Innovation and Productivity of Dutch Firms: A Panel Data Analysis *

M. Vancauteren[†], B. Melenberg[‡], J. Plasmans[§] and R. Bongard[¶]

October 2017

*We gratefully acknowledge support from the Netherlands Organization for Scientific Research (NWO) under their research program "Dynamism of Innovation" and the Research Foundation of Flanders (FWO) for a travel grant. We thank participants of the colloquia on "Dynamism of Innovation" held at Tilburg University (19/12/2007) and at Antwerp University (26/03/2010), a UNU-Merit seminar (University of Maastricht) held on 10/06/2010, the Second Asia-Pacific Innovation Conference held in Singapore (03/05/2011), the Annual Royal Economic Society congress held in Cambridge (26-28/03/2012), the Belgian Entrepreneur Research Day held at Hasselt University (07/05/2012), the CPB seminar, the Hague, held on 11/09/2012, the Institute for Prospective Studies (European Commission), Sevilla (10/04/2013), a seminar held in HEC-Liege (07/06/2013), and the 48th annual conference of the Canadian Economic Association held in Vancouver (28-31/05/2014) for fruitful discussions. In addition, we thank three anonymous referees whose comments helped to improve the paper considerably.

†Corresponding author, mark.vancauteren@uhasselt.be, Hasselt University, (Address: Campus Diepenbeek, Agoralaan Gebouw D, B-3590 Diepenbeek, Belgium, Tel: +32 11 26 86 64 Fax: +32 11 26 87 99) & Statistics Netherlands.

‡CentER, Tilburg University.

§CentER, Tilburg University and University of Antwerp.

¶Swedish Federation of Business Owners.

Innovation and Productivity of Dutch Firms: A Panel Data Analysis

Abstract

We use an extended version of the well-established Crépon, Duguet and Mairesse model (1998, CDM hereafter) to model empirically the innovation-productivity relationship using data for the 2000-2006 period on patent applications by Dutch firms to the European Patent Office. The CDM model disentangles the impact of R&D expenditures on patents and the impact of patents on productivity. A multiple-equation dynamic panel data model of R&D, patent applications or citations and multi-factor productivity (MFP) growth is estimated that suits multiple data distribution properties. We explicitly take into account the role of dynamics and firm-level unobserved heterogeneity in each of the innovation processes and productivity. We find evidence that the output innovation affects productivity positively, which seems to be robust across specifications. We also find that the strong presence of random effects for individual heterogeneity in explaining the R&D-patents relationship is an important driver to innovation. While the estimates of R&D and dynamics depend on whether these unobserved characteristics are taken into account, we find robust evidence on the role of firm size in explaining patent and citation counts.

Keywords: Innovation, Productivity, Panel Data.

JEL codes: C33,D22

1 Introduction

This paper revisits the innovation-productivity relationship based on the empirical research initiated by Crépon, Duguet, and Mairesse (1998) (CDM, hereafter). The basic setup of the CDM model is a structural model generally built as a three-stage econometric model that relates productivity to patent or innovation output, which depends on R&D and other forms of innovation inputs, which in turn is determined by a number of firm, industry, and other market specific factors. CDM type of models are generally estimated using sequential or simultaneous estimators (asymptotic least squares, minimum distance, likelihood-based estimators) offering controls for selectivity and endogeneity and in a recent number of papers the models also account for unobserved heterogeneity. The CDM model has been applied in a number of empirical studies. Based on several variants of the CDM model, these studies generally confirm CDM's finding by showing a positive and robust relationship between innovation and productivity.[1] In this paper, we consider a multiple-equation dynamic panel model of R&D, patent applications or citations and multi-factor productivity (MFP) that suits mixed data distribution properties. In contrast to much of the existing literature, we exploit the importance of dynamics and unobserved heterogeneity in a CDM model taking into account specific features related to sample selection and zero-inflation, in the innovation data. Using the sequential approach in analyzing the CDM model, this has not been investigated yet.

CDM models most often use Community Innovation Survey (CIS, hereafter), for which it was actually designed for, whereby innovation usually takes the form of R&D in the case of inputs and process, product, marketing or organisational innovation in the case of outputs.[2] Indeed, one drawback of CDM studies employing CIS data is that the

[1] A recent survey of results using variants of the CDM model is Hall (2011), Mohnen and Hall (2013) and two CDM special issue papers published in Hall and Kramarz (1998) and Lööf et al. (2016). With a specific reference to Dutch firms, this positive relationship is also found in Klomp and van Leeuwen (2006), Polder *et al.* (2009) and Raymond *et al.* (2010, 2015).

[2] These surveys include CIS I for 1990-1992; CIS II for 1994-1996; CIS III for 1998-2000; CIS IV for

panel dimension is short with information referring to three-year periods to define the qualitative variables, and to the last period of that year for the quantitative variables. Moreover, in different CIS waves not always the same firms are collected. For instance, a possible consequence, when surveys are conducted biannually, is that there is always a one year overlap between two consecutive measures when referring to a three-year period to define the variables.[3] To circumvent these problems, many of the applications of the CDM model have had to use cross-section data and those that use longitudinal data do not always make full use of available panel data estimators.[4] One exception is a recent paper by Raymond *et al.* (2015) who apply the CDM model to a panel based on three biannual waves of the CIS data while accounting for individual effects and persistence effects in innovation and productivity. The authors find persistence in productivity but not for innovation. Results further indicate that controlling for unobserved heterogeneity is important in each equation of the CDM model and that product innovation affects productivity positively. Furthermore, the authors provide evidence of a weak persistence in product innovation but strong persistence in labor productivity levels. While these results are very useful for identifying the causality between product innovation and pro-

2002-2004, and CIS IV.5 for 2004-2006, CIS 2008-14 for the more recent biannually periods. The variables indicating organisational and marketing innovation have been included since 2005. The Oslo Manual defines a product innovation as a new or significantly improved product or service. A process innovation refers to new or significant changes in the way products or processes are produced. Organisatonal innovation refers to new business practices while marketing innovation refers to changes in product design, packaging, placement, and pricing methods.

[3]Not only is the reliability of product and process innovation in CIS under discussion, but also patent data. For instance, a particular critical issue is that not all patent application activities prove to be innovations (for example, Acs and Audretsch, 1989) and, in addition, not all innovations lead to patenting (like in the construction industry, see also Nagaoka et al., 2010). Also, certain technological solutions are likely to be of higher quality than others. For this reason, we use the number of forward citations (i.e., the citations received by a patent) as an indicator of patent quality. We also note that several studies have documented the mutual correlation between product and patents (Brouwer and Kleinknecht, 1999; Duguet and Lelarge, 2006; Vancauteren, 2017; Hall et al., 2013).

[4]For instance, the paper by Klomp and van Leeuwen (2006) model a productivity feedback effect using cross-sectional data. The paper by Lööf and Heshmati (2006) considers dynamics in the model using panel data on two CIS waves but does not control for individual heterogeneity. Similarly, Criscuolo et al. (2009), Hall et al. (2009), Van Leeuwen and Mohnen (2017), Mairesse and Robin (2017) and Hall and Sena (2017) pool several biannual CIS waves but did not incorporate any control for unobserved heterogeneity.

ductivity taking into account unobserved heterogeneity, the proper handling of panel features remains difficult due to the set up of the CIS survey data with the consequence that estimates may be biased in the absence of appropriate correction mechanisms.[5]

While the CDM framework and subsequent studies have provided a solid empirical concept both in a cross-sectional as well as a panel dimension using CIS data, the focus of this study is to empirically enrich a CDM panel model in several ways.

First, we consider patent applications instead of process or product innovation in the CDM model. In fact, the original CDM by Crépon et al. (1998), estimated on French data, used both measures in their analysis. With patent applications, the time of filing or applying for a patent coincides very closely with the time that innovative activities take place within a firm (OECD, 2009). Patents represent the output of the innovation process (for example, due to R&D), measuring a technological invention by which the technological knowledge to other firms is disclosed. This is to be compared with the studies based on the CIS data.[6] We are not aware of any other paper that uses the number of patents and patent citations as innovative output entering a CDM panel data model consisting of a system of equations with endogenous choices of innovation and productivity. One closely related study using patent data is the paper of Hall and Sena (2017). The authors consider a full CDM model on pooled CIS3, CIS4, CIS5, CIS6 and CIS7 data for the UK, whereby in the second stage (modelling the innovation production function as a trivariate probit), the authors consider product and process innovation as well as the role of intellectual property measures as a binary preference choice variable between patenting (formal) and secrecy (informal) appropriation mechanisms at the level

[5]As mentioned, one of the particularities of the innovation survey is that, for each period, product innovation incidence pertains to any year of the three-year period, while the actual share of innovative sales pertains to the last year of the period. As a result, when assessing persistence of innovation, as stressed in Raymond et al. (2015) paper, the lag effect refers to two to four years when the incidence of product innovation is considered and to four years when the share of innovative sales is considered making it difficult to capture precise time lags.

[6]See footnote 3.

of the firm.[7]

Secondly, because of the use of patent applications, we are able to employ a purpose-built data set that is a panel of firms located in the Netherlands with annual data from 2000 through 2006 that includes financial variables, R&D expenditure, patent application counts, forward citations of these patent applications, and patent technology fields. Our sample is constructed such that it includes (almost) all firms, located in the Netherlands, that during the period 2000-2006 applied for one or more patents at the European Patent Office.[8] Our sample also includes as control group firms with zero patent applications.

Third, focusing on the first and second stage of the model, innovation investment may be correlated with the error term in the R&D equation if part of this R&D innovation input is attributed to unobserved firm-specific effects. The presence of unobserved firm-specific effects at this stage of the model can in turn have implications for the linkage between R&D and output innovation. Some aspects of these unobserved firm-specific effects can be corroborated with some aspects of the data that R&D is not necessarily a perequisite for a firm to realize an innovation. In studies using CIS output innovation, a 2010 review study by the OECD (see Roberts and Vuong, 2013) shows that the probability of a firm introducing a new product varies from 40 to 65% across countries while the same probability for zero R&D firms is generally between 10 and 30%. The reporting of R&D investment itself may also concern a strategic firm decision which renders the interpretation that missing R&D can be due to firms choosing not to reveal this information because of strategic, competitive reasons (Bound et al., 1982; Nagaoka et al., 2010; Koh and Reeb, 2015). Yet, these characteristics may be jointly connected by other

[7] Firm-level patent studies that cover partial links within the CDM model in an isolated form, will be discussed in the next section.

[8] The firms in our sample are group enterprises located in the Netherlands, but not necessarily the ultimate parent firm since foreign control is possible. The statistical unit "enterprise" is essential in the construction of a patent sample, because firms may register patents (and R&D) under different names. Generally speaking, the ownership of a patent occurs at the level of an enterprise and it is practically impossible to link this ownership to affiliates or plants.

factors which may not be observable, unknown, or difficult to capture. To get an idea of the importance of these missing R&D firms, for instance, Koh and Reeb (2015) find that among firms that do not report any information on R&D, about 11% of them applied for patents and many of them have a substantial number of patents. Based on a propensity score matched sample, the authors further note that on average, non-reporting R&D firms have patents that are over 27% more influential than the matched zero R&D firms. Our sample shows that 1864 out of 2730 patenting firms do not report any information on R&D. These firms with missing R&D but engaged in patenting activities remains detectable in our CDM model and analysis.

Fourth, an important way in which we depart from the literature on R&D-innovation-productivity is the econometric treatment. By introducing firm-specific unobserved heterogeneity and dynamics within the sequence of our model, we explicitly control for firm-specific factors that may connect the different processes of innovation. In fact, by using the predicted innovation input as an explanatory variable in the patent equation and predicted patent counts as a determinant in the productivity equation, we alleviate the endogeneity arising from the fact that innovation investment, innovation output, and output productivity may be determined simultaneously. Using the sequential approach in analyzing the CDM, our approach is novel and has not been investigated yet. At the stage of the innovation process, the model that we use is an application and generalization of a two part Hurdle model, describing a continuous and count outcome sequence, to allow for independent variables, persistent individuals (random effects), and noise or randomness. The association of the two sequences in the two part model is captured by correlating the random effects of each outcome sequence.[9] In addition, the dynamics helps us to comprehend if there exists a persistence in patenting while accounting for unobserved

[9] Although, there are several panel studies in the area of innovation that have looked at R&D and patents using panel count data (see, among many others, Hausman *et al.*, 1984; Gurmu and Perez-Sebastian, 2008), most of the studies rely on GMM and panel count data models that avoid the assumption of excess zero patent counts.

heterogeneity at the firm level. Regarding the R&D sample selection, we also allow for dynamic random effects. The dynamics works as follows: we include the persistence effect of R&D itself as well past patent behavior so to accommodate with some of the theoretical predictions of the CDM model. Our key result is that the contribution of output innovation, in terms of patenting applications and patent citations, to productivity (growth) is positive and statistically significant. The time-invariant unobserved firm-level characteristics are deemed to be very important for determining firm-level R&D and the number of patents. The estimates of R&D and the dynamics depend on whether these unobserved characteristics are taken into account.

The remainder of the paper is organized as follows. Section 2 presents a brief review of literature dealing with the R&D-patents-productivity relationship. Section 3 describes the data set. Section 4 presents the empirical model. In Section 5 we present the estimation results of the model. Finally, Section 6 concludes.

2 Background

Innovation studies that are based upon the CDM model reconcile two strands of empirical research, namely, first, studies looking at the relationship between different innovation inputs, such as R&D expenditure, and outputs, such as patents measured in counts (Hausman *et al.*, 1984; Pakes and Grilliches, 1984) and, second, studies analyzing the relationship between innovation output and productivity growth (Kortum, 1993; Bloom and Van Reenen, 2002; Balasubramanian and Sivadasan, 2011; Hall and Sena, 2017).[10] Studies based on (variants of) the CDM model share as major advantage that the widespread distinction between patents and R&D expenditure is not treated in an isolated way but enters the model explicitly. Furthermore, as Griliches (1979) already pointed out in his

[10]Recent review papers (Nagaoka et al., 2010; Hall, 2011; Mohnen and Hall, 2013) summarize the pioneering work by Griliches, which attempt to employ patents stock and/or citations to explain the residual growth in productivity.

pioneering work, the relationship between productivity and R&D expenditure embodies in a simplified way two very different and presumably complex processes: the production of innovations starting from R&D activities and the incorporation of these innovations to production. More specifically, the contribution of innovation to productivity (growth) is disentangled into the contribution of R&D input to innovation output and the contribution of innovation output to the overall firm output. This distinction supports the view that it is *innovation output* that matters for productivity (growth) while *innovation input,* such as R&D, only contributes to innovative capabilities within the firm.

Not only the empirical literature cited in the introduction motivates this distinction, but also theory. For instance, in endogenous (equilibrium) models of innovation and growth, this happens in case of a production equation that, in addition to other input variables, depends on innovation output combined with an innovation equation, where R&D is considered as an input in the production of patents or new inventions (see, for instance, Romer, 1990; Aghion and Howitt, 1992; Kortum, 1993). Dynamics are important. Romer (1990) theoretically defines three R&D externalities that yield an interesting interpretation on the further dynamics between R&D and patent applications. First, innovators use their production of (new) innovation in the production of output. This effect is based on the assumption that once a patent is granted, this generates a positive externality on others engaged in R&D.[11] Second, since models such as CDM are based on perfect competition, innovators cannot engage in price discrimination and firms that are licensing the innovation cannot obtain some consumer-surplus. Their effect has a positive externality on R&D. Third, the introduction of a new technology may replace the existing technology, which, in turn, may have a negative effect on the R&D of those

[11]In a broader framework these models rationalize the idea of intertemporal complementarity in innovation: Past experience makes current innovation efforts more productive. Thus, the accumulation of knowledge would induce state dependence in invention flows and, consequently, persistence in innovation, with as result more internal funding that can be used to finance further innovations. Another theoretical explanation considers the sunk costs in R&D investments as an important source for persistence since they create barriers to entry and create engagements to continue innovation.

firms owning the old technology.

In another strand of literature using firm-level structural models, Roberts and Vuong (2013) note that R&D investments is a dynamic decision that is forward-looking since the firm must incur costs in the present period for an anticipated gain in profits in the future which lays out the possible time lag in the R&D-innovation-productivity relationship. Furthermore, as the authors note, a firm's impact of innovation may be persistent which in turn may also influence future investment decisions, which we also highlight in our paper. An important aspect in the model is that it involves unforeseen randomness, uncertainty, unobserved heterogeneity at different stages in a CDM framework that can only be realized after investments are made. Another related paper is Peters *et al.* (2013), which emphasizes the modelling of process and product innovation in a dynamic, structural framework. The firm's productivity is modelled as a stochastic variable that is affected by a firm's past productivity and current process and product innovation. This state variable determines endogenously R&D investments. A firm will invest in R&D if the expected payoff of the R&D-innovation-productivity process is greater than the current investment costs. Recent work by Czarnitzki et al. (2014) and Cui and Li (2016) finds that productivity firms are more likely to patent as this may be considered as a signal to profitability and long term viability of investors.

Focusing on the broader set of studies using microdata when linking the R&D, innovation (patenting), and productivity, Mairesse and Mohnen (2009) note that this relationship also concerns strategic firm decisions which are simultaneously decided and may be jointly connected by other factors which may not be observable, unknown, or difficult to capture as valid instruments. A firm may use its patenting rights in function of market competitive reasons such as exclusion of competitors, strategic licensing, or joint-ventures. Indeed, a firm's decision to patent a certain innovation can be considered as a "strategic decision" because if a firm decides to apply for a patent, it recognizes the potential value of the invention (Jaffe et al., 1993; Cohen et al. 2000; Hall and Sena, 2017; Hall and Harhoff, 2012). This does not mean that the patented knowledge and innovation technology is lower appraised. On the contrary, Nagaoka et al. (2010) illustrate that the pure basic R&D productivity is very significant in generating patentable inventions. Instead, we might expect that patented knowledge is the one most likely to be commercialized. For instance, Arora et al. (2008) found that the incremental value of patented inventions is estimated to be 47% on average. Therefore, one can also consider the advantages of utilizing patents in the strategic analysis within a CDM framework. A paper by Hussinger (2006) provides evidence that firms may use secrecy when developing a new technology, but then apply for a patent when the new product is about to be commercialized.

Not all patents serve only for strategic considerations for some other (un)observed reasons, it still may affect the output and value of patenting firms.[12] Mototashi (2008) reports that about half of the patents are not used, meaning that another half is directly related to production or sales. More than half of the unused patents prevent other firms using the patented technology. Others may be kept for future license negotiations of for future production and sales activities. All of these findings suggest that, taking into account the various motivations of patenting, patents can be considered as a form of knowledge capital entering a production function.

Some further interesting issues arise when looking at partial links within the CDM model which are dealt with in an isolated form. The literature that focuses on a direct link from R&D to productivity (see, for instance, Hall and Mairesse, 1995; Balcombe *et al.*, 2005; Griffith *et al.*, 2006) is not supported in a CDM framework. The theoretical support

[12] Hall (2000) surveys some of the early work relating market-values based measures for firms to their innovation, proxied by R&D and patents. Discrepancies behind market versus book value are expected to be correlated with patenting (Hulten en Hao, 2008). Some recent work using patents as a measure of innovation output in a market value equation confirming this positive relationship is that by Blundell et al. (1999), Toivanen et al. (2002), Bloom and Van Reenen (2002); Hall et al. (2005), using patent citations.

for this type of research is based on defining a production function that, in addition to other input variables, decomposes capital into an R&D component and a remaining physical component. The emerging findings from such studies are somewhat inconclusive: Some studies report an R&D effect on productivity that is essentially zero, whereas others have found a substantial effect. However, most of the estimates lie somewhere between these two extremes and the consensus is that R&D has a significant positive effect on productivity growth (see Hall *et al.*, 2010, for recent evidence on this subject). One major aspect of the results in this type of analysis is the wide variety in the measurement of the R&D variable (for example, assumptions on depreciation rates in the construction of R&D capital, double-counting correction in the labor and capital inputs, etc.) and model specification (for example, panel data, dynamics, etc.), which makes it difficult to identify a precise consensus estimate of the contribution of R&D. However, an important finding that emerges from these studies is that productivity (growth) is better explained by R&D if one takes into account its long-run impact. This statistical finding was first observed by Mansfield (1980), using 1948-1966 U.S. data. In a more recent paper Balcombe *et al.* (2005) find a distributed lag-length between 9 and 10 years on the basis of 1955-2000 time-series data and experimental data on agricultural innovation. An explanation for this is that R&D is likely to yield productivity improvements over longer time horizons. This could have an important implication for public R&D policies.[13]

Micro level studies that look at the dynamics of the patent-R&D relationship show evidence of persistence of innovation. Firms may innovate persistently over time for a couple of reasons as it is an essential feature in endogenous growth models (Romer, 1990; Aghion and Howitt, 1992). With specific reference to the Netherlands, existing studies that investigated the dynamic relationship between R&D and output innovation include

[13]The lag between R&D and patents and the patents' first revenues varies according to industry specificities. We refer to Hall *et al.* (2010) for a review of the literature that looks at the lag distributions of R&D productivity effects.

van Leeuwen (2002) and Raymond *et al.* (2010, 2015), already (partly) discussed in the Introduction.[14] With respect to input innovation, Raymond *et al.* (2010) find that past R&D/sales expenditure affects current R&D activities and this dynamic relationship also holds with respect to the output innovation share of innovative sales in total sales. This result is also confirmed by van Leeuwen (2002) who links innovation input (R&D expenditure/sales) to innovation output (share of innovative sales/total sales) and innovation output to firm performance (revenue/employee). However, a major drawback of the latter study is that individual effects are not accounted for. Only a few papers, such as Peters (2009) and Raymond *et al.* (2010), control for individual heterogeneity and the initial conditions, so as to identify the persistence in innovation. Interesting insights can be gained from applying this approach to our data in a CDM framework.

3 Data

Our data consists of an unbalanced panel of 3030 firms, situated in the Netherlands, during the period 2000-2006, collected from different data sources, representing the population of firms engaged in R&D. Appendix A contains a detailed description of the data collection procedure. Our sample includes 2780 firms that have applied at least for one patent during the years 2000-2006. These patenting firms represent 98 percent of all patent applications during this period. We obtained these patenting firms by matching the entire population of patents applied for at the European Patent Office with all possible subsidiaries in the Netherlands, which are then aggregated to the ultimate parent of the firm. As control group we have (randomly) selected $(3030 - 2780 =)$ 250 firms, stratified by industry and size, which may also be engaged in R&D, i.e., which report R&D expenditure in at least two consecutive years of our sample, but which did not

[14]Other frequently cited empirical studies confirming strong and weak forms of persistence in R&D and output innovation include Geroski et al. (1997), Cefis and Orsenigo (2001), and Peters (2009).

apply for a patent.

We work with the date at which the patent was filed. This date is called the "priority date" or the "effective filing date" and is the date used to establish the novelty of a particular invention. The priority date may be earlier than the actual filing date of an application. If an application claims priority to an earlier parent application, then its priority date may be the same as the parent (OECD, 2009). Nagaoka *et al.* (2010) argue to use the priority dates rather than the grant or application dates, because several applications can be filed at a later stage for the same invention receiving the priority of the original application.

Table 1 presents some descriptive statistics on the R&D and patent behavior of the sampled firms. The column "R&D reported" shows the sample of firms with R&D reported values in any of the sample years, and the column "R&D not reported" presents the sample of firms where R&D is reported as missing throughout the sample period.

—INSERT TABLE 1 APPROXIMATELY HERE—

It is well known from patent application data that a large share of patents is applied for by only a small number of firms (see, for example, Licht and Zoz, 2000). The table shows that this small number of firms are mostly firms that also report R&D. As expected, 20 out of a total of 3030 firms amount to the majority of EPO patents. To shed some light on the sector characteristics of the group of "R&D reported" firms, Tables 6 and 7 in the Appendix reports some additional information on R&D statistics and patent averages broken down into 19 industries. According to these tables, the most important patent applications (measured by patent citations) and R&D activities of these firms are found in physics, food, transportation, electrotechnical equipment, and chemicals.

In the estimations we take account of our stratified sample by using sampling weights. As shown in Table 1, we allocated the total population of patents for the period 2000-

2006 to 2780 enterprises and added a control group of 250 (randomly selected) enterprises with reported R&D expenditure. Each of the 2780 enterprises represents itself in the total population of patenting firms, while each non-patenting firm that is included in the control group represents a fraction from a subpopulation that amounts to 3430 firms from which R&D is reported, excluding patenting firms. In order to select the subsample of the R&D reporting, zero-patenting firms, we used the 2008 CIS survey covering data for the 2006-2008 period. Statistics Netherlands stratifies the CIS survey samples from a total population using size and industry sector criteria. We used this information to weigh each firm to the total population sample. As a result, each enterprise in the control group has weight 3430/250, which is simply the reciprocal of the inclusion probability.

Summary statistics of our key variables (in the transformation used in the analysis) are shown in Table 2a. The correlation matrix of the variables is in Table 2b. The statistics are based on the total sample of firms from the period 2000 to 2006. The unweighted (weighted) average firm is in our sample applies approximately for 1.4 (0.75) patents a year, with an average forward citation count of 0.537 (0.270), spends on average $e^{0.236}$ ($e^{0.420}$) euros (2000 prices) on R&D per number of employee, is involved in approximately 2 (2.5) sectors, has approximately 3.5 (4.8) firms under its control, with a multifactor (log) productivity of 4 (4.2), and a (log) of capital per employee of 2.8 (3.3) euros. On average, 12% (24%) of the panel firms have a foreign mother firm and about 56% (71%) of the firms are part of a consolidated group. The average annual markup is 1.28 (0.79). The distribution of the patent variables are quite skewed, while most of the other variables are more evenly spread.

In the specification of each of the equations we account for a multiple number of firms production and innovation characteristics. Along with the usefulness of their conclusion they might potentially generate multicollinearity problems. Generally, we observe a low correlation coefficients for most of the variables. Low correlation coefficients indicate that multicollinearity is not a problem. Among the explanatory variables, the highest correlation occurs among the log of employment, the number of activities and the number of firms. The correlation coefficients vary between 0.413 and 0.743. These variables are jointly included in the 3rd stage of the model (the MFP stage). As a robustness check, we re-ran regression using the 3rd column specification listed In Table 5 omitting the number of activities and number of firm variables. This omission did not affect the coefficients of the other variables. We note that excluding the Log of employment explaining MFP, which is standard in the productivity literature, would turn into a severe problem of omitted variable bias.

—INSERT TABLES 2a AND 2b APPROXIMATELY HERE—

4 Empirical Implementation

The empirical model consists of three parts that relate (i) firm characteristics to R&D expenditure, (ii) R&D expenditure to patent applications, and (iii) patent applications to total Multi Factor Productivity (Growth) (MFP(G)). We consider each of these parts in the subsequent subsections.

4.1 The R&D Expenditure Equations

We consider sample selection bias in the R&D variable. In CDM models, selectivity is usually captured by an unobserved latent variable that equals zero for firms that do not invest in innovation and one for innovating-investing firms. In the CIS surveys only a subset of innovating firms are by definition also R&D performers.[15] The fact that firms

[15] To have a rough idea on the R&D sample selection in the CDM model, we have collected a pool of CDM based articles which appeared in (i) the two special issues for Economics of Innovation and New

with missing R&D expenditures may be engaged with some form of innovation activity is therefore not accounted for. To investigate missing R&D expenditures we rely on the Tobit II selection model while controlling for unobserved heterogeneity. Because our initial sample comprises firms with patenting activities during our sample period which we link to the CIS and R&D survey data, firms with missing R&D are also included in the sample. Observations with R&D missing values may occur in firms as a result of our sample procedure where we link different data sources.

To model R&D expenditure we use a (dynamic) sample selection model, consisting of two equations, where one is a Probit equation determining the probability that a firm reports its R&D expenditure and the other equation is a regression explaining the amount of R&D invested. We only observe the amount of R&D invested in case the firm reports its R&D expenditure. We extend this Tobit-II type model to a panel data context, following Wooldridge (1995, 2005), which enables us to exploit the unobserved heterogeneity at the individual firm level.[16]

Let a firm be indicated by the subindex i and time by the subindex t. The first equation is specified by a binary variable REP_{it} that is equal to one when R&D is

Technology (2001 and 2016), (ii) in Hall (2011) and Mohnen and Hall (2013) survey articles on innovation and productivity, (iii) as key papers that provided the basis of two bibliometric (Broström and Karlsson, 2017; Notten et al. (2017) as well as all (iv) CDM related papers that are listed in this paper. While a vast majority of these studies employed the CDM model applied to CIS data, the selection issue whereby is based on using sample selection models (generalized Tobit) comprising a regression for the censored innovation expenditure set at zero for the non-innovating firms.

[16]Examples of CDM studies that use similar R&D selection criteria whereby missing R&D expenditures are analyzed in the selectivity are (i) a cross-sectional framework include Griffith et al. (2006), Klomp and Van Leeuwen (2006) and in a panel framework these studies include Criscuolo et al. (2009); Hall and Sena (2017); Van Leeuwen and Mohnen (2017), Hall et al. (2009). The paper of Hall et al. (2009) and Van Leeuwen and Mohnen (2017) are the only studies that employ a Tobit II selection estimator to control for the selectivity bias but did not control for panel effects. At the innovation input stage, rather than employing R&D expenditures, the other papers employ a broader definition of innovation expenditure which includes in addition to in-house R&D expenditure, purchase of external R&D, acquisition of machinery, equipment and software, and other external knowledge. These papers posit the assumption that R&D reflects formal innovation activities so to overcome the selection bias of only dealing with positive R&D expenditures.

reported by firm i in year t and zero otherwise. That is,

$$
\begin{aligned}
REP_{it} &= 1 \text{ if } REP_{it}^{*} = a_{1i} + \beta_1' \mathbf{x}_{1it} + \gamma_1' \mathbf{z}_{1i,t-1} + \varepsilon_{1it} \geqslant 0, \\
&= 0 \text{ otherwise,} \qquad (1)
\end{aligned}
$$

where REP_{it}^{*} is the corresponding underlying latent variable, a_{1i} represents the firm-specific heterogeneity, $\mathbf{x}_{1it}$ is a vector of (firm-related) independent variables, $\mathbf{z}_{1i,t-1}$ is a vector of variables capturing the dynamics, ε_{1it} is an error term (with characteristics to be discussed later), and β_1 and γ_1 are vectors with unknown parameters.

Firm i's R&D expenditure per employee at time t is modeled as (allowing for the possibility that $R\&D_{it} = 1$):

$$
\log(1 + R\&D_{it}) = a_{2i} + \beta_2' \mathbf{x}_{2it} + \gamma_2' \mathbf{z}_{2i,t-1} + \varepsilon_{2it} \qquad (2)
$$

where $R\&D_{it}$ represents firm i's R&D efforts, a_{2i} reflects the firm-specific unobserved heterogeneity, $\mathbf{x}_{2it}$ is a vector of independent variables representing firm i's characteristics, $\mathbf{z}_{2i,t-1}$ is a vector of variables capturing the relevant dynamics, and ε_{2it} is a random error (with characteristics to be discussed later); β_2 and γ_2 are vectors with unknown parameters. The vector $\mathbf{x}_{2it}$ is not equal to $\mathbf{x}_{1it}$, i.e., we allow for an exclusion type restriction, which is typical for a sample selection model (see, for example, Vella, 1998).

We observe

$$
\left(REP_{it}, REP_{it} \times R\&D_{it} \right), \qquad (3)
$$

where observing $(0, 0)$ means that $R\&D_{it}$ is unobserved, because it is not reported, while observing $(1, R\&D_{it})$ means that R&D is reported to be equal to $R\&D_{it}$.[17] We consider

[17] In our sample used we have only five observations with $REP_{it} = 1$ and $R\&D_{it} = 0$. For this reason, we do not model $\log(1 + R\&D_{it})$ by means of a censored regression.

two versions: a static version, corresponding to $\gamma_1 = 0$ and $\gamma_2 = 0$, and a dynamic version, where γ_1 and γ_2 are allowed to be unequal to zero.

With respect to the unobserved errors ε_{1it} and ε_{2it}, we shall assume that, conditional upon a_{1i}, a_{2i}, $\mathbf{x}_{1it}$, $\mathbf{z}_{1i,t-1}$, $\mathbf{x}_{2it}$, and $\mathbf{z}_{2i,t-1}$, they follow a bivariate normal distribution with zero mean, variances σ_1^2 $(= 1)$ and σ_2^2, and covariance $\sigma_{12} = \rho_{12}\sigma_2$, where ρ_{12} is the correlation coefficient between the two error terms ε_{1it} and ε_{2it}. Next, conditional upon $\mathbf{x}_{1it}$, $\mathbf{z}_{1i,t-1}$, $\mathbf{x}_{2it}$, and $\mathbf{z}_{2i,t-1}$, we assume

$$a_{1i} = \alpha_{10} + \delta_{10} \log(1 + R\&D^*_{i0}) + \delta_1' \overline{\mathbf{x}}_{1i} + \xi_{1i}, \tag{4}$$

$$a_{2i} = \alpha_{20} + \delta_{20} REP^*_{i0} + \delta_2' \overline{\mathbf{x}}_{2i} + \xi_{2i}, \tag{5}$$

where α_{10} and α_{20} are constants, $\overline{\mathbf{x}}_{1i}$ and $\overline{\mathbf{x}}_{2i}$ are vectors including the time averages of the variables included in $\mathbf{x}_{1it}$ and $\mathbf{x}_{2it}$,[18] and $R\&D^*_{i0}$ and REP^*_{i0} are initial values,[19] δ_{10}, δ_{20}, δ_1, and δ_2 are the corresponding (vectors of) coefficients to be estimated, and ξ_{1i} and ξ_{2i} may be interpreted as unobserved heterogeneity that is uncorrelated with the regressors.

To explain R&D expenditure we include in the vector $\mathbf{x}_{2it}$ the following independent variables: the number of employees in full-time equivalents ("employment") measured in logs, competitive pressure ("competition") measured in logs, a variable indicating the number of domestic firms under complete control ("number of firms"), and the yearly number of industry segments measured in logs ("number of activities"). We include in the vector $\mathbf{x}_{1it}$, the vector of independent variables in the REP^* selection equation, the same variables as in $\mathbf{x}_{2it}$, with the exception that the variables "number of firms" is replaced by an annual dummy variable reflecting whether the firm is part of a group

[18]Instead of $\overline{\mathbf{x}}_{1i}$, the original estimator uses $\mathbf{X}_{1i} \equiv (\mathbf{x}_{1iT}, \cdots, \mathbf{x}_{1it})$, but time averages allow for a reduction of explanatory variables (see Wooldridge, 2005). When calculating time averages, we follow Hesketh and Skondral (2013), and exclude the initial periods and estimate the model at periods $t = 2, \cdots, T$, given the initial values of innovation and explanatory variables in order to prevent estimation bias. We refer to Hesketh and Skondral (2013) for an empirical verification of this bias.

[19]The initial conditions are values of innovation in the initial year of the observed series $t = 1, \cdots, T$ (Raymond *et al.*, 2015).

("group"). The yearly dummy variable "group" equals 1 if the firm's headquarter is located abroad ("foreign") and/or has a direct ownership of multiple domestic firms.[20]

The variable "employment," representing the size of a firm, is positively related to innovation, following Schumpeter (1942). R&D and economies of scope are likely related because innovation may spill over to different projects; see for example, Filatotchev *et al.* (2003) or Piga and Vivarelli (2004). This research provides evidence that an organizational form that consists of multiple firms may have better access to external finance, thereby suggesting that they are more likely to intensify their R&D. We proxy economies of scope by the number of industry segments ("number of activities") and the number of domestic firms under (complete) control ("number of firms"). The number of industry segments matches for each firm i in year t its business activities that correspond to the number of different 3-digit NACE codes. For example, "number of activities" takes the value 2 if firm i's activities during year t belong to the two activity codes NACE151 and 152. In the selection equation, the yearly group membership dummy ("group") takes the value 1 if the firm is part of a group, and is included on the grounds that reporting R&D may be affected by being part of this group (Hall and Oriani, 2006). As a final control variable we also take into account how competitive pressures ("competition") affect the firm's R&D intensity. We follow Martin *et al.* (2011) and measure the level of competition using a Herfindahl index of industrial concentration being the sum of the quadratic relative firm-sizes,

$$H_{k_it} = \sum_{j \in S_{k_it}} \left(\frac{\text{employees}_{jt}}{\text{employees}_{k_it}} \right)^2,$$

where S_{kt} is the set of firms belonging to industry sector k at time t, and k_i denotes the sector to which firm i belongs. The variable "competition," defined as $\frac{1}{H_{k_it}}$, measures the

[20]We also considered adding industry sector dummies so to capture additional structural effects but we did not achieve ML convergence in this case.

degree of competition firm i of sector k_i faces at time t. Most empirical studies find a positive effect of industrial concentration on R&D spending. However, when we look at studies on the relationship between concentration and innovative output, such as patents and innovative products, the effect is generally found to be non-significant (for example, Vossen, 1999). So, it is expected that firms spend more R&D in more concentrated industries, but this does not necessarily result in more innovative output.[21] [22]

Concerning dynamics, we shall assume that in the R&D equations (1)–(2) $\mathbf{z}_{1i,t-1}$ and $\mathbf{z}_{2i,t-1}$ include information on past patent applications ("past patent applications") and past R&D expenditure ("past R&D"). This assumption can be motivated as follows. First, it is based on the so-called complementarity assumption in knowledge production, where current levels of innovation may be affected by past (levels of) both R&D and patent applications (Klette, 1996; van Leeuwen, 2002). Second, this dynamic formulation is also in line with the R&D-patent externalities discussed in Section 2.

We consider various versions and combinations for capturing past information on patenting activities and R&D. In one version, the past information is captured by two dummy variables, the first one equal to one in case there is a positive outcome on the lagged patent counts or and zero otherwise and the second one equal to one in case the lagged R&D is positive and zero otherwise. In the other version, the past information is captured by the level of the lagged patent application counts and the level of the

[21] We did not include a market share (Nickel, 1996) based on sales data because of data constraints. In the Tobit II R&D sample selection equations and the patent equation (see next subsection), we employ variables that are extracted from the "General Business Register" (Statistics Netherlands) and are recorded for each firm located in the Netherlands. Input (capital, wages) and output (sales) variables, which are employed in the MFP equation (see Section 4.3.) are extracted from the Statistics of Finance of Enterprises and are only available for a smaller selection of firms.

[22] The question of whether foreign controlled firms invest more or less in R&D remains unanswered. For instance, some research on the R&D activities of foreign-controlled firms find evidence that these firms invest less in R&D than domestic firms for the reason that these firms have better access to innovation endowments from the MNE and other subsidiaries (Un and Cuervo-Cazurra, 2008; Ortega-Argilés and Moreno, 2009). On the other hand, these firms may also have better access to capital financing which may induce the subsidiary to invest in more R&D which is also confirmed by a number of empirical studies (see the review in Narula and Zanfei, 2005; Un and Cuervo-Cazurra, 2008).

past R&D per employee (with missings set equal to zero and in log terms after adding one). We consider these two versions, because the estimates can be sensitive to firm-year observations for which the levels (in particular of patent application counts) are unusually high (as mentioned earlier when discussing Table 1). To reduce the possible effect of this variation on the estimated coefficients, we also work with a dummy variable approach for capturing the past information.

4.2 The Patent Equation

The next part of the model explains the innovation output measured by the number of patent applications in a given year. We also consider the number of forward citations received from the resulting patents, in order to get a potentially better quantification of innovation output. Indeed, a number of studies (for example, Hall *et al.*, 2002; Bekkers *et al.*, 2011) have established the existence of a positive correlation between forward citations and technological importance. In both cases we shall refer to the dependent variable as "patent counts." The discreteness of patent data motivates the use of a count model. An important characteristic of our data is that we find for many firms zero patent counts. The zero patent counts occur for firms that have not applied for a patent during our sample period, or, if applied (whether granted or not), the zero citation patent counts also occur for firms that received no forward citations.[23] A firm can decide not to apply for a patent for many reasons such as difficulties in the R&D process, technological and market uncertainty, competition, or one-time technological activities (see, for example, Crépon and Duguet, 1997). To take this excess of zeros into account, we estimate the output innovation equation using a pooled Hurdle model

[23]We observe that 79% of our sample includes panel-year firms with zero patent applications. Similarly, Bound *et al.* (1984) observe for the U.S. that zero patent firms represent 60% of their sample; in Crépon and Duguet (1997), these firms represent 73% of their sample using French data.

allowing for unobserved heterogeneity.[24] Suppose firms engage in patent activities, with positive counts, once a threshold is crossed. Let $PAT^*_{it} = 1$ indicate that firm i passed the threshold in year t, while $PAT^*_{it} = 0$ indicates that firm i did not pass the threshold in year t. We write $\Pr(PAT^*_{it} = 1) = p_{it}$ and $\Pr(PAT^*_{it} = 0) = 1 - p_{it}$. Next, we introduce the negative binomial distribution (NB),

$$P(y, \lambda_{it}) = \frac{\Gamma(\alpha^{-1} + y)}{\Gamma(\alpha^{-1}) \times \Gamma(1 + y)} \left(\frac{\alpha^{-1}}{\alpha^{-1} + \lambda_{it}} \right)^{\alpha^{-1}} \left(\frac{\lambda_{it}}{\alpha^{-1} + \lambda_{it}} \right)^{y}, \tag{6}$$

where $\Gamma(\cdot)$ denotes the gamma function and where $\alpha \geq 0$ captures the deviation from the Poisson distribution. This model reduces to Poisson when $\alpha \to 0$, so that we can use a test on α to discriminate between the NB and the Poisson model.

Let PAT_{it} be the number of firm i's patent counts in year t. Then we model

$$\Pr(PAT_{it} = 0) = (1 - p_{it}) + p_{it} P(0, \lambda_{it}), \ \Pr(PAT_{it} = y) = p_{it} P(y, \lambda_{it}), \ y \geq 1. \tag{7}$$

Thus, conditional on the event that the threshold is crossed, positive outcomes follow the NB distribution (including Poisson as special case), see Cameron and Trivedi (2013).

We model λ_{it} from the NB distribution and the p_{it} probabilities as[25]

$$\ln \lambda_{it} = a_{3i} + \delta_3 \log(1 + R\&D_{i,t-1}) + \beta_3' \mathbf{x}_{3it} + \gamma_3' \mathbf{z}_{3i,t-1} \tag{8}$$

$$\text{logit } p_{it} = a_{4i} + \delta_4 \log(1 + R\&D_{i,t-1}) + \beta_4' \mathbf{x}_{4it} + \gamma_4' \mathbf{z}_{4i,t-1} \tag{9}$$

where a_{3i} and a_{4i} are time-invariant unobserved firm-effects, $\mathbf{x}_{3it}$ and $\mathbf{x}_{4it}$ are vectors of independent variables, $\mathbf{z}_{3i,t-1}$ and $\mathbf{z}_{4i,t-1}$ capture the relevant dynamics other than lagged

[24] Our model with reference to the innovation stage draws heavily from Vancauteren (2017). The author uses a static random effects Hurdle model controlling for zero inflation. As a robustness check a zero-inflated model is also estimated. The unobserved heterogeneity follows the specification in Wooldridge (2005) and includes an additional random firm-effects. Because their model is presented as a static model, the problem of initial conditions is not specified.

[25] We use here $\text{logit}(z) = \ln(z/(1-z))$.

R&D, and δ_3, β_3, and γ_3 together with δ_4, β_4, and γ_4 include the unknown parameters. We use the lag of R&D because it is unlikely that R&D innovation becomes productive immediately.[26] In case the firm's $R\&D_{it-1}$ is unobserved, we take its predicted value from the model described in the previous subsection.

The time-invariant unobserved firm-effects a_{3i} and a_{4i} are assumed to be normally distributed, allowing for a correlation between them. More precisely, as in equations (4)–(5), and following Wooldridge (2005), we model this unobserved heterogeneity as being dependent on the initial conditions and the average of the continuously distributed explanatory variables with additional random firm-effects that are uncorrelated with the regressors.

We consider $\mathbf{x}_{3it}$ and $\mathbf{x}_{4it}$ the same as $\mathbf{x}_{1it}$ in equation (2). The inclusion of these variables in (8) and (9) is motivated by previous patent studies (see, for example, Peters, 2009). Concerning the dynamics ($\mathbf{z}_{3i,t-1}$ and $\mathbf{z}_{4i,t-1}$, with $\mathbf{z}_{3i,t-1} = \mathbf{z}_{4i,t-1}$), we shall assume that the current patent application counts are affected by the past patent application counts. We consider two specifications for capturing the persistence in output innovation, similarly to the R&D equation: One specification where the lagged patent behavior appears as a dummy variable and another specification by including lagged patent application counts.

Given (7) and (8)–(9), with $P_0(y, \lambda_{it})$ modeled by (6), the expectation of PAT_{it} conditional upon $D_{it} \equiv \left(a_{3i}, a_{4i}, R\&D_{i,t-1}, \mathbf{x}'_{3it}, \mathbf{x}'_{4it}, \mathbf{z}'_{3i,t-1}, \mathbf{z}'_{4i,t-1}\right)'$ is given by

$$E\left(PAT_{it}|D_{it}\right) = p_{it}\lambda_{it}. \tag{10}$$

This means that the partial derivative of the log of (10) with respect to, for example,

[26]For a detailed discussion on lag effects we refer to Hall *et al.*, (2011). We tested for the best-fitting lags and we obtained more significant results, using the Bayes Information Criteria, with lagged R&D up to one year. In the results section, further robustness checks on whether the lag structure matters for the estimated coefficients are being discussed.

$\mathbf{x}_{1it} = \mathbf{x}_{3it} = \mathbf{x}_{4it}$ becomes

$$\frac{\partial \log E\left(PAT_{it}|D_{it}\right)}{\partial \mathbf{x}_{1it}} = (\beta_3 + \beta_4) - p_{it}\beta_4 = \beta_3 + (1 - p_{it})\beta_4. \tag{11}$$

The derivative with respect to $R\&D_{i,t-1}$ or $\mathbf{z}_{3i,t-1} = \mathbf{z}_{4i,t-1}$ is similar (just replace β_3 by δ_3 or γ_3 and β_4 by δ_4 or γ_4 in the right hand side of (11)). Thus, to calculate the sign of the marginal effect or the total marginal effect of one of the components of $\mathbf{x}_{1it}$, we have to take the sum of the corresponding component of β_3 and β_4, where the latter has to be multiplied by $1 - p_{it}$ (and similarly for $R\&D_{i,t-1}$ or $\mathbf{z}_{3i,t-1}$, with some obvious modifications).[27]

4.3 The MFP Equation

The third and final part of the model explains the Multi-Factor Productivity (MFP).[28] We measure MFP by estimating a translog value-added production function using the econometric approach proposed by Olley and Pakes (1996) and later modified by Levinsohn and Petrin (2003). With this procedure a markup is derived from an estimated firm-level scale elasticity, multiplied by an output-input ratio that is computed with real data (see Appendix B for the formulation and further details).

[27] If the corresponding components of β_3 and β_4, say β_{3j} and β_{4j}, have the same sign, then the elasticity will also have the same sign, irrespective of the value of p_{it}. If β_{3j} and β_{4j} have opposite signs, and $|\beta_{4j}| < |\beta_{3j}|$, then the elasticity will have a constant sign, equal to the sign of β_{3j}, irrespective of the value of p_{it}. However, if β_{3j} and β_{4j} have opposite signs, and $|\beta_{4j}| > |\beta_{3j}|$, then the sign of the elasticity depends on the value of p_{it}. In the case with a high probability of zero patents (i.e., p_{it} close to 0) the sign of the elasticity will be determined by β_{4j}. However, when the probability of zero patents is low (i.e., p_{it} close to 1), the elasticity will change sign (determined by β_{3j}).

[28] In the above mentioned studies that examine the empirical evidence on innovation and productivity growth in the CDM framework, several important aspects are not considered. In almost all studies, productivity is interpreted as a partial measure; the most common is "labor productivity," measured as output per man-year or hours worked. A notable exception is the study of Parisi et al. (2006) who use a more comprehensive measure of (total or) multi-factor productivity (MFP), i.e., capital, labor, energy, materials, and services are taken into account. MFP is considered to be a more realistic representation of the entire production process (Syverson, 2011).

We estimate the following model

$$\log MFP_{it} = \alpha_{5i} + \delta_5 \log(1 + PAT_{it}) + \beta_5' \mathbf{x}_{5it} + + \gamma_5' \mathbf{z}_{5i,t-1} + \varepsilon_{5it} \tag{12}$$

where MFP_{it} is MFP (measured by equation (15) in Appendix B), α_{5i} is the time-invariant unobserved firm-effect, $\mathbf{x}_{5it}$ is the vector of independent variables, $\mathbf{z}_{5i,t-1}$ captures the relevant dynamics, and ε_{5it} is the error term; δ_5, β_5, and γ_5 include the unknown parameters. We allow for random effects, by assuming that $\alpha_{5i} | PAT_{it}, \mathbf{x}_{5it}, \mathbf{z}_{5i,t-1} \sim N(0, \sigma_\alpha^2)$ and $\varepsilon_{5it} | PAT_{it}, \mathbf{x}_{5it}, \mathbf{z}_{5i,t-1} \sim N(0, \sigma_\varepsilon^2)$, with α_{5i} and ε_{5it} stochastically independent variables, conditional on PAT_{it}, $\mathbf{x}_{5it}$, and $\mathbf{z}_{5i,t-1}$.

The following variables which we defined earlier are included in $\mathbf{x}_{5it}$: industry and time dummies, "employment," "competition," "foreign," the "number of activities," and the number of domestic firms ("number of firms"). We also add annual sector markups and the firm's capital intensity, measured by the log of the capital-labor ratio ("$\log(K/L)$"). We include the firm's capital intensity and "employment" since these have been shown to increase factor productivity (Bernard *et al.*, 2006). We also include the foreign ownership dummy ("foreign") which is motivated by recent literature on firm's heterogeneity (for example, Bernard *et al.*, 2006). We include *two* competition variables. Next to the inverse of the Herfindahl index of industrial concentration ("competition"), measuring the concentration of firms within a market, we also include the "markup" (using equation (13) in Appendix B). The "markup" measures the average profitability of firm i in sector k_i as it reflects a firm's ability to set its prices above marginal costs. While both competition indicators focus on a particular aspect of competition, the Herfindahl index market share variable may point into the wrong direction when competition intensifies. As discussed in Boone *et al.* (2007), an increase in competition may force the least efficient firms out of the market which in turn increases market shares. Thus, the market share fails to pick up a selection effect. For the firm's innovative output PAT_{it}, we take its predicted values from the random ZINB model. We include in $\mathbf{z}_{5i,t-1}$ the time lags of the effects of output innovation on MFP.

5 Estimation Results

In this section we present our estimation results for the three parts of our model, presenting the estimation results of various specifications. We conclude this section with some robustness checks.

The R&D sample selection model—We start with the Tobit II model discussed in Subsection 4.1. Three variants have been estimated, see Table 3. In the first two columns of Table 3 (with estimation results), we present the results of a static model with the individual effects a_{1i} and a_{2i} set equal to zero, using pooled data. The third and fourth columns show the results of the dynamic model, setting the individual effects ξ_{1i} and ξ_{2i} equal to zero, again using pooled data. Columns five and six present the estimation results of the dynamic model focusing on the R&D level equations, including the entire unobserved heterogeneity. We estimate this version of the model following Dustmann and Rochina-Barrachina (2007) who work out the Wooldridge (1995) estimation approach using minimum distance. This means that in a first step we estimate the parameters in (1), using Maximum Likelihood. Next, we consider for each time period the reduced form of (2) for the subsample of firms that report R&D. This reduced form includes an inverse Mills ratio as a sample selection bias correction. Each of these time specific reduced forms is estimated separately, using the first round Probit estimates to estimate the inverse Mills ratios. Finally, we impose the restriction that the regression coefficients of the different years are equal to each other via a minimum distance estimation procedure. These latter results are presented in columns five and six. We consider two specifications for capturing dynamics: (i) a dummy variable capturing 1/0 (i.e., Yes/No past innovation), see column five, and (ii) lagged levels, see column six. We report the estimates of β_1, β_2, and γ_1, γ_2 (if not zero), and δ_{10}, δ_{20}; we do not report the estimates of δ_1, δ_2.[29]

We first focus on the most general model (columns five and six).[30] A firm's engagement in prior R&D affects positively, and significantly so, the level of R&D intensities while the past patent behavior does not seem to have any significant impact. Firm size, measured as the number of employees ("employment"), has a negative and significant effect, meaning that conditional on investing in R&D, smaller firms exhibit a higher R&D investment per employee. Unlike Schumpeters (1942) claim, a robust finding from our results is that instead the employment elasticity is significant and negative ranging between -0.16 and 0.14 if we consider the full model in the level equation. This finding therefore suggests that as firms increase in size, they devote a smaller percentage of their R&D investments per employee. These results, contrary to those identified by earlier studies, can be reconciled with Audrey and Thurik (2009) claim that smaller firms have nowadays better access to new technologies that makes it possible to respond to market conditions as efficiently as large firms. This suggests that in some instances, R&D intensities can also be larger for smaller firms compared to larger firms. Other studies have also found a negative relationship between R&D intensity and size (see Kamien and Schwartz, 1982, for details on these early studies). For instance, Soete (1979), using US data for some 500 firms in 1975 and 1976, found that R&D intensity increased with size in a number of sectors, although this was not always true for the largest firms, and decreased with size in others. Using US firm level data, Park (2011) shows that small firms are more R&D intensive. In a more recent study on Italian manufacturing firms, Hall et al. (2012) also observe that R&D intensity is negatively correlated to firm size.

—INSERT TABLE 3 APPROXIMATELY HERE—

[29]These are available upon request from the corresponding author.

[30]We used the specification of the third column to calculate the Inverse Mills ratios. The corresponding regression results are available upon request.

Among the other controls, firms with more subsidiaries do less R&D when specifying the dynamics by means of lagged levels. This result corresponds to our earlier finding that R&D investment per employee is negatively related to size. We also observe that the effect of the initial level of R&D is significant. In the specification where past innovations is measured via a dummy variable, diversification ("number of activities") is the only variable that seems to have an impact, although with a negative sign. We also note that the potential selection bias is significant: In the row $H_0 : \rho_{21,t} = 0$, all t, we report the Wald test for testing this hypothesis, with corresponding p-value between brackets.

Comparing the estimation results of the fifth and sixth columns with the estimation results of columns two and four reveals that accounting for individual effects, initial conditions, and dynamics seem to affect, at least to some extent, the estimation results. In column one and three, we also reported the selection equation. Focusing on the results in column three, we find evidence that larger firms and group membership are major drivers in the probability of reporting R&D. A firm's engagement in past patent applications affects positively and significantly the probability of R&D.[31] In addition, we also observe a significant negative effect of competition on the probability of doing R&D. The effect of diversification ("number of activities"), as an additional measure of scope, hardly affects R&D probabilities. The potential selection bias is significant: We find a significant correlation between the errors (see the row ρ_{12}, where we report the estimated ρ_{12}, with the corresponding standard error in brackets).

The patent equation—Next, we consider the estimates of the patent equation discussed in Subsection 4.2. First, a likelihood ratio test comparing the Poisson model with the Negative Binomial model reveals that in all cases the Negative Binomial is to be preferred. As shown in Table 4 (row with "−2 Log-lik"), the hypothesis that the overdispersion parameter, denoted by α, equals zero (i.e., $H_0 : \alpha = 0$), is conclusively rejected.

[31] When considering the effect of prior R&D, we encountered convergence problems.

As a consequence, we only report the outcomes based on the Negative Binomial distribution. To fit the pooled hurdle model with random effects, we adopt the approach from Min and Agresti (2005) (allowing for possible correlation between the bivariate normally distributed random effects).

In the first two columns (with estimation results) we report the ML based results of the static model for the number of patent applications, without random effects, including the lagged value of the log of R&D expenditure per employee, the log of "employment" and the log of "competition".[32] In the next four columns we present the full static model with random effects for both the number of patent applications and forward patent citations. In the final four columns we present the full dynamic model, including initial conditions, where the lagged patent behavior is included via a dummy variable.[33] We estimate the hurdle model using the SAS procedure NLMIXED.[34] The dynamic model is estimated for both the number of patent applications and the forward patent citations. The log likelihood values and the number of observations are reported at the bottom of Table 4. In the table we present the results with the lag of R&D per employee included, but our results do not change dramatically once we replace the lagged effect of R&D with an instantaneous effect (not reported).

—INSERT TABLE 4 APPROXIMATELY HERE—

Based on the static results, we find that the parameter estimates for R&D turn out to be an important driver for the patent count part, but not so much so for the propensity patent part of the model. The coefficients of the log of employment are always positive,

[32]We also tried to evaluate a full model where we added the other available explanatory variables (employment, number of activities, group dummy), but because of convergence problems, we cannot report these estimates.

[33]We also investigated a dynamic specification with the persistence of innovation captured by lagged patent applications. However, because of convergence problems, we cannot report these estimates.

[34]We refer to Min and Agresti (2005) for further technical details regarding the calculation of the ML estimates.

indicating that the propensity to patent and the number of patents are increasing with firm size. This innovation-size relationship is not in line with the results reported in Raymond et al. (2015). Although, to make results fully comparable, the authors consider two measures: the propensity to innovate (1/0) and the intensity of output innovation (share of innovative sales). R&D or any form of input innovation is not regressed on any size indicator. The authors find that size does not affect innovation intensity nor its propensity for the Netherlands while size is more likely to be positively related in France. Competition does not seem to play a significant role in explaining the patent behavior. We see that including random effects, capturing the unobserved ability to be innovative, results in a dramatic improvement of the log-likelihood values. However, the coefficients on the R&D variables drop when we include these random effects.

The total effect on patents is given by equation (11). For the static model with patent counts and random effects, Figure 1 illustrates the nonparametric densities of $1-p_{it}$, for each of the years $t = 2001, \cdots, 2006$. Most probability mass is between zero and 0.6, with mean values ranging from around 0.23 to 0.30. As a consequence, for this specification the total estimated effect of log employment according to (11) lies between around 0.15 and 0.35. The total estimated effect of the lagged value of log R&D lies between around 0.065 and 0.072 and the total estimated effect of log competition lies between around -0.05 and 0.06.

—INSERT FIGURE 1 APPROXIMATELY HERE—

When comparing the static results on patent counts with the corresponding dynamic extensions, we see relatively minor changes of the log-likelihood values. Indeed, the evidence of including the innovation persistence in explaining patenting behavior is not present. Firm size, measured by the log of employment, and the unobserved random effects seem to drive the major effects on patenting. No significant impact of the R&D

variable on patent activity is observed. The estimates of the patent citation specification clearly show differences in the coefficients between the number of citations and citation probabilities. Firm size is not a statistically significant determinant of the probability of positive patents but becomes significant in the level equation part. Surprising is the large competition effect found in the propensity to citation part of the model. This reveals that strong rivalry effects between firms seem to affect the patenting quality. We also find a strong presence of random effects, and the covariance between these random effects is ascertained at a one percent significance.

The MFP equation—Finally, we consider our estimates of the MFP equations (discussed in Subsection 4.3) in Table 5. In the first two columns (with estimation results) we report the regression of the level of MFP (see equation (15) in the appendix) on only the lagged predicted value of the log of the number of patent citations and the log of the number of applied patents, respectively. In the next two columns (columns three and four) we also include the other independent variables. In the columns five and six we present the results as in columns three and four, after replacing MFP in levels by the dependent variable MFP growth (see equation (16) in the appendix). Since not all inventions are patented, a direct path from R&D to a firm's productivity will also be considered. The final two columns report the results of the dynamic model with as lagged value the log of R&D per employee (both for MFP in levels and for MFP growth). We estimate the equations by an iterated ML version of the generalized least squares method allowing for random effects. For the firms' output innovation we take the predicted values from the estimated ZINB equations reported in Table 4 (columns 4-5).

—INSERT TABLE 5 APPROXIMATELY HERE—

Most notably, we find significant estimates of the impact of the (predicted) output innovation on both the MFP level and its growth. The elasticity ranges between 0.783

and 1.655 using the number of patent citations and between 0.294 and 0.622 using the number of patent applications. Bloom and van Reenen (2002) found that the contribution of patent stocks to British firms' sales are significant, with an elasticity of 0.03, and lower than in the CDM (1998) study, where a patent elasticity of 0.13 is found. In other studies, where variants of the CDM model have been more directly applied in cross-sectional analysis, Griffith *et al.* (2006) report elasticities of both product and process innovation in the range of .06 to .18 across some of the largest EU countries; the contribution of innovation output (share of innovative sales) to productivity (value-added labor productivity) for the Netherlands, found in Klomp and van Leeuwen (2006) were insignificant while Raymond et al. (2015) report a positive and significant effect using more recent data.

Our results also show that R&D does not have any impact on MFP and MFP growth. This results is somewhat in line with a comprehensive overview of studies by Hall *et al.* (2010) that looks at the direct R&D-MFP relationship. The authors find that the estimates of R&D elasticities are sensitive to the inclusion of additional firm and industry level characteristics and its impact does not consistently differ when expressing MFP growth rates versus levels. Next, we find a positive and significant effect of employment on the level of MFP when considering R&D, but a negative effect in most of the other estimations. A positive effect is in line with many other empirical studies; however, Raymond *et al.* (2015) report a negative elasticity of 0.06 for the Netherlands, suggesting that smaller firms are more productive using a partial labor productivity measure expressed in levels.

The direction of competition on productivity is in the empirical literature not clear *a priori*, which is also confirmed by empirical evidence on the competition-productivity link in the Netherlands (Polder *et al.* 2009). There are several reasons why less competition fosters productivity growth (Syverson, 2011): More competition may constrain or postpone investments which is likely to hamper productivity growth, and more competition may induce incumbent firms to keep entrants out of the market which is a strategy that may negatively affect the productivity growth of the incumbent firms. In our estimation results competition measured by using the inverse of the Herfindahl index is either not significant or statistically negatively related. However, based on the markups and partially for the competition measure, the results suggest that less competition in the output market leads to more MFP and MFP growth.[35] Our results confirm the idea that using a single variable for competition may yield incomplete results.

Finally, we find partial evidence that foreign firms are more productive, as follows from the positive significant estimated coefficients. The effects of the capital/labor ratio are significant but dot not have the expected positive signs. The "number of activities" and also have positive effects, but not always significantly so.

Some robustness checks—We conducted some robustness checks to investigate the sensitivity of our results. A first robustness check involves an alternative distribution where we employ a zero-inflated count model instead of a two-part model. The zero-inflation model rests on the assumption that zero patents and positive patents are ruled by the same mechanism. More specifically, a zero-inflated count model considers a probability distribution that allows for modelling two separate types of zeros: (i) a structural zero resulting from a binary process and (ii) an incidental zero as a realization of the count process of the number of patent applications.[36] On the other hand, in a hurdle model, the decision to apply for a patent is made on the basis of the first invention and the decision to apply for additional patents depends on the outcome of this first decision. So, in a two-part model we would expect different decision criteria concerning the first patent and the additional patents. Overall, the results in terms of the effect of R&D on

[35] The negligible impact of the inverse of the Herfindahl index is robust to omitting markups.

[36] For detailed discussion on zero-inflated count models we refer to Cameron and Trivedi, 2013.

patents, using a zero-inflated count model instead of a hurdle model, are only a little affected. That is, the effect of R&D on patenting and other firm characteristics become very weak or insignificant once unobserved heterogeneity is taken into account in the zero-inflated count model.[37]

A second robustness check involves an alternative measure for output innovation. Instead of working with patent count data, we also employed a "fixed-weighted number of forward citations per patent," where weights are equal to the number of forward citations of each individual patent, divided by the average number of forward citations of patents of the same publication year and technological class (see Hall *et al.*, 2002). For simplicity, we applied the same Hurdle model with the same regressors that appear in the full specification of the static patent model allowing for random effects (see columns 3 and 4 of Table 4 with estimation results). We consider a log-normal model for the weighted patents and estimate the continuous data part of the model by a Tobit model.[38] Our results show that the coefficient on the effect of the R&D coefficient on the "fixed-weighted number of forward citations per patent" equals 0.85 and statistically significant in determining whether or not citations are positive, while we find a negative and significant impact on the level of (weighted) citations with a coefficient of 0.061. Furthermore, when we extend this analysis to the productivity equation, the statistical significance of the patent variable remains robust with an elasticity of 0.046.

One notable extension of this paper is that a richer specification of productivity dynamics can be tested by including additional lags of the innovation output variables. We find statistical evidence of a 4-year distributed lag of citation and patent counts, affecting

[37]More specifically, we find similar evidence as in the static versions of the hurdle model with the competition variable omitted. However, due to convergence problems, we were not able to estimate a dynamic version of the zero-inflated model. Most likely, the relationship between patents and R&D involves many more count distributional features than can be modeled by means of a zero-inflated count model or a hurdle model.

[38]For further details on a two part hurdle model with continuous data, we refer to Cameron and Trivedi (2005, pp. 544-547)

the level of MFP. The time of innovation with the highest impact on productivity occurs at the third lag for patent counts and the first lag for citations. The total patent elasticity is roughly equal to 2 for both citations and counts. The long-run nature of these results stresses the importance of dynamic responses.

As a final robustness check we also re-estimated the TFP equation where not only patents and citations but all variables are treated as endogenous. We therefore re-estimated the MFP equation in dynamic form using the system GMM, which includes fixed effects and tackles endogeneity of the right-hand-side variables by using their lagged values (in first differences and levels) as instruments. Therefore, under the assumption that current random shocks are uncorrelated with past values from firm-level regressors, we use past values from t2 onwards as instruments. Using the full specification as shown in column 3 and 4 of Table 5, the effect of persistence in the level of TFP and the significant effect of the predicted number of patents and citations is still confirmed. The persistence in TFP has an elasticity of 0.42 and 0.65 respectively while the effects of patents or citations on TFP growth remain fairly robust with an elasticity of 0.76 and 1.36 respectively. The effect of the other variables are very similar to the ones presented in Table 5. We note that the validity of the instruments is tested using the Sargan test, yielding the problem of weak instruments; the estimates obtained are sensitive to the instruments used.

6 Conclusion

This paper revisits, at the firm level, the innovation-productivity relationship based on the empirical research initiated by Crépon, Duguet, and Mairesse (1998) using panel data analyzing over 3000 Dutch firms for the period 2000-2006. We estimate a model that disentangles the impact of R&D expenditure, the number of patent applications, and multifactor productivity growth. The model assumes that the effect of R&D on patents is computed using data on firms that do not necessarily report their R&D effort.

The results related to the sample selection of our data reveals some firm characteristics that can be attributed to both their R&D and patent activities. We find that those companies that report their R&D activities also tend to be the largest innovators, measured by their R&D efforts and the number of patents issued. However, this does not rule out that the majority of the firms which are not engaged in patent activities can also be classified as non-R&D firms. Actually, they may take an important role in the innovation debate. More research should be oriented towards attributes that are related to their activities.

Turning to the substantive results in the paper, we find, *first*, that patent applications and patent citations have both an economically and statistically significant impact on firm-level productivity with an elasticity in the range of 0.783 and 1.655, using the number of patent citations, and between 0.294 and 0.622, using the number of patent applications. *Second*, the results suggest that the time-invariant unobserved firm-level characteristics are deemed to be very important for determining the number of patents. While the estimates of R&D and the dynamics depend on whether these unobserved characteristics are taken into account, we find evidence on the role of firm size in explaining patent and citation counts.

After controlling for a variety of firm-characteristics, size (measured by employment) is negative in the R&D equation but positive in the patent equation, confirming only partially the Schumpeterian hypothesis. In addition, our results also suggest that smaller firms are more productive. Both for the R&D and patent counts, the role of competition is now clearer and can be judged more in detail than "minor statistical significance" between competition (measured at the industry level using a Herfindahl index of industrial concentration) and innovation.

We find no convincing evidence on the persistence of innovation. Once we control for the initial innovation levels and firm-level effects, the impact of past innovation on current innovation in terms of patenting cannot be ascertained. On the other hand, there is clear evidence on the effect of past R&D on the current level of R&D activities and on the effect of past patent activities on the R&D probability.

As it is clear from the above discussion, smaller firms tend to subsume an important role in the innovation-productivity debate: smaller firms not only tend to be more productive but are also more R&D intensive compared to larger firms. The role of entrepreneurship and better access to technology, which we capture in the model by unobserved heterogeneity, implies better opportunities for SMEs. An innovation policy in this framework must incorporate new type of instruments which are independent of the size of the firm. For example, to foster productivity enhancement, policy should promote entry and exit and competition in the market. Also, at the early stage of innovation, fundamental R&D is essential. For small firms, this may yield potential barriers to innovation. Policy response such as financial assistance for small firms can be further explored.

As a sensitivity check, we find that our results are affected when we replace the number of forward citations with an adjusted (fixed) number of forward citations per patent because the effect of R&D becomes more prominent. In addition, many of the results are not affected, when we consider a zero-inflated count model as an alternative.

The main caveat of this paper relates to the data we use to implement the model. Although we relax the essential role of R&D in explaining innovation output and its impact on productivity growth, one major assumption in that relationship is that the effect for non-R&D reporting firms is the same as for R&D reporting firms. This may tend to give some bias in the results. Therefore, as a prelude to further research, more investigation should be devoted to the characteristics and the matching procedures of R&D reporting and non-R&D reporting firms.

References

[1] Acs, Z. and D. Audretsch (1989), Patents as a Measure of Innovative Activity, *Kyklos* 42(2), 171–180.

[2] Aghion, P. and P. Howitt (1992), A Model of Growth Through Creative Destruction, *Econometrica* 60(2), pp. 323–351.

[3] Amoroso, S., Melenberg, B., Plasmans, J. and M. Vancauteren (2012), Firm-level Productivity under Imperfect Competition in Output and Labor Markets, Mimeo.

[4] Arora, A., M. Ceccagnoli and M. Cohen (2008), R&D and the Patent Premium, *International Journal of Industrial Organization* 26, 1153-1179.

[5] Audretch, D. and A. Thurik (2001), What's About the New Economy? From the Managed to the Entrepreneurial Economy, *Industrial and Corporate Change*, 10(1), 267–315.

[6] Balasubramanian N. and J. Sivadasan (2011), What Happens When Firms Patent? New Evidence from U.S. Economic Census Data. *The Review of Economics and Statistics* 93, 126-146.

[7] Balcombe, K., A. Bailey and I. Fraser (2005), Measuring the Impact of R&D on Productivity from a Time Series Perspective, *Journal of Productivity Analysis* 24(1), 49–72.

[8] Bekkers, R., R. Bongard and A. Nuvolari (2011), An Empirical Study on the Determinants of Essential Patent Claims in Compatibility Standards, *Research Policy* 40, 1001–1015.

[9] Benavente, J.M. (2006), The Role of Research and Innovation in Promoting Productivity in Chile, *Economics of Innovation and New Technology* 15(4/5), 301–315.

[10] Bernard, A., Jensen B. and P. Schott (2006), Trade Costs, Firms and Productivity, *The Journal of Monetary Economics* 53(5), 917–937.

[11] Bloom, N. and J. Van Reenen (2002), Patents, Real Options and Firm Performance, *The Economic Journal* 112(March), 97–116.

[12] Blundell, R., R. Griffith and J. Van Reenen (1999), Market Share, Market Value and Innovation in a Panel of British Manufacturing Firms, *Review of Economic Studies* 66, 529-554.

[13] Boone, J., J.C. van Ours, and H. van der Wiel (2007), How (Not) to Measure Competition, CentER Discussion Paper No. 2007–32.

[14] Bound, J., C. Cummins, Z. Griliches, B. Hall and A. Jaffe (1984), *Who Does R&D and Who Does Patents?*, in Z. Griliches (ed.), *R&D, Patents and Productivity*, University of Chicago, Press for the National Bureau of Economic Research, Chicago.

[15] Broström, A. and S. Karlsson (2017), Mapping Research on R&D, Innovation, and Productivity: A Study of an Academic Endeavour, *Economics of Innovation and New Technology* 26(1-2), 6–20.

[16] Brouwer E., and A. Kleinknecht (1999), Innovative Output, and a Firm's Propensity to Patent: An Exploration of CIS Micro Data. *Research Policy* 28, 615–24.

[17] Cameron, C. and P. Trivedi (2005), *Microeconometrics*, Cambridge Books, Cambridge University Press.

[18] Cameron, C. and P.Trivedi (2013), *Count Panel Data*, Oxford Handbook of Panel Data Econometrics, Badi H. Baltagi eds., Oxford University Press, (forthcoming).

[19] Cefis, E. and L. Orsenigo (2001), The Persistence of Innovative Activities. A Cross-Countries and Cross-Sectors Comparative Analysis, *Research Policy* 30(7), 1139–1158.

[20] Cohen, W., R. Nelson, J. Walsh (2000), Protecting their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not), NBER Working Paper 7552.

[21] Crépon, B. and E. Duguet (1997), Research and Development, Competition and Innovation: Pseudo Maximum Likelihood and Simulated Maximum Likelihood Methods Applied to Count Data Models with Heterogeneity, *Journal of Econometrics* 79, 355–378.

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

[23] Criscuolo, C, M. Squicciarini. and O. Lehtoranta (2009), R&D, Innovation and Productivity, and the CIS: Sampling, Specification and Comparability Issues, Contributed paper for the 2nd Conference on corporate R&D (CONCORD - 2010)

[24] Cui, J. and L. Xiaogang (2016), Innovation, Technology Scope and Firm Productivity: Evidence from the US Patent Data, Working Paper.

[25] De Loecker, J. and F. Warzynski (2012), Mark-Ups and Firm-Level Export Status, *The American Economic Review* 102, 2437-2471.

[26] Diewert, W.E. and K.J. Fox (2009), On the Estimation of Returns to Scale, Technical Progress and Monopolistic Markups, *Journal of Econometrics* 145, 174–193.

[27] Dustmann, C. and M.E. Rochina-Barrachina (2007), Selection Correction in Panel Data Models: An Application to the Estimation of Females' Wage Equation, *Econometrics Journal* 10, 263–293.

[28] Filatotchev I., C. Piga and N. Dyomina (2003), Network Positioning and R&D Activity: A Study of Italian Groups, *R&D Management* 33, 37–48.

[29] Geroski, P.A., J. van Reenen and C.F. Walters (1997), How Persistently Do Firms Innovate?, *Research Policy* 26(1), 33–48.

[30] Griffith R., E. Huergo, J. Mairesse and B. Peters (2006), Innovation and Productivity Across Four European Countries, *Oxford Review of Economic Policy* 22, 483–498.

[31] Grilliches, Z. (1979), Issues in Assessing the Contribution of Research and Development to Productivity Growth, *Bell Journal of Economics* 10(1), 92-116.

[32] Gurmu, S. and F. Perez-Sebastian (2008), Patents, R&D and Lag Effects: Evidence from Flexible Methods for Panel Data on Manufacturing Firms, *Empirical Economics* 35, 507-526.

[33] Hall, B. (2000), Innovation and Market Value. In Barrell R., Mason G, O'Mahoney M. (eds.), *Productivity, Innovation and Economic Performance*. Cambridge: Cambridge University Press.

[34] Hall, B. (2011), Innovation and Productivity, NBER Working Paper No. 17178.

[35] Hall, B. and J. Mairesse (1995), Exploring the Relationship between R&D and Productivity in French Manufacturing Firms, *Journal of Econometrics* 65, 263–293.

[36] Hall, B. and F. Kramarz (1998), Effects of Technology and Innovation on Firm Performance, Employment and Wages, Economics of Innovation and New Technology, 5 (2-4), 99–108.

[37] Hall, B., A.B. Jaffe and M. Trajtenberg (2002), The NBER-Citation-Data-File: Lessons, Insights and Methodological Tools, in: Jaffe, A.B. and Trajtenberg, M. (2002). *Patents, Citations and Innovations. Cambridge*, MIT Press, 403–459.

[38] Hall B. H., A.B. Jaffe and M. Trajtenberg (2005), Market Value and Patent Citations, *RAND Journal of Economics* 36(1), 16–38

[39] Hall, B. and R. Oriani (2006), Does the Market Value R&D Investment by European Firms? Evidence from a Panel of Manufacturing Firms in France, Germany and Italy, *International Journal of Industrial Organization* 24(5), 971–993.

[40] Hall, B., F. Lotti and J. Mairesse (2009), Innovation and Productivity in SMEs: Empirical Evidence for Italy, *Small Business Economics* 33, 13-33.

[41] Hall, B.,J. Mairesse and P. Mohnen (2010), Measuring the Returns to R&D, in B. Hall and N. Rosenberg (eds.): *Handbook of the Economics of Innovation*, Elsevier.

[42] Hall, B. and D. Harhoff (2012), Recent Research on the Economics of Patents, NBER Working Paper 17773, Cambridge MA.

[43] Hall, B., F. Lotti and J. Mairesse (2013a), 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.

[44] Hall B., C. Helmers, Rogers M., Sena V. (2013b). The Importance (or not) of Patents to UK firms. *Oxford Economic Papers* 65, 603-629.

[45] Hall B. and V. Sena (2017), Appropriability Mechanisms, Innovation and Productivity: Evidence from the UK, *Economics of Innovation and New Technology* 26(1-2), 42–62.

[46] Hausman J., B. Hall and Z. Grilliches (1984), Econometric Models for Count Data with an Application to the Patents-R&D Relationship, *Econometrica* 52 (July), 909–937.

[47] Hesketh, S. and A. Skrondal (2013), Avoiding Biased Versions of Wooldridge's Simple Solution to the Initial Condition Problem, *Economic Letters* 120, 346-349.

[48] Hottenrott, H., Hall, B., Czarnitzki, D. (2016), Patents as Quality Signals? The Implications for Financing Constraints on R&D, *Economics of Innovation and New Technology* 25(3), 197-217.

[49] Huergo, E. and J. Jaumandreu (2004), Firms' Age, Process Innovation and Productivity Growth, *International Journal of Industrial Organization* 22, 541–559.

[50] Hulten, C. and J. Hao (2008), Intangible Capital and the 'Market to Book Value' Puzzle, NBER Working Paper 14548, Cambridge MA.

[51] Hussinger K. (2006), Is Silence Golden? Patents versus Secrecy at the Firm Level, *Economics of Innovation and New Technology* 15(8), 735-752.

[52] Jefferson, G., B. Huamao, G. Xiaojing and Y. Xiaoyun (2006), R&D Performance in Chinese Industry, *Economics of Innovation and New Technology* 15(4/5), 345–366.

[53] Kamien, M.I. and N.L. Schwartz (1982), *Market Structure and Innovation*, Cambridge: Cambridge University Press.

[54] Klette, T. (1996), R&D, Scope Economies and Plant Performance, *RAND Journal of Economics* 27(3), 502–522.

[55] Klomp, L. and G. van Leeuwen (2006), On the Contribution of Innovation to Multifactor Productivty Growth, *Economics of Innovation and New Technology* 15(4), 367–390.

[56] Koh, P.S. and D. Reeb (2015), Missing R&D, *Journal of Accounting and Economics* 60, 73-94.

[57] Kortum, S. (1993), Equilibrium R&D and the Patent-R&D Ratio: U.S. Evidence, *American Economic Review* 83(2), 450–457.

[58] van Leeuwen, G. (2002), Linking Innovation to Productivity Growth Using Two Waves of the Community Innovation Survey, OECD Science, Technology and Industry Working Papers, 2002/8, OECD Publishing.

[59] van Leeuwen, G. and P. Mohnen (2017), Revisiting the Porter Hypothesis: An Empirical Analysis of Green Innovation for the Netherlands, *Economics of Innovation and New Technology* 26(1-2), 63–77.

[60] Levinsohn, J. and A. Petrin (2003), Estimating Production Functions using Inputs to Control for Unobservables, *Review of Economic Studies* 70, 317–340.

[61] Licht, G. and K. Zoz (2000), Patents and R&D - An Econometric Investigation Using Applications for German, European and US Patents by German Companies, in: Encaoua, D; B. Hall, F. Laisney, J. Mairesse, *The Economics and Econometrics of Innovation*, 307–338.

[62] Lööf, H. and A. Heshmati (2006), Knowledge Capital and Heterogeneity in Firm Performance. A Sensitivity Analysis, *Economics of Innovation and New Technology* 15(4/5), 317–344.

[63] Lööf, H., Mairesse J. and P. Mohnen (2017), CDM 20 Years After, *Economics of Innovation and New Technology* 26(1-2), 1–5.

[64] Mairesse, J. and P. Mohnen (2010), A Survey of Innovative Surveys: Taking Stock of a Growing Literature, in B. Hall and N. Rosenberg (eds.): *Handbook of the Economics of Innovation*, Elsevier.

[65] Mairesse, J. and S. Robin (2017), Assessing Measurement Errors in the R&D-Innovation-Productivity Relationship, *Economics of Innovation and New Technology* 26(1-2), 93–107.

[66] Mansfield, C. (1980), Basic Research and Productivity Increase in Manufacturing, *American Economic Review* 70, 863–873.

[67] Martin, Ph., Mayer, T. and F. Mayneris (2011), Spatial Concentration and Plant-Level Productivity in France, *Journal of Urban Economics* 69(2), 182–195.

[68] Min, Y. and A. Agresti (2005), Random Effect Models for Repeated Measures of Zero-Inflated Count Data, *Statistical Modelling* 5, 1–19.

[69] Mohnen, P., J. Mairesse and M. Dagenais (2006), Innovativity: A Comparison Across Seven European Countries, *Economics of Innovation and New Technology* 15(4/5), 391–413.

[70] Mohnen, P. and B. Hall (2013), Innovation and Productivity: An update, *Eurasian Business Review*, 3(1), 2013, 47-65

[71] Mototashi, K. (2008), Licensing or Not Licensing? An Empirical Analysis of the Strategic Use of Patents by Japanese Firms, *Research Policy* 37, 1548-1555.

[72] Nagaoka, S, K. Motohashi and A. Goto (2010), Patent Statistics as an Innovation Indicator, in B. Hall and N. Rosenberg (eds.): *Handbook of the Economics of Innovation*, Elsevier.

[73] Nickell, S. (1996), Competition and Corporate Performance, *Journal of Political Economy* 104(41), 724–746.

[74] Notten, A., J. Mairesse and B. Verspagen (2017), The CDM Framework: Knowledge Recombination from an Evolutionary Viewpoint, *Economics of Innovation and New Technology* 26(1-2), 21–41.

[75] OECD (2009), IECD Patent Statistic Manual, OECD Publishing, Paris.

[76] Olley, S. and A. Pakes (1996), The Dynamics of Productivity in the Telecommunication Equipment Industry, *Econometrica* 64, 1263–1297.

[77] Ortega-Argilés, R. and R. Moreno (2009), Evidence on the Role of Ownership Structure on Firms' Innovative Performance, *Investigaciones Regionales*, 15, 231–250.

[78] Pakes A. and Z. Grilliches (1984), Patents and R&D at the Firm Level: A First Look, in Z. Grilliches (ed.), *R&D, Patents and Productivity*, University of Chicago, Press for the National Bureau of Economic Research, Chicago.

[79] Parisi, M., Schiantarelli, F. and A. Sembenelli (2006), Productivity, Innovation and R&D: Micro Evidence for Italy, *European Economic Review* 50, 2037–2061.

[80] Peters, B. (2009), Persistence of Innovation: Stylized Facts and Panel Data Evidence, *Journal of Technology Transfer* 34, 226–243.

[81] Peters B., Roberts, M., Vuong V. and H. Fryges (2013), Estimating Dynamic R&D Demand: An Analysis of Costs and Long-run Benefits, NBER Working Paper 19374, Cambridge, MA.

[82] Piga, C. and M. Vivarelli (2004), Internal and External R&D: A Sample Selection Approach, *Oxford Bulletin of Economics and Statistics* 66(4), 457–482.

[83] Polder, M., G. van Leeuwen, P. Mohnen and W. Raymond (2009), Productivity Effects of Innovation Modes, Statistics Netherlands, Discussion paper 09033.

[84] Raymond, W., P. Mohnen, F. Palm and S. Schim van der Loeff (2010), Persistence of Innovation in Dutch Manufacturing: Is it Spurious? *The Review of Economics and Statistics* 92(3), 495–504.

[85] Raymond, W., Mairesse, J., Mohnen, P. and F. Palm (2015), Dynamic Models of R&D, Innovation and Productivity: Panel Data Evidence for Dutch and French Manufacturing, *European Economic Review* 78, 78, 285–306.

[86] Roberts, M. and V. Vuong (2013), A Dynamic Empirical Model of R&D Demand, Mimeo.

[87] Romer, P. (1990), Endogenous Technological Change, *Journal of Political Economy* 98(5), 71–102.

[88] Schumpeter, J.A. (1942), *Capitalism, Socialism, and Democracy.* Harper and Row, New York.

[89] Soete, L.L.G. (1979), Firm Size and Inventive Activity: The Evidence Reconsidered, *European Economic Review* 12, 319–340.

[90] Syverson, C. (2011), What Determines Productivity?, *Journal of Economic Literature* 49, 326–365.

[91] Toivanen, O., P. Stoneman, and D. Bosworth (2002), Innovation and the Market Value of UK Firms, 1989–1995, *Oxford Bulletin of Economics and Statistics* 64, 39–61.

[92] Vancauteren, M. (2017), The Effects of Human Capital, R&D and Firm's Innovation on Patents: A Panel Study on Dutch Food Firms, *Journal of Technology Transfer*, DOI 10.1007/s10961-016-9523-2.

[93] Vancauteren, M. and B. Henry de Frahan (2011), Trade policy, Competition and Productivity: The Impact of EU Harmonization in the Dutch Food Processing Industry, *De Economist* 159(4), 483–509. (Special Issue: Productivity and Internationalization: A Micro-Data Approach)

[94] Vella, F. (1998), Estimating Models with Sample Selection Bias: A Survey, *Journal of Human Resources* 33, 127–169.

[95] Vossen, R. (1999), Market Power, Industrial Concentration and Innovative Activity, *Review of Industrial Organization* 15, 367–378.

[96] Wooldridge, J. (1995), Selection Corrections for Panel Data Models under Conditional Mean Independence Assumptions, *Journal of Econometrics* 68, 115–132.

[97] Wooldridge, J. (2005), Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity, *Journal of Applied Econometrics* 20, 39–54.

Appendix A: Data Construction

As underlying population, we consider all firms that are "engaged" in R&D. We define a firm to be engaged in R&D if it applies for one or more patents or if it reports R&D expenditure. To collect the firms that applied for at least one patent, we used the database of the *total* population of patents applied for in Europe (by the European Patent Office (EPO)). This patent data sets give us information about indicators that include (on top of other information) the application number, the patent owner (name of the firm), patent title, name of the inventor, publication year, and location. The ownership criteria to matching patents with firms are essential in the construction of the sample on patent firms. Since firms register patents or report R&D expenditure under different names, the ABR data (i.e., data from the general business register, in Dutch "Algemeen Bedrijven Register"), issued yearly by Statistics Netherlands, retrieves information on firms' ownership structure to find the names and the direct ownership (expressed in percentage) of all their subsidiaries, holding units, and their shareholders. In the sample of firms we define the possible (not necessarily ultimate) parent of the firm (enterprise) as a firm that is located in the Netherlands. The total population of the patent applications between 2000 and 2006 are matched with all possible subsidiaries and then assigned to the ultimate parent firm. This yields 2776 firms. In addition, we used two complementary R&D data sources. First, we extract R&D data from the CIS waves (CIS3, CIS3.5, CIS4 and CIS4.5) and R&D surveys that are collected by Statistics Netherlands. The R&D surveys report R&D expenditure in the odd years while each of the CIS waves measures R&D expenditure in the even years of our sample period 2000-2006. Second, we used annual reports of Dutch firms with more than 50 employees in order to append any R&D data for firms that are not reported in the CIS and R&D surveys. This yields 3430 extra firms, after excluding the firms that applied for one or more grants. From this group of 3430 firms we randomly selected 250 firms, to be included in our sample. This additional control group of 250 R&D non-patenting firms are added to the sample so also to represent a group of firms that may patent at a stage or would never ever patent. The R&D and patent performing firms subsume a high proportion of the sample. We note that a robust analysis reveals minor implications of adding this extra control group for the empirical R&D-patent relationship.

We also include information about forward citations. A forward citation means that a patent is cited by a later patent. For each sample firm, we calculate the total number of forward citations for all patent applications belonging to its enterprise group. Journal citations are excluded from these figures. In order to minimize the problem of truncation (for instance, older patents have a higher probability to be cited compared to new patents), the analysis is restricted to patents granted up to the year 2006 with forward citations until March 2011.

Table 6 shows the number of EPO patent applications and the number of citations received per application period and per International Patent Class on a 1-digit level. This Table also included 2000 patents that are not linked to enterprises. Older patents have a higher probability of receiving a citation. Therefore, the average number of forward citations per patent is (much) higher in 2000-2003, when compared to 2004-2006.

—INSERT TABLE 6 APPROXIMATELY HERE—

We extracted information on output, value-added, net tangible fixed capital assets, sales, depreciation, and wages, from the "Statistics of Finance of Enterprises," provided by Statistics Netherland, all expressed in thousands of Euros. The data on the number of employees, ownership structure, the number of subsidiaries, and the number of industry segments is taken from the ABR (the general business register). The exact industry category assignment scheme which we use throughout this paper, based on NACE codes, is presented in Table 7.

—INSERT TABLE 7 APPROXIMATELY HERE—

To select appropriate input and output deflators, we use the 1999–2006 input-ouput tables. The input-output tables are available in current prices as well as prices from the previous year and classifies 104 industries. This data enables us to generate a (sector) price index deflator based on a chained Paasche index for each of the variables. The value-added deflator is derived from the difference of gross output and consumption of intermediate goods. The capital deflator is derived from investments in fixed (capital) assets. The calculation of the output deflator is based on output expressed in basic prices.

Appendix B: Markups and MFP

We rely on previous research (see, for example, Diewert and Fox, 2009; Amoroso *et al.*, 2012, De Loecker and Warzynski, (2012)) where, from a production or dual cost function, firm-level markups can be expressed as a time-varying, input dependent firm-level elasticity of scale multiplied by an (observable) output-input ratio at the level of the firm. Let

$$\mu_{it} = \frac{p_{it}y_{it}}{\mathbf{w}'_{it}\mathbf{x}_{it}}\theta_{it} \tag{13}$$

where μ_{it} is the firm-level markup of firm i in year t, $p_{it}y_{it}$ measures the firm's output in value (price $\times$ quantity), $\mathbf{w}'_{it}\mathbf{x}_{it}$ denotes the inputs also measured in values (vector of input prices $\times$ vector of input quantities), and θ_{it} is the elasticity of scale. The firm-level output-input ratio is observable using production data while the output elasticities θ_{it}, need to be estimated.[39]

[39] We refer to Vancauteren and Henry de Frahan (2011) for a complete derivation of the markup equation in a production framework.

The only restriction that is imposed on the production function for deriving the expression of the markup is that it is continuous and twice differentiable with respect to its inputs. To estimate output elasticities, we estimate a production function that is represented by a translog approximation (see also De Loecker and Warzynski, 2012). The assumption of a flexible functional form, such as the translog, allows for a more flexible structure on the data to reveal production patterns for heterogeneous firms of widely different sizes. The translog value-added production function is represented as

$$\ln y_{it} = \beta_0 + \beta_k \ln k_{it} + \beta_l \ln l_{it} + \beta_{kk} (\ln k_{it})^2 + \beta_{ll} (\ln l_{it})^2 + \beta_{kl} \ln l_{it} \ln k_{it} + \varepsilon_{it}, \quad (14)$$

where y_{it} is value-added of firm i in year t, k_{it} is capital, l_{it} is employment and ε_{it} is the disturbance term. For the translog production function the elasticities of scale θ_{it} are equal to the sum of the output elasticities, i.e., $\theta_{it} = \theta_{ikt} + \theta_{ilt}$, where

$$\begin{aligned} \theta_{ikt} \equiv \quad & \partial \ln y_{it} / \partial \ln k_{it} = \beta_k + 2\beta_{kk} \ln k_{it} + \beta_{kl} \ln l_{it}, \\ \theta_{ilt} \equiv \quad & \partial \ln y_{it} / \partial \ln l_{it} = \beta_l + 2\beta_{ll} \ln l_{it} + \beta_{kl} \ln k_{it}. \end{aligned}$$

This procedure generates firm-heterogeneous estimated scale elasticities which in turn delivers firm-level markups.

The production approach generates estimates for both mark-ups and MFP. After estimating equation (14), a measure of MFP can be recovered using the method proposed by Olley and Pakes (1996) and later modified by Levinsohn and Petrin (2003). The estimated productivity of firm i is equal to

$$\widehat{MFP}_{it} = \exp(\widehat{\beta}_0 + \widehat{\varepsilon}_{it}). \quad (15)$$

Since a firm's production technology is very likely to vary across sub-sectors, we estimate equation (14) using the Levinsohn-Petrin method by sub-sector.[40] We refer to $\widehat{MFP}_{it}$, as defined in (15), as the *level* of MFP. Next, to the level we also consider the *growth* of MFP, defined as

$$\Delta \ln \left(\widehat{MFP}_{it} \right) = \ln \left(\widehat{MFP}_{it} \right) - \ln \left(\widehat{MFP}_{i,t-1} \right) . \tag{16}$$

Appendix C: Figure and Tables

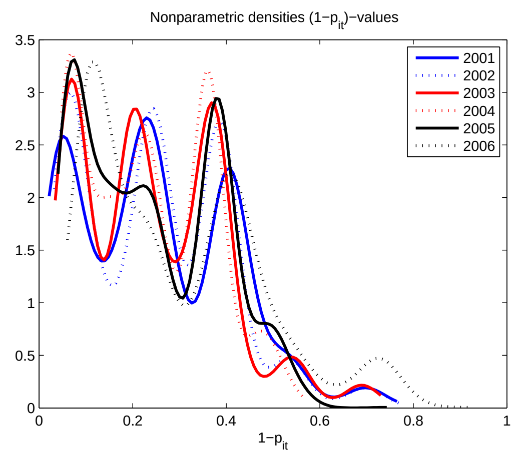

Figure 1: Nonparametric densities of $1 - p_{it}$, for each of the years $t = 2001, \cdots, 2006$.

[40]We also estimated the scale elasticities using the Olley-Pakes (OP) method. The use of the OP method leads to minor changes in the return to scale estimates while both the OP and LP specifications for the markup give very similar results.

Table 1: EPO Patent Applications

	R&D reported	R&D not reported	Total firms	Total EPO patents
total sample of firms	1166	1864	3030	31509
of which:				
patent firms (> 100)	16	4	20	22451
patent firms ($> 1, < 100$)	470	511	981	7279
patent firms ($= 1$)	430	1349	1779	1779
patent firms ($= 0$)	250	0	250	0

Table 2a: Sample Means and Standard Deviations

Summary statistics are of the overall sample of 3030 firms, consisting of 2780 patenting and 250 non-patenting firms. There are 15804 observations for each of the variables while the R&D variable contains 3903 panel firm year observations and the capital, markup and MFP variables count 8990 observations

	Mean	(Weighted)	Std. Dev.	Q1	Median	Q3
Patent application counts	1.485	(0.746)	40.162	0	0	0
Patent citation counts	0.537	(0.270)	8.201	0	0	0
Log R&D per employee	0.236	(0.420)	2.905	-0.882	0.799	2.072
Log Employment	3.304	(4.658)	2.434	1.098	3.091	5.170
Log Competition	2.261	(2.382)	0.896	1.543	2.178	3.059
Number of activities	1.972	(2.532)	2.506	1	1	2
Group (1/0)	0.563	(0.708)	0.495	0	1	1
Foreign (1/0)	0.122	(0.244)	0.327	0	0	0
Number of firms	3.504	(4.793)	10.736	1	1	3
Log M FP	4.009	(4.231)	1.097	3.549	4.111	4.596
Markup	1.284	(0.790)	0.771	0.742	1.145	1.723
log K/L	2.807	(3.339)	2.054	1.842	3.187	4.079

Table 2b: Correlation Matrix.

	Log $\frac{R\&D}{emp.}$	Patents	Citations	Past R&D	Past patents	Log emp.	Log comp.	# activ.	Group	Foreign	# firms	Log MFP	Markup	Log K/L
Log $\frac{R\&D}{emp.}$	1.000													
Patents	0.061	1.000												
Citations	0.120	0.415	1.000											
Past R&D	0.091	0.024	0.035	1.000										
Past patents	0.041	0.060	0.091	0.036	1.000									
Log emp.	-0.153	0.168	0.153	0.252	0.086	1.000								
Log comp.	-0.059	-0.088	-0.048	0.021	-0.082	-0.088	1.000							
# activ.	-0.166	0.067	0.039	0.136	0.129	0.615	-0.028	1.000						
Group	-0.100	0.031	0.010	0.167	0.011	0.416	0.089	0.336	1.000					
Foreign	0.062	-0.034	-0.033	0.085	-0.023	0.105	0.149	-0.168	0.318	1.000				
# firms	-0.142	0.042	0.015	0.104	0.090	0.506	-0.049	0.743	0.193	-0.131	1.000			
Log MFP	-0.100	0.112	0.114	0.087	0.094	0.471	-0.072	0.287	0.131	0.090	0.210	1.000		
Markup	-0.000	-0.004	-0.007	-0.019	-0.028	-0.063	0.038	-0.027	-0.016	0.034	0.003	0.017	1.000	
Log K/L	0.093	0.020	0.049	0.098	0.000	0.107	0.070	0.049	0.087	0.120	-0.049	0.130	-0.216	1.000

Table 3: R&D equations: selection and intensity.

ML-estimates with (robust) standard errors in parentheses. Statistical significance is indicated by stars (*: 10%, **: 5%, ***: 1% significance level). Reported in the Probit equation is the change in probability (that R&D is positive) for a unit change in each of the explanatory variables. Number of firms is 2793.

Independent variables	Probit (Y/N, 1/0)	R&D per empl.	Probit (Y/N)	R&D per empl.	R&D per empl.	R&D per empl.
Past Patent Applications (Y/N)			0.506^{***}	-0.047	0.038	
			(0.042)	(0.468)	(0.040)	
Log (1+Patent Counts$_{t-1}$)						0.0006
						(0.0004)
Past R&D (Y/N)					0.125^{***}	
					(0.040)	
Log (1+R&D per employee$_{t-1}$)						0.340^{***}
						(0.008)
Log(Employment)	0.255^{***}	-0.237^{***}	0.125^{*}	0.707^{***}	-0.158^{***}	-0.134^{***}
	(0.014)	(0.072)	(0.074)	(0.180)	(0.029)	(0.028)
Log(Competition)	-0.012	-0.112	-0.333^{**}	-0.427	0.059	0.142^{***}
	(0.030)	(0.169)	(0.134)	(0.661)	(0.057)	(0.054)
Log(Number of Activities)	-0.011	0.266	-0.051^{*}	0.210^{*}	-0.023^{*}	-0.018
	(0.010)	(0.304)	(0.035)	(0.128)	(0.012)	(0.016)
Number of Firms		-0.011		-0.012^{***}	-0.001	-0.002^{**}
		(0.007)		(0.003)	(0.001)	(0.001)
Group (Y/N)	0.308^{***}		0.353^{***}			
	(0.066)		(0.073)			
Intercept	-2.836^{***}	2.713^{***}	-2.291^{***}	1.455	time-varying	time-varying
	(0.097)	(0.684)	(0.120)	(1.561)		
Initial Reported R&D (Y/N)			0.365^{***}	0.296	0.042	0.073^{**}
			(0.058)	(0.456)	(0.040)	(0.030)
ρ_{12}/Wald Test $H_0 : \rho_{12,t} = 0$, all t	-0.999^{***}		-0.032		22.012^{***}	46.931^{***}
	(0.002)		(0.573)		(0.002)	(0.000)
Log-likelihood	-15665.98		-15191.99			
Estimation method	Tobit II		Tobit II		min dist	min dist
Number of Observations	13797		13797		13797	13797
Not Reported/Reported	10493/3304		10493/3304		10493/3304	10493/3304

Table 4: Two-part Patent Counts equations

ML-estimates with (robust) standard errors in parentheses. Statistical significance is indicated by stars (*: 10%, **: 5%, ***: 1% significance level) Reported in the logit equation is the change in probability (that patenting or citing is positive) for a unit change in each of the explanatory variables. Estimations of the number of the patent counts are based on the Negative Binomial distribution. Number of firms is 2793.

Indep. var-s	logit p_{it}	Patents	logit p_{it}	Patents	logit p_{it}	Ci-tations	logit p_{it}	Patents	logit p_{it}	Ci-tations
P. Pat. (Y/N)							-0.405	-0.852	-1.703	-1.210
							(1.427)	(3.175)	(4.072)	(2.092)
Log(RD$_{t-1}$)	-0.173^{***}	0.538^{***}	-0.009	0.072^{***}	0.184^{***}	0.315^{***}	-0.304	-0.710	-1.500	-1.038
	(0.027)	(0.038)	(0.033)	(0.024)	(0.051))	(0.038)	(1.320)	(3.079)	(4.021)	(2.208)
Log(Empl.)	0.249^{***}	0.073	0.296^{***}	0.139^{***}	0.072	0.034	0.296^{***}	0.143^{***}	0.067	0.146^{***}
	(0.051)	(0.066)	(0.055)	(0.046)	(0.091)	(0.094)	(0.055)	(0.046)	(0.091)	(0.051)
Log(Comp.)	0.089	0.025	0.143	-0.055	0.751^{***}	0.203	0.143	-0.066	0.727^{***}	-0.032
	(0.960)	(0.204)	(0.108)	(0.103)	(0.197)	(0.268)	(0.108)	(0.102)	(0.193)	(0.174)
Interc.	-0.769^{***}	-9.291^{***}	-1.210^{***}	-2.282^{***}	-4.846^{***}	-1.190^{***}	–	–	–	–
	(0.058)	(0.113)	(0.124)	(0.194)	(0.242)	(0.231)	–	–	–	–
In. Pat. (Y/N)	–	–	–	–	–	–	-0.311	-0.472	-1.005	-0.995
							(1.497)	(1.834)	(2.188)	(2.760)
α		0.257^{***}		0.231^{***}		0.910^{***}		0.226^{***}		0.153^{***}
		(0.001)		(0.027)		(0.130)		(0.027)		(0.021)
$\sigma^2_{\xi_1}$				2.818^{***}		5.969^{***}		2.832^{***}		5.754^{***}
				(0.129)		0.433		(0.129)		(0.411)
$\sigma^2_{\xi_2}$				3.519^{***}		1.795^{***}		2.991^{***}		4.505^{***}
				(0.269)		0.302		(0.140)		(0.527)
σ_{ξ_1,ξ_2}				2.976^{***}		1.793^{***}		3.562^{***}		4.794^{***}
				(0.157)		0.241		(0.232)		(0.413)
-2 Log-lik		31334		24214		12102		24233		11483
No of Obs		13797		13797		13797		13797		13797

Table 5: MFP equation.

ML-estimates with (robust) standard errors in parentheses. Statistical significance is indicated by stars (*: 10%, **: 5%, ***: 1% significance level). Year and Industry effects included. Number of firms is 1705.

Independent variables/Dependent variable	Multifactor Productivity							
			Level		*Growth*	*Growth*	*Level*	*Growth*
Log(Patent Citations$_{t-1}$)	0.783***		1.655***		1.134***			
	(0.043)		(0.170)		(0.187)			
Log(Patent Applications$_{t-1}$)		0.518***		0.294***		0.622***		
		(0.038)		(0.098)		(0.123)		
R&D per employee$_{t-1}$							-0.009	−0.022
							(0.016)	(0.018)
Log(Employment)			−0.253***	0.047	−0.318***	−0.107***	0.081***	−0.022**
			(0.048)	(0.049)	(0.053)	(0.021)	(0.021)	(0.008)
Log(Competition)			−0.098**	0.051	−0.072*	0.023	0.062	0.035
			(0.040)	(0.037)	(0.044)	(0.041)	(0.037)	(0.041)
Foreign (Y/N)			0.106**	0.240***	0.005	0.026	0.271***	0.041
			(0.048)	(0.052)	(0.032)	(0.033)	(0.056)	(0.034)
Number of firms			−0.003	−0.003	−0.000	−0.000	−0.003	−0.001
			(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Number of activities			0.004	0.011	0.010**	0.010*	0.011	0.010*
			(0.009)	(0.009)	(0.005)	(0.005)	(0.009)	(0.005)
Log(K/L)			−0.048***	−0.007	−0.032***	−0.018*	−0.038*	−0.014
			(0.016)	(0.018)	(0.010)	(0.010)	(0.022)	(0.010)
Markup			0.016***	0.016***	0.013***	0.013***	0.016***	0.013***
			(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Intercept	5.435***	4.568***	12.033***	5.634***	5.084***	3.454***	4.255***	0.343***
	(1.818)	(1.847)	(0.837)	(0.551)	(0.907)	(0.622)	(0.136)	(0.125)
R^2	0.255	0.167	0.287	0.231	0.054	0.036	0.222	0.022
No of Obs	6703	6703	6472	6472	5891	5891	6597	5891

Table 6: Number of patent applications and the number of citations received.

Application year	2000-2003		2004-2006		TOTAL	
	patents	citations	patents	citations	patents	citations
A: Human Necessities	2417	1797	2046	450	4463	2247
B: Performing Operations, Transporting	1984	1371	1607	427	3591	1798
C: Chemistry, Metallurgy	2886	1276	1899	300	4785	1576
D: Textiles, Paper	187	122	148	22	335	144
E: Fixed Construction	607	427	532	228	1139	655
F: Mechanical, Lighting, Heating	696	537	445	131	1141	668
G: Physics	4930	2579	4441	861	9371	3440
H: Electricity	5230	1390	3093	190	8323	1580
Total	18937	9499	14211	2609	33148	12108

Table 7: Summary Statistics (Total Sample).

Nfirm = number of firms per industry. NRDfirm = number of firms with reported R&D expenditures. AR&D = average R&D (in thousands of euros). Empl = average employment. APat1 = 6-year average patents for firms. APat2 = 6-year average patents for firms with > 1 patents. ACit = 6-year average citations per patent.

Industry	Nfirm	NRDfirm	ARD	AEmpl	APat1	APat2	ACit
Agriculture (01-10)	55	14	7510.98	61.82	0.20	2.33	0.38
Petroleum (11,14)	10	2	2142.22	155.13	0.14	0.16	0.04
Food, beverage (15-16)	70	59	7657.96	1010.65	3.56	3.72	2.06
Textiles, clothing and leather (17-19)	27	15	1536.41	245.32	0.38	0.40	0.19
Paper, pulp (20-22)	63	47	3105.78	668.20	0.27	0.36	0.15
Petroleum prod., chemicals (23-24)	106	80	20365.81	689.69	5.07	22.54	3.20
Rubber and plastic, glass (25-26)	121	73	988.53	2198.71	0.35	0.41	0.39
Metals (27-28)	137	86	2698.32	333.46	0.51	1.90	1.23
Machinery (29)	273	159	9233.42	190.20	0.90	9.30	5.83
Electrical (30-33)	131	59	55639.5	683.43	17.62	35.10	4.21
Transportation (34-35)	64	43	9275.41	425.19	0.76	1.31	1.41
Furniture, recycling (36-37)	48	21	828.20	305.57	0.21	0.22	0.06
Energy, water (40-41)	19	11	7390.44	2399.37	0.25	0.38	0.16
Construction (45)	120	58	2302.80	1035.92	0.32	4.96	1.25
Wholesale, maint. vehicles (50-51)	543	185	5127.89	166.82	0.34	2.09	0.87
Retail, hotels and restaur. (52,55)	65	20	6148.35	1560.12	0.24	0.26	0.25
Transp., storage, telecom (60,67)	178	45	22688.77	3602.83	0.57	0.36	0.31
Financial, Business (70-74)	861	152	4327.03	367.68	0.55	1.30	0.59
Educ., other public (75,80,85,90-93)	139	14	1401.29	1389.29	0.80	0.86	0.36