Birkbeck Working Papers in Economics & Finance

ISSN 1745-8587

Birkbeck
UNIVERSITY OF LONDON

School of Economics, Mathematics and Statistics

BWPEF 0713

Research, Knowledge Spillovers and Innovation

Piergiuseppe Morone
Università degli Studi di Foggia

Carmelo Petraglia
Università degli Studi di Napoli "Federico II"

Giuseppina Testa
Birkbeck, University of London

July 2007

▪ Birkbeck, University of London ▪ Malet Street ▪ London ▪ WC1E 7HX ▪

Research, Knowledge Spillovers and Innovation

Evidence from the Italian Manufacturing Sector[*]

Piergiuseppe Morone [♦a], *Carmelo Petraglia* [b] *and Giuseppina Testa* [c]

This version 15.07.2007

Abstract: In order to assess the relationship between internal and external innovative inputs and innovative output at firm level, a knowledge production function is estimated for a representative sample of Italian manufacturing firms over the period 1998-2003. To account for endogeneity of R&D effort in the knowledge production function, we estimate a Heckman selection model on R&D decisions. Results support the view that R&D intensity is positively linked to firm size, age and human capital endowment as well as to higher exposure to international competitive pressure. Then, the knowledge production function is estimated using a standard probit, where the probability to innovate of each firm depends upon intramural R&D effort, regional and industrial spillovers and on a vector of interaction and control variables. Our measures of external knowledge, which circulates and potentially transfers across firms belonging to the same geographical or industrial spaces, are based on predicted values for R&D effort in the region and industry respectively. Our results suggest a positive relationship between sectoral spillovers and innovation; knowledge diffusion in the regional space positively impacts on the probability to innovate of the recipient firm only if the latter has an appropriate endowment of human capital.

JEL classification: O3, L6, C25
Keywords: Innovation, knowledge, spillovers

---

[*] Data provision from Capitalia and helpful advices from Attilio Pasetto (Area Studi Capitalia) are gratefully acknowledged; we also thank Rosanna Pittiglio for having kindly supplied the information drawn from the Aida database. We thank Ron Smith, Renata Narita and Marion Frenz for useful comments and suggestions. Any errors remain those of the authors.

[♦] Corresponding author: Piergiuseppe Morone (p.morone@gmail.com).
[a] Università degli Studi di Foggia, Largo Papa Giovanni Paolo II, 71100 – Foggia, Italy; email: p.morone@gmail.com
[b] Università degli Studi di Napoli "Federico II", Dipartimento di Economia, Via Cintia, Complesso M. S. Angelo – 80126 – Napoli – Italy; e-mail: carmelo.petraglia@unina.it
[c] School of Economics, Mathematics and Statistics, Birkbeck, University of London, Malet Street, London WC1E 7HX: gtesta@econ.bbk.ac.uk

1. Introduction

The ongoing process of globalisation, which brings into the same market firms located in distant parts of the planet, has posed firms of developed countries under a growing competitive pressure. Specifically, most manufacturing firms located in mature economies are threatened by the growing price competition coming from emerging economies like China and India where firms can rely on much cheaper labour costs.

In light of these facts, several scholars and researchers have maintained that the only viable way for firms in advanced economies to enhance their competitiveness would be to empower their innovative capabilities through investments in knowledge creation and diffusion. This knowledge-based approach is grounded in the idea that the ability to create and transfer knowledge is a crucial step in sustaining competitive advantage (Pinch et al., 2003; Forsman and Solitander, 2003). In other words, firms' long-term competitiveness is highly dependent on their ability to innovate and learn continuously (Florida, 1995; Cooke, 2001; Malmberge and Maskell, 2002).

In this paper we aim to investigate the determinants of innovation in the Italian manufacturing sector. Following the line of research first developed by Griliches (1957, 1979 and 1992), we shall use a "knowledge production function" approach, relating innovative inputs to innovative output. The main innovative input is intramural Research and Development (R&D). However, several studies show that R&D investments are not the only viable way to innovate. This is particularly true for small firms which, though unable to perform intramural R&D, can benefit from investments in research and development made by larger firms and universities which literally "spill-over" for economic exploitation by other firms (see, among others, Jaffe, 1986).

Moving within this framework, we shall attempt to explore the extent to which the existence of spillovers and intramural R&D may affect innovations and the extent to which recipient firms' human capital endowment might play a role in making more effective the diffusion of knowledge potentially available to the firm.

Our estimation strategy is closely related to that of Crépon et al. (1998) and Griffith et al. (2006) that take into account for endogeneity of intramural R&D expenditure. In examining the role of R&D on firms' product/process innovation, we investigate the firm's decision to engage in R&D activities and, for those who do, we estimate the intensity of the effort devoted to such activities. However, the main novelty of our investigation arises from the knowledge production function specification which explains innovation allowing for knowledge spillovers.

The paper is structured as follows. In the next section, we introduce the main motivation, providing a brief review of the literature on the determinants of R&D effort and the impact of knowledge spillovers on innovation. In section 3 we present the empirical model of knowledge accumulation and knowledge diffusion as well as a description of the database. Results are discussed and interpreted in section 4. Section 5 concludes. The appendixes to the paper contain information on data and technical estimation details.

2. Motivation and theoretical background

Most of the empirical works on innovation have focused on the impact of R&D activities and knowledge spillovers on firm performance – measured in terms of productivity growth – introducing some measures of knowledge capital as an additional input in the production function (for the Italian case, see Medda and Piga, 2004; Parisi et al., 2005). This part of the literature focusing on the direct effect of R&D effort on productivity suffers from a methodological criticism, that is, two different phenomena are summarized in a simplified way: the successful transformation of R&D activities in innovations and the impact of the latter on the production process.

Due to the uncertainty associated with the innovation process, not all R&D expenditures result in successful innovations. Moreover, if technological activities finally generate innovations, these can be product innovations, with a positive effect on demand, or process innovations, that allow productivity improvements and cost reductions. Therefore, knowledge capital entering the production function should account only for the part of R&D expenditure related to the generation of process innovations. This implies a measurement problem for those firms that simultaneously obtain product and process innovations, since the R&D expenditures are not divided into the portions associated with each type of innovation. Crépon et al. (1998) provide a solution to this methodological limitation introducing a structural multi-equation model that explains productivity by innovation output and the latter by research investment. Huergo and Jaumandreu (2003) look directly at the impact of process innovations on productivity growth.

As for the Italian case, several studies have addressed the issue of firms' productivity. For instance, Parisi et al. (2005) assume that innovation output (not innovation input) directly affects productivity growth in a single regression framework. Less attention, however, has been paid to identifying the relevant factors affecting innovation, i.e. studying the sources of knowledge which are relevant to the transformation process of innovative inputs – either internal or external to the firm – into innovative output. Recognizing this gap in the literature, our main objective is the estimation of a knowledge production function for Italian manufacturing firms.

In light of the Schumpeterian idea on firms selection process based on successful transformation of R&D investment into new product/process innovations (Schumpeter, 1972), we make an attempt to look into the black box of the Italian firm's knowledge production process. The achievement of higher productivity ultimately results from higher competitiveness standard which in turn reflects enhanced innovative capabilities. Hence, we focus on the innovation production process estimating a knowledge production function at firm level (Griliches, 1979) where innovative inputs account for both internal R&D effort and relevant external sources of knowledge which are endogenously determined.

In our attempt to contribute to this literature, we exploit the results of the empirical research recently achieved in the field of structural models of R&D and innovative output (Crépon et al., 1998), which explored the determinants of the firm's decision and R&D activities and the determinants of its innovation success.

In particular, in assessing the impact of knowledge spillovers on the firm's likelihood to innovate, we maintain that knowledge accumulated by other firms can be exploited by means of spillovers within two different – and sometimes overlapping – "spaces" bordered by either geographical or industrial proximity.

Firms within a geographical unit can benefit from network externality due to knowledge exchanges transmitted through informal channels of communication and embodied in human capital which circulates among firms (Nelson and Winter, 1982). Furthermore, firms located in the same geographical unit can benefit from technological activities of institutions outside the industrial system (namely, local research institutions).

Industrial proximity also plays a role: a firm can benefit from knowledge embodied in R&D investment by another firm belonging to the same industry, via a costless diffusion process from the latter to the former.

Such a twofold perspective is related to our analysis through two extensively explored questions: what are the key factors explaining R&D effort and what are the relevant innovative inputs which enter the knowledge production function. The answers to these questions are – to some extent – related, as factors influencing current effort in shaping innovative inputs will be the ones that will make firms more likely to be innovative in the future. In other terms, future innovative output of a firm depends on its accumulated innovative inputs whose availability depends upon its past decision to engage in R&D effort.

At this stage it is worth noting that, since the seminal contribution by Schumpeter (1942), the link between firm size and innovation has received a great deal of attention. Empirical evidence has established a positive correlation between firm size and its commitment to formal R&D activities. Nevertheless, small firms often contribute significantly to innovative output either in particular sectors or geographical areas (for Italy, see Piergiovanni et al. 1997). A commonly agreed explanation for this evidence is that small firms with low commitment to R&D effort acquire knowledge through co-operation with larger firms or with research institutions such as universities (Acs, Audretsch and Feldman, 1994). Such evidence is particularly relevant in the Italian economy, where the average size of manufacturing firms is relatively small compared with other European countries. The methodology we shall implement allows distinguishing between R&D internal effort determinants and the following innovation performance (determined by both internal and external sources of knowledge). Hence, we expect to be able to provide new evidence on whether or not small firms can innovate although they invest relatively little in internal R&D compared with larger firms.

3. Empirical strategy and data

3.1 Empirical model on knowledge accumulation and diffusion

Inspired by Crépon et al. (1998) and Griffith et al. (2006), our estimations are based on a three-equation specification. The first equation, given in (1), describes the firm's decision whether to undertake or not R&D activities. A binary response model is thus used as follows:

$$h_i^* = z_i'\gamma + v_i \qquad (1)$$

$$h_i = 1 \text{ if } h_i^* > 0$$

where h_i^* represents firm *i*'s propensity to undertake R&D activities, z_i is a vector of the firm's exogenous characteristics and v_i is an unobserved term. The main firm's characteristics that enter its decision whether to engage in R&D activities are size, location, age, sector and export orientation.

We then investigate the effects of the firm's specific characteristics on the amount of intramural R&D expenditure reported by firm *i* on the basis of its previous decision to undertake R&D activities. The specification for this model is:

$$r_i^* = x_i'\beta + u_i \qquad (2)$$

$$r_i^* = r_i \text{ if } h_i = 1$$

where r_i^* is the unobservable underlying process (r_i^* is observed only for firms undertaking R&D activities) and r_i is the data observation. Note here that the variables in x_i are all included in z_i, i.e. $x_i \subset z_i$, since variables in z_i with coefficients γ affect the choice of undertaking R&D activities without affecting directly the effort of such activities. The system of equations (1)-(2) is referred to as a standard generalised tobit model which takes into account the sample selection problem.[1] In this system, the error terms, v_i and u_i are i.i.d. across firms and $(u_{i,} v_i) \approx \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_u^2 & \sigma_{uv} \\ \sigma_{uv} & \sigma_v^2 \end{pmatrix} \right)$.

The econometric method estimates the following equation derived from our system (1)-(2):

$$E\left(r_i \middle| h_i = 1\right) = x_i'\beta + \frac{\sigma_{uv}}{\sigma_v^2} \frac{\phi(z_i'\gamma)}{\Phi(z_i'\gamma)} \qquad (3)$$

As a final step, a binary response model is used to examine the relationship between firm's intramural R&D expenditure, knowledge spillovers and its propensity to

[1] See appendix B for econometric details.

innovate.[2] We estimate firm's propensity to innovate using a standard probit model which is defined as follows:

$$I_i^* = r_i'\beta_1 + w_i'^{xy}\beta_2 + C_i'\beta_3 + \kappa_i'\beta_4 + \varepsilon_i \qquad (4)$$

$$I_i = 1 \text{ if } I_i^* > 0$$

where I_i^* represents firm's *i* propensity to report innovation product and/or process innovation; r_i is the internal knowledge obtained from the estimation of (2); w_i^{xy} is a vector of both industry-specific knowledge spilled-over from other firms operating in the same sector *x* and geographical-specific knowledge spilled-over from other firms located in the same region *y*; κ_i is a measure of physical capital, and finally C_i is a vector of control variables which capture heterogeneity across firms.

In the knowledge production function (4), we assume that firms do not consider spillover benefits at the time of their R&D decision, but that they are able to take advantage of such externalities when innovating.

In (4), we note that there are at least two main sources of endogeneity. First, intramural R&D expenditure (r_i) is likely to be correlated with unobservable terms (ε_i). This correlation may occur because firms that expect to be able to innovate are those that might be more likely to engage in R&D. Second, r_i is likely to be measured with error. This measurement error problem may occur since it is possible that firms which did not report R&D investment, did, in fact, undertake some R&D activities. In our study, both problems are dealt with by using the predicted values of R&D from (2) in (4).

3.2 Measuring R&D Spillovers

If a firm's own accumulated knowledge is defined in a rather simple way and is typically captured by the internal investments in R&D, the problem of providing a

[2] Note here that our estimation strategy is based on the belief that an appropriate specification of any functional relationship relating innovative inputs to innovative outputs measures should carefully consider the outcome of the previous decision faced by firms to engage in R&D activities and, for those who do, on the intensity of the effort devoted to such activities. Hence, we assume that firms first decide whether and how much to invest in internal R&D and then – given a time lag – "produce" their innovative output.

measure of any possible external source of knowledge "available" to the firm to be used in its innovation process does not have an easy solution. In this paper, we assume that innovative inputs used by other firms can, in principle, be exploited by firm *i* by means of spillovers occurring within the two different – and sometimes overlapping – "spaces". Thus, in our analysis, using information on the location of firms along these two spaces, we construct measures of geographical and industrial knowledge spillovers as follows:

$$w_i^s = \sum_{j \neq i} \hat{r}_j^s \qquad (5)$$

where w_i^s denotes total knowledge available in space $s \in [x, y]$, defined as the sum of R&D predicted effort $\hat{r}_j^s$ of any other firm *j* in space *s*. All in all, our analysis takes into account three different measures of knowledge used in the innovating process of firm *i*:

(a) $\hat{r}_i$ - firm *i*'s intramural knowledge coming from own investments in R&D;

(b) w_i^x - industry-specific knowledge spilled-over from knowledge accumulated by other firms operating in *x*;

(c) w_i^y - geographical-specific knowledge spilled-over from knowledge accumulated by other firms located in *y*.

In (b), we are assuming that knowledge is "useful" – to the same extent to insider firms – only within the industry ("industry-specific" knowledge).[3] Further, (c) implies a similar assumption: limitations (and opportunities) to competitiveness due to location are specific to the locality and, as a consequence, so is the information needed to overcome such limitations (geographical-specific knowledge).[4] Note that we neglect intersectoral knowledge flows across regions, following Autant-Bernard and Massard (2004), who show that sectoral diversity is a source of knowledge flows within a geographic unit, whereas sectoral proximity is a good way to take advantage of knowledge generated in distant places. For us, benefiting from knowledge flows calls for proximity in at least one of these two dimensions, whereas intersectoral flows across regions cumulate two hurdles: the geographical and the industrial ones.

[3] Note that we use the 12 sectors provided in our sample as a framework for calculating the sectoral spillover variable.

[4] Note that we use the 19 Italian regions of our sample (Valle D'Aosta and Piemonte are counted as one) as a spatial framework for calculating the regional spillover variable.

3.3 Data

We use a balanced panel of Italian manufacturing firms observed over the period 1998-2003. This panel is constructed by merging two subsequent waves of the Capitalia survey.[5] Using the panel structure of the data we assume that R&D activities occur in the first observed time period, whereas the transformation process of internal and external sources of knowledge into innovative output occurs in the second one.

Table 1. *R&D activities and intensity*

	Firms with intramural R&D, 1998-2000		R&D intensity in 2000	
	Num.	%	Mean	Standard deviation
Sectors				
Low-technology sectors	226	49.89	250907	967365
Science-based sectors	227	50.11	623095	2337893
Location				
North	332	73.29	538349	2071211
South	121	26.71	157604	461559
Firm size (employees)				
Up to 50	241	53.20	73457	94001
More than 50	212	46.80	842132	2555528
Export orientation				
No	60	13.30	118511	399234
Yes	391	86.70	487391	1923504
Num.	451			

Table 1 reports descriptive statistics on intramural R&D activities in 1998-2000 and R&D intensity in 2000 disaggregated by sector, location, size and export orientation. The share of firms engaged in intramural R&D is similar across science-based and low-technology sectors. However, as expected, on average firms operating in science-based sectors spend on R&D more than firms operating in low-technology sectors. As for firm's location, a large proportion of firms engaged in R&D operate in the northern regions of Italy. Similarly, R&D expenditure is, on average, higher in firms located in the northern than in the southern part of Italy.

Table 1 also shows that 53 percent of R&D performers are firms of small-medium size while 47 percent are medium-large sized; further, small-medium sized firms show,

[5] For details on the Capitalia survey and on the merging methodology used to construct the balanced panel, see Appendix A.

on average, a smaller commitment to R&D investment than medium-large sized ones. Indeed, in line with the Schumpeterian hypothesis, R&D investment is positively correlated with size.

Finally, we notice that both the share of firms engaged in R&D activities and the amount of R&D investment are higher in industries which are export-oriented. This finding probably reflects the way international competition is perceived by Italian firms (firms might decide to undertake R&D in order to be internationally competitive).

Table 2. *Firms' innovative behaviour*

	Innovative behaviour (2001-2003)	
	Yes (%)	No (%)
Sectors		
Low-technology sectors	44.33	55.7
Science-based sectors	63.92	36.07
Location		
North	53.01	46.98
South	45.79	54.21
Firm size (num. of employees)		
Up to 50	44.76	55.24
More than 50	65.48	34.52
Firm export orientation		
No	34.74	65.24
Yes	57.35	42.45

Table 2 shows the share of firms that have innovated over the period 2001-2003 with respect to our control variables.[6] It is unsurprising to note that the largest innovators are those located in the northern regions of Italy (more than 53 percent), those operating in the science-based sectors (53 percent) and the export-oriented ones (almost 58 percent).

4. Results

In this section we first report the main results on R&D activities and, if firms decide to carry out R&D, the determinants of their investment. We then turn our

[6] Innovative activities data is based on firms' answer to the question "Have new product and/or process innovations been introduced over the period 2001-2003?"

attention to the knowledge production function, where we investigate the determinants of innovation.

4.1 R&D choice and intensity

Table 3 reports the results on the decision whether or not to invest in R&D activities (lower section) and on the R&D intensity determinants (upper section). Looking at the estimated coefficients of R&D choice equation, we can infer that there is a positive relationship between our dependent variable and human capital, implying that increasing the amount of human capital of the firm (i.e. the number of employees with a university degree) increases the probability of being engaged in R&D activities. Specifically, the marginal effect estimates show that each additional employee with a university degree increases, on average, the probability of being involved in R&D by 1.3 percentage points.

Additionally, there is a positive relationship among the dependent variable and size, export orientation and science-based firms and a negative, but not significant, relationship with those firms located in the south of Italy. This implies that export-oriented larger firms which operate in science-based sectors (according to Pavitt's taxonomy) are more likely to undertake R&D activities.

Ceteris paribus, exporting firms are over 20 percentage points more likely to perform R&D than firms producing solely for the national market. This result is consistent with the existing literature which envisages a strong causality between firms' export attitude and R&D investments.

Specifically, exporting "makes firms more easily aware of potential innovations taking place abroad and they may assimilate these in order to improve their position both in domestic and foreign markets" (Barrios et al., 2003: 476). However, in order to take full advantage of these learning opportunities "firms must first acquire the appropriate knowledge and technological capability, for example, through a firm's own Research and Development (R&D) activities" (2003: 476). In fact, as pointed out by several authors (Cohen and Levinthal, 1989; Lucas, 1993; Hewitt and Wield, 1992; Mody, 1993; and Audretsch, 1995), "R&D not only generates new information, but also enhances the firm's ability to assimilate and exploit existing information. Therefore, the productivity effect of knowledge gained through export experience may depend on the firm's own investment in R&D or worker training" (Aw et al., 2007: 86)

Table 3: *R&D choice and amount equations (estimation technique: Heckman)*

Dependent variable: R&D intensity (referred to year 2000)	Coefficients		Marginal Effects	
	Estimates	P >| Z |	Estimates	P >| Z |
Human capital	0.013	0.001	0.013	0.001
South (1 if the firm is located in the south of Italy)	-0.410	0.055	-0.410	0.055
Size (1 if the firm has more than 50 employees)	0.672	0.002	0.672	0.002
Rho	-0.969	0.000		
Wald test of indep. eqns	16.29	0.0001		
Number of obs	553			
Censored obs	280			
Uncensored obs	273			

Dependent variable: R&D activities (dummy variable referred to the period 1998-2000)	Coefficients		Marginal Effects	
	Estimates	P >| Z |	Estimates	P >| Z |
Human capital	0.035	0.053	0.013	0.056
South (1 if the firm is located in the south of Italy)	-0.042	0.796	-0.016	0.769
Size (1 if the firm has more than 50 employees)	0.396	0.008	0.156	0.007
Export orientation (1 if the firm is involved in exporting activities)	0.556	0.045	0.207	0.000
Age of the firm	-0.007	0.001	-0.002	0.044
Science-based sector (1 if the firm operates in the high-technology and specialised sectors)	0.692	0.000	0.269	0.000

This argument seems to suggest a causal relationship between R&D and exports – i.e., the higher is the internal effort in R&D, the higher the propensity to export will be. However, as pointed out by Smith et al. (2002) "investment in R&D may itself be […] the result of an internationalised firm having a relatively large part of its turnover coming from export. In fact, it might be requisite for R&D activities to have a large market in order to make the investment in R&D pay off". This observation leads the authors to maintain that investments in R&D resulting in an export strategy for the firm might, in turn, boost the firm to further invest in R&D hence creating a benign circle of export and R&D. All in all, the literature provides strong evidence in favour of a reciprocal relationship between R&D and export.

As mentioned above, R&D is positively related also with firm size. Specifically, large firms (i.e. those with more than 50 employees) are, on average, more than 15 percentage points more likely to be engaged in R&D activities. As maintained by Cohen and Klepper (1996), it is commonly agreed that the likelihood of performing R&D rises with firm size. This pattern (and several others relating R&D and firms size) “can be explained by a simple idea advanced long ago as a possible advantage of large firms in R&D, namely R&D cost spreading”.

Finally, our findings show that firms operating in science-based sectors (e.g. chemical and electronic) are, ceteris paribus, almost 27 percentage points more likely to perform R&D. Also this last finding is in line with theoretical results; in this regard it is worth recalling that according to Pavitt’s taxonomy, the main source of technology for science-based firms is internal R&D.

Moving on to the R&D intensity equation (upper section), we assume that R&D intensity is dependent on a sub-sample of the variables used in the R&D choice equation. Specifically, we maintain that once having decided whether or not to be engaged in R&D activities, the intensity of the firm’s investment will depend on its size, on a geographical dummy (which accounts for the dualistic nature of the Italian economy) and on the amount of human capital (measured as the number of employees with an higher education degree) it possesses.

As is shown in table 3 (upper section), all our independent variables are statistically and economically significant. This implies that R&D intensity is positively dependent on the amount of human capital present in the firm as well as on the size of the firm itself. Further, being located in the south has a negative effect upon the decision how much to invest in R&D. Finally, we note that the dependent variable is observed for 273 firms and is missing for the remaining 280 who do not report R&D. We also note that the p-value attached to the Rho estimate, which captures the correlation between the error terms of the R&D choice and amount equations, suggests that there is evidence of a selection bias, supporting, therefore, the methodology used in this step, while the Wald test indicates that all explanatory variables are statistically different from zero.

4.2 The knowledge production function

The estimated knowledge production function includes – among other explanatory variables – the predicted values for R&D obtained from the regression discussed above.[7]

As we can immediately observe in table 4, most of our explanatory variables are statistically significant and correctly signed. Specifically, we can observe that an increase of one unit in the logarithm of estimated R&D effort exerted in the year 2000 increases the probability of innovating over the period 2001-2003 by more than 35 percentage points. Moreover, sectoral knowledge spillovers affect the probability of innovating by almost 7 percentage points.

Both findings are in line with our expectation and the theoretical background discussed in section 2. However, when it comes to regional knowledge spillovers, we can observe a statistically significant negative relationship between this independent variable and the probability of innovating. This finding is rather interesting as it is at odds with our expectations. A possible explanation is that a knowledge competition effect prevails over the knowledge flow effect. In other words, since firms located in the same region draw the supply of skilled labour force from the same pool of workers, then a high level of accumulated knowledge by competing firms might imply that not much knowledge is left for our representative firm. This might have a negative effect on the probability to innovate. This hypothesis is corroborated by the fact that the sign of regional knowledge spillover changes when we consider the interaction between such knowledge spillover and our human capital variable. In the latter case the relationship is positive and significant. This implies that the knowledge present in the region can be better exploited by those firms which have managed to reach a good level of human capital. Note that human capital and the interaction between sectoral spillovers and internal human capital are not statistically significant.

[7] Note here that we have also compared the results of this probit regression with those obtained from a probit regression not correcting for endogeneity problem. The main difference in the results is that the coefficients of the actual values of R&D, knowledge spillovers and their interactions with human capital are lower than those in the model with predicted values, indicating that the measurement error occurring when employing the actual values instead of the predicted ones leads to an underestimation of the effect of these variables on firms' probability to innovate. It is also worth noting that the number of observations is different between the two models. The explanation is that the predicted values would reduce the presence of missing observations, accounting for firms which would not appear to be R&D performers.

Table 4. *Knowledge Production Function (estimation technique: probit)*

Dependent variable: Innovation (dummy variable referred to the period 2001-2003)	Probit coefficients		Marginal Effects	
	Estimates	P >| Z |	Estimates	P >| Z |
Log R&D (hat)	0.953	0.039	0.355	0.040
Log sectoral spillovers (hat)	0.184	0.100	0.069	0.099
Log regional spillovers (hat)	-0.220	0.020	-0.082	0.020
Human capital	-0.124	0.172	-0.046	0.170
HK- sectoral spillovers interaction	-0.005	0.735	-0.002	0.735
HK- regional spillovers interaction	0.028	0.013	0.010	0.011
Log of Physical Capital	0.119	0.008	0.044	0.009
Science-based sector (1 if the firm operates in the high-technology and specialised sectors)	0.219	0.135	0.269	0.130
Size (1 if the firm has more than 50 employees)	-0.603	0.083	-0.231	0.083
Constant	-12.005	0.017		
Number of obs	533			
Pseudo R2	0.068			

Other variables affecting the probability to innovate are the logarithm of physical capital and firm size. Specifically, we can observe that one unit's increase in the logarithm of physical capital exerted over the period 1998-2000 increases the probability to innovate over the period 2001-2003 by one percentage point.

Finally, we observe another unpredicted result when we measure the impact of a firm's size on its probability to innovate. In fact, we can observe a negative and statistically significant correlation, which suggests that the likelihood of innovating is higher for smaller firms.[8] Note that many studies found a positive relationship between innovation and firm size, although there are also claims to support the counter argument. For instance, Holmstrom (1989) argues that larger firms are at a comparative disadvantage in innovating, because of the cost associated with handling efficiently a heterogeneous set of tasks. In this perspective, smaller firms might enjoy a comparative advantage which, in turn, might lead to a higher propensity to innovate.

Note that, at a first glance, the fact that in our sample small firms are more innovative seems to be in contrast with what was observed in the first step of our investigation, where we found a positive relationship between firm size and the probability to be engaged in R&D as well as with its intensity. However, if we look more carefully at these findings, we can see that they are not necessarily in contrast. In

[8] Recall that we defined small firms as those with less than 50 employees. However, our sample has a lower truncation at 10 employees, implying that micro firms are not included in the sample.

fact, it is possible that larger firms have an advantage in acquiring knowledge through internal investments in R&D (due to the usual arguments that larger firms have more resources to invest in R&D), but at the same time it can be the case that small firms are more innovative as they can access knowledge from different sources (i.e. knowledge spillovers).[9]

5. Conclusions

We have estimated a knowledge production function for a panel of Italian manufacturing firms obtained by merging two contiguous waves of Capitalia surveys. The panel structure of this newly constructed dataset has allowed us to account for the sequence in time of two related issues: the determinants of innovative inputs and the transformation of innovative inputs into innovative output.

As for the relevance of knowledge diffusion across firms within industrial and geographical spaces, a first novelty of our approach with respect to the existing literature on the Italian case is the estimation of a knowledge production function instead of the focus on the impact of knowledge spillovers on productivity. Second, and more importantly, we control for a selection bias in evaluating firms' R&D internal effort and the relevant sources of knowledge spillovers. Third, we measure external knowledge which circulates and potentially flows across firms within the same geographical or industrial spaces in such a way to account for the previous decision of firms on R&D activities and intensity.

We conclude that the probability to innovate is positively linked to sectoral spillovers. On the other hand, knowledge diffusion in the regional space positively affects the probability of the recipient firm to innovate only if it has an appropriate endowment of human capital, which we interpret as a relevant means of absorption of external knowledge.

[9] Indeed, further investigation is required here in order to properly understand this relation. Such an examination, however, goes beyond the scope of our paper.

Reference

Audretsch, D. B. (1995), “Innovation and Industry Evolution”. Cambridge: MIT Press.

Acs Z.J., Audretsch D.B. and M.P. Feldman (1994), “R&D Spillovers and Recipient Firm Size”, *The Review of Economics and Statistics*, Vol. 76(2): 336-340.

Autant-Bernard C. and N. Massard (2003), “Innovation and Local Externalities: Evidence and Ambiguities drawn from the Geography of Innovation”, *Working Paper CREUSET.*

Axelrod R. (1984), “The Evolution of Cooperation”. New York: Basic Books.

Aw B.J., Roberts M.J. and T. Winston (2007), “Export Market Participation, Investments in R&D and Worker Training, and the Evolution of Firm Productivity”, *The World Economy*, 30(1): 83-104.

Barrios S., Goerg, H. and E. Strobl (2003), “Explaining Firms’ Export Behaviour: R&D Spillovers and the Destination Market”, *Oxford Bulletin of Economics and Statistics*, 65(4): 475-496.

Burns T. and G. M. Stalker (1961), “The Management of Innovation”. London: Tavistock.

Capitalia (2002), “Indagine sulle Imprese Manifatturiere. Ottavo Rapporto sull’Industria Italiana e sulla Politica Industriale”, Rome.

Capitalia (2005), “Indagine sulle Imprese Manifatturiere. Nono Rapporto sull’Industria Italiana e sulla Politica Industriale”, Rome.

Cohen W.M. and S. Keppler (1996), “A Reprise of Size and R&D”, *Economic Journal*, 106: 925-51.

Cohen W.M. and D.A. Levinthal (1989), “Innovation and Learning: The Two Faces of R&D”, *Economic Journal*, 99: 569-96.

Cooke, P. (2001), “Knowledge Economies: Clusters, Learning & Co-Operative Advantage”. London: Routledge.

Crépon B., Duguet E. and J. Mairesse (1998), “Research, Innovation and Productivity: An Econometric Analysis at a Firm level”, *Economics of Innovation and New Technology*, 7(2): 115-58.

Dosi, G. and L. Marengo (1994), “Some Elements of an Evolutionary Theory of Organisational Competences”, in: R. England (Ed.) “Evolutionary Concepts in Contemporary Economics”. Ann Arbor: University of Michigan Press.

Florida R.L. (1995), “Toward the Learning Region”, *Futures*, 27(5): 527-36.

Forsman M. and N. Solitander (2003), “Knowledge Transfer in Clusters and Networks”, *Journal of International Business Studies – Literature Review*, n.3, (available online at: http://copenhagen.jibs.net/LitReview/2003/2003_3_24.pdf)

Griliches Z. (1957), “Hybrid Corn: An Exploration in the Economics of Technical Change”, *Econometrica*, 25:501–522.

Griliches Z. (1979), “Issues in Assessing the Contribution of Research and Development to Productivity Growth”, *The Bell Journal of Economics*, 10: 92-116.

Griliches Z. (1992), “The Search for R&D Spillovers”, *Scandinavian Journal of Economics*, 94: S29-S47.

Griffith R., Huergo E., Mairesse J. and B. Peters (2006), “Innovation and Productivity across Four European Countries”, NBER Working Paper 12722.

Hewitt T. and D. Wield (1992), “Technology and Industrialization”, in Hewitt T., Johnson H. and D. Wield (eds.), “Industrialization and Development”. Oxford: Oxford University Press.

Holmstrom B., 1989, “Agency Costs and Innovation”, *Journal of Economic Behavior and Organization*, 12, 305-27.

Huergo E. and J. Jaumandreu (2004), “Firms’ Age, Process Innovation and Productivity Growth”, *International Journal of Industrial Organization*, 22, 541-559.

Jaffe A. (1986) “Technological Opportunity and Spillovers of R&D: Evidence from Firms Patents, Profits, and Market Value”, *American Economic Review*, 76: 984-1001.

Lucas R. (1993), “Making of a Miracle”, *Econometrica*, 61, 251–72.

Lundvall B.A. (1988), “Innovation as a Interactive Process: from User-Producer Interaction to the National System of Innovation”, in Dosi G., Freeman C., Nelson R., Silverberg G. and L. Soete (Eds.), “Technical Change and Economic Theory”. London: Pinter Publishers: 349-369.

Malmberg A. (1997), “Industrial Geography: Location and Learning”, *Progress in Human Geography*, 21(4):573-82.

Malmberg A. and P. Maskell (2002), “The Elusive Concept of Localisation Economies: Towards a Knowledge-based Theory of Spatial Clustering’, *Environment & Planning A*, 34:429-449.

Medda G. and C. Piga, (2004), “R&D e spillover industriali: un’analisi sulle imprese italiane”, Working paper CRENoS 2004/06, Centre for North South Economic Research, University of Cagliari and Sassari, Sardinia.

Mody A. (1993), “Learning through alliances”, *Journal of Economic Behavior and Organization*, 20: 151-170

Morgan K. (2001), “The Exaggerated Death of Geography: Localised Learning, Innovation and Uneven Development”, *Paper presented at The Future of Innovation Studies Conference*, Eindhoven University of Technology, 20-23 September 2001.

Nelson R.R. and S.G. Winter (1982), “An Evolutionary Theory of Economic Change”. Cambridge, MA: Belnap Press.

Oerlemans L., Meeus M. and F. Boekeman (1998), “Innovation: Some Empirical Explorations of Spatial Embeddednes”, in van Dijk J. and F.Boekema (eds.), “Innovation in Firms and Regions” (in Dutch). Assen: Van Gorcum: 31-61.

Parisi M. L., Schiantarelli F. and A. Sembenelli (2005), “Productivity, Innovation and R&D: Micro Evidence from Italy”, *European Economic Review* 50(8): 2037-2061.

Piergiovanni R., Santarelli E. and M. Vivarelli (1997), “From Which Source Do Small Firms Derive Their Innovative Inputs? Some Evidence from Italian Industry”, *Review of Industrial Organization* 12(2): 243-258.

Pinch S., Henry N., Jenkins M. and S. Tallman (2003), “From “Industrial Districts” to “Knowledge Clusters”: A Model of Knowledge Dissemination and Competitive Advantage in Industrial Agglomeration”, *Journal of Economic Geography*, 3:373-388.

Schumpeter, J. A. (1972), “History of Economic Analysis”. London: Allen & Unwin.

Schumpeter, J.A. (1934), “The Theory of Economic Development”. London: Oxford University Press.

Schumpeter, J.A. (1942), “Capitalism, Socialism and Democracy”. New York: Harper.

Smith V., Madsen E. and M. Dilling-Hansen, (2002) “Do R&D Investments Affect Export Performance” The Danish Institute for Studies in Research and Research Policy, WP 2002/4.

Appendix A – Construction of panel data, cleaning procedure and definition of variables

Construction of panel data

The 8th and 9th Capitalia surveys cover the periods 1998-2000 and 2001-2003 respectively. The firms included in the surveys were selected by means of a mixed procedure: sample-based for firms with between 11 and 500 employees, and exhaustive for firms with more than 500 employees. The composition of the sample was determined using of a random selection procedure stratified by class of employees, location and sectors.

Given this panel data, we proceeded to evaluate the difference in firms' sectors and size between the balanced panel data and the 8th and 9th waves of Capitalia survey, in order to evaluate if the sectoral and dimensional composition of the initial samples has been respected.

Table A1.

	Panel data 1998-2003 (N=1019)	Capitalia 1998-2000 (N=4289)	Capitalia 2001-2003 (N=4289)
Size	%	%	%
11-20 employees	34.00	39.90	22.10
21-50 employees	37.60	37.10	29.60
51-250 employees	21.80	16.20	26.90
251-500 employees	3.30	3.90	5.10
>500 employees	3.20	2.90	6.10
Location	%	%	%
North West	37.39	37.60	35.90
North East	31.50	27.40	30.10
Center	18.80	20.60	17.70
South	12.27	14.40	16.30
Sectors	%	%	%
Traditional sector	51.20	52.30	51.90
Scale sector	16.80	18.10	16.80
Specialised sector	27.70	24.30	26.70
High-tech sector	4.00	5.30	4.60

From table A1, we can notice that the share of firms of our panel is, on average, in line with the one observed in the two Capitalia samples. However, we should mention that the share of firms in the balanced panel data appears to be slightly underestimated in some cases and slightly overestimated in other cases. More precisely, we can observe that our panel, when compared to both Capitalia survey waves, slightly overestimates the share of firms with 21 to 50 employees and underestimates the share of firms with 251 to 500 employees. Similarly, our sample overestimates the share of firms located in the north-east and underestimates the share of those located in the south. Finally, firms in traditional sectors are slightly underestimated, whereas those in specialised sectors are overestimated.

All in all, we believe the results reported in the table A1 provide a first confirmation of the reliability of our sample.

Cleaning procedure

Our data cleaning procedure consisted of several different stages. First, to refine the firm's constitution year variable, which contains several missing values, we compared the information from the Capitalia questionnaire with information gathered from an independent data source (AIDA database). In doing so, we substituted all missing and erratic observations with AIDA information and, in the case of inconsistency, proceeded to report the oldest year of firm's foundation. The second step was converting into euros the R&D expenditure and the physical capital investment recorded in Italian liras back in 1998.

All mentioned variables were also reported to constant prices by using value added industry output deflators of Southern and Northern areas of Italy (the source of deflator is SVIMEZ).

However, the presence of several missing values in most of the relevant variables obliged us to perform our study on a restricted number of observations.

Definition of variables

R&D intensity: the amount of total expenditure on research and development (R&D) activity reported by the firms deflated by the output price.

Industrial spillover: the sum of R&D expenditure of other firms located in the same industry (sector) as the representative firm minus the R&D stock of the representative firm.

Geographical spillover: the sum of R&D expenditure of other firms located in the same region minus the R&D expenditure of the representative firm.

Size (1/0): 1 if the firm has more than 50 employees.

Age of the firm: the year of the firm's foundation.

Science-based sectors (1/0): 1 if firms operate in high-tech and specialised sectors and 0 otherwise.

Export orientation: 1 if the firm is involved in export activities.

Appendix B - Methodological issues -

Econometric details

Data on our key variable – internal R&D expenditure – is present only for a clearly defined subset of firms, i.e. R&D efforts are not observable for all firms but only for firms that are actually engaged in intramural R&D expenditure. To take into account this selection bias, we estimate the generalised tobit model (3) derived from equations (1)-(2) reported in section 3. The conditional expected R&D expenditure amount, given that a firm does undertake R&D activities, is as follows:

$$E(r_i | h_i = 1) = x_i'\beta + E(u_i | h_i = 1)$$
$$E(r_i | h_i = 1) = x_i'\beta + E(u_i | v_i \succ -z_i'\gamma)$$
$$E(r_i | h_i = 1) = x_i'\beta + \frac{\sigma_{uv}}{\sigma_v^2} E(v_i | v_i \succ -z_i'\gamma) \quad (3)$$

where the indicator variable h_i allows controlling for firm's choice.

Without loss of generality, as equation (1) is a standard probit model, σ_v^2 is equal to 1. Thus, equation (3) becomes:

$$E(r_i | h_i = 1) = x_i'\beta + \sigma_{uv} \frac{\phi(z_i'\gamma)}{\Phi(z_i'\gamma)} \quad (4)$$

where β cannot be estimated consistently by OLS unless the error terms of the two equations (1)-(2) are uncorrelated, i.e. $\sigma_{uv} = 0$.

Once having obtained the β's estimates by running equation (3), we compute the predicted estimates of R&D effort, $\hat{r}_i$. These latter estimates allow us to estimate the effect of intramural R&D expenditure on propensity to innovate for all firms, as they account for firms which did not appear to be R&D performers.

Issues on the choice of z and x

As noted in the Heckman results shown in table 3 in the paper, export orientation, age and science-based sectors, which appear in the R&D choice equation, are all excluded in the R&D

intensity equation. In this view, several caveats on the choice of variables for *z* and *x* are worth mentioning. First, if *x* contains variables incorrectly excluded from *z*, the coefficient relative to the mills ratio, i.e. the covariance between the error terms, may appear significant when it should not be so. Second, if z contains variables that are incorrectly excluded from *x*, the β's estimate could appear significant when it should not be so. As there are no real solutions to these problems, we follow the economic theory's suggestions reported in section 2 for the choice of our explanatory variables.