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

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

Very preliminary version. October 2011

*We gratefully acknowledge support from the Netherlands Organisation 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) and participants of the Second Asia-Pacific Innovation Conference held in Singapore (03/05/2011) for fruitful discussions.

[†]Corresponding author, mark.vancauteren@uhasselt.be, Hasselt University & Statistics Netherlands.
[‡]Statistics Netherlands.
[§]CentER, Tilburg University and University of Antwerp.
[¶]CentER, Tilburg University.
[‖]CentER, Tilburg University and University of Antwerp.

Innovation and Productivity of Dutch Firms: A Panel Data Analysis

Abstract

This paper revisits, at the firm level, the innovation-productivity relationship based on the empirical research initiated by Crépon, Duguet and Mairesse (1998). We estimate a structural, dynamic panel data model that disentangles the impact of R&D expenditure on (European) patents and the impact of patents on Total Factor Productivity growth. We match the entire population of patents issued by the European Patent Office to roughly 3000 firms (group enterprizes), located in the Netherlands, for the 2000-2006 period. We allow for the possibility that a firm does not report its R&D. We find evidence that R&D affects the output innovation positively and significantly which in turn has a positive impact on productivity. The persistence of input and output innovation is statistically confirmed but does not appear to affect the innovation-productivity relationship.

1 Introduction

This paper revisits, at the firm level, the innovation-productivity relationship based on the empirical research initiated by Crépon, Duguet, and Mairesse (1998, CDM hereafter). In the CDM model, the contribution of innovation to productivity growth is disentangled into a contribution of R&D input to innovation output and a contribution of innovation output to the overall firm's output. The CDM model has been applied in a number of empirical studies. Recent examples include Griffith *et al.* (2006), Benavente (2006), Mohnen *et al.* (2006), Lööf and Heshmati (2006), Jefferson *et al.* (2006), Klomp and van Leeuwen (2006), Polder *et al.* (2009), Hall *et al.* (2009), and Raymond *et al.* (forthcoming). 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]

These studies typically employ data provided by the European Community Innovation Surveys (CIS hereafter).[2] The innovation surveys assemble data on innovative sales as well as qualitative indicators on demand pull and technological push indicators that are available in a micro-aggregated form. However, we employ a purpose-built data set based on a panel of firms located in the Netherlands with annual data from 1999 through 2006 that includes financial variables, R&D expenditure, patent counts, forward citations, and patent technological fields. In our sample, the entire population of patents issued by the European Patent Office (EPO) is matched to roughly 3000 firms. These firms are group enterprizes located in the Netherlands, but not necessarily the ultimate parent firm since foreign control is possible. The statistical unit "enterprize" 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 enterprize and it

[1]With a specific reference to Dutch firms, this positive relationship is also found in Polder *et al.* (2009) and Raymond *et al.* (forthcoming).

[2]These surveys include CIS I for 1990-1992; CIS II for 1994-1996; CIS III for 1998-2000 and CIS IV for 2002-2004, CIS IV.5 for 2004-2006.

is practically impossible to link this ownership to affiliates or plants.

Our study is one of the first studies that includes granted patents as an innovation output measure in a CDM framework. Granted 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. Since the CIS data does not contain patent data, a proxy for innovation output should be used for studies based on the CIS data. A typical proxy measure of innovation output is the share in total sales due to innovative products, measuring the success in introducing new products into the market. However, Mairesse and Mohnen (2007) argue that the share of innovative sales measure suffers from reliability.[3] Moreover, in case of patents, a particular critical issue is that not all patent 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). 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.

Furthermore, this study contributes to the innovation-productivity literature in the following ways. First, our model builds upon the CDM model in a panel data context, where R&D availability is not necessarily a prerequisite. Our model does not depend entirely on the role of R&D in explaining innovation output and its impact on productivity growth. In the CDM model, R&D data availability is taken as a starting point, merely because the CIS data classifies innovating firms as those that generate both R&D and output innovation. Our approach allows us to exploit differences between innovators and non-innovators, both at the level of R&D expenditure and patent activities. For instance, non-patenting firms might perform R&D and vice versa. We explore these differences

[3]Critical issues to data reliability are, for example, censoring (the definition of innovators is such that it only relates to firms that do both R&D and output innovation), subjectivity (share variables have to be rounded up; guesses on output by entrepreneurs), time inaccuracy (refers to new products in the last three years while other variables are measured in current years).

between innovators and non-innovators taking into account the following situations. Data availability enables us to identify that firms with zero R&D may actually be involved in patent activities. In this situation, firms could be defined as a non-R&D firm but at the same time also as a patent innovator. In addition, it may well be that a firm that does not report its R&D expenditure is involved in innovation output. In such a situation, firms could be defined as a possible R&D or patent innovator. So, the model assumes that the effect of R&D on patents is computed also using data on firms that do not necessarily report their R&D efforts.

Second, the richness of the panel data set enables us to control for individual heterogeneity and possible dynamics of the innovation process . More specifically, it enables us to test the success-breeds-success hypothesis, i.e., the positive impact of past innovation activities (patents and R&D) on current innovation activities and we are also able to control for past productivity on current productivity. The most appropriate form of data to capture the key features of the dynamics is firm-level panel data which allows to control for unobservable heterogeneity across firms. In practice, many of the applications of the CDM model have had to use cross-section data and those that use longitudinal data do not make full use of available panel data estimators.[4],[5]Some partial exceptions are Peters (2009), Raymond *et al.* (forthcoming) and Raymond *et al.* (2010). The studies of Peters (2009) and Raymond *et al.* (forthcoming) only look at innovation and confirm a persistence while individual effects are accounted for and modeled in a dynamic Tobit II panel selection model. Raymond *et al.* (2010) apply the CDM model to panel data while accounting for individual effects but do not investigate any persistence effects in innovation or productivity.

[4]For instance, Klomp and van Leeuwen (2001) model a productivity feedback effect using cross-sectional data; Lööf and Heshmati (2002) model dynamics using panel data on two CIS waves but do not control for individual heterogeneity.

[5]A major drawback of CDM studies employing CIS data is that the panel dimension is short with overlapping years, data for one out of three years, with big yearly gaps and CIS waves are firm specific, i.e., different sample firms are collected in each survey.

Third, we take a closer look at the fundamental methodological issue to measure productivity growth. In the above mentioned studies that examine the empirical evidence on innovation and productivity growth in CDM framework, several important aspects are overlooked. First, is that in almost all studies, productivity is interpreted as a partial measure; the most common are "labor productivity", measured as output per man-year or hours worked. A notably exception is the study of Parisi *et al.* (2006) who use a more comprehensive measure of (total or) multi-factor productivity (MFP), which is used when inputs are comprehensive (that is, capital, labor, energy, materials, and services are taken into account). This measure is considered to be a more realistic representation of the entire production process (e.g., Vancauteren et al., 2009; Syverson, 2010). Second, if a market is assumed to be characterized by some form of imperfect competition in the output market, this assumption should also be reflected in the construction of a TFP measure to take into account possible biases. Often, productivity is measured under the neoclassical assumptions of perfect competition and constant return to scale. This implies, amongst others, that there is no market power and production units make zero profits. To conform to current research (e.g., Amoroso *et al.* (2010), Vancauteren *et al.* 2009), the empirical evidence shows that perfect competition assumptions are not adequate for Dutch firms. In particular, markups are significantly and fairly larger than one and TFP annual growth rates are statistically and significantly affected when relaxing the assumption of perfect competition.

Fourth, we take a closer look at the role of firm-level characteristics that may affect the innovation process at different stages and TFP growth. By merging data at the firm-level from different sources, we pay specific attention to the firm's ownership structure (group membership, foreign ownership), and economies of scope (range of goods) which a deemed to be important drivers of innovation. Usually these effects (known as technology push and demand pull effects) are captured by survey data on the basis of Likert scales.

The paper is organized as follows. Section 2 presents a quick review of literature dealing with the R&D-patents-productivity relationship. Section 3 presents the empirical model. Section 4 describes the data set. In Section 5, we estimate 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 (Pakes and Grilliches, 1984; Hausman *et al.*, 1984) and, second, studies analyzing the relationship between innovation output and productivity growth (Kortum, 1993; Bloom and Van Reenen, 2002). A major importance in studies based on (variants of) the CDM model, is that the widespread distinction between patents and R&D expenditure is not treated in an isolated way but enters the model explicitly. 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 raises some important issues.

First, it 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 large 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), or Kortum (1993)).

Romer (1990) theoretically defines three R&D externalities that yield an interesting interpretation on the further dynamics between R&D and patents. 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. Second, since models 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 firms owning the old technology. The dynamic specification of our empirical model enables us to identify some of these effects; see Section 5.2.

Second, the literature that focuses on a direct link from R&D to productivity (see, for instance, Griffith *et al.*, 2006, Hall and Mairesse, 1995; Balcombe *et al.*, 2005) is not supported in a CDM framework. The theoretical support 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 recent paper Balcombe *et al.* (2005) find on the basis of 1955-2000 time-series data and experimental data on agricultural innovation, a distributed lag-length between 9 to 10 years. An explanation of 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.

3 Empirical Implementation

The model starts from the assumption that it is patent counts that affects productivity and not R&D. The model builds upon the CDM model in a panel data context, where R&D availability is not necessarily a prerequisite. Therefore, the model does not depend on the role of R&D in explaining innovation output and its impact on productivity growth. Examples of empirical studies that use a similar R&D selection criterion, in a cross sectional dimension, are given by Griffith *et al.* (2006), Klomp and van Leeuwen (2006), and Hall *et al.* (2009). The EPO patent count data is fully observed in our sample.

The model consists of three parts that relate (i) firm characteristics to R&D, (ii) R&D to patenting, and (iii) patenting to total factor productivity growth (TFPG). We consider each of these parts in the subsequent subsections.

3.1 The R&D Equation

We adopt a sample selection Tobit-II model[6] with censoring consisting of two equations, where the first one is a Probit equation determining whether a firm reports its R&D and

[6]See, for example, Amemiya (1985) on Tobit-II.

the second equation is a censored regression explaining the amount of R&D invested, which is censored at zero in case R&D is zero. We only observe the amount of R&D invested in case the firm reports its R&D. We extend this Tobit-II model to a panel data context, following Wooldridge (2005), which enables us to exploit the unobserved heterogeneity dimension at the individual firm level.

Let a firm be indicated by the subindex i and time by the subindex t. Firm i's R&D (expenditure) effort at time t is written as:

$$R\&D_{it}^{*} = a_{1i} + \beta_1' \mathbf{x}_{1it} + \varepsilon_{1it}, \tag{1}$$

where $R\&D_{it}^{*}$ is a latent variable representing the firm's R&D effort, a_{1i} reflects the firm-specific unobserved heterogeneity, $\mathbf{x}_{1it}$ is a vector of independent variables representing firm characteristics, and ε_{1it} is a random error. *Ex ante*, firm i engages in nonzero R&D expenditure in year t if $R\&D_{it}^{*} \geqslant 0$.

The second equation is specified by a binary variable REP_{it} that is equal to one when R&D is reported by firm i in year t and zero otherwise. That is,

$$\begin{aligned} REP_{it} &= 1 \text{ if } REP_{it}^{*} = a_{2i} + \beta_2' \mathbf{x}_{2it} + \varepsilon_{2it} > 0, \\ &= 0 \text{ otherwise}, \end{aligned} \tag{2}$$

where REP_{it}^{*} is the corresponding latent variable, a_{2i} represents the firm-specific heterogeneity, $\mathbf{x}_{2it}$ is a second vector of (firm-related) independent variables, and ε_{2it} is an error term.

Combining (1)–(2), we have:

$$
\begin{aligned}
R\&D_{it} &= R\&D^*_{it} \quad \text{if} \quad REP_{it}=1 \quad \text{and} \quad R\&D^*_{it} \geqslant 0, \\
&= 0 \quad \text{if} \quad REP_{it}=1 \quad \text{and} \quad R\&D^*_{it} < 0, \\
&= \text{unobserved if} \quad REP_{it}=0.
\end{aligned} \tag{3}
$$

Thus, we set the actual R&D expenditure $R\&D_{it}$ equal to $R\&D^*_{it}$ in case $R\&D^*_{it} \geq 0$ *and* R&D is reported ($REP_{it} = 1$). If $R\&D^*_{it} < 0$ *and* R&D is reported, we set $R\&D_{it} = 0$. In all other cases R&D is not observed ($REP_{it} = 0$). The model is completed by assuming that the unobserved errors ε_{1it} and ε_{2it}, conditional upon $\mathbf{x}_{1it}$ and $\mathbf{x}_{2it}$, follow a bivariate normal distribution with zero mean, variances $\sigma_1^2(=1)$ and σ_2^2, and covariance $\sigma_{12} = \rho\sigma_2$, where ρ is the correlation between the two error terms ε_{1it} and ε_{2it}.

To illustrate the relevance of the selection equation in our Tobit-II model, we shall also apply as comparison the standard Tobit-I model for the sub-sample of firms which report R&D.

As independent variables to explain the probability of reporting R&D, we include in the vector $\mathbf{x}_{1it}$ the (log) employment (employment_{it}), industry sector dummies (indicated by α_k), industry sector's competitive pressure in logs (comp_{kt}), a dummy variable representing whether the firm's headquarter is located abroad (foreign_{it}), a variable indicating the number of domestic firms under complete control ($\#\text{firms}_{it}$), and the number of industry segments (P_i). As independent variables in the REP^* selection equation, we include in the vector $\mathbf{x}_{2it}$ the same variables as in $\mathbf{x}_{1it}$, with the exception that $\#\ \text{firms}_{it}$ and foreign_{it} are replaced by a dummy variable reflecting whether the firm is part of a group (group_{it}). The dummy variable group_{it} is a linear combination of foreign_{it} and ownership_{it} and equals to 1 if the firm is foreign-controlled and has a direct ownership of multiple domestic firms.[7]

[7]Since the vector $\mathbf{x}_{2it}$ is not equal to $\mathbf{x}_{1it}$, we allow for an exclusion restriction, which is typical for a

Industry dummies capture technological opportunities or structural effects (see Mairesse and Mohnen, 2002) while size measured by the number of employees is, following Schumpeter (1942), positively related to innovation. R&D and economies of scope are related because innovation may spill over to different projects (for example, Piga and Vivarelli, 2004; Filatotchev *et al.*, 2003). This research provides evidence that an organizational form that consists of multiple firms may have better control to external finance, thereby suggesting that they are more likely to intensify their R&D. We proxy scope economies by the number of industry segments P_i and the number of domestic firms (control number of firms_{it}). The number of industry segments match for each firm i its business activities that corresponds to the number of different 3-digit NACE codes. For example $P_i = 2$ if a firm's activities belongs to the NACE151 and 152 activity codes. In the selection equation, the group membership dummy (group_{it}) 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). The dummy variable foreign_{it}, representing whether the firm is controlled by a foreign firm, is also included in the R&D level equation. 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 easier better access to innovation endowments from the MNE and other subsidiaries (Un and Cuervo-Cazurra, 2008). 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). As a final control variable, we also take into account how competitive pressures affect the firm's R&D intensity, we follow Martin *et al.* (2011) and measure the level of competition using a Herfindahl index of

sample selection model (see for example, Vella, 1998).

industrial concentration:

$$H_{ikt} = \sum_{i \in S_{kt}} \left(\frac{\text{employees}_{ikt}}{\text{employees}_{kt}} \right)^2$$

where S_{kt} is the set of firms belonging to industry sector k at time t. The variable $\text{comp}_{kt} = \ln\left(\frac{1}{H_{ikt}}\right)$ measures the degree of competition a firm i of sector k faces at time t.[8]

Following the approach proposed by Wooldridge (2005), we use maximum likelihood (ML) to estimate the model with individual effects. We assume the distribution of the unobserved individual effects (a_{1i} and a_{2i}) to be modeled as follows,

$$\begin{aligned} a_{1i} &= \alpha_{10} + \delta_{10} R\&D^*_{i0} + \delta_1' \overline{\mathbf{x}}_{i1} + \xi_{1i}, \\ a_{2i} &= \alpha_{20} + \delta_{20} REP^*_{i0} + \delta_2' \overline{\mathbf{x}}_{i2} + \xi_{12}, \end{aligned} \tag{4}$$

where α_{10} and α_{20} are constants, $\overline{\mathbf{x}}_{i1} = \overline{\mathbf{x}}_{i2}$ are vectors including the time averages of the variables e_{it}, s_{it}, and l_{it},[9] $R\&D^*_{i0}$ and REP^*_{i0} are initial values, δ_{10}, δ_{20}, δ_1' and δ_2' are the corresponding (vectors of) coefficients to be estimated, and ξ_{1i} and ξ_{2i} are assumed to be independent error terms following normal distributions $\xi_{1i}|\mathbf{x}_{i1} \sim N(0, \sigma^2_{\xi_1})$ and $\xi_{2i}|\mathbf{x}_{i2} \sim N(0, \sigma^2_{\xi_2})$.[10]

[8] 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, we employ variables that are fully observed for almost the entire sample in order to optimize our results. This data are extracted from the "General Business Register" and are recorded for each firm located in the Netherlands. Input ((capital, wages) and output (sales) variables which are employed in the TFP equation (see Section 3.3.) are extracted from the Statistics of Finance of Enterprizes and are only available for smaller selection of firms.

[9] Instead of $\overline{\mathbf{x}}_i$, the original estimator uses $\mathbf{X}_i \equiv (\mathbf{x}_{1i1}, \ldots, \mathbf{x}_{1it})$, but time averages allow for a reduction of explanatory variables (see Wooldridge, 2005).

[10] We estimate the model using Stata 11 applying the Gllamm program that was developed in Miranda and Rabe-Hesketh (2006). Gllamm uses the adaptive Gauss-Hermite quadrature to approximate the integrals involved, and the default maximization approach is the quasi-(Gauss-)Newton method.

3.2 The Patent Equation

The next part of the model explains the innovation output measured by either the number of patents filed in a given year or the total number of forward citations received from these patents (see Data Section). We first focus on the number of patents filed in a given year. The discreteness of patent data motivates the use of a count model (Nesta and Saviotti, 2005). 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 never been granted with a patent during our entire sample period.[11] 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 zero-inflated count model, allowing for unobserved heterogeneity by means of random effects include, for example, Hall (2000) and Min and Agresti (2005).[12]

Introduce

$$P_0(y, \lambda_{it}) \equiv \exp(-\lambda_{it})\lambda_{it}^{y}/y!, \qquad y \in \{0, 1, 2, ...\},$$

where λ_{it} is the Poisson distribution parameter. Let PAT_{it} be the number of patents or the number of forward citations-weighted patents. We assume that it follows the following

[11] We find that 79% of firms have not been granted with a patent. In similar studies: Bound *et al.* (1984) observe that zero patent firms represent 60% of their sample; in Crépon and Duguet (1997), these firms represent 73% of their sample.

[12] Two types of extra zeros can occur: one arising from the zero state and the other from the ordinary count model such as the Poisson or negative binomial with one that is degenerated at zero (Lambert, 1992).

zero inflated Poisson distribution,[13]

$$\Pr(PAT_{it} = y) = (1 - p_{it})P_0(y, 0) + p_{it}P_0(y, \lambda_{it}), \tag{5}$$

where $1 - p_{it}$ represents the probability of extra zeros. We model $\ln \lambda_{it}$ as

$$\ln \lambda_{it} = (a_{3i} + \gamma R\&D_{it} + \beta_3' \mathbf{x}_{3it})$$

where a_{3i} is a time-invariant unobserved firm-effect, and the vector of additional independent variables $\mathbf{x}_{3it}$ is the same as defined in equation (1). The inclusion of these variables is motivated by previous patent studies (see, for example, Menon, 2010; Peters, 2009). In case the firm's $R\&D_{it}$ is unobserved, we take its predicted value from the model described in the previous subsection. We model the probability p_{it} as a logit model

$$p_{it} = \frac{\exp(\gamma' \mathbf{z}_{it})}{1 + \exp(\gamma' \mathbf{z}_{it})}$$

where $\mathbf{z}_{it}$ is the vector of the zero-inflated covariates and γ is the vector of the zero-inflated coefficients to be estimated. Finally, the time-invariant unobserved firm-effects a_{3i} are assumed to be standard normal distributed (conditionally on $\mathbf{x}_{1it}$ and $\mathbf{z}_{it}$).

Alternatively, PAT_{it} can also be modeled to follow a zero-inflated negative binomial distribution (ZINB), which we obtain from (5) by considering the negative binomial count data model, *i.e.*, $P_0(y, \lambda_{it})$ in (5) is modeled as

$$P_0(y, \lambda_{it}) = \frac{\Gamma(\alpha^{-1} + y)}{\Gamma(\alpha^{-1}) + \Gamma(1 + y)} \left(\frac{\alpha^{-1}}{\alpha^{-1} + \lambda_{it}} \right)^{\alpha^{-1}} \left(\frac{\lambda_{it}}{\alpha^{-1} + \lambda_{it}} \right)^{y}, \tag{6}$$

where $\Gamma(\cdot)$ denotes the gamma function, and where α captures the deviation from the

[13] The zero-inflated Poisson model can be seen as a mixture model having two classes, where the first class, $P_0(y_{it}, 0)$, has a fixed value at 0 (Böhning *et al.*, 1999).

Poisson distribution. This model, which is particularly suited for overdispersed data, reduces to the ZIP when $\alpha = 0$. Thus, we can use a standard test on α to discriminate between the negative binomial and the Poisson model.[14] We estimate this part of the model using Maximum Likelihood (ML).[15]

In Section 5.2. we will also consider as alternative a weighted number of forward citations per firm (see Hall *et al.*, 2002). For the weights, we divide the number of forward citations of each individual patent with the average number of forward citations of patents of the same publication year and technological class (see Data Section).

3.3 The TFP Equation

The third part of the model explains the total factor productivity (TFP) growth. We rely on previous research (for example, Amoroso *et al.*, 2010; Diewert and Fox, 2009) and use a measure of TFP growth that is adjusted to imperfect competition in the output market (markups). Our strategy to measure TFP and markups builds on previous work by Amoroso *et al.* (2010), but with some differences. First, TFP in Amoroso *et al.* (2010) is estimated allowing for imperfect competition in the output and input markets, while in this paper we construct TFP from an exact index number approach using a translog production function. Second, the markup measure in Amoroso *et al.* (2010) is recovered from an estimated, sector-specific coefficient that is made time-varying through an interaction with trend polynomials, while the markup in this paper equals an estimated

[14] It should be noted that misspecification of the model (that is, ZINB or ZIP) is more serious in the zero-inflated context than in the standard situation, since the Poisson assumptions of parameter estimators are not consistent if the actual distribution is negative binomial.

[15] In our random effect zero-inflated count model the random effects are assumed to be standard normal variables multiplied by the standard normal probability density function that enter the loglikelihood function. The complete log-likelihood for the zero-inflated count model with random effects is given by, $\ln L = \sum_{i=1} \log \phi(b_i) + \sum_{i=1} \sum_{t=1} z_i \log(p_{it}) + (1-z_{it}) \log(1-p_{it}) + z_{it} \log[Po\{y_{it}, \exp(b_i + \gamma_1 R\&D_{it} + \beta_3' \mathbf{x}_{i3t})\}]$, where ϕ denotes the standard normal probability density function and z_{it} is an indicator variable which is equal to 1 if $y > 0$, and 0 if $y = 0$ (see Böhning *et al.* (pp. 207-209, 1999) and Hall (2000) for technical details regarding the calculation of the ML estimates). The above calculations are based on the assumption that zero and non-zero patent counts are independent of the explanatory variables.

firm-level scale elasticity times the output-input ratio that is computed with real data at the level of the firm.

We estimate the following model

$$\log TFP_{it} = \alpha_{4i} + \gamma_2 \log PAT_{it} + \beta_4' \mathbf{x}_{4it} + \varepsilon_{i4t} \quad (7)$$

where TFP_{it} is TFP growth (see Appendix A.3 for the formulation), α_{4i} is the time-invariant unobserved firm-effect, ε_{4it} is the error term, assumed to follow a normal distribution, and the vector of additional explanatory variables $\mathbf{x}_{4it}$ includes industry and time dummies, employment_{it}, comp_{kt}, foreign_{it}, log of age (age_{it}), markups (markup_{it}), and a firm's capital intensity, measured by the log of the capital-labor ratio $(K/L)_{it}$, where we include the firm's capital intensity and employment since these have been shown to increase factor productivity (Bernard *et al.*, 2006). Equation (7) also includes two competition variables. The inverse of the Herfindahl index of industrial concentration ratio (comp_{kt}) measures the concentration of firms within a market, whereas the markup (measured by equation (**??**) in the appendix) measures the profitability of firms as it reflects a firm's ability to set its prices above marginal costs. While both 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. We also include plant age, in order to account for life-cycle effects, and a foreign ownership dummy (foreign_{it}) which is motivated by recent literature on firm heterogeneity (for example, Bernard *et al.*, 2006). For the firm's innovative output PAT_{it}, we take its predicted values from the random ZIP or ZINB model. We estimate equation (7) by an iterated ML version of the generalized least squares method allowing for random effects.

3.4 Extension to Dynamics

In this section, we discuss extensions to a dynamic framework. Generally, micro level studies that look at the dynamics of the patent-R&D relationship show evidence of persistence of innovation. The persistence of output (input) innovation is defined as a significant and positive effect of past output (input) innovation on current output (input) innovation activities. Firms may innovate persistently over time for a couple of reasons. As highlighted in Section 2, the dynamics of a firm's innovation behavior is an essential assumption for endogenous growth models (Romer, 1990; Aghion and Howitt, 1992). These models rationalize the idea of intertemporal complementarity in innovation: past experience makes current innovation efforts more productive. Further theoretical underpinnings of such persistence innovation behavior can be explained. There is the so called "success breeds success" hypothesis. This hypothesis postulates that firms become more prosperous through successful innovation due to broader technological opportunities. Thus, the accumulation of knowledge would induce state dependence in invention flows and consequently persistence in innovation. As a 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 in important source for persistence since they create barriers to entry and create engagements to continue innovation.

With specific reference to the Netherlands, existing studies that have investigated the dynamic relationship between R&D and patenting include van Leeuwen (2002) and Raymond *et al.* (forthcoming).[16] These studies use respectively the [1994-1996, 1996-1998] and [1994-1996, 1996-1998, 1998-2000], [1994-1996, 1996-1998, 1998-2000, 2000-2002 and 2002-2004] biannual waves of the Dutch CIS data. Both studies confirm a persistence of innovation. In the Raymond *et al.* (forthcoming) paper, individual effects are accounted

[16] Other frequently cited empirical studies confirming strong and weak forms of persistence in R&D and patents include Cefis and Orsenigo (2001), Geroski et al. (1997) and Peters (2007).

for and modeled in a dynamic Tobit II panel selection model similar to the one outlined in Section 3.1. The authors found that with respect to the input innovation past R&D/sales expenditures affect 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 expenditures/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.

We shall assume that in the R&D input equation, the past level of both the stock of patents and R&D expenditures affect current R&D. This assumption can be justified as follows. First, it is based on the so called complementarity assumption in the knowledge production, where current levels of innovation may be affected by past levels of both R&D and patents (Klette, 1996; van Leeuwen, 2002). Second, this dynamic formulation is also in line with the R&D-patent externalities discussed in Section 2. More specifically, once patents are issued, this may generate positive or negative externalities on current R&D activities. In case of the patent equation we shall assume that the current stock of patents is affected by its past stock and current level of R&D (taken at its predicted value). In the TFP equation, we shall include the lagged number of patents or patents weighted by citations.

4 Data

Our data consists of an unbalanced panel of 3026 firms, situated in the Netherlands, during the period 2000-2006, collected from different data sources. The Appendix contains a detailed description of the data collection procedure. Our resulting sample contains 2786 firms that have applied at least for one patent during the years 2000-2006. These firms

Table 1: Innovation Data Sample

	R&D reported	R&D not reported	Total firms	Total EPO patens
total firms	1166	1860	3026	31509
of which:				
patent firms (> 100)	20	3	23	22451
patent firms ($> 1, < 100$)	466	518	1007	7279
patent firms ($= 1$)	430	1349	1779	1779
patent firms ($= 0$)	250	0	250	0

represent 98 percent of all patent applications during this period. As control group we have 250 firms with no patent activities, whose R&D is reported in any of the years of our sample. Table 1 presents some descriptive statistics on the R&D and patent behavior of the sampled firms. The column "R&D reported" shows the firms with R&D reported values in any of the sample years while the column "R&D not reported" presents the firms where R&D is reported as missing throughout the sample period. For the purpose of this paper, we make a distinction among firms whose R&D is reported by "zero" or "missing."

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, 23 out of a total of 3026 firms amount to the majority of EPO patents. To shed some light on the sector characteristics of the group of "R&D reported" firms, Table A1 in the Appendix reports some additional information on R&D statistics and patent averages broken down into 19 industries. According to this table, the most important patents and R&D activities of these firms are found in food, machinery & equipment, chemicals, and electrotechnical equipment.

5 Estimation Results

In this section we present the estimation results. First, we look at the results based on the static and dynamic models, with number of patents or number of number of forward citations weighted patents, and in the second subsection we look at the sensitivity of the model using the fixed-weighted number of forward citations per patent.

5.1 Results of the Basic Model

The R&D equation—We start with the Tobit II model discussed in Section 3.1. Three variants have been estimated, see Table 2. In the first column of table 2, we present the estimation results of a model without the random effects initial conditions. The second columns shows the results of a model with random effects and the initial conditions, while column 3 presents the estimation results of a dynamic model, including the preceding innovation status (a binary indicator). Accounting for individual effects, initial conditions, and dynamics does not largely affect the estimation results. As one can see from the correlation between the errors (ρ), the selection bias is significant.

The first point to observe is that R&D activities and the effect of the number of patents issued in the previous year affect positively and significantly both the probability and level of R&D intensities. Firm size, measured as the number of employees, has a positive and significant effect both on the probability to do R&D and the R&D intensity which supports the Schumpeterian hypothesis. The significant impact of size on the probability to engage in R&D may be due to the data selection problem, since essentially large firms report this effort. Our result that firm size increases the R&D intensity is also confirmed in Raymond *et al.* (2010) for the Netherlands. In that study, R&D is linked to innovative sales in a pooled model in 1994-1996, 1998–2000 and 2002-2004.

Among the other controls, competition and diversification in activities have a positive

Table 2: R&D equations: selection and intensity. ML-estimates with (robust) standard errors in parentheses. Reported in the probit equation is the change in probability (that R&D is greater than or equal to zero) for a unit change in each of the explanatory variables. Chi-industry reports the p-value of the joint significance of the industry dummies. Sigma_$a_1(a_2)$ shows the variance of the $a_1(a_2)$ random effects. ρ is the estimated correlation coefficient between the sample selection (probit) and R&D expenditure (Tobit) equations. The last column of the table reports the Tobit I estimates as a robustness check.

Dependent variable	Probit (R&D =0/1)	Log of R&D expenditure per employee	Probit (R&D =0/1)	Log of R&D expenditure per employee	Probit (R&D =0/1)	Log of R&D expenditure per employee	Log of R&D expenditure per employee
Past R&D Yes/No (1/0)					.837 (.046)	1.036 (.368)	.480 (.526)
Past Patents Yes/No (1/0)					.520 (.045)	.562 (.157)	.209 (.145)
Log(Employment)	.372 (.009)	.474 (.031)	.302 (.009)	.652 (.140)	.191 (.036)	.380 (.073)	-.383 (.036)
Log(Competition)	.114 (.053)	.158 (.118)	.120 (.051)	.457 (.430)	-.010 (.063)	.127 (.323)	.009 (.102)
Number Activities	.034 (.007)	.106 (.090)	.023 (.008)	.018 (.194)	.002 (.010)	.060 (.050)	-.023 (.020)
Foreign Yes/No (1/0)		.233 (.127)		.419 (.521)		.178 (.086)	.285 (.092)
Number of Domestic Firms	-.003 (.001)	-.017 (.015)	-.002 (.002)	-.013 (.070)	-.002 (.002)	-.011 (.061)	-.003 (.004)
Group	.338 (.033)		.264 (.035)		.369 (.037)		
Intercept	-2.742 (.092)	3.781 (2.901)	-2.607 (.091)	6.691 (3.708)	-3.285 (.133)	3.856 (2.950)	2.673 (.293)
Initial log of R&D per employee (R&D$_{i0}$)	-	-		.023 (.061)		.096 (.035)	
R&D Yes/No (z_{i0})	-	-	.831 (.031)		.011 (.044)		
Sigma_a$_1$			.754 (.235)		.450 (.110)		
Sigma_a$_2$				.317 (.151)		.282 (.104)	
$\rho\varepsilon_1\varepsilon_2$	.912 (.000)		.651 (.325)		.678 (.017)		
Log-likelihood	-5441.653		-5098.68		-4343.87		-4412.62
N Observation Censored Uncensored	14422 10536 3886		14422 10536 3886		12547 9258 3289		3289 2046 1243
Estimation Method	**Tobit II**						**Tobit I**

and significant impact on the probability, but not on the level of doing R&D. In addition, once we control for dynamics, these variables become insignificant. Next, we find that group membership is a major driver in the probability of reporting R&D. In the dynamic model, we find evidence that foreign owned firms do more R&D while the number of activities, as a measure of scope, does not seem to explain innovation in R&D.

In the last column of Table 2, we run the same regression, but only for a sample of firms where R&D is always reported (Tobit I). The results are quite different from the other models, indicating the importance of incorporating our selection equation. In case of Tobit I, the only significant findings are that less employment and foreign-owned firms have a positive impact on the R&D level.

The patent equation—We consider next the estimates of the patent equation discussed in Subsection 3.2. Maximum Likelihood-based estimation results for the ZINB model are reported in Table 3 (columns 1-5). The likelihood-ratio test (not reported) comparing the ZINB model with the ZIP model reveals that in all cases, the ZINB model is preferred. In order to test the zero-inflation of the model, we used the Vuong test (Vuong, 1989) which identifies which of the conditional models (ZINB versus negative binomial; ZIP versus Poisson) is the closest to the "true" distribution. The model is estimated for both the actual and he citation-weighted patent counts.

R&D (per employee) turns out to be a major determinant in generating new knowledge. As expected, the R&D intensity variable (which is the corresponding elasticity of the number of patents with respect to R&D), taken at its predicted values from the Tobit II equations (Table 2, column 3), has a significant and positive impact on the number of the actual and citation-weighted patents. The magnitude ranging between 0.15 and 0.19 indicates a robust result throughout each of the specifications. These results do not change once we replace the instantaneous effect of R&D with a one-year lag. In the CDM literature, elasticities in the range of $.4 - 1.1$ are typically found. Our estimate of the

Table 3: Patent equations. ML-estimates with (robust) standard errors in parentheses.

Dependent variable	Number of EPO patents	Number of EPO patents	Number of citation-weighted patens	Number of EPO patents	Number of citation-weighted patens
Past Patents Yes/No (1/0)				1.061 (.081)	1.312 (.138)
Log(R&D per employee)	.148 (.047)	.195 (.031)	.184 (.048)	.167 (.028)	.161 (.042)
Log(Employment)		.294 (.026)	.225 (.056)	.235 (.022)	.173 (.038)
Log(Competition)		-.037 (.077)	-.1354 (.266)	-.084 (.074)	-.037 (.201)
Number Activities		.066 (.033)	.014 (.051)	.067 (.029)	-.007 (.042)
Foreign Yes/No (1/0)		-.016 (.144)	-.198 (.338)	.005 (.136)	-.227 (.226)
Number of Domestic Firms		-.000 (.010)	-.003 (.016)	-.000 (.012)	.004 (.012)
Group		.185 (.112)	.265 (.214)	.074 (.096)	.056 (.155)
Intercept	-1.700 (.192)	-3.113 (.421)	.579 (7.644)	-2.961 (.396)	-32.640 (.933)
Log-likelihood	-14004.46	-12694.65	-6658.64	-12443.63	-6575.79
N Observations	13515	12820	12820	12820	12820
Industry-effects Year-effect	Yes Yes	Yes Yes	Yes Yes	Yes Yes	Yes Yes
Estimation Method	ZINB-RE				

R&D is very similar to the 0.18 elasticity which was reported in Raymond et al. (2010).

In the patent equation, the persistence of output innovation is also ascertained at the 5% level, and we still observe a statistical and economic significance on the effect of contemporaneous R&D intensity on the actual and number of citations-weighted patents. We also find that by taking into account the past level of patent stocks, the R&D elasticity remains at a similar magnitude in comparison to our static results. This result does not confirm the conjecture found in Klette (1996) stating that the current level of innovation is due to the past history of innovation and, therefore, the "innovation opportunities of the most recent R&D may be small."

Concerning the role played by the other control variables, we find that the elasticity of size also exhibits a significant and positive effect. This can be interpreted towards a higher productivity of innovation activities of larger firms. Controlling for firms within the enterprize group also exhibits a positive effect on patenting in the static equations, implying that as a firm belongs to a group, it produces more patents. The remaining variables are not significant. These results are also in line with the R&D Tobit II estimation results. We should also note that, despite a large correlation between R&D and patents, these coefficients change only slightly when we remove the other variables one-by-one.

The TFP equation—Finally, we consider our estimates of the TFP equations (see Subsection 3.3) in Table 4. We estimated the (log-log) equations by an iterated ML version of the generalized least squares method allowing for random effects. Most notably, we find significant estimates of the impact of the one year-lagged (predicted) output innovation on TFP growth with an elasticity of 0.16. When we consider the instantaneous effect, the results are less convincing. The negative and significant effect of employment on productivity is not in line with many other empirical studies; however, Raymond *et al.* (2010) also report a similar negative elasticity of 0.06 for the Netherlands, suggesting that smaller firms are more productive. Competition measured by using the inverse of the

Herfindahl index is not significant. When we include markups, the results suggest that less competition in the output market lead to more TFP growth. The impact of the inverse of the Herfindahl index is robust to omitting markups. Our results confirm the idea that using a single variable for competition may yield incomplete results. 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, 2010): 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. Finally, the effects of the capital/labor ratio and age are significant and have the expected positive sign.

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 .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 (2002) were insignificant while Raymond et al. (2010) report a positive and significant effect using more recent data.

5.2 Sensitivity Analysis

In this section we investigate the sensitivity of the CDM model using the fixed-weighted number of forward citations per patent rather than the count of forward citations per patent. As a generalization of the Tobit model applied to continues data, we perform

Table 4: TFP growth equations. ML-estimates with (robust) standard errors in parentheses.

Dependent variable	TFP growth	TFP growth	TFP growth	TFP growth
Log(1+Pat_Citations)	-.036 (.034)	.059 (.034)		
Lagged Log(1+Pat_Citations)			.165 (.035)	
Lagged Log(1+Pat_Count)				.170 (.043)
Log(Employment)		-.014 (.006)	-.027 (.006)	-.044 (.007)
Log(Competition)		.077 (.041)	.066 (.038)	.052 (.041)
Foreign Yes/No (1/0)		.016 (.027)	.020 (.027)	.031 (.027)
Log(Age)		.010 (.009)	.012 (.009)	.016 (.009)
Log(K/L)		.158 (.017)	.168 (.017)	.176 (.017)
Markups		.119 (.014)	.119 (.013)	.120 (.013)
Intercept	.164 (.023)	.259 (.145)	-.426 (.152)	-.371 (.152)
Log-likelihood	-7074.53	-6618.42	-6608.79	-6600.03
N Observations	6135	6040	6040	6040
Industry-effects Year-effect	Yes Yes	Yes Yes	Yes Yes	Yes Yes
Estimation Method	MLE			

a ML-estimation of a two-part model allowing for random effects. The two-part model, sometimes referred to as the Hurdle model, rests on the assumption that the zero and positive values are generated by different mechanisms and allows for a more flexible specification than the standard Tobit model.[17] The first part of the model consists of an estimation of a discrete choice model (Probit in our case) for the patent's probabilities using all observations. The second part of the model estimates a regression equation using only observations when the patent counts are positive.

The economic rationale for applying the Hurdle model rests on the assumption that the decision to apply for a patent and the decision to apply for additional patents might be ruled by different processes. The decision to protect patentable inventions is usually made on the basis of a first invention and the decision to apply for additional patentable inventions is based on this first decision. So, we might expect different decision criteria concerning the first patent and additional patents.[18] Regression results for the two-part Tobit model are reported in Table 5 where now we work with the logarithm of the fixed-weighted number of forward citations per patent as the dependent variable. To compare with our previous results, we also apply the Hurdle model to the number of forward citations per patent (citation-weighted patents).

Overall, we find remarkable differences on the effect of R&D on patents, when we compare our results to the zero-inflated count model estimations (see Table 3). The two-step estimates show that differences in coefficients of variables between the two parts are apparent. The elasticity of the R&D expenditures' contribution is statistically significant and explains positively the probability of positive patents, but becomes negative in the log-linear regressions. In addition, the effect of past innovation behavior negatively affects both parts. Furthermore, when we extend this analysis to the productivity equation, the

[17]We refer to Cameron and Trivedi (2005, Chapter 16.4) for a detailed discussion on the Hurdle model.

[18]It is important to note that all firms in our sample are assumed to decide whether to patent or not, since we do not have any information that allows us to distinguish between firms that would never be involved in patenting and others that are involved.

Table 5: Patent equations: Two-Part Model. ML-estimates with (robust) standard errors in parentheses. $PAT = 0$ is an indicator for whether or not patents are positive, estimated by a probit equation allowing for random effects. The $PAT > 0$ equation uses the natural logarithm of the adjusted number of forward citations per patent as the dependent variable. This equation is estimated by the ML procedure of the generalized least squares model allowing for random effects. We report the joint log-likelihood for the two-parts as the sum of the two-part equations' log likelihoods.

Dependent variable	**Log of fixed citation-weighted patent**		**Citation-weighted patent**			**Log of TFP growth**
	TWO-PART		TWO-PART			
	PAT=0	PAT>0	PAT=0	PAT>0		
Equation 1					**Equation 3**	
Log (R&D per employ.)	.034 (.013)	-.012 (.051)	.123 (.007)	.155 (.036)	Log of fixed citation-weighted patents (t-1)	-.096 (.068)
Log-likelihood	-5384.98		-9935.53			-6566.38
Equation 2					**Equation 4**	
Log (R&D per employ.)	.039 (.013)	-.015 (.071)	.074 (.008)	.151 (.032)	Log of Citation-weighted patents (t-1)	.032 (.017)
Past Patents Yes/No	-.396 (.052)	-.178 (.038)	.217 (.039)	.233 (.184)		
Number of observations	14422	1253	14422	1253		6040
Log-likelihood	-5344.91		-9337.45			-6614.43

statistical significance of the patent variable disappear. Overall, it is shown that similar conclusions can be made from the two-step Tobit model when we use the count of forward citations per patent.

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 1997-2005. We estimate a structural model that disentangles the impact of R&D expenditures, the number of patents issued, and total factor 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 one-year lag patents have had an economic and statistical impact of firm-level productivity with an elasticity of 0.16. This estimate is consistent with those of other country studies that are merely based on a cross-sectional analysis. *Second*, our investigation on the contribution of R&D on patents is in line with previous research where a positive relationship is found. However, our estimates suggest that firms seem to have equal success in appropriating (current)

R&D activities into higher patent propensities, once we take into account the persistence of innovation which is also statistically confirmed. This result does not confirm the conjecture found in Klette (1996) stating that the current level of innovation is due to the past history of innovation.

After controlling for a variety of firm characteristics, size (measured by employment) is positively in both the R&D and patent equation, confirming the Schumpeterian hypothesis, but on the other hand our results suggest that smaller firms are more productive. Both for the R&D and patent counts, we find no statistical significance between between competition (measured at the industry level using an Herfindahl index of industrial concentration) and innovation. A general intuitive argument is that these effects may be picked up by the individual heterogeneity. When we consider only the sample of firms with R&D reported values, results change, and the persistence of R&D innovation becomes less significant. As a sensitivity analysis, we find that the positive impact of innovation on productivity disappears if we use the adjusted (fixed) number of forward citations per patent rather than the count of forward citations per patent.

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 R&D non-reporting firms.

References

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

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

[3] Aghion, P., N. Bloom, R. Blundell, R. Griffith and P. Howitt (2005), "Competition and Innovation: An inverted U relationship", *The Quarterly Journal of Economics* 120(2), pp. 701-728.

[4] Amemiya, T. (1985). *Advanced Econometrics.* Harvard University Press.

[5] Amoroso, S., Kort, P., Melenberg, B., Plasmans, J. and M. Vancauteren (2010), "Firm level productivity under imperfect competition in output and labor markets", CESifo Working Paper Series No. 3082.

[6] 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), pp. 49-72.

[7] Bekkers, R., R. Bongard and A. Nuvolari (forthcoming), "Essential Patents in Industry Standards: The Case of UMTS", *Research Policy.*

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

[9] Bernard, A., Jensen B. and P. Schott (2006), "Trade costs, firms and productivity", *The Journal of Monetary Economics*, 53(5), pp. 917-937.

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

[11] Böhning, D., D. Ekkehart, P. Schlattmann, L. Mendonça and U. Kircher (1999), "The Zero-Inflated Poisson Model and the Decayed, Missing and Filled Teeth Index in Dental Epidemiology", *Journal of the Royal Statistical Society*, Series A 162, pp. 195-209.

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

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

[14] Cameron, C. and P. Trivedi (2005). *Microeconomics: Methods and Applications.* Cambridge University Press, New York.

[15] Caves, D., L. Christensen, and E. Diewert (1982), Multilateral comparisons of output, input, and productivity using superlative index numbers", The Economic Journal, 92 (365), pp. 73-86.

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

[17] Crépon B., E. Duguet and J. Mairesse (1988), "Research investment, innovation and productivity", in *Economics of Innovation and New Technology,* B.H. Hall and F. Kramarz (eds.), National Academy Press Book.

[18] 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, pp. 355-378.

[19] Diewert , W.E. (1976), "Exact and superlative index numbers", *Journal of Econometrics*, vol. 4, pp. 114-145.

[20] Diewert, W.E. and K.J. Fox (2009), "On the estimation of returns to scale, technical progress and monopolistic markups", *Journal of Econometrics*, 145, pp. 174-193.

[21] Diewert, E. (1993), "Fisher Ideal Output, Input and Productivity Indexes Revisited" in *Essays in Index Number Theory,* E. Diewert and A. Nakamura (eds.), Elsevier Science Publishers.

[22] Filatotchev I., C. Piga and N. Dyomina (2003), "Network Positioning and R&D Activity: a Study of Italian Groups", *R&D Management* 33, pp. 37-48.

[23] Geroski, P.A., J. van Reenen and C.F. Walters (1997), "How Persistently do Firms Innovate?", *Research Policy* 26(1), pp.33-48.

[24] Gilbert, R. (2006), "Competition and Innovation", *Journal of Industrial Organization Education* 1(1), pp. 1-23.

[25] Griffith R., E. Huergo, J. Mairesse and B. Peters (2006), "Innovation and Productivity across four European Countries.", *Oxford Review of Economic Policy* 22, pp. 483-498.

[26] 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), pp. 971-993.

[27] 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, pp.403-459.

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

[29] Hall, B., F. Lotti and J. Mairesse (2009), "Innovation and productivity in SMEs. Empirical Evidence for Italy", Banca D'Italia, Working Paper 718.

[30] Hall, B., J. Mairesse and P. Mohnen (2010), "Measuring the Returns to R&D", Chapter prepared for the Handbook of the Economics of Innovation, B. H. Hall and N. Rosenberg (eds.), UNU-MERIT Working Paper 2010-006

[31] Hall, D. (2000), "Zero-Inflated Poisson and Binomial Regression with Random Effects: A Case Study", *Biometrics* 56, pp. 1030-1039.

[32] 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), pp. 909-937.

[33] 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), pp. 345-366.

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

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

[36] Konings, J. and H. Vandenbussche (2002), "Does Antidumping Protection raises Market Power? Evidence from EU firm level data", CEPR discussion paper, DP3751.

[37] Konings, J., P. Van Cayseele and F. Warzynski (2001), "The Dynamics of Industrial Markups in Two Small Open Economies: Does National Competition Matter?", *International Journal of Industrial Organization* 19(2), pp. 841-859.

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

[39] Lambert, D. (1992), "Zero-inflated Poisson Regression, with an Application to Defects in Manufacturing", *Technometrics* 34, pp. 1–14.

[40] 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.

[41] 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.

[42] 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), pp. 317-344.

[43] Mairesse J. and P. Mohnen (2002), "Accounting for Innovation and Measuring Innovativeness: An Illustrative Framework and an Application", *American Economic Review* 92(2), pp. 226-230.

[44] Mairesse, J. and P. Mohnen (2007), "A Survey of Innovative Surveys: Taking Stock of a Growing Literature", paper prepared for the CEPR-Banque de France Conference on Innovation.

[45] Mansfield, C. (1980), "Basic Research and Productivity Increase in Manufacturing", *American Economic Review* 70, pp. 863-873.

[46] Martin, Ph., Mayer, T. and F. Mayneris (2011), "Spatial concentration and plant-level productivity in France", *Journal of Urban Economics*, 69(2), pp. 182-195.

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

[48] Miranda, A. and S. Rabe-hesketh (2006), "Maximum Likelihood Estimation of Endogenous Switching and Sample Selection Models for Binary, Ordinal, and Count Variables", *The Stata Journal* 6(3), pp.285-308.

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

[50] Narula, R. and A. Zanfei (2005), "Globalization and Innovation: the Role of Multinational Enterprises", in Fagerberg, J., Mowery, D.C., Nelson, R. (Eds.), Oxford Handbook of Innovation, Oxford University Press, New York.

[51] Nesta, L. and P. Saviotti (2005), "Coherence of the Knowledge Base and the Firm's Innovative Performance: Evidence from the U.S. Pharmaceutical Industry", *Journal of Industrial Economics* LIII(1), pp.123-142.

[52] Nickell, S. (1996), "Competition and Corporate Performance", *Journal of Political Economy* 104(41), pp. 724-746.

[53] 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.

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

[55] 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.

[56] Plasmans, J. and R. Lukach (2009), "The Patterns of Inter-firm and Inter-industry Knowledge Flows in the Netherlands", Working Paper, University of Antwerp.

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

[58] Raymond, W., P. Mohnen, F. Palm and S. Schim van der Loeff (forthcoming), "Persistence of Innovation in Dutch Manufacturing: Is it Spurious?" *The Review of Economics and Statistics.*

[59] Raymond, W., P. Mohnen, F. Palm and S. Schim van der Loeff (2007), "The Behavior of the Maximum Likelihood Estimator of Dynamic Panel Data Sample Selection Models", Cirano Working Paper 2007s-06.

[60] Raymond, W., Mairesse, J., Mohnen, P. and F. Palm (2010), "On the Link Between R&D, Innovation and Productivity: Panel Evidence for Dutch and French Manufacturing", Mimeo.

[61] Romer, P. (1990), "Endogenous Technological Change", *Journal of Political Economy* 98(5), pp. 71-102.

[62] Un, C. A. and A. Cuervo-Cazurra (2008), "Do Subsidiaries of Foreign MNEs Invest More in R&D than Domestic Firms?", *Research Policy* 37, pp. 1812-1828.

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

[64] Syverson, C. (forthcoming), "What Determines Productivity?", *Journal of Economic Literature*, forthcoming.

[65] Vancauteren, M., Balk, B.M., van den Bergen, D. and E. Veldhuizen (2009), "Measures of productivity change: Which outcome do you want?", Bulletin of the International Statistical Institute, Proceedings of the 57th Session of the International Statistical Institute.

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

[67] Vuong, Q. (1989), "Likelihood Ratio Tests for Model Selection and Non-nested Hypotheses", Econometrica 57, pp. 307-333.

[68] Wooldridge, J. (2005), "Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity", *Journal of Applied Econometrics* 20, pp. 39-54.

A Data Construction

We extracted information on output, value-added, net tangible fixed capital assets, sales, depreciation, and wages, all expressed in thousands of Euros. The data is obtained from the "Statistics of Finance of Enterprises" which are provided by Statistics Netherlands. The data on the number of employees, ownership structure, the number of subsidiaries, and the number of industry segments are extracted from the general business register (in Dutch "Algemeen Bedrijven Register," ABR). The exact industry category assignment scheme which we use throughout this paper, based on NACE codes, is presented in Table A1 in the Appendix.

We follow Konings *et al.* (2001) and approximate the capital stock by net fixed capital assets that are valuated at book value, i.e., tangible fixed assets at historic costs minus depreciation. To measure the rental price of capital and its depreciation rate, we follow Konings and Vandenbussche (2002), who set $r_{it} = p_{kt}(\delta_{it} + rr_{kt})$, with r_{it} the rental price of capital of firm i at time t, rr_{kt} the real interest rate measured at industry-level k at time t, δ_{it} the depreciation rate at the firm-level, and p_{kt} the price index of investment goods at the industry level. To calculate the depreciation rate, we divide total depreciation by tangible fixed assets for each firm i. The price index of investment goods is calculated at the industry-level and measured by the capital deflator that is derived from investments in fixed (capital) assets. The real interest rate is calculated by the nominal interest minus the inflation rate, assuming the Fisher hypothesis. We calculate an industry-specific real interest rate by subtracting an industry-specific inflation rate from an economy-wide nominal interest rate. The industry-specific inflation rate is derived from the gross producer price indices. The industry-specific determinants in constructing the rental price of capital are derived from the annual input-output tables provided by Statistics Netherlands (CBS).

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.

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 expenditures in the odd years while each of the CIS waves measures R&D expenditures in the even years of our sample period 2000-2006. Second, we used annual reports of on-line data 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.

We measure the innovative performance of the firm by the number of patents. To construct the patent counts by firm i for each year t, we used a database of the *total* population of patents granted in Europe (issued by the EPO). The 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.

In order to attach the technological importance to each patent, we include information about forward citations. A number of studies (for example, Hall *et al.*, 2002; Bekkers *et al.*, forthcoming; Plasmans and Lukach, 2009) have established the existence of a positive correlation between forward citations and technological importance. For each sample firm, we multiply the number of forward citations for each patent belonging to the enterprize group. 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 of the most recent patents (March 2011).

As an alternative robust measure for the innovative performance of a firm, we construct a fixed-weighted number of forward citations per firm (see Hall *et al.*, 2002). We do so by dividing the number of forward citations of each individual patent with the average number of forward citations of patents of the same publication year and the same international patent classification.

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 expenditures under different names, the ABR data, 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 (enterprize) as a firm that is located in the Netherlands. The total population of the patent counts issued between 2000 and 2006 are matched with all possible subsidiaries and than aggregated to the ultimate patent of the firm.

Table A2 shows the number of patent applications and the number of citations received per application period and per International Patent Class on 1-digit level. It is noted that all patents in this study are published after application. The dataset is restricted to applications in 2000-2006 at the European Patent Office (EPO) by applicants from the Netherlands. Only applications that were granted and hence published are considered. A forward citation means that a patent is cited by a later patent. Journal citations are excluded from these figures. Older patents have a higher probability of receiving a citation. Therefore, the average number of forward citations per patent is higher in 2000-2003 compared to 2004-2006.

Table A1: Summary Statistics I (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	#Firms	#RD firms	AR&D	AR&D Employ	Employ	APat1	APat2	ACit
(01, 05, 10) – Agriculture & mining.	55	14	4551.19	16.78	1330.12	.22	1.33	.63
(11,14) – Crude petrol. & other	10	7	3371.88	4.91	262.41	.18	0	.22
(15-16) – Food prod. & bev.	68	60	5533.21	5.68	1057.33	6.64	21.76	.86
(17-19) – Textiles, clothing & leather	27	14	1233.83	2.56	191.25	.60	3.05	.64
(20-22) – Paper & pulp	67	49	2276.53	4.18	590.77	.43	2.16	.63
(23- 24) –Petroleum prod. & Chemicals.	98	78	13878.44	28.25	738.80	7.08	27.28	.59
(25-26) –Rubber, plastic, glass	118	72	803.75	56.05	252.79	.44	2.43	.70
(27-28) – Basic & fabric. metals	134	90	1772.03	3.12	322.93	.70	3.34	.78
(29) – Mach. & equip.	270	155	4839.65	8.46	178.39	1.07	8.97	.59
(30-33) Electrotech. Equipm.	126	60	33183.14	20.87	589.78	1.66	8.2	.64
(34-35)- Transportation	60	41	5270.12	5.17	412.5	1.06	6.11	.81
(36-37) – Furniture & recylcing	45	22	808.12	3.29	287.00	.25	0	.33
(40-41) – Energy & water	18	16	4605.28	3.99	1841.76	.70	3.16	.09
(45) – Construction	116	55	1303.99	1.28	1033.11	.47	4	.54
(50-51) – Wholesale & maint. vehicles	566	190	2595.90	8.45	128.63	.40	3.26	.60
(52, 55) – Retail trade, hotels & restaurants	71	18	2739.64	1.45	965.05	.30	2.5	.42
(60-67) – Transp., storage, telecom	161	43	15392.72	2.13	3036.21	.78	6.29	.42
(70-74) – Financial & business	875	159	3753.811	65.53	274.30	.65	5.92	.40
(75, 80, 85, 90-93) – Educ. & other public	145	18	2490.13	23.86	1083.97	1.07	4.59	.34

Table A2: Number of patent applications and the number of citations received.

Number of published EPO patents (applied by Dutch firms) and received citations by application year and International Patent Class (1-digit)

	application year	*2000 - 2003*		*2004 - 2006*		*total*	
		patent applications	*forward citations*	*patent applications*	*forward citations*	*patent applications*	*forward citations*
IPC1	*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 Constructions	607	427	532	228	1139	655
	F: Mechanical Engineering, Lighting, Heating, Weapons	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

B Firm-Level Total Factor Productivity

The measure of productivity used in this paper is total factor productivity growth (TFP) allowing for monopolistic behaviour in the output market (mark-ups). The TFP mesure is derived in a production framework. In particular, each firm $i \in \{1, ...N)$ faces the following production function for period t:

$$y_{it} = a_{it} f_i(\mathbf{x_{it}}) \qquad i = 1, 2, \ldots N \ ; \ t = 0, ..., T, \tag{B1}$$

where y_{it} measures firm i's output, $\mathbf{x_{it}} \equiv (x_{i1t}, x_{i2t}, \ldots, x_{iJ_it})'$ denotes the vector of J_i non-negative factor inputs, $f_i(\cdot)$ is the core of the (differentiable) production function and a_{it} is total factor productivity (TFP) measured as Hicks-neutral disembodied technical change.

To derive the TFP measure, we assume that f_i can be approximated by a translog functional form. In line with this assumption, we can use the exact index number approach to the measurement of productivity growth (see Caves *et al.* (1982) and Diewert (1976)). Under this approach, a measure of TFP growth is equal to the difference between the relative change in real outputs and the relative changes in inputs without estimating any parameters.

Under the assumption that f_i is translog, we can apply Diewert's (1976, pp.118) quadratic approximation identity[19] to expression (**??**), relating differences in production

[19] Diewert (1976, pp.118) shows that if and only if a function $f(z) \equiv \alpha_0 + \sum_m a_k z_k + \frac{1}{2} \sum_k \sum_l b_{kl} z_k z_l$ is differentiable and quadratic in the vector of variables $z \equiv (z_1, z_2, \ldots z_M)'$ then the following identity is true, $f(z^1) - f(z^0) = \frac{1}{2} \left[\bigtriangledown_z f(z^1) + \bigtriangledown_z f(z^0) \right] (z^1 - z^0)$, where $\bigtriangledown_z f(z)$ denotes the gradient of $f(z)$ w.r.t. z. This is called the quadratic approximation lemma.

between period t and some base period 0 as

$$\begin{aligned}\Delta \ln y_{it} &\equiv \ln y_{it} - \ln y_{i0} \\ &\equiv \ln f_i(\mathbf{x_{it}}) - \ln f_i(\mathbf{x_{i0}}) + \Delta \ln a_{it} \\ &= \frac{1}{2}\sum_{k=1}^{J_i}\left[\frac{\partial \ln f_i(\mathbf{x_{it}})}{\partial \ln x_{ikt}} + \frac{\partial \ln f_i(\mathbf{x_{i0}})}{\partial \ln x_{ikt}}\right][\ln x_{ikt} - \ln x_{ik0}] + \Delta \ln a_{it}.\end{aligned}$$

The logarithmic differentiation of $\ln f_i(\mathbf{x_{it}})$ with respect to $\ln x_{ikt}$ can be expressed as[20]

$$\frac{\partial \ln f_i(\mathbf{x_{it}})}{\partial \ln x_{ikt}} = \mu_{it} s_{ikt} = \overbrace{\theta_{it}\frac{y_{it}p_{it}}{\mathbf{w'_{it}x_{it}}}}^{\mu_{it}}\overbrace{\frac{w_{ikt}x_{ikt}}{p_{it}y_{it}}}^{s_{ikt}} \tag{B2}$$

where w_{ikt} denotes the firm's k input price, p_{it} represents the output price, θ_{it} is equal to the elasticity of scale (see below), s_{ikt} denotes the share of input value k in the total production value of firm i, and μ_{it} are markups. Substituting this expression in equation (**??**) yields:

$$\Delta \ln y_{it} \equiv \ln y_{it} - \ln y_{i0} = \frac{1}{2}\sum_{k}[\mu_{it}s_{ikt} + \mu_{i0}s_{ik0}][\ln x_{ikt} - \ln x_{ik0}] + \Delta \ln a_{it}, \tag{B3}$$

where the mark-up ratio, μ_{it}, may vary over time. We use value-added as a measure of output and relate k to the inputs capital and labor. TFP growth ($\Delta \ln a_{it}$) is measured by the difference between the relative change in real outputs and the relative changes in inputs. In equation (**??**), all firm-level output and input variables are observable using production data while the output elasticities θ_{it}, needs to be estimated.

The translog production function is given by

$$\ln y_{it} = \beta_k \ln k_{it} + \beta_l \ln l_{it} + \beta_{kk}(\ln k_{it})^2 + \beta_{ll}(\ln l_{it})^2 + \beta_{kl}\ln l_{it}\ln k_{it} + \varepsilon_{it}, \tag{B4}$$

[20]See, for example, Diewert and Fox, 2009 and Amoroso et al., 2010. Diewert and Fox (2009) start from a translog cost function assuming multiple outputs where u_{it} can be written as $\mu_{it} = \theta_{it}^{-1}\mathbf{w'_{it}x_{it}}/y_{it}p_{it}(y_t)$.

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: $\theta_{ikt} + \theta_{ilt} = \theta_{it}$ where

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

We estimate equation (**??**) for each industry. Estimation of (**??**) raises questions about simultaneity bias. Since firms choose input and output simultaneously, unobserved firm-level characteristics may cause the error term for (11) to be correlated with the input factors of the production function. As a result, this violates the orthogonality of the error term, so standard Ordinary Least Squares (OLS) techniques will be biased and inconsistent. We rely on semi-parametric methods developed by Levinsohn and Petrin (2003) to control for simultaneity. This method solves the simultaneity problem by using intermediate inputs to proxy unobserved productivity shocks (assumed to follow a first-order Markov process) that are observed by the firm.