

IZA DP No. 7671

**Succeeding in Innovation:
Key Insights on the Role of R&D and Technological
Acquisition Drawn from Company Data**

Andrea Conte
Marco Vivarelli

October 2013

Succeeding in Innovation: Key Insights on the Role of R&D and Technological Acquisition Drawn from Company Data

Andrea Conte

JRC, European Commission

Marco Vivarelli

*Università Cattolica, Milano,
SPRU-Sussex University and IZA*

Discussion Paper No. 7671
October 2013

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Succeeding in Innovation: Key Insights on the Role of R&D and Technological Acquisition Drawn from Company Data^{*}

This paper discusses the relationship between a company's investment in innovation and its success in introducing new product and/or process innovations. In doing so, this analysis departs from the standard approach which puts forward a homogenous R&D-based knowledge production function by introducing different types of innovation investments (R&D and technology acquisition) for different sets of companies. Using the Community Innovation Survey (CIS) dataset comprising more than 3000 Italian manufacturing companies, the econometric analysis adopts a set of techniques which allows to control for the sample selection, endogeneity and simultaneity problems which arise when dealing with CIS data. The main findings are summarised as follows: (1) beyond the acknowledged effect of R&D in increasing the probability of success of product innovation, a larger-than-expected role is played by technology acquisition in the innovation process; (2) the relative importance of R&D and technology acquisition varies significantly across different types of companies where crucial dimensions of analysis are company size and the technological domain of a sector.

JEL Classification: O31

Keywords: R&D, product innovation, process innovation, embodied technical change, sample selection, SUR, community innovation survey

Corresponding author:

Andrea Conte
Edificio Expo
c/ Inca Garcilaso 3
41092 Sevilla
Spain
E-mail: andrea.conte@ec.europa.eu

^{*} The authors would like to thank Giovanni Seri and the ADELE Laboratory at ISTAT in Rome for the provision of CIS data. Previous versions of this paper have been presented at CNR-Rome, University of Turin, Max Planck Institute of Economics, Jena. The authors would like to thank participants at these seminars and two anonymous referees for very useful comments.

1 Introduction

Innovation is not randomly distributed among companies, but is rather the outcome of successful strategies, investment, and relationships which companies actively seek and develop in the market. Using a sample of innovative companies, this paper discusses the determinants of innovation at company level by looking at R&D activities and the acquisition of technology (TA), whether embodied (machinery and equipment) or disembodied. R&D refers to "*all creative work undertaken on a systematic basis in order to increase the stock of knowledge and the use of this stock to devise new applications, such as new and improved products (goods/services) and processes (including software research)*". Technology acquisition indicates the purchase of embodied technology - in the form of "*advanced machinery and computer hardware*" - and disembodied knowledge - in the form of "*purchase of rights to use patents and non-patented inventions, licenses, know-how, trademarks, softwares and other types of knowledge from others for use in the enterprise*" (Oslo, 2005).

The "Knowledge Production Function" (KPF) (Griliches, 1979) represents the most adopted tool for describing the transformation process that runs from innovative inputs to innovative outputs. While most previous research has focused on the R&D-Innovation-Productivity chain (see next section), few studies have explicitly discussed the role of TA and possible differences in the estimation of such augmented KPF across two fundamental dimensions: sectors and company size.

By using microdata from the third Italian Community Innovation Survey (CIS 3), a first novelty of this paper lies in checking whether the distribution of companies' innovative expenditure between these two inputs - R&D and TA - leads to significant differences in the composition and magnitude of their innovative outputs measured by the introduction of product innovation, process innovation and the share of sales due to innovative products.

A second novelty lies in the econometric strategies adopted to account for the peculiar nature of each dependent variable, possible problems of endogeneity, simultaneity and sample selection biases. Indeed, the nature of the CIS database requires two (sequential) controls for sample selection. First, we take into account the divide between innovative and non-innovative companies. Indeed, a seemingly unrelated regression (SUR) is used to explain the introduction of product and process innovation in the sub-sample of innovative companies, taking into account both this source of sample selection and the possible correlation between the two decisions. Second, we control for a subsequent source of sample selection that occurs when we investigate the role of R&D and TA in determining the commercial success of product innovations - an issue relevant only for a smaller sample of companies. Moreover, R&D and TA are instrumented controlling for companies' characteristics, thus solving the issue of endogeneity which commonly arises in this analytical framework (Crepon *et al.*, 1998). Finally, the possibility that different KPF relationships emerge in firms of different sizes and sectors is investigated, the basic hypothesis being that innovative small companies in traditional sectors (mainly characterised by process innovation) should rely more on TA, while their larger counterparts in high-tech sectors (mainly characterised by product innovation) should rely more on formal R&D, once it has been controlled for a set of company-specific characteristics.

The remainder of the paper is organised as follows: a discussion of the theoretical framework on which this work is based (Section 2) is followed by the description of data and indicators used in the empirical analysis (Section 3). Subsequently, the empirical outcomes derived from the descriptive analysis (Section 4) and the econometric estimates (Sections 5 and 6) are discussed. Section 7 concludes the paper by briefly summarising the main findings obtained and discussing possible avenues for further research.

2 The literature

Previous literature adopted R&D and patents as the starting point for the analysis of innovative activities across countries, industries and companies. In line with the growth-accounting conceptual framework, an input-output relationship emerged in the innovation literature as a feasible analytical tool to measure returns to innovation and its contribution to growth. In this stream of literature, Griliches (1979 and 1990) put forward explicitly the innovative input-output relationship, through a three-equation model where one of the equations is called Knowledge Production Function (KPF), a function intended to represent the transformation process leading from innovative inputs (R&D) to innovative outputs (patents). The other two equations in Griliches' simultaneous model represent the production function (augmented by the innovation term) and the determinants of R&D investment (Hall and Mairesse, 1995; Mairesse and Mohen, 2002; Harhoff *et al.*, 2003). Similarly, the KPF is also included in the four-equation models provided by Crepon *et al.* (1998) and Loof and Hesmati (2001).

Historically driven by a relatively higher data availability with respect to other measures of innovation, the relationship between a company's R&D investment and patenting activity has been at the core of much empirical research so far. Nowadays, this approach is complemented by a more comprehensive approach to the determinants of a company's innovativeness by means of an extended set of input and output measures of innovation. For instance, innovation surveys permit a better identification of the presence of any kind of innovative activity, and a distinction between product and process innovation. Moreover, different innovation outputs - including measures of the commercial success of innovations - can be seen as the outcomes of several innovation inputs and not only as the consequence of R&D investments¹.

Available literature suggests that more complex product innovation generally relies on formal R&D (Parisi *et al.*, 2006). On the other hand, process innovation, for which it is not so easy to distinguish between pure innovation, diffusion and imitation, is much more related to technological acquisitions (TA), through both the "embodied technical change" - acquired by investment in new machinery and equipment - and the purchasing of external technology incorporated in licences, consultancies and know-how (Freeman, 1982). Moreover, the embodied nature of technological progress and the effects related to its spread in the economy were originally underlined by Salter (1960; see also Jorgenson, 1966); since then, vintage capital models have been put forward to describe how the replacement of old equipment is the main way through which companies introduce process innovation (Hulten, 1992; Greenwood *et al.*, 1997; Mukoyama, 2006).

For the particular purpose of this paper, most of the previous empirical studies suffer from two main limitations. First, the relationship between innovation inputs and innovation outputs is not their main focus but rather a secondary equation, ancillary to the authors' main purpose of proposing an extended version of a production function where capital and labour are augmented by the innovation term. Second, the KPF is simplified as a link between R&D and patents (and/or product innovation). Exactly these two elements characterise, for instance, the approach adopted by the OECD (2009) for its multi-country innovation assessment.

An additional point worth investigating is the bi-directional relationship which occur when the decision to perform R&D and TA is taken simultaneously. Different results emerge in the empirical literature. For instance, Lach and Schankerman (1989) suggest a positive two-way causality effect of investing in R&D and capital among firms in science-based industries. A similar result is found by Chiao (2001) which, however, stresses that the granger causality between capital investment and R&D is bi-directional only in the short run. On the contrary, De Jong (2007) shows that R&D and

¹This broader perspective is well represented by the shift from the R&D-focused Frascati Manual ("Guidelines for the collection of R&D data", first published in 1963) to the Oslo Manual in the 1990s (OECD, 1st edn. 1997).

capital investments are co-integrated and the causality runs in both directions only in the long-run. Finally, Lach and Rob (1996) find that R&D granger-causes capital investment while the opposite effect does not appear supported by empirical evidence.

As a follow-up of the review made so far, this paper represents an attempt to open up a broader perspective by taking into account the simultaneous occurrence of a number of different innovative inputs and outputs. In doing so, a *first hypothesis* to test in the paper is whether the dichotomy R&D \Rightarrow product innovation vs TA \Rightarrow process innovation will be supported by our empirical analysis. In addition, we pave the way to the analysis of company and sector peculiarities in the KPF.

The evergreen debate on the role of size and sectors in affecting innovation performance has grown over the years by including a long list of important contributions². For the purpose of this paper, it is important to remind that most of the analysis has dealt empirically and theoretically with the differences in R&D and patenting behaviours across these groups of firms while, at the same time, neglecting other possible innovative behaviours which could provide innovative returns not directly accountable to R&D investments.

The idea that a company's size affects innovative activities dates back to Schumpeter (1942). Several arguments sustain the view that larger companies are in a better position to invest in R&D activities. First, larger companies are not affected by liquidity constraints since they have both easier access to external finance and larger internal funds to support expensive R&D projects; second, large corporations are characterised by a higher degree of diversification that helps them to deal with the uncertainty of R&D investment; finally, scale and scope economies make larger companies keener to embark on large R&D projects (Cohen and Klepper, 1996; Mairesse and Mohen, 2002). However, not all innovative companies are large corporations. Indeed, the economic literature supports the hypothesis that small companies face a different technological and economic environment from large companies with respect to innovative activities (Winter, 1984; Acs *et al.*, 1994; Brouwer and Kleinknecht, 1996). In particular, small companies tend to rely less on R&D (Kleinknecht, 1987), while being in a better position to incorporate both embodied and disembodied external knowledge (Audretsch and Vivarelli, 1996). In turn, this should lead innovative small companies to favour innovation adoption through TA, rather than by promoting new products based on large-scale R&D investment where the disadvantages discussed above are much more obvious (see Pavitt, 1984; Dhawan, 2001). While the financial and competitive reasons discussed above can hamper an R&D-based innovative strategy for small companies, it seems much more effective for them to rely on the market and choose 'to buy' instead of 'to make' technology (Acs and Audretsch, 1990). This argument leads to the *second hypothesis* to be tested in this study, namely that innovation performance in small companies relies more on TA than on formal R&D.

Industries are characterised by different patterns in the selection process and interaction among companies³. Indeed, companies face sector-specific technological opportunities and appropriabilities which respectively 'push' and 'pull' their innovative activity (Nelson and Winter, 1982; Winter, 1984; Dosi, 1988). Consequently, companies adapt their innovative strategy to their own particular economic environment by choosing the most effective combination of innovative inputs and outputs. Hence, the *third hypothesis* tested in this work is that in traditional sectors characterised by low technological opportunities, cost-cutting process innovation and a relevant presence of small companies, one would expect TA⁴ to have a dominant role, while in high-tech sectors one would expect formal R&D to have a leading role.

²For an extensive review of the evolution of the literature on technological change, see Conte (2006).

³See theories on the so-called "Schumpeterian patterns of innovation" and "sectoral systems of innovation" (Breschi *et al.*, 2000; Malerba *et al.*, 2008).

⁴Santarelli and Sterlacchini (1990) pointed out the important role of embodied technological change in small Italian firms operating in traditional and mature sectors such as textiles, wood and furniture, shoes, etc.

3 Dataset and Indicators

This paper uses company-level data drawn from the third Italian Community Innovation Survey (CIS 3). This survey, conducted by ISTAT (Italian Institute of Statistics) in 2002, supplies data on companies' features and their innovative inputs and outcomes. The entire sample amounts to 15,512 companies, of which 9,034 (58.24%) in the manufacturing sector, and is representative of the entire Italian population of companies with more than 10 employees⁵.

This paper focuses on a sub-sample of companies obtained through the following procedure. First, the exclusion of companies not belonging to the manufacturing sector in order to minimise possible discrepancies due to a different interpretation of i.e. what may constitute product / process innovation in manufacturing vs. services. Second, the exclusion of companies that were either newborn or had recorded an output variation of at least 10% due to mergers and acquisitions (M&A) since they may break the link between innovative inputs and outputs (a link that must be studied within the borders of a single firm).

At this point, the dataset included 8,219 innovative and non-innovative companies. An important characteristic of the CIS questionnaire is that requires companies to declare their innovative inputs – such as R&D and TA – and to provide other information on their innovative activities *only* if they have introduced innovation outputs (process and/or product innovation), or started innovative projects (then abandoned or still-to-be-completed). As a consequence, information on innovative investments is recorded only for a sample of 3,138 innovative companies, here defined as the sub-sample of companies having invested in innovative activities (regardless of whether they had already generated an innovative output or had instead been abandoned or were still-to-be-completed).

Such a data structure affects our empirical analysis in two respects. First, the overall sample of 8,219 manufacturing companies is used to identify the company-specific non-technological characteristics which make a company innovative⁶. Second, this analysis implies the adoption of a Heckman correction procedure to account for the possible sample selection arising when estimates are run using the sub-sample of innovative companies.

3.1 Innovative Outputs

Innovative outputs can be distinguished with respect to both their qualitative nature (i.e. product vs process innovation) and their relative position along the innovation process. For instance, patents are defined as the outcome of inventive efforts while product innovations represent the result of market-oriented innovative activity. Moreover, even though product innovation is driven by demand considerations, it represents a pre-market result. An ex-post result is given by the share of sales deriving from innovative products since this accounts for the market reaction to new products introduced by a company (Barlet *et al.*, 1998, Mairesse and Mohnen, 2002).

This paper uses three output indicators for the empirical analysis: two dichotomic variables - the introduction of product and process innovation - and the share of sales derived from innovative products. While the first two indicators can be modelled to explain the qualitative composition of companies' innovative activity, the third one will be used to assess the intensity of innovation.

⁵Statistical weights are used to make the sample representative of the underlying population according to three dimensions: size, sector and (macro) regions by correcting for different response rates across groups. Statistical checks introduced in the survey (such as sampling weights, treatment of missing values etc.) follow the guidelines of the OSLO manual (1st edn. 1997). On average, the response rate to this survey was 62% - among the highest in Europe.

⁶Indeed, CIS data structure does not permit the investigation of the role of R&D and TA in making a firm innovative (for a study which is based on a comparison between innovative and non-innovative Italian firms - using a different database - see Parisi *et al.* 2006).

3.2 Innovative Inputs

Two innovative input measures are used in this paper: formal Research and Development (R&D) and Technology Acquisition (TA). The former includes both internally-performed R&D (intra muros) and R&D activities outsourced to other companies or research institutes (extra muros R&D). The latter is the sum of embodied technological change (innovative investment in advanced machinery and computer hardware) and the acquisition of external (disembodied) technology in the form of "*purchase of rights to use patents and non-patented inventions, licenses, know-how, trademarks, softwares and other types of knowledge from others for use in the enterprise*" (Oslo Manual, 2005)⁷.

Table 1 briefly summarises the innovative variables adopted for this study.

Table 1: List and definitions of the innovative variables

Definition of Innovative companies	
INNOVATIVE	Dummy = 1 if the company has invested in innovative activities (1998-2000) and has realised a product or process innovation or has undertaken an innovative project (later abandoned or still to be completed as at December 31st, 2000).
Innovative input variables	
R&D	Logarithm of total R&D expenditures in 2000 ('000 of Euros).
TA	Logarithm of total expenditures in Technological Acquisitions in 2000 ('000 of Euros).
Innovative output variables	
Product innovation	Dummy=1 if company has realised product innovations.
Process innovation	Dummy=1 if company has realised process innovations.
Innovative Sales	Share of a company's total sales related to new products

3.3 Control Variables

The CIS questionnaire provides further information on companies beyond their innovative activity and thus it allows to control for the role played by company-specific characteristics in affecting both the input and output side of the innovative process we are interested in. This information, presented in Tables A1 and A2 in the Appendix, can be grouped into two sub-categories: 'generic company characteristics' and "innovation-relevant information".

The former is collected in the first section of the questionnaire over the full sample of 8,219 manufacturing companies, since its observability is not conditional on having performed some kind of innovative activity (i.e. companies' size, group and sectoral belonging, exports). The latter covers all that information which is directly related to the innovative process (i.e. cooperative technological agreements) and its observability is thus restricted to the innovative sub-sample only.

Both categories of variables should be taken into account when explaining the input-output relationships the innovative activities are based on. For example, Mairesse and Mohen (2002) underline the expected innovative benefits due to the effect of intra-group knowledge spillovers for companies that are members of business groups. Similarly, several papers discuss the importance of cooperation agreements – either between companies or with universities and public labs - for the promotion of innovative activity (Jaffe, 1989; Cassiman and Veugelers, 2002, Piga and Vivarelli, 2004).

⁷In principle, current and lagged values of R&D expenditures should be included in the regressions. Unfortunately, this is not possible due to the cross-section nature of CIS data. Both R&D and TA data are flow variables in the CIS questionnaire and as such they are used in the empirical analysis. Crepon *et al.* (1998) underline the high cross-sectional correlation between stock and flow measures of innovative activity.

4 Descriptive Statistics

Preliminary evidence suggests obvious differences in the input-output relationships, whether we consider product or process innovation (see Table 2). R&D intensity appears to be five times higher for companies that have introduced only product innovation compared to those companies with only innovative processes. On the contrary, TA investment is 3.5 times higher for the latter compared to the former category of companies.

Table 2: Average values of R&D and TA intensity (expenditure / turnover) for the sub-samples identified by Product and Process innovation

	Tentative (N=(160))	Product only (N=677)	Process only (N=676)	Product & Process (N=1625)	Tot (N=3138)
R&D Intensity	1.18	1.96	0.40	1.29	1.17
TA Intensity	1.92	1.99	7.02	4.09	4.38

Further proof of the difference in innovative behaviour across sample groups is provided by Table 3, which gives the average values (normalised by turnover) of the innovative inputs, and the percentages of innovative outputs by sector and size. The analysis is focused on two sectoral groups (low-tech and high-tech) and two size groups (small and large, where ‘small’ means companies with less than 50 employees - see Table A1).

Table 3: Average values of R&D and TA intensity and percentage of companies introducing product and process innovations by size and sector

R&D Intensity	Size			TA Intensity	Size		
Sector	Small	Large	TOTAL	Sector	Small	Large	TOTAL
Low-Tech	0.68	0.83	0.71	Low-Tech	5.83	2.70	5.26
High-Tech	2.41	1.92	2.27	High-Tech	2.45	1.81	2.27
Total	1.14	1.26		Total	4.93	2.35	
Product	Size			Process	Size		
Sector	Small	Large	TOTAL	Sector	Small	Large	TOTAL
Low-Tech	60.13	69.69	61.86	Low-Tech	81.17	77.22	80.45
High-Tech	78.11	87.16	80.68	High-Tech	64.73	66.61	65.26
Total	64.93	76.59		Total	76.78	73.03	

As expected on the basis of the discussion in Section 2, R&D and TA show contrasting behaviour according to sector, R&D being significantly higher in the high-tech sectors while the opposite is true for TA. A similar dichotomy emerges on the output side of innovation, with small, low-tech companies being characterised by the highest share of process innovation (and the lowest of product innovation) while it is to the large, high-tech sub-sample that the highest number of product innovators belong.

On the whole, all four innovative variables (both on the input and output side) appear to differ greatly between the four sub-samples identified by size and sector, with the only exception being process innovation and R&D intensity, which does not seem to be significantly affected by size (see the mean comparison tests in Table 4).

Table 4: Mean comparison tests of the innovative variables over the four sub-samples identified by Size and Sector

Mean Comparison Test ^φ					
	Size (S-L)	Sector (LT-HT)		Size (S-L)	Sector (LT-HT)
R&D Intensity	0.94 (0.349)	-12.27*** (0.000)	Product	-5.36*** (0.000)	-13.43*** (0.000)
TA Intensity	9.06*** (0.000)	8.80*** (0.000)	Process	0.42 (0.673)	7.39*** (0.000)
P-values in brackets * significant at 10%, ** significant at 5%, *** significant at 1%					
^φ Two-sample t test of $H_o : \mu_1 = \mu_2$ where the two groups are defined by size (S and L) and by sector (HT and LT).					

To sum up, descriptive statistics show the innovative process to be something more than a unique, homogeneous input-output chain, the nature of this relationship being strongly related not only to the specific (product or process) outcome targeted by the innovative expenditure, but also to companies' size and sectoral belonging (see Table 5 for the sample distribution along these two dimensions). In particular, a kind of dichotomy seems to emerge between a product-innovation model mainly driven by R&D and followed by large, high-tech companies, and a process innovation model, mainly related to TA and realised by small, low-tech companies. These preliminary results, which appear to be consistent with the hypotheses suggested in Section 2, will be fully investigated in the following sections taking into account different sources of sample selection and controlling for company's characteristics other than just size and sector.

Table 5: Sample distribution over size and technological level

		Company Size		TOTAL
		Small	Large	
Sector	Low-Tech	926 (29%)	872 (28%)	1798 (57%)
	High-Tech	612 (20%)	728 (23%)	1340 (43%)
TOTAL		1538 (49%)	1600 (51%)	3,138
In parentheses: shares over the total innovative sample				

5 The Empirical Strategy

The structure of the CIS questionnaire and the research questions discussed in this paper lead to the adoption of an empirical strategy based on three different steps - each of them focused on addressing specific econometric issues such as sample selection, simultaneity as well as the issue of endogeneity of some of the adopted regressors. All econometric estimates adopt CIS sampling weights in order to correct for biases in response rate and, thus, they provide representative results for the underlying population of companies.

5.1 Sample selection, endogenous regressors and simultaneity (Step 1)

Equations (1) represents the starting point of our empirical analysis. The structure of the CIS questionnaire identifies a group of innovative companies out of the total sample of companies on the basis of two main criteria: the introduction of product and/or process innovation. Only for this subset of companies, innovation expenditures are reported. In order to take into account the sample selection associated with the exclusion of non-innovative companies, a preliminary step consists therefore in the computation of two inverse Mills ratios (IMR-product and IMR-process) based on the following specification:

$$(1) \quad \begin{cases} PROD_i^* = \alpha_1 + \gamma PROC + \beta_{1i} X_i' + \varepsilon_{1i} & PROD_i = 1 \text{ if } PROD_i^* > 0, 0 \text{ otherwise} \\ PROC_i^* = \alpha_2 + \varphi PROD + \beta_{2i} X_i' + \varepsilon_{2i} & PROC_i = 1 \text{ if } PROC_i^* > 0, 0 \text{ otherwise} \end{cases}$$

where $PROD_i^*$ and $PROC_i^*$ are the latent variables underlying the two binary outcomes we observe and also influencing each other; X_i' summarises a set of exogenous explanatory factors other than innovative expenditures, including companies' characteristics and sectoral specificities.

The coexistence of two different targets in innovative activity (product and process innovation) and the likely possibility that the two decisions are taken simultaneously (see Rosenkranz, 2003) have important implications in terms of econometric strategy. Barlet *et al.* (1998) underline the positive correlation between innovative outputs, and a similar finding applies to CIS data. Indeed, table 6 shows the distribution of companies with respect to the introduction of product and/or process innovation. As can be seen, 80% of companies in the total sample (8219 firms) either do not innovate or introduce both types of innovation; similarly, more than 50% of the innovative sample (3138 firms) turn out to be involved in both product and process innovation.

Table 6. Distribution of companies according to product and process innovation

Total sample of companies				Innovative companies			
	Process				Process		
Product	0	1	Total	Product	0	1	Total
0	4808 (58.5%)	804 (9.8%)	5612	0	160 (5.1%)	676 (21.5%)	836 (26.6%)
1	843 (10.2%)	1764 (21.5%)	2607	1	677 (21.6%)	1625 (51.8%)	2302 (73.4%)
Total	5651 (68.7%)	2568 (31.2%)	8219	Total	837 (26.7%)	2301 (73.3%)	3138

Estimating the simultaneous decision to introduce a process and/or product innovation requires a seemingly unrelated regression (SUR) approach with correlated disturbances. In a similar vein, we

compute two inverse Mills ratios (IMR-R&D and IMR-TA) to take into account the sample selection associated with the decision (probit-type) to introduce R&D and/or TA expenditures (equation 2). These IMRs will be used in the next step of the analysis to instrument the original R&D and TA values.

$$(2) \quad \begin{cases} R\&D_i^* = \alpha_1 + \gamma TA + \beta_{1i} X_i' + \varepsilon_{1i} & R\&D_i = 1 \text{ if } R\&D_i^* > 0; 0 \text{ otherwise} \\ TA_i^* = \alpha_2 + \varphi R\&D + \beta_{2i} X_i' + \varepsilon_{2i} & TA_i = 1 \text{ if } TA_i^* > 0; 0 \text{ otherwise} \end{cases}$$

In order to estimate properly the innovative input-output relationship, it is necessary to deal with the issue of endogeneity which affect our innovative input variables. Indeed, the decision to invest in R&D and TA is affected by a company's specific characteristics, such as size, sectoral belonging and so on. Hence, R&D and TA have to be instrumented by their fitted values computed on the basis of the following estimate:

$$(3) \quad \begin{cases} R\&D_i^* = \alpha_1 + \gamma TA + \beta_{1i} X_i' + \eta_{1i} Z_i' + IMR_{R\&D} + \varepsilon_{1i} & \text{if } R\&D_i^* > 0 \\ TA_i^* = \alpha_2 + \varphi R\&D + \beta_{2i} X_i' + \eta_{2i} Z_i' + IMR_{TA} + \varepsilon_{2i} & \text{if } TA_i^* > 0 \end{cases}$$

This SUR specification is similar to equation (2) in dealing with the simultaneity and cross-effects of R&D and TA. However, there are two important differences. First, equation (3) is estimated along the positive values of R&D and TA for the subset of innovative companies (while equation 2 was dealing with the binary choice of investing or not in R&D and TA for all companies). Second, this equation is augmented by an innovation-related vector of variables Z_i' (only available for innovative companies) and by the *IMRs* estimated in equation (2). In turn, this makes equations (2) and (3) as the first- and second-part equation of the standard "generalised selection model" (the 'type 2 Tobit model' in the taxonomy proposed by Amemiya, 1985)⁸.

5.2 The Augmented Knowledge Production Function (Step 2)

Once we focus on the innovative companies only, the determinants of product and process innovation are modelled using the inverse Mills ratios and the fitted values obtained from step 1. The adopted seemingly unrelated (probit) regression (SUR) approach with correlated disturbances implies the following specification:

$$(4) \quad \begin{cases} PROD_i^* = \alpha_1 + \gamma PROC + \beta_{1i} X_i' + \eta_{1i} Z_i' + \mu R\&Df_i + \nu TAf_i + \lambda RD^*TAf_i + IMR_{PRD} + \varepsilon_{1i} \\ PROC_i^* = \alpha_2 + \varphi PROD + \beta_{2i} X_i' + \eta_{2i} Z_i' + \mu R\&Df_i + \nu TAf_i + \lambda RD^*TAf_i + IMR_{PRC} + \varepsilon_{2i} \end{cases}$$

where $PROD_i^*$ and $PROC_i^*$ are the latent variables underlying the two binary outcomes we observe; X_i' summarises a set of exogenous explanatory factors related to a company's characteristics and its sectoral specificities; Z_i' indicates an innovation-related set of variables; $R\&Df$ and TAf indicate, respectively, the fitted values of R&D and TA computed in the previous Step; (RD^*TAf) describes the interaction between these two variables and *IMR* stands for the inverse product and process Mills ratios estimated in equation (1) correcting for the selection of innovative companies only. This specification will be further generalised by replacing the R&D and TA regressors with sector- and size-specific regressors in order to obtain R&D and TA coefficients specific to the group of firms under investigation (see Section 6).

⁸This model is discussed more extensively in Section 5.3

5.3 The Innovative Intensity Equation: A Two-Part Model (Step 3)

Finally, the last step of analysis consists of the estimation of the "share of innovative sales" (innovative turnover) equation. This is the only continuous measure of innovative intensity having positive values for companies that introduced product innovation. It follows that when it is applied to the full sample of 3138 innovative companies, this measure shows a censored distribution with a positive mass in zero, due only minimally to companies that failed to complete their innovative projects (160 firms), the largest part of this being due to only-process innovators whose innovative performance is not captured by innovative sales (676 firms, see Table 7).

Table 7: Share of innovative sales over total turnover by innovative output

%	Process		
Product	0	1	Total
0	0 (160)	0 (676)	0 (836)
1	34.80 (677)	44.84 (1625)	42.02 (2302)
Total	27.41 (837)	28.61 (2301)	28.32 (3138)
In parentheses: number of observations			

A possible way of dealing with the censored nature of the dependent variable would be to estimate a Tobit model; however, while this solution accounts mathematically for the existence of a large mass of probability in zero, it assumes that the same model accounts for both the zeros and the positive values, which is clearly not the case here. In fact, in our context, the zeros do not represent the corner solution of a maximisation problem, but rather a company's choice to target a different form of innovation (namely, process innovation). However, keeping only the 2,302 companies that declared they had introduced product innovation generates an obvious sample selection bias. A two-part maximum likelihood framework is thus needed, accounting for the probability of innovative turnover being larger than zero ($P(\text{turninn}_i > 0)$) and the mean of the sales-weighted measure of innovation $E(\text{turninn}_i \mid \text{turninn}_i > 0)$ being at least partially determined by a different model⁹.

The "generalised selection model" (Amemiya, 1985) represents a standard reference for econometric estimation in such a setting. In particular, since the probability of innovative turnover being larger than zero is equivalent to the probability of a company having introduced product innovation, the participation equation is represented by the first SUR equation computed in equation (4). From this SUR estimate, an additional inverse Mills ratio (IMR-turnover) is computed and introduced as a regressor in estimating innovative turnover (equation 5) on the sub-sample of companies having realised product innovation (either alone or in combination with process).

$$(5) \text{ Innosales}_i = \alpha_1 + \mu R\&Df_i + \nu T Af_i + \lambda RD^* T Af_i + \gamma \text{InnoExp} + \beta_{1i} X'_i + \eta_{1i} Z'_i + \text{IMR}_{TURN} + \varepsilon_{1i}$$

In doing so, this methodology allows to take simultaneously into account: 1) the first sample bias associated with the selection of the 3,138 innovative companies; 2) the revealed correlation between product and process innovative decisions; 3) the second sample bias associated with the selection of the 2,302 companies introducing product innovation and exhibiting a positive value of innovative sales. Finally, it is important to note that the revealed significance of IMR-turnover confirms the greater suitability of this method with respect to the standard Tobit one (see Table 10).

⁹An econometric procedure for testing a Tobit model (Tobin, 1958) against a two-part Heckman-type model in the context of CIS data is provided by Conte (2009).

6 Econometric Results

Table 8 reports the SUR estimations based on the first three equations described in Section 5. These are used to calculate the IMRs and the fitted values of R&D and TA introduced, subsequently, in the augmented version of the Knowledge Production Function (see Table 9).

The first two probit-based estimations describe the probability of introducing (1) product / process innovation and (2) R&D / TA respectively for the entire sample of companies. The third SUR provides results on the determinants of the level of R&D and TA investments made by innovative companies.

Results indicate the strong positive simultaneous relationship between the two innovative inputs and the two outputs. In line with the evidence provided in Table 5, the probability of companies to be innovative seems to imply both types of innovative investments (R&D and TA) and both types of results (product and process innovation). Among innovative companies, the amount of investment in TA affects the R&D effort of a company while the opposite does not hold. In turn, this may suggest a complementary role played by TA in supporting the creation of knowledge internal to the company via R&D.

As expected, results indicate that the bigger a company, the higher its probability of being innovative and investing relatively more on both types of innovation. Being part of a business group in Italy seems not to affect the investment choice but rather the volume of innovative investment in R&D and TA conducted by the firm with positive effects more significant in the case of product innovation only. National boundaries seem to play a role in this context. In particular, although being part of foreign group affects the volume of a company's R&D investments, this does not appear to have an effect on the probability of being innovative - with even a negative sign in the case of process innovation. Export intensity and patent property strongly orientate a company's innovative process towards R&D and product innovations. Among innovative companies, cooperation in innovation activities with both research institutions and market agents (such as customers, suppliers or competitors) affect positively a company's involvement in R&D activities. Public support appear overall quite effective in favouring both types of innovative investments. Finally, the revealed significance of the two IMRs confirms the greater suitability of this method with respect to the standard Tobit one.

Table 9 describes the Knowledge Production Function (KPF) augmented to include the relationship between the two innovative outputs and (fitted) inputs¹⁰. Unlike Table 8, the cross-dependency between product and process innovation disappears in the KPFs presented in Tables 9 and 10 since the joint occurrence of the two outputs is taken into account by the different types of innovative investments (R&D and TA). Indeed, consistently with the first hypothesis discussed in Section 2, R&D plays a significant role only in determining the probability of introducing product innovation, while TA increases the probability of realizing process innovation. The interaction term between R&D and TA (RDF*TAf) does not indicate an additional effect due to the joint effect of these two types of expenditures¹¹. While the Schumpeterian hypothesis (see Section 2) is proved correct in affecting the likelihood of being innovative (see Table 8), once attention is limited to the selected sample of innovative companies, company size turns out to affect negatively product innovation while it is not significant in the case of process innovation. This is a consistent result across the

¹⁰The correlation between actual and fitted values is 0.75 and 0.60 for R&D and TA, respectively. Both correlations are statistically significant at 99%.

¹¹The interaction between R&D and TA is introduced as a further check into the regression. Indeed, it is important to remind that the fitted values of R&D and TA already take into account the cross-causality linkage which may affect these two inputs (see Table 8).

different specifications described by tables 9, 10 and 11. Taking together such evidence, this result suggests a threshold effect; namely, size represents a strong hampering factor for company being innovative or aiming at investing in innovation activities whereas, at the same time, smaller companies appear to be more oriented towards product innovations *within the group of innovative companies*. Not surprisingly, patent property and market cooperation in innovative activities exert a positive impact on, respectively, product and both types of innovations (see the discussion in Section 2). Public support affects the level of innovative investments (Table 8) whereas it does not seem to have an effect on the results of the innovative process (Tables 9 to 11).

Table 10 offers a generalization of the Knowledge Production Function by introducing differentiated R&D and TA coefficients by size and sector groups. The replication of the SUR econometric analysis in this way appears to confirm the general pattern identified in the aggregate estimate and it is consistent with the hypotheses put forward in Section 2. In particular, all four R&D coefficients (for large firms, small firms, high-tech and low-tech sectors) are statistically significant and have a greater effect on product innovation than process innovation. On the contrary, TA is always significant in determining process innovation whereas it has no effect in the case of product innovation. Although R&D intensity is higher for large companies and high-tech sectors (see Table 3), the corresponding coefficient is quite similar across size and sector groups (even larger for low-tech companies). In turn, this means that the threshold effect discussed for the size of a company applies also to the case of R&D investment. Indeed, sector and size are relevant dimensions in affecting the probability and the level of a company's R&D investment; however, these elements do not affect the capacity of R&D to translate into products afterwards. In contrast, a different pattern emerges in the case of TA investments. Indeed, consistently with the theories discussed in Section 2, embodied technological change represents both the major investment by small companies and low-tech sectors (see Table 3) while at the same time exhibiting a larger coefficient in affecting process innovation compared to the other groups of firms. The interaction between R&D and TA ($R\&D \cdot TA$) is significant only for small firms. In turn, this stresses the importance than these two complementary sources of knowledge play when the knowledge available within company boundaries is more limited and the choice between buying and/or producing technology more relevant (see Section 2). As in the case of Table 9, market cooperation in innovation activities does affect positively a company's innovative results.

Table 11 discusses the determinants of innovation intensity, measured as the share of sales related to the introduction of product innovations to the market. Both R&D and TA coefficients turn out to be positive and statistically significant in affecting the degree of innovativeness of a company. In other words, while the two inputs appear to bias ex-ante companies' innovative decisions towards different qualitative outcomes (either product or process innovation), they both positively contribute to increasing innovation intensity. Moreover, not surprisingly the R&D coefficient turns out to be larger in magnitude than the TA one. In line with the second and third hypotheses put forward in Section 2, the R&D/product innovation relationship shows increasing returns for larger firms and high-tech sectors. Indeed, the impact is more significant and larger in magnitude among those groups of companies that invest more in R&D expenditures. Conversely, TA coefficient is relatively higher for small companies and low-tech sectors. The innovation intensity equation includes innovative expenditures other than R&D and TA (such as training and marketing of innovative products) as an additional regressor. Indeed, these expenditures occur in a later stage of the innovative process, very close to the commercialisation phase, when product innovation generates innovative sales. Indeed, also in the case of table 11, the revealed significance of both the additional regressor and the IMRs confirms the greater suitability of this method with respect to the standard Tobit one.

Table 8: Selection Equations and the Determinants of R&D and TA

Variable	SUR Estimation		SUR Estimation		SUR Estimation	
	Product innovation	Process innovation	R&D	TA	Log R&D	Log TA
	Coefficient (Std. Err.)					
Process innovation	1.479*** (0.062)					
Product innovation		1.516*** (0.062)				
TA			1.194*** (0.069)			
R&D				1.296*** (0.073)		
Log TA					0.101*** (0.020)	
Log R&D						0.041 (0.038)
Employment Size	0.078** (0.037)	0.208*** (0.033)	0.260*** (0.038)	0.235*** (0.034)	0.840*** (0.032)	0.742*** (0.053)
Industrial Group in Italy	0.238*** (0.092)	0.013 (0.077)	0.143 (0.101)	0.051 (0.099)	0.230*** (0.074)	0.246** (0.115)
Industrial Group Abroad	0.056 (0.141)	-0.260* (0.136)	-0.018 (0.123)	-0.444*** (0.111)	0.250*** (0.094)	-0.048 (0.198)
Export Intensity	0.734*** (0.152)	-0.242* (0.136)	0.750*** (0.147)	-0.211* (0.123)	0.600*** (0.141)	0.055 (0.186)
Patent Property	0.705*** (0.081)	-0.066 (0.081)	0.552*** (0.080)	0.233*** (0.078)	0.357*** (0.079)	0.119 (0.098)
Cooperation - Research					0.226* (0.135)	-0.271 (0.167)
Cooperation - Market					0.239** (0.111)	0.182 (0.137)
Public Support					0.454*** (0.067)	0.365*** (0.089)
IMR - RD					0.999*** (0.144)	
IMR - TA						0.552*** (0.214)
Constant	-1.870*** (0.301)	-1.564*** (0.286)	-3.160*** (0.321)	-1.431*** (0.249)	-1.671** (0.771)	1.108** (0.520)
Industry Dummies	V	V	V	V	V	V
N	8216	8193	8216	8193	3,062	3,062
Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1						

Table 9: The Augmented Knowledge Production Function (KPF) - Full Sample

Variable	SUR Estimation	
	Product innovation	Process innovation
	Coefficient (Std. Err.)	Coefficient (Std. Err.)
Process Innovation	-0.192 (0.493)	
Product Innovation		-0.163 (0.987)
RDF	0.201*** (0.048)	0.003 (0.042)
TAF	0.042 (0.048)	0.227*** (0.044)
RDF*TAF	0.009 (0.011)	0.003 (0.009)
Size	-0.240*** (0.062)	-0.144 (0.118)
Export Intensity	0.267 (0.272)	-0.409** (0.206)
Industrial Group in Italy	0.200 (0.153)	0.057 (0.113)
Industrial Group Abroad	0.155 (0.218)	0.011 (0.207)
Research Cooperation	0.187 (0.171)	0.137 (0.152)
Market Cooperation	0.600*** (0.191)	0.398*** (0.150)
Patent Property	0.532** (0.210)	-0.157 (0.105)
Public Support	-0.108 (0.093)	0.015 (0.091)
IMR - Product	0.116 (0.532)	
IMR - Process		0.180 (0.980)
IMR - Turnover		
Constant	0.335 (1.152)	0.524 (1.829)
Industry Dummies	V	V
N	3,118	3,118
R ²	0.20	0.16
Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1		

Table 10: The Augmented KPF - Size and Sector Differences in R&D and TA

Variable	SUR Estimation - Size		SUR Estimation - Sector	
	Product innovation	Process innovation	Product innovation	Process innovation
	Coefficient (Std. Err.)	Coefficient (Std. Err.)	Coefficient (Std. Err.)	Coefficient (Std. Err.)
Process Innovation	-0.101 (0.484)		-0.204 (0.480)	
Product Innovation		-0.154 (0.963)		-0.380 (0.998)
RDf (L)	0.145*** (0.048)	0.070 (0.042)		
RDf (S)	0.144** (0.061)	-0.071 (0.054)		
TAf (L)	0.031 (0.048)	0.187*** (0.042)		
TAf (S)	-0.013 (0.054)	0.217*** (0.051)		
RDf*TAf (L)	0.004 (0.010)	-0.002 (0.008)		
RDf*TAf (S)	0.048** (0.019)	0.025 (0.017)		
RDf (HT)			0.189*** (0.064)	-0.085 (0.058)
RDf (LT)			0.215*** (0.061)	0.063 (0.054)
TAf (HT)			0.063 (0.073)	0.195*** (0.073)
TAf (LT)			0.044 (0.057)	0.254*** (0.053)
RDf*TAf (HT)			0.004 (0.015)	0.015 (0.015)
RDf*TAf (LT)			0.008 (0.014)	-0.007 (0.012)
Size	-0.199** (0.088)	-0.187 (0.123)	-0.243*** (0.062)	-0.176 (0.119)
Export Intensity	0.307 (0.268)	-0.396** (0.202)	0.272 (0.269)	-0.355* (0.210)
Industrial Group in Italy	0.245 (0.152)	0.058 (0.112)	0.198 (0.153)	0.061 (0.113)
Industrial Group Abroad	0.215 (0.211)	-0.026 (0.201)	0.166 (0.215)	0.059 (0.207)
Research Cooperation	0.181 (0.166)	0.130 (0.154)	0.191 (0.170)	0.143 (0.151)
Market Cooperation	0.627*** (0.187)	0.401*** (0.154)	0.602*** (0.189)	0.410*** (0.152)
Patent Property	0.568*** (0.210)	-0.174* (0.105)	0.536** (0.208)	-0.135 (0.106)
Public Support	-0.081 (0.094)	0.005 (0.093)	-0.110 (0.093)	0.019 (0.091)
IMR - Product	0.212 (0.525)		0.110 (0.524)	
IMR - Process		0.185 (0.958)		-0.031 (0.991)
Constant	0.224 (1.155)	0.722 (1.776)	0.346 (1.148)	0.793 (1.846)
Industry Dummies	V	V	V	V
N	3,118	3,118	3,118	3,118
R ²	0.20	0.17	0.20	0.17

Robust standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1

Table 11: Heckman-corrected estimates of innovative sales

TOT		Size		Sector	
Variable	Heckman Coefficient (Std. Err.)	Variable	Heckman Coefficient (Std. Err.)	Variable	Heckman Coefficient (Std. Err.)
RDF	0.045*** (0.012)	RDF (L)	0.055*** (0.012)	RDF (HT)	0.048*** (0.015)
Taf	0.028*** (0.008)	RDF (S)	0.036*** (0.014)	RDF (LT)	0.041*** (0.014)
		Taf (L)	0.023*** (0.009)	Taf (HT)	0.022* (0.014)
RDF*Taf	-0.002 (0.002)	Taf (S)	0.028*** (0.010)	Taf (LT)	0.028*** (0.010)
		RDF*Taf (L)	-0.003* (0.002)	RDF*Taf (HT)	-0.001 (0.003)
		RDF*Taf (S)	-0.000 (0.004)	RDF*Taf (LT)	-0.002 (0.002)
Other Innovative Expenditure Size	0.014*** (0.005) -0.109*** (0.015)		0.014*** (0.005) -0.118*** (0.018)		0.014*** (0.005) -0.108*** (0.015)
Export Intensity	0.061* (0.035)		0.063* (0.035)		0.059* (0.035)
Industrial Group in Italy	0.069** (0.028)		0.068** (0.027)		0.068** (0.027)
Industrial Group Abroad	0.094*** (0.030)		0.091*** (0.030)		0.093*** (0.029)
Research Cooperation	0.063** (0.032)		0.062* (0.032)		0.062* (0.032)
Market Cooperation Patent Property	0.097*** (0.035) 0.049** (0.025)		0.095*** (0.036) 0.046* (0.025)		0.094*** (0.035) 0.047* (0.025)
Public Support	-0.018 (0.020)		-0.018 (0.020)		-0.018 (0.020)
IMR - Turnover	0.392*** (0.091)		0.391*** (0.091)		0.386*** (0.091)
Constant	0.304** (0.145)		0.343** (0.151)		0.311** (0.147)
Industry Dummies	V		V		V
N	2,247		2,247		2,247
R ²	0.207		0.210		0.208

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

7 Concluding Remarks

In this paper we have discussed the determinants of three alternative measures of innovative output by looking at companies' formal R&D activities and at the acquisition of external technology in its embodied and disembodied components (TA). These input-output relationships have also been considered after distinguishing between small and large companies and those belonging to low-tech and high-tech sectors.

Based on previous theoretical and empirical literature, a first hypothesis tested in this study was the possible dichotomy linking R&D mainly with product innovation and TA mainly with process innovation. Aggregated and disaggregated results are fully consistent with this hypothesis: R&D expenditures turn out to be statistically significant in enhancing product innovation, while TA significantly impacts the likelihood of engaging in process innovation. However, on the aggregate, both inputs significantly increase overall company innovative performance, measured as sales from innovative products.

A second hypothesis put forward in this study is the allegedly crucial role of TA in small companies, while their larger counterparts are expected to be more R&D-based. This hypothesis is confirmed as far as TA is concerned: in impacting the likelihood of exhibiting process innovation and in increasing innovative sales, the TA coefficient is larger in the small company sub-sample. On the other hand, R&D turns out to be crucial in fostering product innovation in both small and large companies. However, the impact of R&D over innovative turnover turns out to be larger and more significant in the large companies sub-sample, as expected.

Finally, weaker support was found for the third hypothesis put forward in this work, namely that R&D should be particularly crucial in the high-tech sectors. Indeed, Tables 10 and 11 do not reveal striking sectoral differences, although – in affecting innovative performance – the R&D coefficient turns out to be larger in the high-tech sectors (while the TA coefficient is higher and more significant in the low-tech sectors, as expected).

Beyond the evidence here provided, our results pave the way for different avenues of further research. For instance, it might be interesting to apply a similar analysis to companies belonging to the service sectors, using data explicitly designed to capture the particular nature of innovative inputs and innovative outputs within services. By the same token, the availability of larger datasets or representative case studies might allow in-depth analyses of innovative activities within finer subcategories of companies, such as for example SMEs in some particular sectors. Finally, future availability of time series data with the same level of detail as the CIS data might allow researchers to investigate the dynamic relationships between and within different innovative inputs and different innovative outputs.

Appendix

Table A1: List and definitions of the explanatory variables

Generic company characteristics	
Size	Number of company's employees at December 31st, 2000.
⇒ Small companies	Companies with 50 or less employees
⇒ Large companies	Companies with more than 50 employees
Export Intensity	Company's export / turnover in 2000.
Industrial Group in Italy	Belonging to an industrial group with headquarters in Italy (dummy)
Industrial Group Abroad	Belonging to an industrial group with foreign headquarters (dummy)
Patent Property	Property of any valid patent at the end of 2000
Industry dummies	80 Industry dummies according to the three-digit ATECO 91 classification
⇒ High Tech Sectors	High- and medium-high technology industries according to OECD, Science, Technology and Industry Scoreboard 2007
⇒ Low Tech Sectors	Low- and medium-low technology industries according to OECD, Science, Technology and Industry Scoreboard 2007
Innovative-relevant information	
Cooperation - Market Agents	Cooperation agreements in innovative activities with customers, suppliers or competitors (dummy variable)
Cooperation - Research	Cooperation agreements in innovative activities with universities and/or research institutes (dummy variable)
Other Innovative Expenditure	Logarithmic transformation of residual innovative expenditures in 2000, once R&D and TA have been excluded

Table A2: Descriptive statistics of the explanatory variables

N	8219	3138
	Mean	Mean
Generic company characteristics		
Size (Number of employees)	37	59
Export / Turnover	13%	18%
Industrial Group in Italy	7%	12%
Industrial Group Abroad	1%	3%
Patent Property	12%	23%
Innovative-relevant information		
Other Innov. Expenditure / Turnover	—	0.82%
Cooperation - Research	—	5%
Cooperation - Market Agents	—	6%

References

- ACS, ZOLTAN J., & AUDRETSCH, DAVID B. 1990. *Innovation and Small Firms*. Cambridge (Mass.): MIT Press.
- ACS, ZOLTAN J., AUDRETSCH, DAVID B., & FELDMAN, MARYANN P. 1994. R&D Spillovers and Recipient Firm Size. *Review of Economics and Statistics*, **Vol. 4**(2), pp. 336–340.
- AMEMIYA, TAKESHI. 1985. *Advanced Econometrics*. Cambridge (Mass.): Harvard University Press.
- AUDRETSCH, DAVID B., & VIVARELLI, MARCO. 1996. Firms Size and R&D Spillovers: Evidence from Italy. *Small Business Economics*, **Vol. 8**(3), pp. 249–258.
- BARLET, CORINNE, DUGUET, EMMANUEL, ENCAOUA, DAVID, & PRADEL, JACQUELINE. 1998. The Commercial Success of Innovations: an Econometric Analysis at the Firm Level in French Manufacturing. *Annales d'Economie et de Statistique*, **Vol. 49/50**, pp. 457–478. Published also in D. Encaoua et al. (eds.), *The Economics and Econometrics of Innovation*, Boston: Kluwer Academic Publishers, 2000, pp. 435–456.
- BRESCHI, STEFANO, MALERBA, FRANCO, & ORSENIGO, LUIGI. 2000. Technological Regimes and Schumpeterian Patterns of Innovation. *Economic Journal*, **Vol. 110**(463), pp. 388–410.
- BROUWER, ERIK, & KLEINKNECHT, ALFRED. 1996. Firm Size, Small Business Presence and Sales of Innovative Products: A Micro-econometric Analysis. *Small Business Economics*, **Vol. 8**, pp. 189–201.
- CASSIMAN, BRUNO, & VEUGELERS, REINHILDE. 2002. R&D Cooperation and Spillovers: Some Empirical Evidence from Belgium. *American Economic Review*, **Vol. 92**(4), pp. 1169–1184.
- CHIAO, C. 2001. The Relationship between R&D and Physical Investment of Firms in Science-based Industries. *Applied Economics*, **Vol. 33**(No. 1), pp. 23–35.
- COHEN, WESLEY M., & KLEPPER, STEVEN. 1996. A Reprise of Size and R&D. *Economic Journal*, **Vol. 106**(437), pp. 925–951.
- CONTE, ANDREA. 2006. *The Evolution of the Literature on Technological Change over time: A Survey*. Max Planck Institute of Economics, Discussion Paper No. 107.
- CONTE, ANDREA. 2009. Mapping Innovative Activity Using Microdata. *Applied Economics Letters*, **Vol. 16**(18), pp. 1795–1799.
- CREPON, BRUNO, DUGUET, EMMANUEL, & MAIRESSE, JACQUES. 1998. Research, Innovation and Productivity: An Econometric Analysis at the Firm Level. *Economics of Innovation and New Technology*, **Vol. 7**(3), pp. 115–156.
- DE JONG, PIETER J. 2007. The Relationship between Capital Investment and R&D Spending: A Panel Cointegration Analysis. *Applied Financial Economics*, **Vol. 17**(Issue 11), pp. 871–880.
- DHAWAN, RAJEEV. 2001. Firm Size and Productivity Differential: Theory and Evidence from a Panel of US Firms. *Journal of Economic Behavior & Organization*, **Vol. 44**(3), pp. 269–293.
- DOSI, GIOVANNI. 1988. Sources, Procedures, and Microeconomic Effects of Innovation. *Journal of Economic Literature*, **Vol. 26**(3), pp. 1120–1171.

- FREEMAN, CHRISTOFER. 1982. *The Economics of Industrial Innovation*. 2nd edn. London: Pinter.
- GREENWOOD, JEREMY, HERCOWITZ, ZVI, & KRUSELL, PER. 1997. Long-Run Implications of Investment-Specific Technological Change. *American Economic Review*, **Vol. 87**(3), pp. 342–362.
- GRILICHES, ZVI. 1979. Issues in Assessing the Contribution of Research and Development to Productivity Growth. *Bell Journal of Economics*, **Vol. 10**(1), pp. 92–116.
- GRILICHES, ZVI. 1990. Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature*, **Vol. 28**(4), pp. 1661–1707.
- HALL, BRONWYN H., & MAIRESSE, JACQUES. 1995. Exploring the Relationship between R&D and Productivity in French Manufacturing Firms. *Journal of Econometrics*, **Vol. 65**(1), pp. 263–293.
- HARHOFF, DIETMAR, SCHERER, FREDERIC M., & VOPEL, KATRIN. 2003. Citations, Family Size, Opposition, and the Value of Patent Rights. *Research Policy*, **Vol. 32**(8), pp. 1343–1364.
- HULTEN, CHARLES R. 1992. Growth Accounting When Technical Change Is Embodied in Capital. *American Economic Review*, **Vol. 82**(4), pp. 964–80.
- JAFFE, ADAM B. 1989. Real Effects of Academic Research. *American Economic Review*, **Vol. 79**(5), pp. 957–970.
- JORGENSEN, DALE W. 1966. The Embodiment Hypothesis. *Journal of Political Economy*, **Vol. 74**(1), pp. 1–17.
- KLEINKNECHT, ALFRED. 1987. Measuring R&D in Small Firms: How Much are we Missing? *Journal of Industrial Economics*, **Vol. 36**(2), pp. 253–256.
- LACH, SAUL, & ROB, RAFAEL. 1996. R&D, Investment, and Industry Dynamics. *Journal of Economics & Management Strategy*, **Vol. 5**(2), pp. 217–249.
- LACH, SAUL, & SCHANKERMAN, MARK. 1989. Dynamics of R&D and Investment in the Scientific Sector. *Journal of Political Economy*, **Vol. 97**(4), pp. 880–904.
- LOOF, HANS, & HESHMATI, ALMAS. 2001. *On the Relationship between Innovation and Performance: a Sensitivity Analysis*. Eindhoven Centre for Innovation Studies (ECIS), Working Paper No. 446.
- MAIRESSE, JACQUES, & MOHNEN, PIERRE. 2002. Accounting for Innovation and Measuring Innovativeness: An Illustrative Framework and an Application. *American Economic Review*, **Vol. 92**(2), pp. 226–230.
- MALERBA, FRANCO, NELSON, RICHARD, ORSENIGO, LUIGI, & WINTER, SIDNEY. 2008. Public Policies and Changing Boundaries of Firms in A "History-friendly" Model of the Co-evolution of the Computer and Semiconductor Industries. *Journal of Economic Behavior & Organization*, **Vol. 67**(2), pp. 355–380.
- MUKOYAMA, TOSHIHIKO. 2006. Rosenberg's "Learning by Using" and Technology Diffusion. *Journal of Economic Behavior & Organization*, **Vol. 61**(1), pp. 123–144.

- NELSON, RICHARD R., & WINTER, SIDNEY. 1982. *An Evolutionary Theory of Economic Change*. Cambridge (Mass.): Harvard University Press.
- OECD. 2005. *Oslo Manual: Guidelines for Collecting and Interpreting Innovation Data*.
- OECD. 2009. *Innovation in Firms. A Microeconomic Perspective*.
- PARISI, MARIA LAURA, SCHIANTARELLI, FABIO, & SEMBENELLI, ALESSANDRO. 2006. Productivity, Innovation and R&D: Micro Evidence for Italy. *European Economic Review*, **Vol. 50**, pp. 2037–2061.
- PAVITT, KEITH. 1984. Sectoral Patterns of Technical Change: Towards a Taxonomy and a Theory. *Research Policy*, **Vol. 13**(6), pp. 343–373.
- PIGA, CLAUDIO A., & VIVARELLI, MARCO. 2004. Internal and External R&D: A Sample Selection Approach. *Oxford Bulletin of Economics and Statistics*, **Vol. 66**(4), pp. 457–482.
- ROSENKRANZ, STEPHANIE. 2003. Simultaneous Choice of Process and Product Innovation When Consumers Have a Preference for Product Variety. *Journal of Economic Behavior & Organization*, **Vol. 50**(2), pp. 183–201.
- SALTER, W. E. G. 1960. *Productivity and Technical Change*. Cambridge University Press.
- SANTARELLI, ENRICO, & STERLACCHINI, ALESSANDRO. 1990. Innovation, Formal vs. Informal R&D, and Firm Size: Some Evidence from Italian Manufacturing Firms. *Small Business Economics*, **Vol. 2**, pp. 223–228.
- SCHUMPETER, JOSEPH. 1942. *Capitalism, Socialism and Democracy*. New York: Harper and Brothers.
- TOBIN, JAMES. 1958. Estimation of Relationships for Limited Dependent Variables. *Econometrica*, **Vol. 26**(1), pp. 24–36.
- WINTER, SIDNEY G. 1984. Schumpeterian Competition in Alternative Technological Regimes. *Journal of Economic Behavior and Organization*, **Vol. 5**(3-4), pp. 287–320.