QUADERNI DEL DIPARTIMENTO DI SCIENZE
ECONOMICHE E SOCIALI

DO INNOVATIVE INPUTS LEAD TO DIFFERENT INNOVATIVE OUTPUTS IN MATURE AND YOUNG FIRMS?

Gabriele Pellegrino, Mariacristina Piva

Serie Rossa: Economia – Quaderno N. 97 marzo 2014

UNIVERSITÀ CATTOLICA DEL SACRO CUORE
PIACENZA

DO INNOVATIVE INPUTS LEAD TO DIFFERENT INNOVATIVE OUTPUTS IN MATURE AND YOUNG FIRMS?

Gabriele Pellegrino[a b], Mariacristina Piva[b *]

Abstract

This paper investigates the determinants of the choice of different types of innovative input (R&D and technological acquisitions) and their relationship with different innovative outputs (product and process innovation), distinguishing between firms of different ages (mature *vs* young). In order to do so we apply a nonlinear structural model estimated on the third and fourth waves of the Italian Community Innovation Survey (CIS).

We find that firm and market characteristics play a distinct role in boosting different types of innovation activities for firms of different ages. In particular, while methods of appropriability and international market exposure are relevant for both forms of innovative input, cooperation in innovation activities appears to be important for increasing the level of investment in R&D but not for technological acquisition. Moreover, young firms show a higher level of sensitivity than their mature counterparts to sources of information regarding innovation when we consider the magnitude of their innovative effort. On the contrary, factors such as methods of appropriability and support for innovation appear to be more important for enhancing the level of investment in both R&D and technological acquisitions for the mature firms only. Finally, the two innovative inputs appear to be equally important in determining both forms of innovative output for the two sub-samples of firms.

Keywords: R&D, technological acquisitions, innovative outputs, young firms.
JEL Classification: O31.

[a] Barcelona Institute of Economics - University of Barcelona, Barcelona
[b] Università Cattolica del Sacro Cuore, Piacenza and Milano

*** Corresponding author:**
Mariacristina Piva
Università Cattolica del Sacro Cuore
Dipartimento di Scienze Economiche e Sociali
Via Emilia Parmense, 84
29122 Piacenza
Italy
e-mail: mariacristina.piva@unicatt.it

1. Introduction

Historically, technological innovation has been recognised as one of the major sources of growth and development (Solow, 1956; Romer, 1986). Since the seminal work by Crepon *et al.* (1998), many researchers have tried to explain economic growth as being due to technological output, this in turn due to technological effort. In general, results have shown a clear-cut positive relationship between R&D and innovation on the one hand, and different measures of economic growth on the other (see Hall *et al.*, 2010 for a recent review of the subject).

However, most of these studies have omitted to take into account the high degree of heterogeneity associated with firm innovation. Apart from internal formalised R&D, firms can also rely on different external sources of innovation, such as technological acquisition, with particular reference to the part of technological change embodied in acquired goods (as in the case of acquisition/substitution of machinery and equipment).

In addition, as shown by recent evidence (see Santamaría *et al.*, 2009; Ortega-Argilés *et al.*, 2009 and 2010), specific firm and market characteristics play a vital role in determining different firms' innovative strategies, in terms of both innovative inputs and outputs. Among these peculiarities, a firm's size and age can certainly be considered as the most important (Acs and Audretsch, 1987; Audretsch, 1995; Huergo and Jaumandreu, 2004). Traditionally however more attention has been devoted to analysing possible differences in the innovative behaviour of small and large firms than of mature and young ones. Nevertheless, young firms in general, and innovative young firms in particular, are often seen as key actors in economic growth and job creation (Birch, 1979; Acs and Audretsch, 1990; Brüderl and Preisendorfer, 2000). Foster *et al.* (2001) show that one third to one half of aggregate productivity growth in US manufacturing is directly attributable to the creation of new firms, reallocation between firms, and the disappearance of unsuccessful ones. Other studies have focused attention on the relative disadvantage of Europe as regards the birth and growth of the so-called Young Innovative Companies (YICs), due to limited creation of the necessary conditions. In this respect, some evidence suggests that EU start-ups face higher barriers to entry, innovation and growth compared to their US counterparts (see Bartelsmann *et al.* 2004, Philippon and Véron, 2008). Accordingly, over the last few years, several EU Member States have been promoting intervention policies aimed at encouraging the establishment, consolidation and development of YICs (Schneider and Veugelers, 2010; Moncada-Paternò-Castello, 2011).

In this context, the aim of this paper is twofold. Firstly, we analyse the determinants of firms' innovative effort (distinguishing between R&D and technological acquisitions - TA) and the link between this effort and the different outcomes (in terms of product and process innovation) that it produces. Secondly, we try to shed some light on how these particular relationships differ between young and mature incumbent firms.

With this aim and using the third and the fourth waves of the Italian Community Innovation Survey (CIS), we run a recursive model, that can be seen as an extension of the Crépon, Duguet and Mairesse's (1998) partial structure model (hereafter CDM model), to analyse these relationships[1]. Apart from the distinction we make between firms of different ages, this is one of the first studies to include technological acquisition (TA) as an additional innovation input besides R&D in a CDM model. Moreover, in contrast with most previous studies using CIS data, we implement an empirical strategy that takes into account the division between innovative and non-innovative firms in order to correct for the well-known problem of sample selection.

Following this introduction, the next section provides a discussion of the theoretical framework on which our work is based. Section 3 outlines the econometric methodology adopted, and Section 4 describes the database and the variables used in our analysis. Section 5 discusses our empirical results and Section 6 highlights our main conclusions.

2. The literature

In his seminal contribution, Griliches (1979) suggests a model of technological change according to which innovative outputs are seen as the product of knowledge-generating inputs. More specifically, the author proposes a three-equation model in which one of the equations is a function (the so-called Knowledge Production Function (KPF)) that links a measure of innovative input (namely R&D) with a measure of

---

[1] The lack of data on firms' productivity has prevented us from estimating the last equation in the classic CDM model (see next section for a more detailed description of the model).

innovative output (namely patents)[2]. Following this insight, Crèpon, Duget and Mairesse (1998) developed a more comprehensive model based on three distinct, but interrelated relationships: 1) the innovation input linked with its determinants; 2) the KPF that connects innovation input to innovation output; 3) the productivity equation, in which innovative output leads to productivity growth.

These two seminal works have paved the way for the emergence of a relatively recent field of research aimed at analysing the peculiarities of the innovative process (both at macro and micro level) and its contribution to economic growth. In this respect, a distinction has to be made between those studies based on an application of a CDM fully-structured model (i.e. that takes into account all three stages of the model) and those based on a CDM partial structure model (i.e. that consider at least one link between the three stages). Taking into account our main research aim, this work can be included in the latter category. Accordingly, in this brief survey we will focus on the first two stages of the CDM model, namely the innovation inputs linked with their determinants, and the KPF.

Historically, and due partly to the lack of other measures of innovation, most studies mainly focused attention on the determinants of R&D activity and its link with a measure of innovative output, most notably patents. However, such an approach appears to be oversimplified and too restrictive. In this respect, as Stoneman (1995:5f) suggests:

"R&D is not the only source of technological improvement. A firm may generate its own technology through R&D. It may also generate technological advance through learning of various kinds, design, reverse engineering and imitation [….]. New process technologies may also be acquired from the suppliers of capital goods. The relevant importance of these different sources will depend upon the nature of the firm, its industrial sector and its technological base."

Moreover, as pointed out by Kleinknecht *et al.* (2002), R&D accounts for just a quarter of the total expense aimed at obtaining product innovation.

Turning our attention to the innovative output measures, patents appear to be a very rough proxy of innovation for different reasons, as suggested by several studies. Firstly, firms generally prefer other ways of protecting their innovation, such as industrial secrecy with no disclosure of successful innovation (see Levin *et al.*, 1987).

[2] The other two equations in the model represent the determinants of R&D investment and the production function (augmented by the innovation term).

Secondly, firms with different characteristics (i.e. small *vs* large) and operating in different sectors (i.e. high-tech *vs* low-tech) show different propensities to patent (see Archibugi and Pianta, 1992; Patel and Pavitt, 1993). Finally, patents differ greatly in their importance as far as they might have different degree of applicatione and potential success.

Accordingly, in recent years, and thanks also to the availability of more comprehensive and precise innovation surveys, some authors have tried to extend the classic approach used to study firms' innovative processes to include other measures of innovation activities. In this respect the work by Conte and Vivarelli (2005) is a notable example which can be seen as one of the first attempts to extend the classic approach of the KPF by considering, besides R&D, the important role played by technological acquisitions (investment in new machinery and equipment, and external technology incorporated in licences, consultancies, and know-how) and their impact in determining different types of innovative outputs (product and process innovation). They found that R&D is strictly related to product innovation, while technological acquisition is crucial for process innovation. Moreover, their analyses also show that small firms and firms belonging to low-tech sectors are more likely "to buy" instead of "to make" technology, while large firms and firms operating in the high-tech sectors are much more R&D-based. This is in line with Santamaria *et al.* (2009), who find that the impact of non-R&D activities is particularly important in low and medium-tech firms. Similarly, Pellegrino *et al.* (2012) test an augmented KPF, trying to detect some differences between firms of different ages. The results of their analyses suggest that although in-house R&D appears to be important in enhancing the propensity to introduce product innovation in both mature and young firms, innovation intensity in the group of young firms is mainly dependent on embodied technical change from external sources.

A common trait of the above-mentioned works is the fact that they focus on just one stage of the CDM model, that is the KPF. While this approach allows for the testing of the relationship between different measures of innovative inputs and outputs at the same time, it completely ignores the process underlying a firm's innovative decision (i.e. the first stage of the CDM model). This aspect, linked to the way in which most of the data on innovation are collected (in particular CIS data) can compromise the reliability of the results[3]. Thus, the trade-off here is between applying an approach that leads to consistent results but that takes into account just one measure of innovative input (mostly R&D; classical CDM model approach), or ignoring possible sample selection problems in favour of a more detailed analysis of the firm's input-output innovative relationship.

Recently, some authors have proposed an approach that takes into account both of these aspects. In this respect it is worth mentioning the work of Polder *et al.* (2010), who estimate a CDM fully-structured model considering two different measures of innovative input (R&D and the amount of investment in information and communication technology (ICT)) and three different measures of innovative output (product, process and organizational innovation). They find a significant positive effect of ICT on the three measures of innovative output, while R&D turns out to be important only for enhancing the propensity to introduce product innovation. Hall *et al.* (2013) extend this approach further, considering two different measures of organizational innovation (organizational change associated with product and process innovation). Based on a large unbalanced panel of Italian manufacturing firms, they find that both R&D and ICT are important drivers of innovation activity, although R&D appears to be more relevant for product and process innovation.

In the spirit of these contributions and as previously mentioned, in this work we rely on an extension of a CDM partial structure model including investments in TA as an additional innovation input besides R&D, and two different measures of innovative outputs (product and process innovation). To the best of our knowledge, this is one of the first studies to include an indicator of technological acquisition in a CDM-type model. Another innovative aspect is that we analyse the existence of possible differences between mature and young firms (see Section 4 for a more precise definition of these two categories) in terms both of R&D and TA drivers and of peculiarities of the KPF. No existing literature has provided evidence relating to these particular research questions. However, some interesting and useful insights can be gained by considering the main results of some recent studies that have looked at the peculiarities of young companies' innovative processes.

[3] The source of bias stems from a problem of sample selection that arises when the non-innovative firms are excluded from the analyses (for a more articulated discussion see Mairesse and Mohnen, 2010).

Garcia-Quevedo *et al.* (2011), drawing on an unbalanced dataset of more than 2,000 Spanish manufacturing firms, look at the R&D drivers of young and mature firms. The results of their econometric estimations show that different firms and market characteristics play different roles in determining the innovative decisions of mature and young firms. In particular, if on the one hand factors like market concentration and the degree of product diversification are more important in fostering the innovative activity of mature firms only, on the other hand young firms' spending on R&D seems to be more sensitive to demand pull variables, suggesting that the presence of credit constraints for this particular type of firm plays a determinant role.

In a very recent study and in contrast with the above-mentioned work, Criscuolo *et al.*(2012) concentrate attention on the output side of innovative activity. In particular, using a large sample of UK firms, they try to explore possible differences between start-ups and established firms in terms of innovative performance, looking at both manufacturing and service sectors. They find that being a new firm increases the probability of introducing a radical product or process innovation in the service sector, while in the manufacturing sectors newly-established firms tend to be less innovative than established firms. This latter result is in line with the previously-mentioned study by Pellegrino *et al.* (2012), who use data from the Italian CIS to show that young innovative companies are less R&D-based and perform worse in terms of innovative turnover than their mature counterparts.

3. Model and Econometric Methodology

We apply an extended version of a classic CDM partial structure model. More specifically, we follow an approach initially proposed by Griffith *et al.* (2006) and subsequently also used by Mairesse and Robin (2009), who enrich the basic CDM model by considering product as well as process innovation as innovative outputs[4]. We augment their model, including a further equation for technological acquisition. Accordingly, our approach is formalised in 6 equations: (1) the firms' decision to engage in R&D activity; (2) the firm's decision regarding the amount of resources to be invested in R&D activity; (3) the firms' decision to invest in TA; (4) the firm's decision regarding the amount of resources to be invested in TA; (5) and (6) the knowledge

---

[4] Both studies are based on a fully-structured CDM model.

production function, in which we consider two different innovative outputs (product and process innovation).

Another important peculiarity of our empirical strategy is that in contrast with most previous studies, we do not focus our attention only on the cohort of innovative firms, but perform our analysis considering the whole sample of firms. In particular, the KPF (steps (5)-(6)) is estimated using the predicted values for all firms obtained from the estimations of steps (1) - (4) based on reported R&D and TA figures. This approach reflects the assumption that all firms carry out innovative activities, although some of them do not report any innovative investment; for example a firm's workers may spend a certain amount of their workday trying to find a more efficient way to carry out the production process in which they are involved. The same could apply for personnel employed in other firms that provide external technology (investment in new machinery and equipment and purchasing of external technology incorporated in licences, consultancies and know-how). In both cases, if the effort does not exceed a certain threshold it will not be reported by the firm as investment in R&D or TA.

Having delineated the main characteristics of our empirical strategies, in the following two subsections we describe the econometric methodologies and the specifications used for the estimations of the model's 6 equations.

3.1 Innovation inputs: R&D and technological acquisitions

We identify two different types of innovation input: R&D expenditures (both *intramural* and *extramural*) and technological acquisitions (both embodied and disembodied components). As well-documented in the empirical literature dealing with CIS surveys (see discussion in Section 4), these variables are subject to selectivity: only those firms that have claimed to be involved in product or process innovation (completed/ongoing/abandoned) report data on innovative investments. Furthermore, since both types of innovative activity can be performed informally, these two variables may also be censored. However, as noted in the previous section, if this innovative effort does not reach a certain threshold the firm will not report it as expenditure. Consequently, both variables include a certain number of zero and missing values. Econometrically, this mixed pattern of zero/missing and positive values naturally leads to a Tobit II model (see Amemiya, 1984), defined as follows.

Let i=1,..........,N index firms. The firms' two innovative decisions are defined by the two binary variables RDT_d_i and TAT_d_i, which take a value of 1 when R&D and TA respectively are observed and 0 otherwise. We link RDT_d_i and TAT_d_i with the two latent variables $RDT\_d_i^*$, $TAT\_d_i^*$ such that:

$$(1) RDT\_d_i = \begin{cases} 1 \; when \;\; RDT\_d_i^* = \alpha_1' x_{1i} + \varepsilon_{1i} > 0 \\ 0 \; when \;\; RDT\_d_i^* = \alpha_1' x_{1i} + \varepsilon_{1i} \leq 0 \end{cases}$$

and

$$(2) TAT\_d_i = \begin{cases} 1 \; when \;\; TAT\_d_i^* = \alpha_2' x_{2i} + \varepsilon_{2i} > 0 \\ 0 \; when \;\; TAT\_d_i^* = \alpha_2' x_{2i} + \varepsilon_{2i} \leq 0 \end{cases}$$

We indicate RDT_i the amount of a firm's turnover invested in both *intramural* and *extramural* R&D, and TAT_i the amount of a firm's turnover invested in technological acquisitions. Denoting the corresponding latent variables with RDT_i^* and TAT_i^*, we have:

$$(3) RDT_i = \begin{cases} RDT_i^* = \beta_1' w_{1i} + u_{1i} \;\; when \; RDT\_d_i = 1 \\ 0 \;\; when \;\; RDT\_d_i = 0 \end{cases}$$

and

$$(4) TAT_i = \begin{cases} TAT_i^* = \beta_2' w_{2i} + u_{2i} \;\; when \; TAT\_d_i = 1 \\ 0 \;\; when \;\; TAT\_d_i = 0 \end{cases}$$

For each firm i, x_j and w_j (with j = {1, 2}) are vectors of explanatory variables, some of which could be common to both vectors. Assuming that each pair of error terms, ε_1 and u_1, and ε_2 and u_2, is bivariate and normally distributed with correlation coefficients $\rho_{\varepsilon 1 u 1}$ and $\rho_{\varepsilon 2 u 2}$, we estimate equations (1) - (3) and (2) - (4) with the Heckman two-step procedure (Heckman, 1979).

Since our analysis is focused on the whole sample of firms and not only on innovative firms, to model the firms' innovative decisions (equations (1) and (2)) we can use only the limited information available for all firms (see next section). Bearing in mind this important aspect and the primary objective of rendering the microdata from CIS3 and CIS4 datasets fully comparable, the choice of the explanatory variables has been made following both the original framework proposed by Crépon *et al.* (1998) and the extensions put forward by Griffith *et al.* (2006) and Mairesse and Robin (2009). For the sake of symmetry, we decided to estimate the 2 pairs of equations (equations (1)-(3) and (2)-(4)) using the same specifications.

Starting from the selection equations (1) and (2), we use an indicator of whether the firm is part of an enterprise group or not, an indicator of whether the international market is the firm's most significant market (in order to measure international competition), and two indicators of whether the firm makes use respectively of patents and strategic methods (registration of design, trademarks, copyrights) to protect its innovations[5]. Moreover, following the Schumpeterian tradition we include a set of industry dummy variables (based on the 2-digit ATECO codes[6]) to capture market conditions, and a variable reporting the log of the total number of employees as a measure of firm size.

In modelling the firms' propensity to invest in R&D and TA (equations (3) and (4)), we can rely on additional information being available only when firms are innovative; it may therefore be useful for characterizing the R&D - TA (see discussion in Section 4). Together with the regressors used in the selection equations, in accordance with previous evidence that shows the importance of cooperation agreements in determining the level of investment in innovative activities (Cassiman and Veuglers, 2002; Piga and Vivarelli, 2003, 2004), we also consider a dummy variable that identifies firms that had some cooperative agreements regarding innovation activities during the three-year period. Moreover, in order to test for the supposed positive impact of public funding on a firm's innovative activity (see Busom, 2000; Gonzales *et al*., 2005), a binary variable indicating whether the firm has received some (local/national/EU) public financial support for innovative activities is included. In addition, we consider two binary variables that take on a value of 1 if the firm has used any type of internal and/or external sources of information for its innovative activity. In this respect, a recent stream of literature emphasises the important role played by both

[5] Previous studies generally show a clear-cut positive link between these factors and firms' innovative activity (see Levin *et al.*, 1987; Salomon and Shaver, 2005; Liu and Buck, 2007; Raymond *et al.* 2009).

[6] To a large extent the Italian industrial classification codes (ATECO) correspond to the European NACE taxonomy.

internal and external sources of information in determining the innovation choices of a firm (see Amara and Landry, 2005).

For reasons of identification the econometric method adopted requires an exclusion restriction. Accordingly, we decide to exclude from equations (3) and (4) the firm size variable and the variable that indicates whether the international market is the firm's most significant market. For the former, the choice was primarily motivated by the fact that the dependent variables, being expressed in intensities, are implicitly scaled for size, and is further supported by the results of previous studies. For example, Griffith *et al.* (2006) find that in several European countries, firm size significantly affects the probability of engaging in R&D but not the level of R&D investment. Similarly, for the latter, several contributions have shown a positive and significant causal effect of different international competition indicators on the firm's probability to innovate but not on the level of investment in R&D activities (see Salomon and Shaver, 2005; Griffith *et al.* 2006; Liu and Buck, 2007).

3.2 Innovation outputs: product and process innovation

We model the KPF considering two types of innovative output: process and product innovation. Formally, the two equations can be written as follows:

$$(5)\ \ \mathrm{PROD}_i^* = \alpha_3'\widehat{\mathrm{RDT}}_i + \beta_3'\widehat{\mathrm{TAT}}_i + \gamma' x_{3i} + \varepsilon_{3i}$$

$$(6)\ \mathrm{PROC}_i^* = \alpha_4'\widehat{\mathrm{RDT}}_i + \beta_4'\widehat{\mathrm{TAT}}_i + \pi' x_{4i} + \varepsilon_{4i}$$

where $\widehat{\mathrm{RDT}}_i$ and $\widehat{\mathrm{TAT}}_i$ represent the predictions of the dependent variables of equations (3) and (4), conditional on the firm's decision to engage in innovation activities. In this case too, we do not observe the level of knowledge generated by the firm, having information only on whether the firm has realised product and/or process innovation. Accordingly, if we indicate PROD_i PROC_i the two dummy variables that single out the realization of these events, we will have:

$$(5a)\ \Pr[\mathrm{PROD}_i = 1] \quad = \Pr\left[\alpha_3'\ \widehat{\mathrm{RDT}}_i + \beta_3'\widehat{\mathrm{TAT}}_i + \gamma' x_{3i} + \varepsilon_{3i} > 0\right]$$

$$(6a)\ \Pr[\mathrm{PROC_i} = 1] \quad = \Pr\left[\alpha_4' \,\widehat{\mathrm{RDT_i}} + \beta_4' \widehat{\mathrm{TAT_i}} + \pi' x_{4i} + \varepsilon_{4i} > 0\right]$$

Assuming that the two error terms ε_{3i} and ε_{4i} follow a bivariate normal distribution and are correlated with correlation coefficient $\rho_{\varepsilon_{34}}$, equations (5) and (6) define a bivariate Probit model, and are jointly estimated by maximum likelihood in Stata. Apart from firm size (expressed in the logarithm) and the set of industry-specific dummies, the vectors γ and π include two dummies denoting those firms that have realised managerial, strategic or organisational innovation ('IORG'), and those that have implemented changes in marketing concepts or strategies ('IMARK'). The occurrence of other forms of innovation should be complementary to the two innovative inputs considered in the specification (see Bresnahan *et al.*, 2002; Hitt and Brynjolfsson, 2002; Piva *et al.*, 2005).

As already stated, by using the predictions for innovation inputs instead of the reported values, we are able to estimate the knowledge production function using the whole sample. In this way, the number of observations is increased and selectivity bias is avoided. Moreover, as long as the variables in equations (1)-(3) and (2)-(4) are exogenous, such an approach allows us to control for any possible endogeneity of the innovative inputs. In this respect, it is very likely that unobservable characteristics included in the error terms ε_{3i} and ε_{4i} are important in increasing both the firms' innovative efforts and their propensity to introduce new innovations. This would cause an upward-biased estimate of the parameters α_3, β_3 and α_4, β_4 because of their positive correlation with ε_{3i} and ε_{4i} respectively.

4. Data

This work uses firm-level data drawn from the third and fourth (CIS3 and CIS4) waves of the Italian CIS. The CIS is a harmonized survey that is carried out by national statistical agencies (ISTAT in Italy) in all 27 EU Member States, and is coordinated by Eurostat. CIS3 was conducted in 2001 and provides information for the three-year period 1998-2000, while CIS4 was conducted in 2005 and provides information for the three-year period 2002-2004. These surveys are representative at both sector- and firm-size level of the entire population of Italian firms with more than 10 employees. In conducting the surveys, ISTAT adopts a weighting procedure that relates the sample of firms interviewed to the entire population[7] (ISTAT, 2004).

As previously mentioned, the way in which the Italian CIS questionnaire is structured allows us to have only a limited amount of information regarding all the interviewed firms. In particular, all firms are requested to answer some questions providing general information, such as number of employees, main industry of affiliation, whether they belong to a group and whether they have undertaken innovation activities (completed/ongoing/abandoned) or not. Only those firms declaring themselves innovative are asked to answer a much larger set of additional questions covering among other things their innovativeness, the effects of innovation, participation in cooperative innovation activities and access to public funding. Most of the information is available in both datasets, although some differences between the two questionnaires can be detected; specifically, with respect to the variables that we have used for the estimations, while CIS3 gathers detailed information regarding the formal methods of protecting innovation, CIS4 provides information only on whether a firm has applied for a patent or not.

The original CIS3 and CIS4 databases were made up of respectively 15,512 (CIS3) and 21,854 (CIS4) firms operating in all the sectors of economy. After dropping those firms not operating in the manufacturing sectors, those employing more than 5,000 employees and those declaring a level of R&D expenditures and/or TA higher

---

[7] Firm selection was carried out using a 'one step stratified sample design'. The sample in each stratum was selected with equal probability and without re-inclusion. The stratification of the sample was based on the following three variables: firm size, sector, regional location. Technically, in the generic stratum h, the random selection of n_{h} sample observations among the N_{h} belonging to the entire population was realized through the following procedure:
- a random number in the 0-1 interval was attributed to each N_h population unit;
- N_h population units were sorted by increasing values of the random number;
- units in the first n_h positions in the order previously mentioned were selected.
The estimates obtained from the selected sample are very close to the actual values in the national population. The weighting procedure follows Eurostat and the Oslo Manual (OECD, 1997) recommendations: weights indicate the inverse of the probability that the observation is sampled. Therefore, sampling weights ensure that each group of firms is properly represented and correct for sample selection. Moreover, sampling weights help reduce the heteroskedasticity commonly arising when analysis focuses on survey data. It is important to note that this sample weighting was carried out *ex-ante* by ISTAT in the process of providing the original data, and therefore does not imply any cleaning procedure by the authors.

than 50% of the total turnover[8], we ended up with 7,185 (CIS3) and 7,329 (CIS4) innovating and non-innovating firms.

In accordance with the specific aim of this paper it was necessary to single out a given age threshold in order to select the two sub-samples of young and mature firms from the total samples. In line with previous works (see Garcia-Quevedo *et al.* 2011 and Pellegrino *et al.* 2012) and following the general criteria that the European Commission used to single out the YICs, we opted for an 8-year threshold[9]. Table 1 shows the sectoral composition of the total sample, distinguishing between young and mature firms. The overall impression is that no striking distribution differences between the various samples (both total samples and sub-samples of young and mature firms) emerge across the different industry categories[10]. As far as size is concerned, young firms appear to be, with a slightly different degree in the two CIS, smaller than their mature counterparts[11].

< INSERT TABLE 1 >

Table 2 gives the summary statistics (mean and standard deviation) for the dependent variables and the regressors used in the model (see Table A1 in the Appendix for a detailed definition of the variables and Tables A2 and A3 for correlation matrix). In this case too, CIS3 and CIS4 samples generally look very similar, the only notable difference being in the higher percentage of CIS3 firms introducing product innovation (28% *vs* 19%). However, on looking at the two sub-samples of young and mature firms,

---

[8] We assume that young firms cannot hire more than 5,000 employees. Furthermore, it is quite unreliable that companies invest more than 50% of their turnover in R&D and/or TA. Therefore, we prefer not to include potentially biased observations.

[9] According to the European Commission's State Aid rules, Young Innovative Companies are defined as being less than 6 years old, among other requirements. However, in adopting the European Directive some European countries have extended this threshold (i.e. 8 years for France and Estonia). The choice of 8 years allows us to reach a good degree of representativeness of the sub-sample of young firms, without increasing the age threshold too much. However, we performed several robustness checks, assuming the alternative thresholds of 6, 7, 9 and 10 years; results – available upon request – are consistent, both in terms of sign and statistical significance of the estimated coefficients, with those discussed in Section 5.

[10] To aggregate the industry categories in accordance with the 2-digit NACE classification, we follow Griffith *et al.*(2006)

[11] Looking at CIS4 values, one would expect, also in this wave, young firms to be much smaller than mature counterparts. In this case, additional robustness checks were performed using a CIS4 sample cleaned of potential borderline outliers as far as young firms are concerned. The results – available upon request – are in line (in terms of both sign and statistical significance) with those discussed in Section 5.

more evident differences can be detected. In particular, young firms appear to be less innovative with respect to both the intensity of the innovative effort and the firm’s capacity to realise process and product innovations. Furthermore, it seems that the use of appropriability means (both formal and strategic) increases with age as well as the degree of international market exposure.

< INSERT TABLE 2 >

5. Results

In the following two sub-sections we comment on the estimation results of the 6 equation models outlined in Section 3 (and Tables A2 and A3 for correlation matrix. For each step of the model we present the results for the entire samples (CIS3 and CIS4) and for the four sub-samples of mature incumbents and young firms. Accordingly, in discussing the results, we will consider possible differences both between sub-samples of firms belonging to different datasets, and between mature and young firms belonging to the same dataset.

Before moving on to the discussion, it is important to note that our estimations are based on cross-sectional data, and most of the regressors used are simultaneously determined; therefore interpretation of the results has to be undertaken with caution.

5.1 Innovation inputs

Tables 3 and 4 show the estimation results for, respectively, R&D (equations (1) and (3)), and TA (equations (2) and (4)). We first concentrate our attention on possible differences regarding the role played by factors in determining the innovation choice (R&D and TA) of the firms, in terms of both whether or not to engage in innovative activities, and how intensively to invest in the same innovative activities. More specifically, we look at the results of the selection and main equations for the two different innovative inputs (R&D and TA), focusing our attention only on the total samples of the two datasets (columns 1 – 2 and 7 – 8 in Tables 3 and 4).

A first notable result is that generally the sign and the significance of the coefficients are quite similar across the two different waves. This means that our results are robust across different samples of firms over different time periods. If we compare the results of the two input equations, the most evident difference is related to the level of significance of the variable ‘COOP’. Indeed, it appears that those firms that take part into cooperative activities are more likely to increase the intensity of their investment in R&D activities but not in TA. This result could reflect the vital role played by some cooperation partners (in particular universities, private and public research institutes) in determining the firm’s R&D effort. Apart from this result, no other relevant differences can be detected between the two equations. In particular, looking at the other factors that are exclusively included in the level equations, the use of any type of information source regarding innovation (both internal and external) turns out to be insignificant in determining firms’ levels of investment in both innovative input categories. On the contrary, it appears that those firms that benefit from any type of support for their innovation activities are more likely to spend more on R&D and TA.

As for the factors included in both selection and level equations, we can see that being part of a group does not seem to be an important driver of either R&D or TA activities. Indeed, the coefficients of the variable ‘IG’, with the exception of the R&D selection equation referred to CIS4 (column 7 in Table 3), turn out to be insignificant in all the models. On the contrary, those firms that have made use of appropriability means (both formal and strategic), seem to be more likely to engage in both types of innovative activity. Moreover, the formal methods to appropriability (variable ‘PATDEP’) appear to have an important role, again enhancing the level of investment in both R&D and TA. Finally, looking at those variables included only in the selection equations, we can see that larger firms, and firms that are more oriented towards international markets, are also more likely to engage in both types of innovation activity.

We now move on to the comparison of mature and young firms, here describing the estimation results of the remaining columns (3 – 4 – 5 – 6 – 9 – 10 – 11 and 12) in Tables 3 and 4. Firstly, with the exception of some slight differences (i.e. variable ‘COOP’ in the R&D equation, significant for the CIS4 sample of young firms but not for their CIS3 counterparts), the results again are pretty much consistent across the different samples/sub-samples of firms over different time periods. Moreover, looking at the two different sub-samples of mature and young firms, some results are in line with those regarding the total samples. More specifically, the variable ‘IG’, with the exception of the selection equation in CIS4 (where the coefficient is positive and slightly significant), does not affect the two different form innovative decisions.

Furthermore, firm size and international market exposure appear to be important factors boosting a firm's probability of engaging in both R&D and TA, regardless of the age of the firm[12]. In addition, both mature and young firms (but especially mature firms) that cooperate in innovation activity are likely to spend more on R&D but not on TA. Moreover, in line with previously discussed results, the variable 'SUPPORT' appears to play an important role in determining the level of R&D investment in both sub-samples and for both datasets. However, this variable turns out to be still highly significant in the TA equations for the sample of mature firms only. This result, which holds across the two different datasets, could suggest the need to design different policy measures to support different innovative activities (R&D *vs* TA) for different cohorts of firms (mature *vs* young). Another important difference in terms of the relevance of innovative drivers between the two sub-samples is related to the sign and significance of the two dummy variables denoting those firms that make use of any type of internal and external source of information for innovation activities. Indeed, looking at the R&D equation, in both datasets the variable 'INFO_IN' has an important role in boosting the intensity of young firms' investments, but in the case of mature firms it appears to be irrelevant. As for the variable identifying those firms that make use of external sources of information for their innovation activities, as can be seen from Table 3 young firms in CIS3, in contrast with their mature counterparts, seem to be negatively affected by this factor with respect to their R&D intensity decision. Instead, turning our attention to the TA equation (Table 4), this variable appears to increase significantly the level of investment in TA among young firms, but not among mature ones[13]. All in all, this important evidence suggests that: 1) young firms tend to show a higher level of sensitivity to different sources of information regarding innovation with respect to their mature counterparts when they have to decide how much to invest in innovative activities (both R&D and TA); 2) different sources of information (internal *vs* external) have a distinct impact in determining the level of investment in R&D and TA as far as young firms are concerned.

Finally, as for the means of appropriability, while the variable 'PROT' (strategic method of protection) has almost no impact on the amount of a firm's investment in

---

[12] The only result that appears to contrast regards the insignificance of the variable EXP_d in the TA equation for the sample of young firms from CIS3.
[13] This result holds true only with reference to CIS3.

innovation, the use of formal methods of protection (variable ‘PATDEP’) turns out to be highly significant in both the main equations and across the two datasets for mature firms only.

Both the high values of the correlation coefficients (Rho) between the selection and the main equations and the statistical significance of the Lambda Mills ratio in 11 out of 12 models (see the lower parts of Tables 3 and 4) confirm the validity of the choice of this Heckman-type specification.

< INSERT TABLES 3 AND 4 >

5.2 Innovation outputs

Table 5 reports the econometric results of the KPF, considering both product and process innovation. Specifically, as for the two input equations, we report the results for the three different samples (total, mature and young) for both datasets: CIS3 (first 6 columns) and CIS4 (last 6 columns). The numbers reported are marginal effects evaluated at the sample means. The use of predicted variables ($\widehat{\mathrm{TAT}}$ and $\widehat{\mathrm{RDT}}$) as regressors makes the usual standard errors invalid. Accordingly, in Table 5 we report the t-statistics calculated using the bootstrapped standard errors. Following the structure of the previous subsection, we first concentrate our attention on the general results (total samples) and then on possible differences between mature and young firms.

The first important result, in line with most of the related literature (see Section 2), is that R&D appears to be more important for product innovation than for process innovation. This result is particularly evident with respect to CIS4. As can be seen (columns 7 – 8), the effect of the variable $\widehat{\mathrm{RDT}}$ is highly significant for product innovation and not significant for process innovation. Instead, this evidence is less clear in CIS3, where the impact of the variable is equally statistically significant for both innovative outputs. However, as can be seen, the magnitude of its effect is much stronger on product than on process innovation (0.81*vs* 0.33).

On the other hand, investment in TA is important for both types of innovation and in both the CIS3 and the CIS4 samples, the variable $\widehat{\mathrm{TAT}}$ being always highly significant. However, looking at the magnitude of marginal effects, we can see that this particular innovative input appears to be more important for process innovation.

Looking at the other regressors, we notice that the two dummy variables (‘IORG’ and ‘IMARK’) identifying those firms that realised ‘wider’ innovation activities, turn out to be always positive and significant, with the variable ‘IMARK’ appearing more important for product innovation. This result is in line with our expectations, since the implementation of marketing concepts is more related to the realisation of product innovation than process innovation. Finally, the sign and the level of significance of the marginal effects of the variable ‘LSIZE’ suggest that larger firms are more likely to engage in both product and process innovation.

Turning our attention to the 4 sub-samples of young and mature firms, the overall impression is that the estimate results are pretty much in line with those previously discussed for both groups of firms. The only noteworthy evidence is that the variable $\widehat{\mathrm{RDT}}$ in one case (CIS4 dataset) is important in increasing the likelihood of process innovation for young firms only. This result could be related to the fact that young firms, being less experienced than their mature counterparts and possibly less specialised with respect to their innovative process, are more able to exploit the interaction between different innovative inputs in order to pursue the realisation of different innovative outputs at the same time. However, this speculation is not fully supported by our results, since the evidence on which it is based does not hold true for the CIS3 dataset. In the latter case, the variable $\widehat{\mathrm{RDT}}$ appears to be highly significant for both mature and young firms.

As far as the impact of the variable $\widehat{\mathrm{TAT}}$ is concerned, from the estimation results it is quite evident that the level of investment in TA is equally important for both types of innovation without any particular difference between mature and young firms. The impact of this variable again appears to be more important in determining the realisation of process innovation, and this is particularly evident with respect to the CIS4 sample. Similarly, the marginal effects and the level of significance of the remaining variables (‘IORG’, ‘IMARK’ and ‘LSIZE’) are in line with those of the total sample. In this case too, for both young and mature firms the realisation of changes in marketing or strategies (variable ‘IMARK’) is more important for product than for process innovation.

Finally, from the lower part of Table 5 it emerges clearly that the two equations are always highly correlated via the errors terms, the level of the rho ranging between 0.46 and 0.74. This aspect, which suggests the existence of a certain degree of complementarity between the two innovative output, supports the adoption of a Biprobit model.

< INSERT TABLE 5 >

6. Conclusion

Based on an extension of a traditional partial-structure CDM model, this study has analysed the determinants of firms' innovative effort and the results of this effort in terms of innovative outputs, by looking at R&D/TA and PROC/PROD and distinguishing between mature and young firms. Using data from the third and fourth Italian Community Innovation Surveys we have estimated a structural model that allows for the fact that some firms may undertake innovation without reporting it as R&D and/or TA. We capture some interesting results, which turn out to be robust across different samples of firms over different time periods. We sum them up to emphasize them:

1) regarding the impact of the different drivers in determining firms' decisions to invest or not in R&D and TA, no particular differences between mature and young firms can be detected. More specifically, apart from the variable denoting those firms that belong to an industrial group, all the other factors (appropriability conditions, international market exposure and size) turn out to be important in increasing the probability of investing both in R&D and TA for both sub-samples of firms;
2) different firm and market characteristics have different impacts on the level of investment in R&D/TA, both in general and for mature vs. young firms. In this respect, while the variable SUPPORT plays an important role in increasing the level of investment in R&D in both sub-samples and for both datasets, in the TA equation this variable turns out to be still highly significant only for the group of mature firms. Another important result is related to the fact that young firms show a higher level of sensitivity to internal and external sources of innovation with respect to their mature counterparts when they have to decide how much to invest in the two innovative inputs. Moreover, it seems that these two different sources of information have different impacts in determining the level of investment in

R&D and TA as far as young firms are concerned. Finally, the variable that indicates the use of formal methods of protection of innovation activities turns out to be highly significant in both R&D and TA equations and across the two datasets for the mature firms only.

3) No particular differences between young and mature firms emerge in the KPF. Although R&D and TA both appear to be important in increasing the likelihood of introducing both product and process innovation, when looking at the marginal effects, it appears that R&D is more linked to product innovation and TA to process innovation.

Overall, considering the distinction between different innovative inputs and how they impact on product and process innovation (KPF), the split among young and mature firms allows to better grasp partially different dynamics between the two groups, suggesting YICs have a greater intrinsic attention to sources of information/innovation, meanwhile they are less prone to formal property rights than their mature counterparts. As far as CIS3 and CIS4 are considered, it seems that there is not a cycle-effect in the time-periods considered.

References

Acs, Z.J., Audretsch, D.B., 1987. Innovation, Market Structure, and Firm Size. The Review of Economics and Statistics 69, 567–574.

Acs, Z.J., Audretsch, D.B., 1990. Innovation and Small Firms. MIT Press, Cambridge, MA.

Amara, N., Landry, R., 2005. Sources of Information as Determinants of Novelty of Innovation in Manufacturing Firms: Evidence from the 1999 Statistics Canada Innovation Survey. Technovation 25, 245–259.

Amemiya, T., 1984. Tobit Models: A Survey. Journal of Econometrics 24, 3–61.

Archibugi, D., Pianta, M., 1992. The Technological Specialization of Advanced Countries: A Report to the EEC on International Science and Technology Activities, Kluwer Academic Publisher. Boston.

Audretsch, D.B., 1995. Innovation, Growth and Survival. International Journal of Industrial Organization 13, 441–457.

Bartelsman, E.J., Haltiwanger, J., Scarpetta, S., 2004. Microeconomic Evidence of Creative Destruction in Industrial and Developing Countries. Tinbergen Institute Discussion Paper 04-114/3.

Birch, D., 1979. The Job Generation Process: Final Report to Economic Development Administration. Cambridge, MA: MIT Program on Neighborhood and Regional Change.

Bresnahan, T.F., Brynjolfsson, E., Hitt, L.M., 2002. Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence. The Quarterly Journal of Economics 117, 339–376.

Brüderl, J., Preisendörfer, P., 2000. Fast-Growing Businesses: Empirical Evidence from a German Study. International Journal of Sociology 30, 45–70.

Busom, I., 2000. An Empirical Evaluation of the Effects of R&D Subsidies. Economics of Innovation and New Technology 9, 111–148.

Cassiman, B., Veugelers, R., 2002. R&D Cooperation and Spillovers: Some Empirical Evidence from Belgium. The American Economic Review 92, 1169–1184.

Conte, A., Vivarelli, M., 2005. One or Many Knowledge Production Functions? Mapping Innovative Activity Using Microdata. IZA Discussion Paper 1878.

Crepon, B., Duguet, E., Mairesse, J., 1998. Research, Innovation and Productivity: an Econometric Analysis at the Firm Level. Economics of Innovation and New Technology 7, 115–158.

Criscuolo, P., Nicolaou, N., Salter, A., 2012. The Elixir (or Burden) of Youth? Exploring Differences in Innovation between Start-ups and Established Firms. Research Policy 41, 319–333.

Foster, L., Haltiwanger, J., Krizan, C., 2001. Aggregate Productivity Growth: Lessons from Microeconomic Evidence, in: Ernst, Ganiatsos, T., Mytelka, L. (Eds), New Contributions to Productivity Analysis. University of Chicago Press, Chicago.

Garcia-Quevedo, J., Pellegrino, G., Vivarelli, M., 2011. R&D Drivers in Young Innovative Companies. IZA Discussion Paper 6136.

González, X., Jaumandreu, J., Pazó, C., 2005. Barriers to Innovation and Subsidy Effectiveness. The RAND Journal of Economics 36, 930–950.

Griffith, R., Huergo, E., Mairesse, J., Peters, B., 2006. Innovation and Productivity across Four European Countries. Oxford Review of Economic Policy 22, 483–498.

Griliches, Z., 1979. Issues in Assessing the Contribution of Research and Development to Productivity Growth. The Bell Journal of Economics 10, 92–116.

Hall, B.H., Lotti, F., Mairesse, J., 2013. Evidence on the Impact of R&D and ICT Investments on Innovation and Productivity in Italian Firms. Economics of Innovation and New Technology 22, 300–328.

Hall, B.H., Mairesse, J., Mohnen, P., 2010. Measuring the Returns to R&D, in: Hall, B.H., Rosenberg, N. (Eds), Handbook of the Economics of Innovation. National Bureau of Economic Research, Amsterdam and New York.

Heckman, J.J., 1979. Sample Selection Bias as a Specification Error. Econometrica 47, 153–161.

Hitt, L.M., Brynjolfsson, E., 2002. Information Technology, Organizational Transformation, and Business Performance. A Transatlantic Perspective., in: Greenan, N., L'Horty, Y., Mairesse, J. (Eds), Productivity, Inequality, and the Digital Economy. MIT Press, Cambridge, MA.

Huergo, E., Jaumandreu, J., 2004. How Does Probability of Innovation Change with Firm Age? Small Business Economics 22, 193–207.

ISTAT, 2004. Statistiche sull'Innovazione delle Imprese: Settore Industria. Anni 1998 - 2000. Rome.

Kleinknecht, A., Van Montfort, K., Brouwer, E., 2002. The Non-Trivial Choice between Innovation Indicators. Economics of Innovation and New Technology 11, 109–121.

Levin, R.C., Klevorick, A.K., Nelson, R.R., Winter, S.G., Gilbert, R., 1987. Appropriating the Returns from Industrial Research and Development. Brookings Papers on Economic Activity 3, 783–831.

Liu, X., Buck, T., 2007. Innovation Performance and Channels for International Technology Spillovers: Evidence from Chinese High-tech Industries. Research Policy 36, 355–366.

Mairesse, J., Mohnen, P., 2010. Using Innovations Surveys for Econometric Analysis. NBER working paper 15857.

Mairesse, J., Robin, S., 2009. Innovation and Productivity: a Firm-level Analysis for French Manufacturing and Services Using CIS3 and CIS4 Data (1998–2000 and 2002–2004). Mimeo.

Moncada-Paterno-Castello, P., 2011. Companies' Growth in the EU: What is Research and Innovation Policy's Role? Institute for Prospective and Technological Studies Working Paper 2011-03.

OECD, 1997. Oslo Manual: the Measurement of Scientific and Technological Activities. Proposed Guideline for Collecting and Interpreting Technological Innovation Data. OECD, Paris.

Ortega-Argilés, R., Piva, M., Potters, L., Vivarelli, M., 2010. Is Corporate R&D Investment in High-Tech Sectors More Effective? Contemporary Economic Policy 28, 353–365.

Ortega-Argilés, R., Vivarelli, M., Voigt, P., 2009. R&D in SMEs: a Paradox? Small Business Economics 33, 3–11.

Patel, P., Pavitt, K., 1993. Patterns of Technological Activity: Their Measurement and Interpretation, in: Handbook of the Economics of Innovation and Technological Change. Blackwell Publishers, Oxford.

Pellegrino, G., Piva, M., Vivarelli, M., 2012. Young Firms and Innovation: a Microeconometric Analysis. Structural Change and Economic Dynamics 23, 329-340.

Philippon, T., Veron, N., 2008. Financing Europe's Fast Movers. Bruegel Policy Brief 2008/01.

Piga, C.A., Vivarelli, M., 2003. Sample Selection in Estimating the Determinants of Cooperative R&D. Applied Economics Letters 10, 243–246.

Piga, C.A., Vivarelli, M., 2004. Internal and External R&D: a Sample Selection Approach. Oxford Bulletin of Economics and Statistics 66, 457–482.

Piva, M., Santarelli, E., Vivarelli, M., 2005. The Skill Bias Effect of Technological and Organisational Change: Evidence and Policy Implications. Research Policy 34, 141–157.

Polder, M., Leeuwen, G. van, Mohnen, P., Raymond, W., 2010. Product, Process and Organizational Innovation: Drivers, Complementarity and Productivity Effects. MPRA Paper No. 23719.

Raymond, W., Mohnen, P., Palm, F.C., Loeff, S.S.V. der, 2009. Innovative Sales, R&D and Total Innovation Expenditures: Panel Evidence on Their Dynamics. CIRANO Working Papers 2009s- 29.

Romer, P.M., 1986. Increasing Returns and Long-Run Growth. Journal of Political Economy 94, 1002–1037.

Salomon, R.M., Shaver, J.M., 2005. Learning by Exporting: New Insights from Examining Firm Innovation. Journal of Economics & Management Strategy 14, 431–460.

Santamaría, L., Nieto, M.J., Barge-Gil, A., 2009. Beyond Formal R&D: Taking Advantage of Other Sources of Innovation in Low- and Medium-Technology Industries. Research Policy 38, 507–517.

Schneider, C., Veugelers, R., 2010. On Young Highly Innovative Companies: Why They Matter and How (not) to Policy Support Them. Industrial and Corporate Change 19, 969–1007.

Solow, R.M., 1956. A Contribution to the Theory of Economic Growth. The Quarterly Journal of Economics 70, 65–94.

Stoneman, P., 1995. Introduction, in: Stoneman, P. (Ed), Handbook of the Economics of Innovation and Technological Change. Blackwell Publishers, Oxford.

Table 1. Sectoral composition (2-digit manufacturing sectors) and average employment; all firms - mature firms; young firms (CIS3 - CIS4)

	CIS3						CIS4					
	TOT		MATURE		YOUNG		TOT		MATURE		YOUNG	
	N	%	N	%	N	%	N	%	N	%	N	%
Food	492	6.9	446	7.2	46	4.9	638	8.7	550	9.2	88	6.6
Textile	1,191	16.6	995	16.0	196	20.7	1,172	16.0	895	15.0	277	20.7
Wood/Paper	986	13.7	882	14.1	104	11.0	907	12.4	756	12.6	151	11.3
Chemicals	494	6.9	435	7.0	59	6.2	460	6.3	382	6.4	78	5.8
Plastic/Rubber	415	5.8	361	5.8	54	5.7	310	4.2	266	4.4	44	3.3
Non-metallic Min.	471	6.6	414	6.6	57	6.0	504	6.9	432	7.2	72	5.4
Basic metals	853	11.9	741	11.9	112	11.8	1,306	17.8	1,054	17.6	252	18.8
Machinery	551	7.7	483	7.8	68	7.2	566	7.7	472	7.9	94	7.0
Electrical	826	11.5	700	11.2	126	13.3	671	9.2	544	9.1	127	9.5
Vehicles	395	5.5	325	5.2	70	7.4	364	5.0	286	4.8	78	5.8
Misc.	511	7.1	454	7.3	57	6.0	431	5.9	351	5.9	80	6.0
Total	7,185	100	6,236	100	949	100	7,329	100	5,988	100	1,341	100
Av. Emp.	85		90		56		102		102		98	

Table 2. Descriptive statistics: mean and standard deviation of the variables; all firms- mature firms- young firms (CIS3 –CIS4)

	CIS3			CIS4		
	TOT.	MATURE	YOUNG	TOT.	MATURE	YOUNG
RDT_d	0.21 (0.41)	0.22 (0.41)	0.15 (0.36)	0.25 (0.44)	0.26 (0.44)	0.22 (0.41)
RDT	0.048 (1.78)	0.049 (1.79)	0.041 (1.75)	0.066 (2.42)	0.067 (2.43)	0.061 (2.36)
TAT_d	0.29 (0.45)	0.30 (0.46)	0.21 (0.41)	0.34 (0.47)	0.34 (0.47)	0.30 (0.46)
TAT	0.011 (3.70)	0.011 (3.63)	0.011 (4.15)	0.012 (3.76)	0.012 (3.67)	0.013 (4.13)
PROD	0.28 (0.45)	0.29 (0.45)	0.23 (0.42)	0.19 (0.39)	0.19 (0.40)	0.17 (0.38)
PROC	0.29 (0.45)	0.30 (0.46)	0.25 (0.43)	0.30 (0.46)	0.31 (0.46)	0.28 (0.45)
IG	0.18 (0.38)	0.18 (0.39)	0.17 (0.37)	0.23 (0.42)	0.23 (0.42)	0.23 (0.42)
PATDEP	0.12 (0.32)	0.12 (0.33)	0.08 (0.27)	0.12 (0.32)	0.12 (0.33)	0.09 (0.29)
PROT	0.21 (0.41)	0.22 (0.41)	0.15 (0.35)	0.18 (0.39)	0.19 (0.39)	0.15 (0.36)
COOP	0.05 (0.23)	0.06 (0.23)	0.04 (0.19)	0.06 (0.24)	0.06 (0.24)	0.06 (0.24)
SUPPORT	0.20 (0.40)	0.21 (0.40)	0.16 (0.36)	0.18 (0.38)	0.18 (0.39)	0.15 (0.36)
INFO_IN	0.13 (0.33)	0.13 (0.34)	0.09 (0.29)	0.14 (0.35)	0.14 (0.35)	0.12 (0.33)
INFO_EX	0.18 (0.38)	0.19 (0.39)	0.13 (0.33)	0.20 (0.40)	0.20 (0.40)	0.19 (0.39)
EXP	0.66 (0.47)	0.68 (0.47)	0.58 (0.49)	0.53 (0.50)	0.55 (0.50)	0.43 (0.50)
LSIZE	3.64 (1.09)	3.69 (1.10)	3.33 (0.94)	3.67 (1.15)	3.71 (1.15)	3.50 (1.15)
IORG	0.47 (0.50)	0.47 (0.50)	0.45 (0.50)	0.35 (0.48)	0.35 (0.48)	0.34 (0.47)
IMARK	0.47 (0.50)	0.47 (0.50)	0.42 (0.49)	0.21 (0.41)	0.22 (0.41)	0.20 (0.40)
Obs.	7,185	6,239	949	7,329	5,988	1,341

Standard deviation in brackets

Table 3. Estimation results for the R&D equations (CIS3 - CIS4)

	CIS3						CIS4					
	TOT.		MATURE		YOUNG		TOT.		MATURE		YOUNG	
Dep. Var.	RDT_d	RDT	RDT_d	RDT	RDT_d	RDT	RDT_d	RDT	RDT_d	RDT	RDT_d	RDT
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
IG	0.04	-0.02	0.03	0.02	0.16	-0.61	0.13***	-0.03	0.12**	-0.16	0.21*	0.62
	(0.78)	(-0.08)	(0.48)	(0.09)	(1.01)	(-0.76)	(2.83)	(-0.12)	(2.28)	(-0.64)	(1.68)	(0.98)
PATDEP	0.70***	1.08***	0.67***	1.00***	0.95***	1.04	0.73***	1.51***	0.66***	1.47***	1.17***	1.20
	(11.23)	(3.93)	(10.32)	(3.56)	(4.27)	(0.92)	(12.55)	(4.72)	(10.63)	(4.37)	(7.17)	(1.31)
PROT	0.35***	0.39*	0.35***	0.42*	0.34*	-0.08	0.29***	0.46*	0.30***	0.32	0.27**	1.26*
	(6.77)	(1.69)	(6.42)	(1.77)	(1.84)	(-0.09)	(6.04)	(1.87)	(5.61)	(1.23)	(2.08)	(1.94)
COOP		0.86***		0.94***		-0.07		1.16***		0.90***		2.37***
		(4.18)		(4.42)		(-0.10)		(4.98)		(3.57)		(4.13)
SUPPORT		1.26***		1.12***		2.91***		0.90***		0.76***		1.43***
		(7.92)		(6.72)		(5.23)		(4.82)		(3.79)		(3.17)
INFO_IN		0.19		0.08		1.61***		0.25		0.05		1.09**
		(1.16)		(0.48)		(2.79)		(1.32)		(0.25)		(2.43)
INFO_EX		0.06		0.15		-1.15**		-0.05		-0.01		-0.53
		(0.36)		(0.87)		(-2.07)		(-0.30)		(-0.05)		(-1.16)
EXP	0.41***		0.42***		0.30**		0.48***		0.47***		0.55***	
	(7.90)		(7.56)		(2.21)		(11.94)		(10.57)		(5.52)	
LSIZE	0.33***		0.33***		0.29***		0.23***		0.24***		0.20***	
	(15.45)		(14.67)		(4.39)		(12.07)		(11.33)		(4.22)	
_cons	-2.73***	-3.00***	-2.83***	-2.93***	-2.02***	-2.32	-2.21***	-2.78***	-2.23***	-2.16***	-2.16***	-4.77**
	(-24.97)	(-4.64)	(-23.89)	(-4.30)	(-6.62)	(-1.18)	(-24.12)	(-3.77)	(-22.28)	(-2.73)	(-9.10)	(-2.57)
Lambda	2.34***		2.29***		1.84		2.20***		1.90***		3.20***	
	(7.09)		(6.66)		(1.51)		(5.46)		(4.34)		(3.34)	
Rho	0.66		0.66		0.57		0.52		0.46		0.73	
N	7,185	1,513	6,236	1,366	949	147	7,329	1,859	5,988	1,565	1,341	294

t- statistics in brackets: * Significant at 10%; **5%;***1%. All regressions include industry dummies (results available upon request).

Table 4. Estimation results for the Technological Acquisitions equations (CIS3- CIS4)

	CIS3						CIS4					
	TOT.		MATURE		YOUNG		TOT.		MATURE		YOUNG	
Dep. Var.	TAT_d	TAT	TAT_d	TAT	TAT_d	TAT	TAT_d	TAT	TAT_d	TAT	TAT_d	TAT
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
IG	-0.05	-0.43	-0.06	-0.35	0.05	-0.42	0.07	0.17	0.07	-0.01	0.05	0.60
	(-1.04)	(-1.20)	(-1.17)	(-0.94)	(0.34)	(-0.26)	(1.49)	(0.50)	(1.42)	(-0.04)	(0.45)	(0.59)
PATDEP	0.45***	1.23**	0.41***	1.20**	0.88***	2.25	0.42***	1.49***	0.37***	1.20***	0.71***	3.05*
	(7.53)	(2.39)	(6.55)	(2.38)	(4.12)	(0.69)	(7.51)	(3.24)	(6.15)	(2.63)	(4.70)	(1.86)
PROT	0.34***	0.32	0.37***	0.38	-0.01	-0.19	0.32***	0.97**	0.32***	0.87**	0.30**	0.56
	(7.21)	(0.72)	(7.44)	(0.83)	(-0.03)	(-0.10)	(7.05)	(2.45)	(6.51)	(2.16)	(2.53)	(0.44)
COOP		-0.02		0.04		-1.05		-0.07		-0.35		1.27
		(-0.05)		(0.10)		(-0.59)		(-0.22)		(-1.01)		(1.35)
SUPPORT		0.97***		1.14***		-0.59		1.18***		1.24***		0.82
		(3.81)		(4.42)		(-0.56)		(5.27)		(5.21)		(1.34)
INFO_IN		0.12		0.26		-1.11		0.31		0.16		0.64
		(0.41)		(0.89)		(-0.93)		(1.32)		(0.65)		(1.01)
INFO_EX		0.18		-0.03		2.29**		0.13		0.09		0.33
		(0.68)		(-0.13)		(2.10)		(0.58)		(0.40)		(0.55)
EXP	0.16***		0.16***		0.17		0.36***		0.35***		0.44***	
	(4.00)		(3.60)		(1.58)		(10.08)		(8.77)		(5.07)	
LSIZE	0.24***		0.24***		0.18***		0.16***		0.17***		0.11***	
	(12.52)		(11.85)		(3.02)		(8.88)		(8.64)		(2.58)	
_cons	-1.68***	-2.47**	-1.70***	-2.19*	-1.39***	-6.58	-1.46***	-4.91***	-1.48***	-3.40***	-1.43***	-10.67***
	(-18.52)	(-2.02)	(-17.56)	(-1.77)	(-5.02)	(-1.11)	(-18.02)	(-4.41)	(-16.74)	(-3.02)	(-6.80)	(-2.99)
Lambda	5.34***		5.21***		7.14*		6.42***		5.42***		9.58***	
	(6.42)		(6.15)		(1.77)		(8.53)		(7.02)		(4.31)	
Rho	0.74		0.75		0.78		0.85		0.79		0.98	
N	7,185	2,080	6,236	1,880	949	200	7,329	2,458	5,988	2,054	1,341	937

t- statistics in brackets: * Significant at 10%; **5%;***1%. All regressions include industry dummies (results available upon request).

Table 5. Knowledge Production Function: Product and Process Innovation (CIS3 - CIS 4)

	CIS3						CIS4					
	TOT.		MATURE		YOUNG		TOT.		MATURE		YOUNG	
Dep. Var.	PROD	PROC	PROD	PROC	PROD	PROC	PROD	PROC	PROD	PROC	PROD	PROC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\widehat{RDT}$	0.81***	0.33***	0.83***	0.35***	0.54***	0.30***	0.41***	0.02	0.42***	0.05	0.46***	0.20**
	(9.11)	(5.93)	(8.41)	(6.80)	(3.55)	(3.03)	(7.82)	(0.40)	(7.57)	(1.14)	(3.77)	(2.02)
$\widehat{TAT}$	0.79***	0.84***	0.73***	0.77***	0.32**	0.37***	0.98***	1.27***	0.86***	1.09***	0.42**	0.56***
	(9.09)	(11.33)	(9.22)	(11.22)	(2.28)	(3.11)	(11.52)	(14.61)	(9.81)	(15.46)	(2.43)	(4.16)
IORG	0.37***	0.45***	0.38***	0.44***	0.34**	0.52***	0.34***	0.47***	0.34***	0.47***	0.41***	0.56***
	(8.79)	(11.35)	(8.94)	(10.23)	(2.25)	(4.11)	(8.52)	(14.60)	(7.12)	(11.14)	(3.24)	(5.59)
IMARK	0.63***	0.25***	0.62***	0.24***	0.83***	0.37***	0.61***	0.34***	0.60***	0.33***	0.74***	0.47***
	(13.85)	(6.81)	(12.88)	(5.28)	(6.28)	(3.03)	(14.68)	(8.14)	(10.28)	(7.44)	(5.29)	(3.97)
LSIZE	0.18***	0.17***	0.17***	0.16***	0.22***	0.16**	0.27***	0.27***	0.24***	0.22***	0.27***	0.26***
	(9.06)	(8.44)	(7.73)	(8.81)	(2.70)	(2.28)	(14.07)	(15.93)	(10.24)	(12.77)	(4.79)	(5.97)
_cons	-3.00***	-2.60***	-2.91***	-2.53***	-2.55***	-1.89***	-3.57***	-3.01***	-3.40***	-2.76***	-3.07***	-2.44***
	(-19.93)	(-17.90)	(-21.25)	(-19.91)	(-8.55)	(-6.29)	(-23.39)	(-29.44)	(-20.63)	(-24.46)	(-9.20)	(-10.11)
Rho	0.62		0.61		0.74		0.46		0.47		0.54	
N	7,185		6,236		949		7,329		5,988		1,341	

t- statistics in brackets: * Significant at 10%; **5%;***1%. All regressions include industries dummies (results available upon request).

Appendix

Table A1. The variables: acronyms and definitions

Innovative input variables	
RDT_d	Dummy = 1 if firm's R&D expenditures (both intramural and extramural) are positive
RDT	Total firm's R&D expenditures (both intramural and extramural), normalized by total turnover
TAT_d	Dummy = 1 if firm's expenditures for Technological acquisitions (investment in new machinery and equipment and purchasing of external technology incorporated in licences, consultancies and know-how) are positive
TAT	Total firm's expenditures for technological acquisitions, normalized by total turnover
Innovative output variables	
PROD	Dummy = 1 if the firm has introduced new or significantly improved products
PROC	Dummy = 1 if the firm has introduced new or significantly improved processes
Firm's general characteristics	
IG	Dummy = 1 if the firm belongs to an industrial group
Innovation-related information	
PATDEP	Dummy = 1 if the firm has applied for patents
PROT	Dummy = 1 if the firm adopts instruments of protection of innovation activities other than patents (trademarks, copyright, registration of design)
COOP	Dummy = 1 if the firm takes part in cooperative innovative activities
SUPPORT	Dummy = 1 if the firm has received public support for innovation
INFO_IN	Dummy = 1 if the firm has used any type of internal source of information for its innovation activities
INFO_EX	Dummy = 1 if the firm has used any type of external source of information for its innovation activities
EXP	Dummy =1 if the firm has traded in an international market during the three-year period; 0 otherwise
LSIZE	Log of the total number of firm's employees
IORG	Dummy = 1 if the firm has realized managerial, strategic or organizational innovation
IMARK	Dummy = 1 if the firm has implemented changes in marketing concepts or strategies (e.g. packaging or presentation changes to a product in order to target new markets)

Table A2. Correlation matrix (CIS3; overall sample:7,185 firms)

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
(1)	RDT_d	1																
(2)	RDT	0.52	1															
(3)	TAT_d	0.49	0.22	1														
(4)	TAT	0.09	0.05	0.47	1													
(5)	PROD	0.63	0.34	0.57	0.19	1												
(6)	PROC	0.44	0.19	0.70	0.38	0.52	1											
(7)	IG	0.25	0.09	0.16	-0.01	0.21	0.14	1										
(8)	PATDEP	0.40	0.20	0.27	0.02	0.37	0.21	0.25	1									
(9)	PROT	0.33	0.14	0.25	0.01	0.32	0.19	0.24	0.55	1								
(10)	COOP	0.36	0.25	0.26	0.06	0.31	0.24	0.23	0.24	0.18	1							
(11)	SUPPORT	0.49	0.34	0.51	0.27	0.46	0.49	0.13	0.25	0.20	0.27	1						
(12)	INFO_IN	0.42	0.25	0.35	0.13	0.39	0.36	0.16	0.22	0.19	0.19	0.34	1					
(13)	INFO_EX	0.44	0.24	0.47	0.20	0.48	0.46	0.17	0.24	0.20	0.25	0.39	0.45	1				
(14)	EXP	0.25	0.13	0.17	0.00	0.23	0.14	0.22	0.21	0.27	0.13	0.15	0.14	0.16	1			
(15)	LSIZE	0.38	0.11	0.28	0.01	0.29	0.24	0.51	0.34	0.34	0.27	0.24	0.24	0.26	0.39	1		
(16)	IORG	0.33	0.16	0.32	0.10	0.34	0.30	0.24	0.22	0.24	0.17	0.25	0.20	0.25	0.23	0.33	1	
(17)	IMARK	0.28	0.13	0.25	0.08	0.34	0.23	0.12	0.23	0.32	0.14	0.20	0.18	0.20	0.27	0.23	0.45	1

Table A3. Correlation matrix (CIS4; overall sample:7,329 firms)

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
(1)	RDT_d	1																
(2)	RDT	0.47	1															
(3)	TAT_d	0.63	0.28	1														
(4)	TAT	0.19	0.24	0.46	1													
(5)	PROD	0.69	0.37	0.54	0.17	1												
(6)	PROC	0.55	0.23	0.82	0.42	0.41	1											
(7)	IG	0.26	0.09	0.17	0.00	0.22	0.17	1										
(8)	PATDEP	0.37	0.22	0.25	0.03	0.39	0.19	0.23	1									
(9)	PROT	0.25	0.12	0.21	0.03	0.27	0.18	0.19	0.44	1								
(10)	COOP	0.38	0.29	0.30	0.09	0.36	0.28	0.21	0.26	0.19	1							
(11)	SUPPORT	0.51	0.31	0.54	0.29	0.41	0.53	0.13	0.24	0.18	0.29	1						
(12)	INFO_IN	0.45	0.24	0.46	0.22	0.41	0.42	0.06	0.17	0.13	0.18	0.31	1					
(13)	INFO_EX	0.47	0.23	0.59	0.29	0.41	0.55	0.11	0.18	0.15	0.27	0.42	0.36	1				
(14)	EXP	0.30	0.12	0.23	0.02	0.28	0.20	0.24	0.26	0.25	0.14	0.19	0.16	0.17	1			
(15)	LSIZE	0.36	0.11	0.26	-0.03	0.31	0.25	0.56	0.35	0.29	0.24	0.24	0.14	0.20	0.39	1		
(16)	IORG	0.30	0.15	0.29	0.09	0.29	0.30	0.17	0.19	0.23	0.19	0.21	0.18	0.24	0.17	0.24	1	
(17)	IMARK	0.27	0.11	0.24	0.05	0.29	0.24	0.10	0.18	0.32	0.13	0.17	0.15	0.17	0.19	0.15	0.39	1