Forni M. & Paba S., 2002. Spillover and the growth of local industries, *Journal of Industrial Economics* 50 (2), 151–171.

Frenken K., Van Oort F. & Verburg T. (2007), Related Variety, Unrelated Variety and Regional Economic Growth, *Regional Studies*, V. 41(5): 685-697.

Glaeser E.L., Kallal H., Scheinkman J. & Schleifer A. (1992), Cities and Growth, *Journal of Political Economy*, V. 100: 1126-1152.

Gordon, I.R., & P. McCann (2001), Industrial clusters: Complexes, agglomeration and/or social network. *Urban Studies* 37(3): 513–532.

Gordon, I.R., & P. McCann (2005), Clusters, innovation and regional development. *Journal of Economic Geography* 5(5): 523–543.

Griffith, R., Redding S. & Van Reenen J. (2004), Mapping the two faces of R&D: Productivity growth in a panel of OECD industries, *Review of Economics and Statistics*, 86(4), 883-895.

Hall, B. & J. Mairesse (1995), Exploring the relationship between R&D and productivity in French manufacturing firms. *Journal of Econometrics* 21: 89–119.

Harhoff, D. (1998), R&D and productivity in German manufacturing firms, *Economics of nnovation and New Technology* 6: 29–49.

Heckman J. (1976), The common structure of statistical models of truncation, sample selection, and limited dependent variables and a simple estimator for such models, *Annals of Economic and Social Measurement*, 5, 475-492.

Heckman J. (1979), Sample selection bias as a specification error, *Econometrica*, 47, 153-161.

Henderson, V. J., (2003), Marshall's scale economies, *Journal of Urban Economics* 53 (1), 1–28.

Henderson, V. J., Kunkoro, A., Turner, M., (1995), Industrial development in cities, *Journal of Political Economy* 103 (5), 1067–1090.

Iammarino, S., & P. McCann (2006), The structure and evolution of industrial clusters: Transactions, technology and knowledge spillovers. *Research Policy* 35(7): 1018–1036.

Langlois, R., & P.L. Robertson (1995) *Firms, markets and economic change. A dynamic theory of business institutions*. London: Routledge.

Levinsohn J. & Petrin A. (2003), Estimating Production Functions Using Inputs to Control for Unobservables, *Review of Economic Studies,* V. 70(2): 317-341.

Mameli F., Faggian P. & McCann, P., (2007), Agglomeration economies and employment growth in Italian local labour systems. In: Joint Congress of the European Regional Science Association and Association de Science Régionale de Langue Française. Paris.

Marshall, A., (1920), *Principles of Economics*, 8th Edition. MacMillan, London.

Martin P., Mayer T. & Mayneris F. (2008). Spatial concentration and firm-level productivity in France, Discussion Paper Series 6858, CEPR.

Melitz M. (2003), The impact of trade on intra-industry reallocations and aggregate industry production, *Econometrica*, V. 71: 1695-1725.

Paci R. & Usai S., 2006. Agglomeration economies and growth. the case of Italian local labour systems, 1991-2001. Working Paper 2006/12, CRENOS.

Parisi M.L., Schiantarelli F. & Sembenelli A. (2006), Productivity, innovation and R&D: micro evidence for Italy, *European Economic Review*, V. 50, 2037-2061.

Roberts M. J. & Tybout J. (1997), The decision to export in Colombia: An empirical model of entry with sunk costs, *American Economic Review*, 87(4), 545-564.

Signorini, L.F. (1994), The price of Prato, or measuring the industrial district effects, *Papers in Regional Science* 73: 369–392.

Usai, S., & Paci, R., (2003), Externalities and local economic growth in manufacturing industries. In: Fingleton, B. (Ed.), *European Regional Growth*. Springer-Verlag, Berlin.

Wagner J. (2005), Export and productivity: A survey of the evidence from firm level data, Working Paper Series in Economics 4, University of Luneburg, Institute of Economics.

Wakelin, K. (2001), Productivity growth and R&D expenditure in UK manufacturing firms, *Research Policy* 30: 1079–1090.

Wang, J.C., & K.H. Tsai (2003), *Productivity growth and R&D expenditure in Taiwan's manufacturing firms,* NBER Working Paper Series, n. 9724, NBER, Cambridge, MA.