Discussion Paper No. 08-046

The Knowledge Production of 'R' and 'D'

Dirk Czarnitzki, Kornelius Kraft,
and Susanne Thorwarth

ZEW

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

Discussion Paper No. 08-046

The Knowledge Production of 'R' and 'D'

Dirk Czarnitzki, Kornelius Kraft,
and Susanne Thorwarth

Non technical summary

Research on the relationship between input and output of a “knowledge production function” is an important contribution towards the understanding on how firms produce innovations. Frequently, figures on R&D expenditures are used as input to the innovation process and patents are the (intermediate) output. Several studies have repeatedly examined the relationship between firms’ R&D expenditures and patent applications. Not surprisingly, they typically found a positive relationship between patenting and R&D activity.

These frameworks, however, neglect the possible existence of a reverse causality link. Successful research leads to patents and to the development of the ideas towards products and processes. Hence a significant part of R&D expenditures may arise after the patent application has taken place and perhaps also after granting. The empirical study on the relation between input and output of a knowledge production function would in this case be imprecise. Given that the development part of R&D is mostly the larger one, the impreciseness could be substantial, and estimated patent-R&D elasticities would be biased towards zero if development is irrelevant for the patent production.

This study an conducts analysis with data from the Flemish R&D survey which are linked with patent data from the European Patent Office in order to investigate the separate effects of research expenditure and development expenditure on the number of patents. It is shown that research has a significant positive impact on patents, but development expenditures have no effect.

These results suggest the conclusion that previous studies dealing with the knowledge production functions suffer from measurement error. Rather than R&D investment, research investment seems to be the relevant determinant for patenting, and consequently the returns to such activities have been underestimated in the past.

Das Wichtigste in Kürze (Summary in German)

Untersuchungen des Zusammenhangs zwischen Input- und Outputfaktoren einer Wissensproduktionsfunktion (knowledge production function) tragen zum Verständnis des Innovationsverhaltens von Unternehmen bei. Häufig werden zu diesem Zweck FuE-Aufwendungen als Input und die Anzahl der Patente als Outputvariable (bzw. leicht messbares Zwischenprodukt) verwendet. In der Vergangenheit haben sich bereits mehrere Forschungsarbeiten mit dieser Thematik beschäftigt und typischerweise einen positiven Zusammenhang zwischen Patentanmeldungen und FuE-Aufwendungen gefunden.

Dieser Ansatz vernachlässigt jedoch ein mögliches Auftreten von Rückkopplungseffekten. Patente werden aus erfolgreicher Forschungsaktivität generiert, und es ist plausibel, dass sowohl Forschungsaufwendungen und Patentanmeldungen zu Entwicklungsaktivitäten führen, die schließlich in neuen Produkten oder Produktionsverfahren münden. Dies legt die Vermutung nahe, dass ein nicht unerheblicher Teil der FuE-Aktivitäten erst nach der Patentanmeldung stattfindet. Daher können in diesem Fall empirische Untersuchungen, die die gesamten FuE-Aufwendungen im Rahmen einer Patentproduktionsfunktion verwenden, das Input-Output Verhältnis nur verzerrt abbilden. Da die Entwicklungsausgaben von Unternehmen typischerweise größer sind als die für Forschung, wären die geschätzten Elastizitäten gegen Null verzerrt, falls der Entwicklungsprozess nicht für die Generierung von Patenten relevant ist.

Die vorliegende Studie untersucht mittels flämischer FuE-Daten sowie Patentdaten des Europäischen Patentamtes die unterschiedlichen Effekte von Forschungs- versus Entwicklungsausgaben auf die Anzahl der jährlichen Patentanmeldungen von Unternehmen. Es wird gezeigt, dass Forschungsaufwendungen einen signifikant positiven Einfluss auf die Anmeldung von Patenten hat, während die Entwicklungsausgaben jedoch keine Rolle spielen. Dieses Ergebnis lässt die Schlussfolgerung zu, dass bisherige Untersuchungen Messfehler aufweisen.

The Knowledge Production of 'R' and 'D'

Dirk Czarnitzki [a,c,d], Kornelius Kraft [b,d] and Susanne Thorwarth [a,c,d]

[a] *K.U.Leuven, Dept. of Managerial Economics, Strategy and Innovation*
[b] *Technical University of Dortmund, Dept. of Economics*
[c] *Steunpunt O&O Indicatoren at K.U.Leuven*
[d] *Centre for European Economic Research (ZEW), Mannheim*

June 2008

Abstract

Many studies investigate the relationship between R&D expenditures as an input and patents as an intermediate product or output of a knowledge production function. We suggest that the productivity of research in patent production functions has been underestimated in the literature, as scholars typically use information about R&D, i.e. the sum of research expenditure and development expenditure, due to data availability. However, in most industries only (applied) research will lead to patentable knowledge, and development happens after the initial research phase that may have led to a patent. Instead of using data on R&D, we separate the knowledge creating process into `R' and `D'. This data stems from R&D surveys of Belgian firms. It turns out that only the `R' part of R&D expenditure has a significant effect on patents and that development expenditure are insignificant. Thus previous literature relying on R&D expenditure suffers from a measurement error, such that the coefficient of R&D is biased towards zero, as R&D includes a large fraction of irrelevant expenditure, i.e. development expenditure, with respect to patenting.

Keywords: Patents, Research, Development, Knowledge Production Function

JEL-Classification: O31, O32

Contact:

Dirk Czarnitzki, K.U.Leuven, Dept. of Managerial Economics, Strategy and Innovation, Naamsestraat 69, 3000 Leuven, Belgium;
E-Mail: dirk.czarnitzki@econ.kuleuven.be; Phone: +32 16 326 906; Fax: +32 16 326 732

Kornelius Kraft, Technical University of Dortmund, Dept. of Economics, Vogelpothsweg 87, 44227 Dortmund, Germany;
E-Mail: kornelius.kraft@uni-dortmund.de; Phone +49 231 755 3152; Fax: +49 231 755 3155

Susanne Thorwarth, K.U.Leuven, Steunpunt O&O Indicatoren, Dekenstraat 2, 3000 Leuven, Belgium;
E-Mail: susanne.thorwarth@econ.kuleuven.be; Phone: +32 16 325 735; Fax: +32 16 325 799

1 Introduction

Research on the relationship between input and output of a “knowledge production function” is an important contribution towards the understanding of how firms produce innovations. Usually figures on R&D expenditures are used as input to the production process and patents are the output. The different studies also quantify the relationship and find decreasing marginal returns in most cases. The estimates for the elasticities vary and depend on firm size (Griliches 1990), with larger firms showing a lower R&D productivity.[1]

Griliches (1990) points to the possibility of a reverse causality link. Successful research leads to patents and to the development of the ideas towards products and processes. Hence a significant part of R&D expenditures may arise after the patent application has taken place and perhaps even after the patent grant. The empirical study on the relation between input and output of a knowledge production function would in this case be imprecise. Given that the development part of R&D is mostly the larger one, the impreciseness could be substantial, and estimated patent-R&D elasticities would be biased towards zero if development is irrelevant for the patent production.

We report the results of an empirical study on the separate effects of research expenditure and development expenditure on the number of patents. Using panel count data models on Belgian firm-level data, we find that research has a significant impact on patents, but that development expenditures have no effect. Thus, previous results in the literature on the knowledge production functions suffer from, possibly large, measurement error. Rather than R&D investment, research investment seems to be the relevant determinant for patenting, and consequently the returns to such activities have been underestimated in the past.

2 The relationship between Patents and R&D

The understanding of the economic process that leads to product and process innovation is of high interest, as an economy’s wealth and growth may crucially depend on technological progress. At least in highly industrialized countries, technological innovation is seen as key input for long term employment and growth.

[1] This could be the result of selectivity and systematic differences in recording and reporting between small and large firms.

However, the identification of innovative output is not trivial, as there is no single, undisputed variable measuring innovation success comprehensively. Patents are the most commonly used indicator for invention output in the economy although this has been criticised by several authors (e.g. Griliches et al. 1986, Griliches 1990). It is typically emphasized that patents do not fully represent all of R&D output since only a fraction of the knowledge creating process leads to patentable inventions. Even if the knowledge is patentable, inventors may refrain from doing so due to additional cost of patenting or because they favour other means of intellectual property protection. Among others, Cohen et al. (2000) point out that firms frequently follow other strategies, e.g. secrecy or lead time ahead of competitors. A reason to patent, though, may be the blocking potential with respect to rivals, as the patent grants the inventor an exclusive right to exploit the technology for a fixed time period.

Despite several shortcomings of patent information, patent data is the most frequently used innovation output indicator, e.g. for assessing the technological potential of countries or sectors. Patents are an easily available source of information, as patent applications are systematically recorded by (national) patent offices and are, thus, available for researchers. Furthermore, it should be noted that most of the crucial inventions of the last century were patented.

Several studies have examined the relationship between firms' R&D expenditures and patent applications. Not surprisingly, they typically found a positive relationship between patenting and R&D activity. The pioneering work in the use of patent statistics was conducted by Scherer (1965), Mueller (1966) and Schmookler (1966). Pakes and Griliches (1980) found a statistically significant relationship between the R&D expenditures of a firm and the number of patents received at the cross-sectional level, across firms and industries. This finding was also observed by Hall et al. (1984) who applied panel estimators for count data and they report an elasticity between 0.3 and 0.6 in a panel of U.S. firms (see also Bound et al., 1984, for further evidence). Cincera (1997) measures the impact of technological factors on patenting activity indicating a positive impact of R&D and technological spillovers on a firm's innovation output. Crépon and Duguet (1997a, b) use a panel of French manufacturing firms to investigate the relationship between investment in R&D and patents, and also find a positive and significant effect of R&D. This is also confirmed by Licht and Zoz (1998) who look at the patent–R&D relationship using German survey data. More recent results which also report a positive patent-R&D relationship can be found in Hall and Ziedonis (2001) for a sample of 95 U.S. semiconductor firms and in Blundell et al. (2002).

2.1 Research and Development

Research and development encompasses several kinds of activities. According to the definition of the Frascati Manual (1993) which frames the methodology for collecting and using statistics about R&D in OECD countries, the term R&D covers three activities: basic research, applied research and experimental development. However, most surveys, evaluations and reports conducted in this area do not differentiate between these activities.

Usually basic and applied research is aggregated to one component, namely research, and can be described as any activity which is undertaken to generate new knowledge. Experimental development is systematic work, drawing on existing knowledge gained from research that is used as an input for developing new products or production methods.

The Frascati Manual (1993, p. 78) outlines the importance of (applied) research towards patenting: *"[...] Applied research gives operational form to ideas. The knowledge or information derived from it is often patented but may be kept secret."* This core statement was already made in a seminal work conducted by Hall et al. (1986) who attempt to characterize the lag structure in the productivity of R&D and who derive the conclusion that *"it seems reasonable to suppose that successful research leads both to a patent application and to a commitment of funds for development."*

This leads to the hypothesis that mainly the research component rather than the sum of research and development influences the propensity to patent. An earlier work of Link (1982) analyzes the determinants of inter-firm differences on the composition of R&D spending, namely basic and applied research as well as development. However, he does not link R&D to patents, but is interested in the determinants of basic and applied research as well as development, and regresses these R&D components on profits, diversification, ownership structure and subsidies. Nevertheless, all previous microeconomic, empirical studies on the relationship between patenting and R&D have treated R&D investment as a single, homogeneous activity, though. Thus, it is possible that previous studies underestimate the effect of research activities as the development expenditure is typically larger than research expenditure in the business sector. Consequently, it would be reasonable to split internal R&D expenditure into its two components, research and development, to investigate this hypothesis. As already mentioned above most databases do not allow to explore this issue since usually only numbers for aggregated R&D are available. However, in the OECD R&D surveys, firms are asked for the breakdown of internal R&D expenditure into research versus development.

Each of the surveyed firms is provided with the OECD definitions of research and development and an explanation of these terms.

2.2 Empirical Framework

In order to explore the determinants of a firm's patenting process and their change over time, we use a patent production function which was first introduced by Pakes and Griliches (1980). This type of knowledge creation function describes the relationship between the number of successful patent applications made by a firm in a given year and its R&D spending as well as other firm level control variables. The number of patent applications is restricted to non-negative integer values and is also characterized as a variable with many zero values since many firms do not apply for patents regularly. This led scholars to use count data models to investigate patent-R&D relationships. Typically Poisson and Negative Binomial models are used.

Let PAT_{it} be the dependent count variable which describes the number of patent applications by firm i at time t, and is typically assumed to be Poisson distributed with mean $\lambda_{it} > 0$. The link between patents, $R\&D_{it}$ and a set of controls X_{it} is usually assumed to be of an exponential function form:

$$\lambda_{it} = E\left[PAT_{it} \mid R\&D_{it,}X_{it}, c_i\right] = exp\left(\alpha R\&D_{it} + X_{it}^{'}\beta + c_i\right) \quad (1)$$

where α and β are the parameters to be estimated, and c is a firm-specific effect reflecting time-constant unobserved components, such as management quality or a firm's attitude towards patenting, in case panel data are available.

As we hypothesized earlier, α may be biased downwards if development activity is irrelevant for patenting and research activity is the relevant type of investment for creating patentable knowledge. Therefore, we estimate the slightly modified model as

$$\lambda_{it} = E\left[PAT_{it} \mid R_{it}, D_{it}, X_{it}, c_i\right] = exp\left(\alpha_0 R_{it} + \alpha_1 D_{it} + X_{it}^{'}\beta + c_i\right), \quad (2)$$

where R refers to research expenditure, and D to development expenditure. We expect that $\alpha_1 = 0$ *and* $\alpha_0 > 0$, and in addition that $\alpha_0 > \alpha$. The latter can be tested using a Wald test to see whether $\alpha_0 = \alpha_1$.

As neither of the two specifications above, allows to model potential complementarities between R and D, we also estimate a third equation as a robustness check. There, an interaction term of R and D is included.

$$\lambda_{it} = E\left[PAT_{it} \mid R_{it}, D_{it}, X_{it}, c_i\right] = exp\left(\alpha_0 R_{it} + \alpha_1 D_{it} + \alpha_2 R_{it} \times D_{it} + X_{it}'\beta + c_i\right). \tag{3}$$

Note that the test for complementarity is not simply conducted by interpreting whether α_2 is significantly different from zero, as the regression model is non-linear. Instead, one has to calculate the cross-derivative d*PAT*/(d*R*d*D*) to correctly estimate the magnitude and the standard errors of the interaction effect (see e.g. Ai and Norton, 2003).

3 Data, econometric methods, and results

3.1 Database

In order to investigate the patent production function described in the previous section we link three different databases: the Flemish R&D Survey, the BELFIRST database and patent data from the European Patent Office (EPO). The Flemish R&D Survey is conducted every second year, and our analysis is based on six waves of the Flemish R&D survey data which covers the period between 1993 and 2003.

The R&D survey data are supplemented with information from the BELFIRST database which contains annual account data of Belgian firms. Furthermore, the Flemish companies are linked to patent data from the European Patent Office (EPO) which covers all patent applications filed at the EPO since 1978.

We only include firms in the analysis that applied at least for one patent during the observation period (1993 – 2003). After elimination of data with missing values in variables of interest our final sample consists of 103 firms in an unbalanced panel that has 596 firm-year observations. The dependent variable in the empirical analysis, *PAT,* is the number of patents filed by a firm at the European Patent Office. The regressors of main interest are, of course, R&D spending (*R&D*) or, more precisely, its two components research (*R*) and development (*D*) expenditure.

As common in firm level studies we also control for firm size. Larger firms may be able to realize economies of scope in their innovation process, that is, they may be able to apply fundamental research results to more inventions than smaller firms. Thus, they may be more productive in yielding inventions (Scherer, 1983). Since small firms usually do not have a special unit dealing with patents or property rights and often hesitate to apply for a patent because of the large patent litigation costs, the marginal costs for patent applications are expected to be higher for small firms than for large firms. Previous studies (e.g. Arundel and

Kabla, 1998; Licht and Zoz, 1998) provide some evidence which demonstrated that small firms have a lower propensity to patent their innovations than large firms. Following, among others, Hall and Ziedonis (2001), our basic specification includes firm size measured as the logarithm of employment, ln(*EMP)*. Our variables on R&D also enter the model as logarithms, but are divided by employment to avoid multicollinearity [ln(*R&D/EMP*), ln(*R/EMP*), ln(*D/EMP*)].

Additionally to firm size we also include the firms' age measured as the logarithm of the difference between the current and the founding year of the firm, ln(*AGE*). Very young firms may show a higher likelihood to patent because typically spin-offs from larger firms or research institutions involve innovative ideas which are then protected by intellectual property rights. Contrary to this, older firms may be more likely to patent because they have more experience in managing the application procedure which may raise their efficiency in patenting activities (Hall and Ziedonis 2001). We also add capital intensity, ln(*KAPINT),* measured as the logarithm of the ratio tangible assets to employment as an explanatory variable.

The variable *GROUP* has unit value, if a firm is a member of a group of companies. Research output may have more than one patent application possibility. A group of companies is able to make use of an invention in more than just one way. An alternative way of considering possible effects of *GROUP* is the internalisation of positive externalities. The value of a composition of several firms to a group is among other factors determined by the use of spillovers from innovation activities. The existence of spillovers is well documented and a group can make use of such effects by internalising the externalities. In contrast, group membership may also result in less patent applications if the firm in question is a subsidiary. It is possible that research conducted at a subsidiary may lead to patents that are taken out by the parent company, as it manages the intellectual property rights for the group.

The group dummy is supplemented by a collaboration dummy (*COLL*). Firms that engage in collaboration may also be able to internalize external effects, and thus file more patents. In addition to pure spillover effects, collaboration may also lead to more patents, as the firms simply engage in contractual agreements on the usage of intellectual property. In order to ensure that all collaboration partners have access to the knowledge produced within the collaboration network, patenting may become more likely, as, for instance, co-assigneeship will unambiguously secure that collaboration partners have access to the patented knowledge.

Finally, 11 sector dummies on basis on the European standard industry classification NACE should capture different technological opportunities, and a full set of time dummies captures shocks over time common to all firms.

All variables in monetary units are measured in thousand Euros in prices of the year 2000. We used the GDP deflator for price adjustment. Table 1 shows the descriptive statistics for all variables. Average firm size is about 696 employees, and average age about 27 years. Note, however, that the size distribution is highly skewed, since the median firm size is about 168 employees in our sample. Regarding R&D expenditures firms spent around 14 million Euros per year where approximately 10 million were spent for development, and about 4 million for research. In total, all firms of our sample applied for a total amount of 1222 patents, however, the distribution of this variable is rather skewed since about half of the companies only applied for one patent and only about 10 percent filed for more than 10 patents during the whole observation period. On average, a sample company applied for 4 patents per year. About 69% use a cooperation network for conducting research and development and 68% belong to a group.

Table 1 Descriptive Statistics (596 observations)

Variable	Label	Mean	Std. Dev.	Min	Max
Patents	*PAT*	3.958	17.316	0	193
R&D expenditure	*R&D*	13743.39	45243.07	5	471353.2
Research expenditure	*R*	3759.919	14848.65	0	217581.4
Development expenditure	*D*	9983.45	35799.6	0	424217.9
Employment	*EMP*	695.497	1317.274	1	7470
Age	*AGE*	26.728	25.750	0	121
Capital Intensity	*KAPINT*	54.232	173.461	0.408	2113.098
Collaboration dummy	*COLL*	0.688	0.464	0	1
Group dummy	*GROUP*	0.676	0.468	0	1

Note: Time and industry dummies not presented.

Earlier work which also focused on the patent–R&D relationship experimented with lag structures of R&D. Hall et al. (1986) provide evidence that a contemporaneous relationship between patents and R&D can be assumed. Longer lag structures of R&D did not improve the explanatory power of their regression models. For this reason and since we only have short time-series data for the majority of our companies in the sample, we follow this line of literature and also specify a contemporaneous relationship. We experimented with one-year lag structures, but it did not improve the results.

3.2 Estimation method and results

Commonly used count data models are the Poisson and the Negative Binomial model (NegBin). Traditionally, scholars have applied the random or fixed effects estimators introduced by Hausman et al. (1984). A major drawback of these, however, is that the parameter estimates are only consistent under the strict exogeneity assumption, that is, it rules out any feedback from patenting in period t to future values of research and development expenditure. This will be clearly violated in our analysis, as one should expect that patents lead to future development activities. Therefore, we apply the linear feedback model suggested by Blundell et al. (1998) which relaxes the strict exogeneity assumption. The model approximates the fixed effects by including the log of patents from a pre-sample period as regressor in a standard pooled cross-sectional count model set-up, ln(*PRE_PAT*). In case, the firm had no patents in the pre-sample period a dummy is used to capture the "quasi-missing" value in log of patenting in the pre-sample period, *NO_PRE_PAT*.

For estimating (pooled) cross-sectional count models, scholars have frequently used negative binomial regression models (NegBin), as a basic assumption of the Poisson model is the equality of the conditional mean and the conditional variance which is typically violated in applications, and overdispersion is found. Although the NegBin relaxes this assumption it would only be consistent (and efficient) if the functional form and distributional assumption of the variance term is correct. It has been shown, though, that the Poisson model is consistent even in the case of overdispersion (always assuming that the conditional mean function is correctly specified). The only drawback is that the standard errors will be biased if the equidispersion assumption is violated. However, this can be corrected by using a fully robust covariance matrix estimator (see Wooldridge, 2002, for example).

Table 2 shows the quasi fixed effects panel regressions using the estimator proposed by Blundell et al. (1998). We also show pooled cross-sectional estimates assuming that no fixed effects are present in the appendix (see Table 3).

Models 1 and 2 estimate the knowledge production function with aggregated R&D. First, we exclude all time invariant regressors as typical in traditional fixed effects models (model 1). Since we depart from the traditional estimation method with the quasi fixed effects model, it is possible to include time invariant variables in order to reduce the error variance (model 2).

Table 2 Panel Regressions on the Number of Patent Applications (596 observations)

Variable	(1)	(2)	(3)	(4)	(5)	(6)
ln(*R&D/EMP)*	0.131 *	0.125 **	-	-	-	-
	(0.077)	(0.062)				
ln(*R/EMP*)	-	-	0.235 ***	0.188 ***	0.219 **	0.217 **
			(0.078)	(0.062)	(0.110)	(0.104)
ln(*D/EMP*)	-	-	-0.118	-0.070	-0.131	-0.044
			(0.081)	(0.082)	(0.093)	(0.098)
ln(*R/EMP*)	-	-	-	-	-0.015	-0.019
*ln(*D/EMP*)					(0.327)	(0.031)
RDUMMY	-	-	0.038	0.054	0.045	0.030
			(0.252)	(0.231)	(0.256)	(0.237)
DDUMMY	-	-	-0.922 ***	-0.907 ***	-0.878 **	-0.983 **
			(0.331)	(0.313)	(0.403)	(0.390)
ln(*EMP*)	0.258 ***	0.289 ***	0.223 ***	0.275 ***	0.223 ***	0.277 ***
	(0.078)	(0.073)	(0.064)	(0.065)	(0.064)	(0.065)
ln(*AGE*)	-0.576 **	-0.526 ***	-0.435 ***	-0.425 ***	-0.436 ***	-0.423 ***
	(0.098)	(0.101)	(0.094)	(0.100)	(0.095)	(0.099)
ln(*KAPINT*)	0.165 *	0.099	0.134	0.107	0.125	0.122
	(0.098)	(0.084)	(0.084)	(0.080)	(0.086)	(0.090)
COLL	-	-0.076	-	-0.041	-	-0.049
		(0.216)		(0.207)		(0.215)
GROUP	-	-0.334 **	-	-0.411 **	-	-0.415 **
		(0.165)		(0.167)		(0.167)
ln*(PRE_PAT)*	0.955 ***	0.889 ***	0.908 ***	0.866 ***	0.910 ***	0.861 ***
	(0.078)	(0.065)	(0.066)	(0.063)	(0.067)	(0.061)
No_PRE_PAT	-1.692 ***	-1.583 ***	-1.645 ***	-1.531 ***	-1.649 ***	-1.522 ***
	(0.347)	(0.276)	(0.230)	(0.232)	(0.229)	(0.227)
INTERCEPT	0.415	0.451	0.579	0.252	0.624	0.166
	(0.697)	(0.689)	(0.454)	(0.643)	(0.470)	0.679
Joint significance of year dummies	$\chi^2(10)$ =69.53***	$\chi^2(10)$ =75.14***	$\chi^2(10)$ =96.52***	$\chi^2(10)$ =68.65***	$\chi^2(10)$ =91.12***	$\chi^2(10)$ =76.95***
Joint significance of industry dummies	-	$\chi^2(10)$ =32.56***	-	$\chi^2(10)$ =18.94**	-	$\chi^2(10)$ =20.00**
Log Pseudo likelihood	-1007.604	-954.265	-962.471	-932.156	-962.377	-931.861
Wald test on $\alpha_0 = \alpha_1$			6.40 **	3.83 **	6.08 **	3.83 **

Notes: Standard errors in parentheses are clustered to capture within firm correlations.
*** (**, *) indicate a significance level of 1% (5%, 10%).

The results of model (1) and (2) report the regression results using the sum of *R* and *D* expenditure. The estimated patent-R&D elasticity amounts to 14% [=exp(0.131)-1] in the model that excludes the time-invariant covariates, and to 13% [=exp(0.125)-1] in the model with time invariant covariates. Models (3) and (4) relax the assumption that *R* and *D* have the same coefficient. Once we separate R&D into R and D, there are a few firms in the sample that have either zero R or zero D spending. Since we have to take the log of the variables, and the log of zero if not defined, we set the variables ln(*R*) and ln(*D*) to zero is these cases. As commonly done in the literature (e.g. Hall and Ziedonis 2001), we capture the arising bias from that by two dummy variables (*RDUMMY* and *DDUMMY*) that capture the zero values in R and D. As a result, we do not have to discard these observations. Note that the estimated coefficients of these dummies have no interpretation in itself. Instead of zero, we could also have imputed -9999, for instance, and the estimates of the slopes of *R* and *D* would be numerically identical. Then the dummies would just have different coefficients due to the arbitrary choice of the imputation value.[2]

In these models, we find interesting differences: the estimated patent-research elasticity goes up to 26% in model (3) and 21% in model (4). The coefficient of development expenditure is not significant, though. Thus, we can conclude that the patent production function underestimates research productivity if R&D is used as the relevant measure. Note that we report the Wald test on the difference of the coefficients of *R* and *D* that indicate a significant difference. Once we separate *R* and *D*, the estimated knowledge productivity of research basically doubles, and development is irrelevant for patenting.

The models (5) and (6) include the interaction term ln(*R/EMP*)*ln(*D/EMP*) to model complementarity between *R* and *D* with respect to patenting. However, the estimated coefficient is zero. As this does not necessarily imply no complementarity, since we estimate a non-linear model, we also computed the cross-derivative d*PAT*/(d*R*d*D*) and calculated its standard error using the delta method. However, the results do not improve. While we certainly believe that complementarities exist between research and development activities in general, we do not find any with respect patenting. As said above, it indeed seems that research expenditure is the relevant component for patenting, but not development. Instead, research and patenting will trigger future development activities, which finally leads to new product introductions and new processes.

The results on the control variables are interesting as well. As expected, larger firms are more likely to patent, which confirms arguments on economics of scope and related arguments mentioned earlier. Interestingly, younger firms, all else constant, file more patents than older firms. This supports the argument that (high tech) spin off companies involve more innovative

[2] As a robustness check, the regressions were also conducted without observations that had either zero R or zero D spending. All results are robust, but somewhat less pronounced. For instance, in the fully specified model, the coefficients of R and D did only differ at the 10% significance level, possibly due to smaller sample size and thus higher correlation between R, D and their interaction term.

ideas for commercialization. The capital intensity and the collaboration dummy are insignificant in all models. The firms associated with a group are less likely to patent. This indicates that subsidiaries may not take out the patent themselves, but that the parent company does so. Note that many of the larger firms in Flanders are foreign owned so that we do not have the parent company itself in our data.

Finally, the regressions also indicate that there are significant firm-specific effects as measured by the pre-sample patent measures. Note, however, that the pooled cross-sectional models yield very similar results (see Table 3 in the Appendix). We also find that patenting is heterogeneous across industries, as the industry dummies are jointly significant. The year dummies are also jointly significant, which indicates the presence of macroeconomic shocks that affect patenting behavior of all companies.

4 Conclusion

We report the results of a study on the separate effects of research and development expenditures on patent productivity. We argue that research productivity has often been underestimated as scholars typically have only data on R&D expenditure available, but not on research expenditure and development expenditure separately.

Employing our Flemish firm-level panel database, we find that the patent-R&D elasticity is about 13% when using R&D in the regressions. Once we separate research and development spending, it turns out that development expenditure has no impact on patenting, but the estimated patent-research elasticity amounts to 21% and 26% depending on the model specification. Thus, the estimated research productivity is about twice as high compared to the R&D productivity.

Our estimations should only be seen as some first evidence on the different contribution of research and development to patenting. As our sample size is quite small, we cannot estimate separate equations for different industries. While we find that only research contributes to patenting but development is irrelevant, there may well be industries where development contributes to some extent to the patent process. Consequently, it would be desirable to replicate our regressions for a larger sample of firms, where scholars could group the sample by industry. While our available database is small as it only concerns Flanders, the survey data we use are actually available for many OECD countries, for some even since the late 1970s. Surprisingly there is little research on the OECD R&D surveys. While other surveys, such as the Community Innovation Survey are widely used for research, it seems that the

R&D surveys are difficult to access for scholars in other countries. If this barrier could be overcome, there would be the opportunity to replicate or extend our study with large firm-level panels from larger countries.

References

Ai, C. and E. C. Norton (2003), Interaction Terms in Logit and Probit Models, *Economic Letters 80*, 123-129.

Arundel, A. and I. Kabla (1998), What Percentage of Innovations are Patented? Empirical Estimates for European Firms, *Research Policy 27*, 127–141.

Blundell, R., R. Griffith and J. van Reenen (1995), Dynamic Count Data Models of Innovation, *Economic Journal 105*, 333-345.

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.

Blundell, R., R. Griffith and J. van Reenen (2002), Individual Effects and Dynamics in Count Data Models, *Journal of Econometrics 108*, 113-131.

Bound, J, C. Cummins, Z. Griliches, B. H. Hall and A. Jaffe (1984), Who Does R and D and Who Patents in Z. Griliches (ed.): R&D, Patents, and Productivity, Chicago, University of Chicago Press.

Cincera, M. (1997), Patents, R&D and Technological Spillovers at the Firm Level: Some Evidence from Econometric Count Models for Panel Data, *Journal of Applied Econometrics 12*, 265-280.

Cohen, W. M., R. R. Nelson and J. P. Walsh (2000), Protecting their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not), *NBER Working Paper No. 7552*, Cambridge, MA: National Bureau of Economic Research.

Crépon, B. and E. Duguet (1997a), Estimating the Innovation Function from Patent Numbers: GMM on Count Data Panel, *Journal of Applied Econometrics 12*, 243-263.

Crépon, B. and E. Duguet (1997b), 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.

Griliches, Z., A. Pakes and B. H. Hall (1986), The Value of Patents as Indicators of Inventive Activity, *NBER Working Paper 2083*, Cambridge, MA: National Bureau of Economic Research.

Griliches, Z. (1990), Patent Statistics as Economic Indicators: A Survey, *Journal of Economic Literature 28(4)*, 1661-1707.

Hall, B. H., J. A. Hausman and Z. Griliches (1986), Patents and R and D: Is There a Lag, *International Economic Review 27*, 265-283.

Hall, B. H. and R. H. Ziedonis (2001), The Patent Paradox Revisited: an Empirical Study of Patenting in the U.S. Semiconductor Industry, 1979-1995, *RAND Journal of Economics, 32(1)*, 101-128.

Hausman, J. A., B. H. Hall and Z. Griliches (1984), Econometric Models for Count Data with an Application to the Patents-R&D Relationship, *Econometrica 52*, 909-938.

Link, A. N. (1982), An Analysis of the Composition of R&D Spending, *Southern Economic Journal, 49(2)*, 342-349.

Mueller, D. C. (1966), Patents, Research and Development, and the Measurement of Inventive Activity, *The Journal of Industrial Economics 15(1)*, 26-37.

OECD (1993), Frascati–Manual, Proposed Standard Practice for Surveys of Research and Experimental Development, Organisation for Economic Co–Operation and Development, Paris.

Pakes, A. and Z. Griliches (1980), Patents and R&D at the Firm Level: A First Report, *Economics Letters 5*, 377-381.

Scherer, F.M. (1965), Firm Size, Market Structure, Opportunity and the Output of Patented Inventions, *American Economic Review 55*, 1097-1125.

Scherer, F.M. (1983), The propensity to patent, *International Journal of Industrial Organization* 1, 107-128.

Schmookler, J. (1966), Invention and Economic Growth, Harvard University Press, Cambridge.

Schumpeter, J. A. (1934), The theory of economic development, Cambridge, MA.: Harvard University Press.

Schumpeter, Josef A. (1942), Capitalism, Socialism and Democracy, New York: Harper and Row, (Reprinted 1960).

Winkelmann, R. and K. F. Zimmermann (1995), Recent Developments in Count Data Modelling: Theory and Application, *Journal of Economic Surveys 9*, 1-24.

Appendix

Table 3 Pooled cross-sectional Poisson models on patent applications (596 observations)

Variable	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
ln(*R&D/EMP)*	0.458 ***	0.478 ***	-	-	-	-
	(0.120)	(0.109)				
ln(*R/EMP*)	-	-	0.752 ***	0.444 ***	0.834 ***	0.735 ***
			(0.194)	(0.127)	(0.226)	(0.149)
ln(*D/EMP*)	-	-	-0.251	-0.015	-0.183	0.260
			(0.170)	(0.127)	(0.215)	(0.171)
ln(*R/EMP*) *ln(*D/EMP*)	-	-	-	-	-0.115	-0.062
					(0.074)	(0.046)
RDUMMY	-	-	0.100	0.095	0.085	-0.160
			(0.419)	(0.356)	(0.427)	(0.410)
DDUMMY	-	-	-1.300 **	-1.301 ***	-1.488 **	-1.769 ***
			(0.550)	(0.499)	(0.661)	(0.516)
ln(*EMP*)	1.191 ***	1.111 ***	1.051 ***	0.996 ***	1.043 ***	0.988 ***
	(0.305)	(0.150)	(0.167)	(0.135)	(0.168)	(0.133)
ln(*AGE*)	-0.339	-0.387 ***	-0.002	-0.138	0.005	-0.988
	(0.210)	(0.147)	(0.165)	(0.137)	(0.166)	(0.133)
ln(*KAPINT*)	0.019	-0.470 ***	-0.183	-0.454 ***	-0.132	-0.291 **
	(0.191)	(0.125)	(0.156)	(0.119)	(0.166)	(0.135)
COLL	-	0.089	-	0.162	-	0.024
		(0.250)		(0.267)		(0.274)
GROUP	-	-0.595 **	-	-0.569 *	-	-0.685 **
		(0.283)		(0.304)		(0.311)
INTERCEPT	-6.535 ***	-4.680 ***	-5.435 ***	-4.499 ***	-5.639 ***	-5.164 ***
	(1.485)	(0.837)	(0.985)	(0.875)	(1.035)	(0.919)
Joint significance of year dummies	$\chi^2(10)$ =54.32***	$\chi^2(10)$ =119.53***	$\chi^2(10)$ =34.69***	$\chi^2(10)$ =74.74***	$\chi^2(10)$ =38.46***	$\chi^2(10)$ =103.20***
Joint significance of industry dummies	-	$\chi^2(10)$ =64.73***	-	$\chi^2(10)$ =54.76***	-	$\chi^2(10)$ =57.93***
Log Pseudo likelihood	-2176.233	-1560.191	-1804.207	-1497.452	-1800.399	-1458.610
Wald test on $\alpha_0 = \alpha_1$	-	-	8.51 **	3.95 **	8.85 ***	4.76 **

Note: Standard errors in parentheses are clustered to capture within firm correlations.
*** (**, *) indicate a significance level of 1% (5%, 10%).