The market value of patents and R&D: Evidence from European firms
Bronwyn H. Hall, Grid Thoma, and Salvatore Torrisi

JEL No. D24,G32,L86,O31,O34

ABSTRACT

This paper provides novel empirical evidence on the private value of patents and R&D in European firms during the period 1991-2004. We explore the relationship between firm's stock market value, patents, and "quality"-weighted patents issued by the European Patent Office (EPO) and the US Patent and Trademark Office (USPTO). We find that Tobin's q is positively and significantly associated with R&D and patent stocks, but that only those patents taken out in both patent offices or at the USPTO alone seem to be valued. Either forward citations or a composite quality indicator based on forward citations, family size and the number of technical fields covered by the patent are modestly informative for value. Software patents account for a rising share of total patents in the USPTO and EPO. Moreover, some scholars of innovation and intellectual property rights argue that software and business methods patents on average are of poor quality and that these patents are applied for merely to build portfolios rather than for protection of real inventions. We found that such patents are considerably more valuable than ordinary patents, especially if they are taken out in the U.S. However their quality indicators are no more valuable than those of other patents, suggesting that their primary purpose may be to increase the size of the patent portfolio.

Bronwyn H. Hall
Department of Economics
549 Evans Hall
UC Berkeley
Berkeley, CA 94720-3880
and NBER
bhhall@nber.org

Grid Thoma
University of Camerino, Camerino and
CESPRI, Bocconi University
Via Sarfatti 25
20136
Milano
Italy
grid.thoma@unibocconi.it

Salvatore Torrisi
Department of Management
University of Bologna
Via Capo di Lucca 34
40126 Bologna, Italy
torrisi@unibo.it

The market value of patents and R&D: Evidence from European firms

Bronwyn H. Hall, Grid Thoma, and Salvatore Torrisi

1. Introduction

Measuring the private returns to investment in innovation or knowledge assets is important both to firms and to economists who wish to assess and compare firm performance in this area. At the aggregate firm level, the primary methods for obtaining quantitative measures of these returns relate profits, revenue, or the market value of the firm to observable measures of innovation investment such as R&D or patents. This paper contributes to this literature by providing novel empirical evidence on the value of a number of different measures based on the patenting activities of European firms, both in Europe and in the United States.

In addition to the goal of measuring innovative assets in European firms, our investigation is motivated by an interest in several issues related directly to the patents themselves. First, we hope to gain a deeper understanding of the 'patent paradox', that is, the fact that the number of patent applications to the USPTO and the EPO continues to grow despite the weakness of patents as an instrument for protecting innovation, documented in various surveys of innovators in a number of different industries and countries (Levin et al 1987; Cohen, Nelson and Walsh, 2000; Arundel 2001, 2003). Previous studies have demonstrated that the distribution of patent technical and economic value is very skewed with only a few patents yielding a significant value to their owners (Harhoff *et al*. 1999). Some argue that the lower barriers to patenting are responsible for an increasing number of low quality patents, that is, patents that have a low inventive step, overly broad claims, or that should not have been issued under existing legal frameworks. If so, it is desirable to explore whether this is reflected in indicators of individual patent value. In this paper we look at how a firm's stock of patents and different indicators of its 'quality' are priced by the financial markets. We use a number of indicators of technical and economic value: forward citations adjusted for citation truncation, technological scope, measured by the number of technological fields, and family size (the number of different patent systems in which protection for a single invention is sought).

Another motivation of this paper arises from the differences between the US and the European patent systems. Unlike the US system, the European system is very fragmented. Applicants to the EPO systems have to specify the EU countries where the inventions should be protected. If granted, the patent must be defended in national courts because there is at present no European-wide court dealing with patent litigation. The same patented invention then may yield varying private values to its owner depending on the enforcement power offered by the national courts in which the invention is protected. Recently the European Commission (EC) has proposed a new treaty, the European Patent Litigation Agreement (EPLA) that would establish a new European Patent Court. It is unclear whether such a move would represent a significant step towards a "community-wide" patent. However, the proposal testifies to the great concern of the EC about the costs of patenting in Europe and the application of uniform standards in patent examination and enforcement.

In theory the absence of a centralized European patent system, which increases both the application and enforcement costs of EPO patents as compared with US patents, should discourage patent applications to the EPO. However, the EPO examination system appears to be more rigorous than the USPTO (see, for example, Quillen et al. 2002) and this should reduce the expected post-grant litigation costs, especially given the availability of the lower cost opposition system for challenges to newly-issued patents. On the other hand, until the year 2000, patent applications to the USPTO were not published until (and if) they were granted. New applicants then could not know whether their patents were infringing a pending patent. After the year 2000, the US system adopted a variation of the EPO system rule and patent applications are now published after 18 months unless the applicant has sworn not to file in any other jurisdiction. Other differences between the two patent systems pertain to citation of prior art and patentable subject matters. These differences may affect the economic value of patents in the two systems.

This paper looks carefully at the implications of these differences for the economic value of patents by comparing the market value of patents granted by the USPTO and by the EPO. Some European firms protect their inventions in both patent systems and some rely on only one patent system. The choice to protect in one or the other system or in both systems can result from at least two sources: patents on more valuable inventions may be taken out in more jurisdictions (Lanjouw et al. 1998) and firms may differ in their patenting strategies or exposure to international competition. Although we cannot distinguish these two hypotheses precisely in the absence of an appropriate instrument for the choice, we are able to determine whether patents from different jurisdictions yield significantly different consequences for the market value of the firm, or indeed whether measures based on the different patents from the two different systems have different predictive power.

In the last part of the paper, we focus on a specific technological field, software, so that we can distinguish the differences between the two systems from other factors specific to the patent system. Software is of particular interest because it is treated differently in the EPO and the USPTO. A few key decisions taken by the Courts of Appeals for the Federal Circuit (CAFC) in 1994-1995 led the USPTO to release new guidelines for software patentability in 1996 which allowed the patenting of any software embodied in physical media. In 1998 an important decision of the US Federal Circuit removed most of the exceptions to the patentability of software 'as such', i.e., independently of its links with a physical device. Not surprisingly, the number of software patents granted by the USPTO has increased dramatically during the 1990s.

The treatment of software in the EPO is different. According to the European Patent Convention (EPC) (Art 52) computer programs "as such" are excluded from the patentable subject-matter. The EPO recognizes the patentability of computer-implemented inventions (CII), that is "inventions whose implementation involves the use of a computer, computer network or other programmable apparatus, the invention having one or more features which are realized wholly or partly by means of a computer program" (EPO, 2005:3). A further test applied by the EPO relates to the subject matter of any CII, effectively excluding those related to business methods or otherwise "nontechnical" in nature. This distinction has proved difficult to make in practice, but it does lead to rejection of a number of patent applications whose equivalents are granted at the USPTO. The European Commission released a proposed Directive on the Patentability of CIIs in 2002 which effectively codified EPO practice in this area, but the Directive was rejected by the European Parliament in 2005 after considerable amendment of various kinds.

As a preliminary test of the consequences of the different legal treatment of software in the two patent systems we have analyzed EPO patents and found an increasing number of software-related patents during the 1990s.[2] This suggests that, despite the different legal environment, barriers to software patents have fallen somewhat in Europe as well. It is important

[2] For a detailed analysis of software-related patent applications and the search methodology used to identify this category of patents, see Thoma and Torrisi (2005)

to note, however, that the majority of software patents in the EPO are probably 'software-related' patents, that is patents granted to computer programs that are implemented in physical devices, rather than "pure" software patents.[3]

Our examination of the market value of patents draws on a body of studies which have addressed the issue of measuring the private returns or value of innovation investments using data on the firm's valuation in public financial markets. Most of these studies use R&D expenditures and patent counts as measures of technological activity (e.g., Griliches 1981; Hall 1993). More sophisticated indicators of technological assets such as citation-weighted patents have also been experimented with in the literature to account for the great dispersion in the value distribution of patents (Hall, Jaffe, and Trajtenberg 2005). In the absence of more direct measures of the economic value of patents, these studies provide a useful methodological setting to explore the relationship between technological importance and the profitability of patented inventions. These studies have mostly used data for US firms and UK firms.

Research that compared indicators of individual patent value such as citations with survey-based direct measures of profits from the associated invention has found a positive and significant association between them (Harhoff *et al.* 1999). More recently, Gambardella *et al.* (2005) have adopted the same approach as Harhoff *et al.*, but using a new survey of European inventors and found similar results. However, to our knowledge, there are only few studies focusing on European firms which analyze the economic value of R&D or patents using firm market value and most of these are for the UK only: Blundell *et al.* (1995), Toivanen *et al.* (2002), Bloom and Van Reenen (2002), and Greenhalgh and Rogers (2006). The only exception is Hall and Oriani (2006), who look at the market value of R&D (but not of patents) for three continental European countries: France, Germany, and Italy.

Several of these market valuation studies rely on measures of R&D expenditure, which is usually considered a measure of innovation input rather than innovation output or 'success' of innovative activities. However, in the case of European firms, data on R&D expenditures are often missing because reporting these expenditures is not required by accounting and fiscal regulations across most European countries. The UK is probably the only European country where an explicit recommendation of accounting practice encourages firms to disclose their

---

[3] This assertion has been confirmed by Bergstra and Klint (2007), who looked closely at 32 of the patents defined as software using the union of the two methods described later in the paper and concluded that only two were "pure" software.

R&D expenditures.[4] Nevertheless in this paper we rely on a sample of European firms for which R&D data is available, covering about 70 per cent of European business sector R&D in the year 2000, and then augment this panel with patent data. Patents as a measure of innovation have their own drawbacks but, as Griliches (1990: 1661) has remarked, 'in this desert of data, patent statistics loom up as a mirage of wonderful plenitude and objectivity.'

The paper is organized as follows. The next section describes the method for estimating the private value of R&D and patents using financial data. Section 3 presents the data and describes the main variables while Section 4 reports the main results. Section 5 discusses the results and closes the paper.

2. Estimating the economic value of innovation assets

There are two streams of the literature that attempt to evaluate the economic returns to innovative activities.[5] The first relates innovation as proxied by R&D and patents to total factor productivity or profitability, in most cases capturing a measure of private returns, although in principle the productivity approach can also yield social returns if prices are properly accounted for. The second, into which the present paper falls, measures the private returns to innovation using a forward looking measure of firm performance, its valuation in the stock market. Each of the two approaches has both merits and weaknesses.

Total factor productivity (TFP) is simply the ratio of outputs to inputs both expressed in real terms. Assuming only two inputs (capital K and labor L) and taking the natural logs of all variables the TFP of a firm can be expressed as follows:

$$log(TFP) = log(S) - \alpha \cdot log(L) - \beta \cdot log(K) \qquad (1)$$

This is an appropriate measure of productivity under conditions of constant returns to scale and competition in the markets for inputs and outputs.[6] Several studies have showed the importance of technology, measured by R&D expenditures, for the growth of total factor productivity at the firm level (e.g., Mansfield 1968, Gold 1977, and Griliches 1980).

[4] This recommendation dates from 1989 (see Toivanen *et al.*, 2002).

[5] See Hall (2006) for an analytical overview of econometric approaches to measuring the returns to R&D. Mairesse and Sassenou (1991) and Hall (2000) provide surveys of empirical results using the first and second methodologies respectively.

[6] Note that it is possible to relax the assumptions of constant returns and perfect competition in the output market and derive a version of this equation that will still yield a measure of productivity (or profitability) that can be related to innovation inputs.

Besides the strong assumptions necessary for TFP estimation, a major problem with this approach is the fact that the lag between R&D and its impact on productivity or profits is usually long and difficult to predict. Since this gives rise to serious measurement problems when the data are not available in long time series and when the process relating input and output is not stationary, much empirical work turns to alternative methods of measurement. In addition, the productivity approach that relies on accounting data often fails to allow for the effects of differences in systematic risk, temporary disequilibrium effects, tax laws and accounting conventions.

Some of these limitations are less important with the market value approach, which combines accounting data with measures of the value of the firm on the financial markets (Lindenberg and Ross, 1981; Montgomery and Wernerfelt, 1988). The market value approach draws on the idea, derived from the hedonic price models, that firms are bundles of assets (and capabilities) that are difficult to disentangle and to price separately on the market. These assets include plants and equipment, inventories, knowledge assets, customer networks, brand names and reputation. The assumption is that financial markets assign a valuation to the bundle of firms' assets that is equal to the present discounted value of their future cash flows. This approach has been used in several studies to calculate the marginal shadow value of knowledge assets across a range of firms (Griliches 1981; Griliches *et al.* 1991; Hall 1993; Hall *et al.* 2005; Hall 2006).

The general functional form of the value function for an intertemporal maximization program with several capital goods is difficult to derive and does not have a closed form in most cases (Wildasin, 1985). In most econometric studies this difficulty has been tackled by assuming that the market value equation takes a linear or log-linear (Cobb-Douglas) form. The typical linear market value model, which we use here, relies on the assumption that a firm's assets enter additively:

$$V_{it}(A_{it}, K_{it}) = q_t(A_{it}, + \gamma_t K_{it})^{\sigma t} \qquad (2)$$

where A represents the physical assets and K the knowledge assets of firm i at time t. Under constant returns to scale ($\sigma_t = 1$) equation (2) in log form can be written as

$$logV_{it} = logq_t + logA_{it}, + log(1+\gamma_t K_{it}/A_{it}) \qquad (3)$$

or

$$logQ_{it} = logV_{it}/A_{it} = logq_t + log(1+\gamma_t K_{it}/A_{it})$$

(4)

The left hand side of equation (4) is the log of Tobin's q, defined as the ratio of market value to the replacement cost of the firm, which is typically measured with the replacement value of firm's physical assets. On the right hand side, γ_t is the marginal or shadow value of the ratio of knowledge capital to physical assets at a given point in time. It measures the expectations of the investors over the effect of the knowledge capital relative to physical assets on the discounted future profits of the firm. The intercept ($log\ q_t$) represents the average logarithm of Tobin's q for the sample firms while $q_t\gamma_t$ is the absolute hedonic price of the knowledge capital.

As in Hall *et al.* 2005, equation (4) will be estimated by non-linear least squares. Most earlier research, beginning with Griliches (1981), have approximated the $log(1+\gamma_t K_{it} /A_{it})$ with $\gamma_t K_{it} /A_{it}$ and have estimated the market value equation by ordinary least squares; as the ratio of knowledge assets to ordinary assets has increased over time in many firms, this approximation has become less and less appropriate.[7] To ease interpretation of coefficient estimates for variables measured in widely differing units (dollars, euros, or counts) we computed the elasticity of Tobin's q with respect to each of the main regressors and displayed it in the tables below the coefficients.

$$\frac{\partial \log Q_{it}}{\partial \log X_{it}^{j}} = \frac{\gamma_j X_{it}^{j}}{1+\gamma_1 (RD_{it} / A_{it}) + \gamma_2 (P_{it} / RD_{it}) + \gamma_3 (CIT_{it} / P_{it})} \tag{5}$$

where X^j_{it} is the regressor of interest - R&D stock/physical assets, patent stock/R&D stock (total or software patents) and citation stock/patent stock. We computed these elasticities and their standard errors using the "delta" method for each observation in the dataset and then averaged them. The tables show the average elasticity and its average standard error.

Note that in general, shadow prices are equilibrium prices resulting from the interaction between the firm's demand and the market supply of capital for a specific asset at a given point in time. This implies that no structural interpretation should be attached to estimates of the market value equation. However, the values obtained by estimation of the market value equation are still informative, in the sense that they do measure the average marginal shadow values of an additional euro spent in R&D or an additional patent filed at a particular point in time.[8]

---

[7] We have also used OLS for comparison but the results are not reported here due to lack of space.

[8] For a more detailed discussion of various problems concerning the estimation and interpretation of the market value equation, see Hall (2000, 2006).

The market value approach rests on the restrictive hypothesis of capital market efficiency and therefore it can be used only for firms quoted in well-functioning and thickly traded stock markets. In fact, financial markets are not always perfect and there are persistent institutional differences across countries which may result in different evaluations of intangible assets.[9] To have an idea of the differences in the level of development of the stock market across countries, we looked at a somewhat imperfect measure, the ratio of stock market capitalization (aggregate market value of equity) to GDP (IMF 2006) in the right hand column of Table 2.

This ratio ranges from about 1.37 in the UK (very close to 1.36 of the US) through 0.73 of France, 0.43 in Germany, 0.45 in Italy to 0.11 in emerging Eastern European countries like Poland, Hungary and Ukraine.[10] The differences in financial development across European countries persist over time despite the rapid overall growth of the European financial markets during the 1980s and the 1990s (Rajan and Zingales, 2003) and tthese differences in financial market development could have a confounding effect on our estimates of the market value of intangible assets. For example, Hall and Oriani (2006) found that financial markets in France and Italy placed little value on the R&D performed in firms where the largest shareholder owned more than 30 per cent of the firm. However, over time the globalization of financial institutions (e.g., IMF, 2007) probably reduces the differences in asset valuation across countries with different financial development. In the regressions that follow, we control for country-specific differences in valuation (which are significant), but it would be desirable to probe this question more deeply in future work.

Not surprisingly, most empirical studies that follow the market value approach rely on data from the US and the UK, where the stock markets are larger and more thickly traded than in other countries. For related reasons, studies based on data from these countries also benefit from the availability of large sets of firm-level panel data. These studies find that R&D stocks are significantly valued by financial markets in addition to physical assets. The empirical evidence for the US also shows that patent counts have an additional, albeit weaker, impact on market

---

[9] We should note that other indicators of patent value have their own drawbacks. For example, survey data obtained by interviewing inventors may suffer from retrospective response bias. Data on patent renewal as an indicator of patent value do not provide information on the upper tail of the value distribution, where the most valuable patents are located.

[10] The large numbers for Switzerland and Spain presumably reflect the global nature of the financial sector in those countries, relative to the size of these country's economies.

value after controlling for R&D. Finally, Hall *et al.* 2005 find that citation-weighted patents are more informative than mere patent counts about the market value of innovation.

A series of studies based on European datasets have used the varying indicators of innovation (R&D, patents and patent citations) to confirm that, by and large, innovative assets impact significantly upon the firm market value (see Table 1 for a list of these studies).

[Table 1 about here]

3. Data

3.1. Sample

To construct our sample we started with 10,218 publicly-traded firms headquartered in 33 European countries over the period 1980-2005. Our sample includes a large variety of countries with different levels of financial development and accounting regulations, ranging from the UK, a common-law country with an active equity market, to emerging Eastern European countries with a very small market capitalisation-to-GDP ratio. Only 2,197 firms reported data on R&D expenditures for one or more of the sample years. For these firms we collected data on patents and found that 575 were granted at least one patent and 165 at least one software patent by the EPO during the period 1985-2005.

Data on corporate structure (date of incorporation, ownership structure, ultimate parent company, subsidiaries) and balance sheet were obtained from the Bureau van Dijk's Amadeus database. Changes in corporate structure were checked for the years 1998 to 2006 by drawing on different issues of Amadeus.[11] Data on market capitalization at the end of each year were obtained from Thomson Financial's Datastream. R&D data were obtained from Amadeus, Thomson Financials' Global Vantage and the UK Department of Industry's R&D Scoreboard. More precisely, we extracted from Amadeus all quoted companies reporting positive R&D expenditures for at least one year between 1980 and 2005 and filled in any missing R&D numbers for these firms using data from the other sources.

Firms' patent counts in all technological classes were obtained by matching the name of the assignee from the PATSTAT patent database with the company name in Amadeus. Patent citations and the number of IPC classes were also extracted from the PATSTAT database, available under license from the EPO-OECD Taskforce on Patent Statistics (PATSTAT 2006).

[11] Information on corporate structure for previous years has been retrieved manually from other sources, such as Who Owns Whom, Hoovers, etc..

For companies with subsidiaries, the patents of the ultimate parent company have been consolidated on the basis of the 1998-2006 ownership structure reported in Amadeus. Further information on corporate structure was collected from Hoovers, Who Owns Whom, and company websites for the period before 1998. Holding companies have been reclassified manually according to the main line of business or their most important subsidiaries using additional information from Amadeus, Hoovers, and company websites.

After dropping a few observations with extreme outliers in the patent data, very small firms and those with unconsolidated data, our final sample consisted of 1,060 firms for the period 1991 through 2002. The choice of time period was dictated by the fact that the patent quality measures are based on forward citations, and we required at least three years in which to observe them following the patent application (that is, 2003-2005). We consolidated some of the countries with small numbers of firms into larger groupings in order to reduce the number of dummies needed (e.g., all Eastern European countries form one group, and Spain and Greece another).

As Table 2 shows, over 90 per cent of the sample of firms for which both R&D and market capitalization are available generally consists of medium to large firms (over 5 million sales and 100 employees according to the Eurostat definition). About two-thirds of the sample is composed of firms with over 100 million sales (the Eurostat definition of a large firm is one with more than 20 million sales).

[Table 2 about here]

Tables A.1 and A.2 in Appendix A report the distribution of the firms in the sample by market capitalization and the main stock markets involved. About half the firms in this sample have a market capitalization less than 100 million euros and only about 10 per cent of firms have a capitalization above 5 billion euros, with over half of these very large firms having been established before 1970. On the other end of the distribution, about one third of firms with a capitalization less than 1 billion euros were incorporated since 1990; 20 per cent of those with capitalization between 1 and 5 billion euros are also new firms. This latter fact is in part the result of restructuring, liberalization and privatization of formerly state-owned corporations in many European continental countries during the 1990s. Another reason is the entry of software and “internet economy” companies such SAP, Business Objects, Infineon Technologies and O2.

The R&D-reporting firms in our sample are in a large number of sectors (see Table A.3 in the Appendix) and about half of these firms hold EPO or US patents. However, although about

20 per cent have US software patents, only 30 firms have EPO patents that we identify as "pure" software, reflecting the fact that such patents are not usually granted at the EPO.[12] Of these 30 firms, two-thirds are in computing hardware and software, telecommunications, electronics, and electrical machinery.

The distribution of patents and R&D expenditures across industries is reported in Table A.4. The most important sectors in terms of R&D expenditures are pharmaceuticals and chemical products, motor vehicles, electronic instruments & communications equipment, and electrical machinery. These are also the most important sectors in terms of total US and EP patents. As expected software patents are more concentrated in few industries, with electrical machinery alone accounting for half of EP software patents and 32 per cent of US software patents. Electrical machinery, electronics and communication equipment, and telecommunications services together account for over 85 per cent of EP software patents and 65 per cent of US software patents.

We should note that the high concentration of software patents in few sectors is due in part to the exclusion of non-European firms from the sample. For example, IBM accounts for about 10 per cent of total EPO software patents granted to business enterprises, followed by Siemens and Canon (about 4 per cent each). Other large electronics firms are also relatively large software owners – e.g., Philips (3.4 per cent) and Sony (2.5 per cent). The largest software firm among the top owners of EPO software patents is Microsoft with a one per cent share.

3.2. Variables

Our dependent variable is Tobin's q for the firm, that is, the ratio of the firm's market value to tangible assets. Firm's market value is defined as the sum of market capitalization (price multiplied by the number of outstanding shares at the end of the year) and non current liabilities less a correction for net current liabilities plus inventories.[13] Tangible assets are the net costs of tangible fixed property and inventories used in the production of revenue, and are obtained as the sum of gross fixed assets plus inventory stocks less depreciation, depletion, and amortization (accumulated), investment grants and other deductions.[14]

---

[12] The exact definition of a software patent used here is given in section 3.5 of the paper.

[13] Outstanding shares include both common shares and preferred shares.

[14] All values expressed in domestic currencies have been converted into euros by using annual average exchange rates reported by EUROSTAT.

Corporate finance scholars have developed alternative, more complex estimations of the Tobin's q which rely on estimated market value of the firm compared with that used in this paper (e.g., Perfect and Wiles, 1994). These alternative approaches to Tobin's q measurement produce more precise estimations but are computationally costly. Moreover, their greater precision is traded off by a larger selection bias. DaDalt et al. (2003) have used the Compustat dataset and found that using the Perfect and Wiles' approach produces a 20 per cent loss in sample size. It is important to note that DaDalt et al (2003) have estimated that simple methods, like that used here, and complex ones, like that of Perfect and Wiles, agree in approximately 90% of cases for values of q below 0.8 and above 1.2. As Table 3 clearly shows, for most firms in our sample the Tobin's q value is above 1.2.

The R&D expenditure history of each firm was used to compute R&D stock. R&D spending includes amortization of software costs, company-sponsored research and development, and software expenses. As mentioned earlier, European firms are not required or recommended to disclose information on their R&D expenditures, implying that the availability of data on R&D expenditures is potential source of sample selection bias. Reporting R&D is then an endogenous variable since the decision whether or not to disclose this information rests upon the discretion of the firm. Hall and Oriani (2006) found that selection was not a factor for most of the countries they considered. We treat this issue in Section 4 of the paper.

Since the stock of our key regressors cannot be measured directly from the firm books we rely on proxies obtained from current and past flows of R&D and patent-related variables. R&D stocks (*KRD*) were obtained using a declining balance formula and the past history of R&D spending:

$$KRD_t = R\&D_t + (1-\delta)KRD_{t-1}$$

where δ is the depreciation rate. We chose the usual 15 per cent depreciation rate for easy comparison to earlier work. Our starting R&D stock was calculated for each firm at the first available R&D observation year as $KRD_o = RD_o/(\delta+g)$. This assumes that real R&D has been growing at a constant annual growth prior to the sample; we used a growth rate g of 8 per cent. Patent stocks were obtained using the same methods, except that the initial available patent counts were not discounted to obtain an initial capital stock because we have a longer pre-sample history of patenting (back to 1978) than for R&D, so the impact of the initial stock is minimal.[15,16]

Our controls include firms' annual sales, which account for scale effects in the market value equation, industry dummies, country dummies and year dummies.[17] Firms' R&D and sales have been depreciated by the annual GDP deflator extracted from the AMECO-EUROSTAT web directory. In future work, we will control for differences in ownership structure (see also Hall and Oriani, 2006).

3.3. Patent variables

For each of our firms, we have data on their EPO patents (and European national patents) as well as data on their USPTO patents. We have also constructed the 'family' of each patent as described in the next section.[18] We then identified the categories of patents shown in the table following:

(i) EP	All EPO patents (labelled European in the tables)
(ii) EP only	EPO patents only (i.e., EPO patents without US equivalents, although they may have equivalents elsewhere in the world or in the European national offices)
(iii) EPUS	EPO patents with at least one US equivalent (i.e., patents whose family includes at least one US patent)
(iv) US	All USPTO patents (labelled US in the tables)
(v) US only	USPTO patents only (i.e., USPTO patents with no EPO or European national office equivalents, although they may have equivalents elsewhere in the world)
(vi) USEP	US patents with at least one EPO or European national office equivalent

[15] Because our patent data begin in 1978 and the first year we use in the regressions is 1991, the effects of omitted initial conditions will be small ($0.85^{14} = 0.10$).

[16] Our approach to the construction of patent stocks follows the methodology in Hall, Jaffe, and Trajtenberg (2005), in order to facilitate comparison with that paper's results.

[17] We use sales rather than assets to reduce measurement error bias arising from the fact that assets also appear on the left hand side of the equation.

[18] We used the PATSTAT database to find priority links between patent documents in different the EPO and USPTO.

Note that (i) is the disjoint sum of (ii) and (iii) and (iv) is the disjoint sum of (v) and (vi), but that (i) and (iv) may contain many of the same inventions. Note also that in principle (ii)+(v)+(iii) should be approximately equal to (ii)+(v)+(vi) and that either should cover all inventions that are patented in the US and/or the EPO.[19] In the regressions presented in section 4 we have explored the significance of these different measures by including (i) and (iv) separately and then breaking these into their constituents.

3.4. Patent quality measures

Research on the economic importance of individual patented inventions have demonstrated that the distribution of patent value is very skewed (e.g., Harhoff *et al.* 1999). The large majority of patents have an extremely limited commercial value and only few represent an important source of revenues to the assignee. It is therefore desirable to make use of patent stock measures that are adjusted for the quality of the patents they contain. We make use of two such quality weights, both of which have been used in prior empirical investigations: forward citations (as in Hall, Jaffe, and Trajtenberg 2005) and an index derived from a factor model based on three indicators, as suggested by Lanjouw and Schankerman 2004. The indicators we use are forward citations, number of IPC classes, and family size). We describe each of these two quality measures in more detail below.

Forward citations received by a patent indicate that the information in an invention has served as a basis for a future invention. Citations, i.e., citations of 'prior art' that is relevant to a patent, serve an important legal function, since they delimit the scope of the property rights awarded to the patent. Thus, if patent B cites patent A, it implies that patent A represents a piece of previously existing knowledge upon which patent B builds, and over which B cannot have a claim. Citations to other patents then can be considered as evidence of spillovers or knowledge flows between patented inventions.

However, the usefulness of citations as a proxy for knowledge spillovers is limited by the fact that citations are not always added by the inventor (Jaffe *et al.* 2000). In the US, the applicant is required to disclose her knowledge of the prior art, although in fact, references to prior art are often found by the inventor's patent attorneys, rather than the inventor, and the

[19] The counts are not identical between whether one starts with EP or US patents, for two reasons: 1) the equivalence correspondence may be one-to-many in either direction; and 2) our name-matching of patents to firms may not have picked up all the US subsidiaries of the European firms. However, the two stocks are correlated 0.96 so that the error from (2) is fairly small,.

decisions regarding which patents to cite ultimately rests with the patent examiner, who is supposed to be an expert in the area and hence to be able to find prior art that the applicant misses or conceals.

In the case of EPO patents, inventors are not required to cite prior art and therefore references to earlier patents are usually added by patent examiners. This suggests that patent citations to EPO patents may be even less useful as a measure of spillovers. However, compared to the USPTO, citations contained in EPO patents tend to be more consistent and objective because they are assigned by a single team of patent examiners. Unlike that at the USPTO, EPO citation practice also tends to minimize the number of citations per patent. For more information on the meaning of European patent citations, see Harhoff, Hoisl, and Webb (2006).

In order to make citations to EPO and USPTO patents as comparable as possible given these differences, we have to take into account another important difference between the two patent systems. Unlike US patents, a large share of EPO patents are cited indirectly through their non-EPO equivalents, i.e., different 'incarnations' of the same inventions in other patent systems such as the European national patent offices and the USPTO. For this reason Harhoff et al. (2006) suggest that citation links to EPO patents should include also citations received by their equivalents. To account for this difference in citation patterns we counted direct and indirect citations to both EPO and USPTO patents.

We used PATSTAT (release of September 2006) to retrieve data on citations counts, which reports around 63 million citing correspondences up to December 2005. US patents received directly about 42.6 million or 68 per cent of all world citations contained in the PATSTAT dataset (for comparison, US patent applications were about one quarter of worldwide applications during the 2001-2004 period, according to the WIPO statistical database). 5.5 million US patents have received at least one cite.

After excluding patent applications that were not yet granted, we retrieved information on the publication dates of the citing patents. When the publication date of the citing patent was missing or it was antecedent to the date of the cited patent (approximately 2.7 million citations, about 7 per cent of the total number), the citation was not included in the analysis. Our final sample consisted of approximately 4.7 million U.S. patents having at least one citation prior to December 2005, 3.1 million patents having at least one citation within 5 years from the publication date, and 2.45 million having at least one citation within 3 years from the publication date.

EPO patents, which are about 8 per cent of worldwide applications during the same period, receive far fewer citations directly. EPO patents and their non-EPO equivalents overall receive about 1.8 million citations (2.8 per cent of the total) and 529,161 EPO granted patents have at least one cite. Restricting the citation lag to three years gives 460,142 citation links, of which about half are accounted for by citations to non-EPO equivalents of EPO patents.

For comparability we used the same search strategy for both EPO and USPTO patents, including citations to their equivalents. In particular, for EPO patents we considered as a citation link to an EPO patent the direct citation to a direct equivalent of that EPO patent. For example if the EPO patent X had two direct equivalents Y and Z respectively in two other patent offices, the citation count of X included not only the direct cites to X but also the direct cites to Y and Z (with duplicate cites removed). The same search strategy was followed for USPTO patents. For more details on this methodology see Harhoff *et al.* (2006).

Previous studies have also used backward citations as a measure of the quality of the citing patent. Some scholars have suggested that large numbers of citations to others reveal that a particular invention is likely to be more derivative in nature and, therefore, of limited importance (Lanjouw and Schankerman 2004). However, a large number of backward citations may also indicate a novel combination of existing ideas. This is probably the reason why Harhoff *et al.* (1999) have found that backward citations are positively correlated with patent value. Because of this ambiguity we do not use this variable in our analysis.[20]

Our second measure of patent quality was based on three indicators of patent value rather than one; in addition to forward citations, we used family size (the number of jurisdictions or countries the patent has been applied for) and the number of different technological classes assigned by patent examiners to a given patent.[21]

Our measure of family size was obtained as follows. We identified all priorities for the EPO patents in our sample firms (recall that there is a many-to-many correspondence between patents and priorities). Using this information, we found the non-EPO patents that reported the same priority. This first step gives a lower bound on the family size. The second step was to find all applications (EPO and non-EPO) that report an EPO application from one of our firms as a

[20] Our results do not change substantially when backward citations are used along with other indicators of patent quality.

[21] Other studies have also used the number of claims which delimit the scope of the invention as a measure of patent quality; this variable was not available to us in PATSTAT.

priority.[22] After removing any double counting, the number of patent applications thus identified plus those from the first step constitute the size of the patent family. The same procedure was followed to obtain the family size of US patents. Note that our definition is the same as the middle of the three definitions (equivalent, family, and extended family) suggested by Harhoff *et al.* (2006).

The number of technological classes have been shown to be an indicator of technological "quality" similar to the number of citations by Lerner (1994). To guarantee a reasonable level of precision, we use the number of eight-digit IPC classification codes reported in the patent document. The number of IPC classes can be viewed as a measure of technological scope or generality of the patent even though, as noted by Guellec and Pottelsberghe de la Potterie (2000), it may be also a measure of ambiguity reflecting the difficulty of the examiner in locating the invention in the technological space.

These three indicators were combined into a composite index of patent 'quality' derived from a common factor model in an approach developed by Lanjouw and Schankerman (2004). The common factor explains as much as possible the total variance of each indicator while minimizing its idiosyncratic component. The methodology is briefly described in Appendix B. The three component indicators are all strongly correlated with each other at the 1% level of significance.

3.5. Correcting for citation truncation

Patent citations suffer from several potential sources of biases, the most obvious of which is truncation. The number of citations to any patent is truncated in time because only citations received until the end of the dataset are observed. The observed number of citations to any given patent may also be affected by differences across patent cohorts, technological fields and patent offices. The observed citations then have to be adjusted or normalized for this multiplicity of effects. For this purpose we have adopted the approach developed by Caballero and Jaffe (1993) and Hall *et al.* (2005) – hereafter referred to as the HJT method- which is based on the estimation of a semi-structural model where the citation frequency is explained by cited patent-year effects, citing patent-year effects, technological field effects and citation lag effects. The estimated parameters of this model can be used to correct observed citation rates. Appendix B reports a

[22] EPO patents which refer to earlier EPO patents as their priority are classified as divisional patents by the EPO and correspond to continuations in the USPTO system.

brief description of the HJT method and the distribution of the weights used by technology field. The inverse of the numbers in Tables B.1 and B.2 gives the proportion of the lifetime citations that are predicted to occur in the time window observed. Actual citations are multiplied by the numbers given to correct for truncation.

3.6. Software patents

One of the goals of the research reported here was to get a picture of the use and valuation of software patents in European firms. More precisely, comparing the existence of and valuation of software patents in the US and the EP patent systems may shed light on some differences between these two patent systems. Software represents an interesting technology for our purposes because of the growing attention to software patents amongst business practitioners, scholars and policy-makers. Critics claim that software patents have an average poor quality and are applied for mainly for 'strategic' reasons rather than for protecting real inventions, whereas advocates maintain that software inventions are technological inventions like any other and should be entitled to patentability. Scholars looking at software-related patents have found evidence consistent with the hypothesis that strategic patent portfolio building in the ICT sector lies behind the increase in software patents. Studies using different definitions of software patents all find that the number of USPTO software patents is large and growing and that the holders of these patents are large hardware rather than software firms (Bessen and Hunt 2004; Graham and Mowery 2003; Hall and MacGarvie 2006). Bessen and Hunt (2004) have pointed out that IBM alone accounts for over 20% of software patents held by US firms. Hall and MacGarvie (2006) find that the widespread introduction of software patenting in the U.S. via court decisions was initially negative for software firms, but that these patents have become more privately valuable than other patents in the recent past. At the same time, their "quality" as measured by citations does not matter for hardware firm value, which suggests that adding an additional patent to the portfolio is more important than the patent *per se*.

Even in the U.S., it is difficult to find a simple definition of a software-related patent that can be used for statistical purposes, that is, does not require the reading of individual patents. In Europe it is even more difficult, because the international patent classification system does not actually recognize their existence. We therefore chose to rely on the methods used in the earlier studies on USPTO data, which are based on keyword searching as well as identifying class/subclass combinations in which pure software firms patent. The three main alternatives are those used by Graham and Mowery (2003), Bessen and Hunt (2004), and Hall and MacGarvie (2006).

Graham and Mowery identify as software patents those that fall in particular International Patent Classification (IPC) class/subclass/groups. Broadly defined, the classes are "Electric Digital Data Processing" (G06F), "Recognition of Data; Presentation of Data; Record Carriers; Handling Record Carriers" (G06K), and "Electric Communication Technique" (H04L).[23] Graham and Mowery selected the subclasses from these classes in which six large U.S. software producers patented between 1984 and 1995. They found that patents in these classes account for 57% of the patents assigned to the hundred largest firms in the software industry.[24]

An alternative definition is that adopted by Bessen and Hunt who define software patents as those that include the words "software" or "computer" & "program" in the patent document description. Patents that meet these criteria and also contain the words "semiconductor", "chip", "circuit", "circuitry" or "bus" in the title are excluded under the assumption that they refer to the device used to execute the computer program rather than the program itself.

Hall and MacGarvie (2004) suggest a third algorithm to define software patents that identifies all the U.S. patent class-subclass combinations in which fifteen "pure" software firms patent, yielding 2,886 unique class-subclass combinations. Patents falling in the classes and subclasses combinations obtained from this search method are defined as software patents. The definition preferred by Hall and MacGarvie combines this definition with that of Graham and Mowery and then takes the intersection of the result with the Bessen-Hunt sample. Hall and MacGarvie report that their results for the market value of software patents are not significantly affected by the choice of definition.

We followed a combination of the search methods above to identify software patents at the EPO. First, we searched the title, abstract, claims and description of patents in the EPO dataset by relying on the same keywords used by Bessen and Hunt in their 2002 study of US software patents: ((software) OR (computer AND program)) AND NOT (chip OR semiconductor OR bus OR circuit OR circuitry <in> TI) AND NOT (antigen OR antigenic OR chromatography). To obtain keywords and classification for the patents we relied on the

---

[23] The detailed class/subclass groups included are G06F: 3,5,7,9,11,12,13,15; G06K: 9,15; H04L: 9.

[24] Graham and Mowery (2003), p. 232. The firms are Microsoft, Adobe, Novell, Autodesk, Intuit, and Symantec.

Delphion dataset (www.delphion.com), which gives access to the full-text of the patent document, including the application date, the technological classes and the address of the assignee. This procedure yielded 11,969 patents (in 7,117 different IPC classes-subclasses) (the *keyword method* hereafter).

Second, we analyzed the IPC (International Patent Classification) classes of the patent portfolios of the world's 15 largest specialized software firms (the IPC method hereafter). We expanded the set of firms used in earlier studies to obtain a representative sample of specialized software firms including European companies.[25] The firms we used account for over 30% of the world software market ($227 billion according to European Information Technology Observatory estimates). They have been granted 373 patents in 3,518 different technological classes-subclasses (117 if one considers only the main IPC codes in each patent).

As in Hall-MacGarvie (2006), we defined a software patent as one that fell in the intersection of the two sets of patents defined by the keyword and IPC methods.[26] As one might expect, this method yielded very different results for patents issued by the two patent systems: in the US, 6.7 per cent of the granted patents applied for during the 1991-2002 period by firms in our sample are software patents by this definition, whereas at the EPO, only 0.4 per cent of issued patents are software patents, a total of 286 patents. Of these, one third of the sample software patents are held by Siemens, and 75 per cent by the top five firms (Siemens, BT, Philips, Oce, and Alcatel). The largest software firm, SAP, holds 5. Two conclusions can be drawn from these facts: first, the EPO has been mostly successful at holding the line against "pure" software patents; and second, to the extent they exist, they are mostly held by hardware rather than software firms, as in the case of USPTO software patents.

---

[25] The top European software patenters over 1978-2004 are Microsoft, Oracle, Peoplesoft, Veritas, Symantec, Adobe Systems, Novell, Autodesk, Intuit, Siebel Systems, Computare, BMC Software, Computer Associates, Electronic arts (Japan), and SAP (Germany), whereas the top U.S. software patenters during the 1980-2000 period are Microsoft, Oracle, Peoplesoft, Veritas, Symantec, Adobe Systems, Novell, Autodesk, Macromedia, Borland, Wall Data, Phoenix, Informix, Starfish, and RSA Security. Only half the firms are common between the two lists, and only two firms are not U.S.-based.

[26] By relying on the intersection between the two methods we reduce the Type I-error (excluding a patent that we should have included among software patents) and high Type-II error (classify as software patent a patent that is not related to software). Preliminary work by Bergstra and Klint (2007) suggests that there is fair amount of Type-II error in EPO software patents when the union of the keyword and the IPC method is adopted. Using the intersection of the two methods we find few EPO patents that qualify as pure software, which suggests that the EPO is successful in restricting patenting in this area (many pure software patents do not qualify for patentable subject matter because, according to the EPC, they do not produce any technical effect or are not capable of industrial application).

3.7. *Descriptive statistics*

Tables 3a and 3b show some descriptive statistics for the final sample of 1061 firms, an unbalanced panel with 5,312 observations (from 1 to 12 years per firm). Table 3a gives statistics for the continuous variables and Table 3b for the various patent measures. The firms in the sample are large, with median sales of 306 million euros and median employment of 1423. They are fairly R&D intensive, with a median R&D to tangible asset ratio of 0.25, and this is reflected in their median Tobin's q of 1.7, which is well above unity.

Table 3b reports descriptive statistics for granted patents (by their priority date), the stock of granted patents, the ratio of patent stocks to R&D stocks, and the ratio of citation stocks to patent stocks, for all patents and for software patents separately. In this table the statistics for all of the variables are based on the entire sample, but the number of non-zero observations is reported for each variable. For the patent flows, we report statistics on the six types of patents described in the previous section: all EPO patents, all US patents, EPO only, US only, EPO with US equivalents, and US with EPO or European national office equivalents. For the sake of brevity, only the statistics for patent grants include those for EPO only and US only patents, as these can generally be derived for the difference between the total and the equivalents (see the discussion in section 3).

[Table 3 about here]

This table reveals that the firms in our sample take out twice as many USPTO as EPO patents (13 per firm year versus 26 per firm per year) and that this is reflected in a much larger share of inventions for which protection is sought only in the US and not in Europe (about 50 per cent) as compared to the reverse situation (about 25 per cent). The average firm that spends one million euros on R&D obtains 0.3 EPO patents and 0.44 USPTO patents, but of course the distributions are very skew, with medians of 0.08 and 0.15 respectively. USPTO patents receive far more citations (corrected for truncation) than EPO patents (12 versus 3), probably reflecting differences in the two patent systems.[27]

[27] In principle, the differences in citation behavior should affect the citing, not the cited patent, but to the extent that search is local to a patent office, and also to particular technologies, these differences will also affect the patents being cited.

4. Results

Tables 4 through 7 contain the results of our estimations. The equation estimated is based on equation (4) and is estimated by nonlinear least squares: [28]

$$\log Q_{it} = q_t + \lambda_k + \delta_j + \beta_s \log S_{it} + \log\left[1 + \sum_l \gamma_l X_{it}^l\right] \tag{1}$$

where *i, t, k, j,* and *l* index firms, years, countries, industries, and variables respectively. *S* is a control for size (the current sales or turnover of the firm); the size coefficient was invariably small and positive, and had little impact on the rest of the equation. The X_{it}^j are the various measures of R&D, patent, and citations stock ratios.

Table 4 contains our basic results using the R&D-assets ratio and various patent stock-R&D stock ratios. Table 5 adds information on software patents, and Table 6 includes information on the two patent value indicators. Table 7 reports results with the value indicators for software patents also included separately. Each table displays coefficient estimates and their robust standard errors in the top panel, and the average elasticities implied by the coefficients in the bottom panel. Below we discuss each of the tables in turn.

4.1. Estimation of the basic model without citations

The results for the basic model that includes R&D stocks, total patent stocks, and software patent stocks are shown in Table 4. For this model only, we also show the coefficients on dummies for zero patent stocks; these were included in all the models to control for possible differences in non-patenting firms or errors in matching, but they are not readily interpretable.

The first and by far the most significant and robust result in all the tables is that the ratio between R&D stock and physical assets is positively and significantly related to Tobin's *q* across different specifications of the market value equation. The magnitude of the coefficient (slightly less than unity) is consistent with most of those reported in earlier works on single or multiple countries (e.g., Hall 2000; Blundell *et al.* 2002; Toivanen *et al.* 2002; Hall and Oriani 2005; Greenhalgh and Rogers 2006). The estimated elasticity is even more robust across all the specifications in Tables 4 and 5, taking values within a small interval around 0.20 in almost all

[28] OLS estimates of the log approximation to equation (4) produced similar results and are therefore not shown. We should also note that approximating the log (1+x) with x reduces substantially the accuracy of estimates. For instance, approximating the log (1+R&D/assests) with R&D/assets=0.2 yields a 10 per cent measurement error (log(1+0.2)=0.18). With a R&D/asset ratio equal to 0.55 the error amounts to about 25 per cent.

cases. The average R&D-assets ratio is 0.51 with a standard deviation of 0.74, so that these estimates imply that firm which is one standard deviation above the mean has a market value that is 30 per cent higher than the average firm.

Second, in all specifications a firm's patent stocks are significantly related to value, above and beyond the R&D stock that generated them, but with some interesting detail, depending on the jurisdictions in which the patent was taken out. As discussed earlier, we have six possible (overlapping) patent measures: (i) EPO, (ii) EPO only, (iii) EPO with US equivalents, (iv) US, (v) US only, and (vi) US with European equivalents. In columns (2) and (3), we compare the use of all EPO and all US patents in the equation and find that both are significantly related to market value, with US patents having a slightly higher coefficient and elasticity (0.05 versus 0.03).

In models (4) and (5) we break up these two measures into patents with equivalents in the other jurisdiction and without. In models (6) and (7) we include the three indicators that should exhaust the information available, first using EP patents with US equivalents and then using US patents with EP equivalents.[29] The message is fairly clear and persists throughout these tables with only a few exceptions: Patents taken out at the EPO only are not valued by the financial markets once we control for equivalent patents taken out in the US. In addition, patents taken out in both jurisdictions are clearly more valuable than those taken out only in the U.S. An additional US patent with European equivalent per million euros of R&D leads to a 20 per cent increase in market value, whereas an additional EPO patent with a US equivalent leads to a 30 per cent increase in market value. An additional US patent without an equivalent per million euros of R&D leads only to a 12 per cent increase in market value, but an additional patent taken out only in Europe adds an insignificant amount to value. Clearly, there is a substantial premium to geographical scope for EPO patents, even when controlling only for patenting in the US and not for the rest of the world. Financial markets place a positive value on EPO patented inventions owned by European firms only when patent protection is also acquired in the United States.

[29] It is worth noting that about two-thirds of our firms regularly patent in the USPTO and about one third never patents there. The number of EP patents is smaller than US patents for various reasons. First, the EP system is younger than its US counterpart. Second, the examination-granting lag is larger in the EP. Finally, many firms in our sample carry out R&D activities in the US and therefore they may file their first patent application to the USPTO to establish the priority of an invention and then use the PCT system to obtain protection in the European national countries. Finally, German firms (and UK as well) tend to apply more to their national patent system and the USPTO than the EPO. We took account of this fact when defining equivalents to US patents, but data constraints prevent us from including these patents themselves.

The average elasticities reported in the bottom panel of Table 4 show that EPO and US patents have a similar impact with an elasticity of 4 and 3 per cent respectively, but with some of the US impact coming from patents taken out only in the US. A one standard deviation increase in the stock of EPO patents with US equivalents relative to R&D is associated with about an 11 per cent increase of market value, and similarly a one standard deviation increase in the stock of US patents with EPO equivalents (relative to R&D stock) yields a 10 per cent increase in market value.

[Table 4 about here]

The coefficients for EPO and US patents in Table 4 are substantially higher than the coefficient obtained by Hall *et al.* 2005 using the same methodology for U.S. firms and U.S. patent data during the 1980s: between 0.16 and 0.18 as compared with 0.03 for the earlier period and data. However, they are closer to the estimate obtained by Hall and MacGarvie 2006 for a sample of US information and communication technology (ICT) firms during the late 1990s, which was 0.15. Note that the estimates here are probably the first set of estimates using patents for firms from continental European countries and they seem to suggest that the incremental value of EPO patents above and beyond the R&D that generated them is roughly the same as that of US patents, but only if these patents have equivalents in the US system.

Given the results in Table 4, which shows that patents taken out in only one jurisdiction have little if any association with firm market value, in Table 5, which looks at software patents, we focus on the specifications that break patents up into those that have equivalents in the other jurisdiction and those that do not. This table repeats the regressions of Table 4, adding separate patent stock-R&D stock ratios for software patents. The coefficients of the software patents stock-R&D ratios are to be interpreted as premia or discounts for patents that fall into the software class. However, the elasticities shown are the total elasticities for software patents rather than premia, for ease in interpretation. Because the very small number of EPO only software patents (fewer than 10 per year) and because EPO only patents are generally not value-relevant, we have omitted this variable from the regression in Table 5.

[Table 5 about here]

The results in Table 5 for patents in general are similar to those in Table 4, with the only patents that are informative for market value are those taken out in both jurisdictions and to a lesser extent, those taken out in the US only. US software patents with EP equivalents and EP software patents with US equivalents are both valued at a considerable premium over other patents. Although the coefficients appear very large, it has to be remembered that the variables themselves (software patent-total R&D ratios) are very small so that the elasticities are small. The US software patent-R&D ratio has an elasticity of around one per cent, implying that a doubling of software patent yield per R&D would increase market value by one per cent. Although the coefficient on US software patents is smaller than that on EPO patents the average elasticity of EPO software patents is very close to zero. This is because the higher coefficient for EPO patents is inversely correlated with their smaller numbers. Thus, although each EPO software patent is more valuable than a USPTO patent, the same per cent increase in either stock produces a much smaller impact on market value in the case of EPO patents.

Note also that software patents taken out only in the U.S., which are actually more numerous than those with equivalents in Europe, are no more valuable than other US only patents. These are presumably patents on inventions that are not eligible for patenting at the EPO, and it is interesting that they are not as valuable to European firms as software patents that can be taken out in both jurisdictions.

4.2. Sample selection bias

As mentioned earlier in the paper, the disclosure of R&D expenditures is an endogenous variable and this gives rise to potential sample selection bias. To see whether sample selection biases our results, we first calculated the share of total R&D in the population of manufacturing and utility firms accounted for by our sample. Country-level R&D expenditures were taken from the OECD STAN dataset. As Table A.5 shows, the ratio of total R&D in our sample to the country-level industrial R&D varies across countries. For example, the ratio was 99.5 per cent in France, 98 per cent in Germany and over 100 per cent in the UK and Switzerland.[30] Apparently, the problem of sample selection is potentially relevant for firms from Spain and Italy while it is less important for other firms in our sample. Overall, the high coverage of national R&D expenditures demonstrates that in Europe, as in the US, most of the business R&D activity is conducted by large, publicly traded firms. Moreover, our sample accounts for around 15.9% of overall patenting activity and 6.7% of software patenting activity at the EPO. These shares are quite large given that our sample does not include firms from the United States and Japan.

---

[30] The fact that the share is above unity is explained by the R&D activity of their foreign subsidiaries abroad.

To check for sample selection bias we estimated a sample selection model using the Heckman two step method. For this purpose we collected accounting data for 3,773 publicly-listed firms that report data on R&D and for a matching sample of 3,194 publicly-listed firms from the same countries but which had reported no R&D data over the period 1991-2002.[31] The non-R&D doing firms are smaller, less labour-intensive, have higher leverage, and lower Tobin's *q*.

Our selection equation includes leverage (the ratio of current + non-current debt to tangible fixed assets), capital intensity (the ratio of tangible fixed assets to sales), and labor intensity (the ratio of labor cost to sales), as well as the share of the firm held by the main shareholder to account for observable firm characteristics that can affect its decision whether or not to reveal R&D expenditures. To account for 'environmental' factors we also included industry and year dummies in the equation. The inverse Mills' ratio obtained from the first stage estimation obtained by a probit model was then entered in the market value equation (see Maddala, 1983 and Hall, 1987). Our results show that there is little evidence of sample selection.[32] This result is consistent with that of Hall and Oriani (2006) for firms in France, Germany, and Italy.

4.3. Accounting for patent quality

Tables 6 and 7 report estimations that include our patent value indicators (forward citations and the composite 'quality' index) for total patents and software patents. In these tables we restrict the specifications to two: one that includes EPO only patents, US only patents, and EPO patents with US equivalents, and one that includes EPO only patents, US only patents, and US patents with European equivalents. The first two columns of Table 6 report the results of specification including the average forward citation/patent stock ratios, and the second two columns report those including the average factor index/patent stock ratios. Table 7 reports the same thing including software patents, but only for the second specification (US patents with EPO equivalents) because of the paucity of EPO software patents.

[31] The sample includes all publicly listed firms in the sample countries whose accounting data are available in Amadeus company directory.

[32] The estimated coefficient on the inverse Mills' ratio does not enter significantly in the market value equation at the 10 per cent level . Firms from Austria and Ireland were dropped because of the small number of observations. The results of these estimations are available upon request.

The results in these tables for R&D and patents are similar to those in the previous tables. Citations yield an additional albeit small premium to either the EPO or the US patent counts. It is worth to note that the 'quality' of EPO only and US only patents, whether measured by citations or the factor index, does not yield any significant impact on the market value of the firm. Recall that both EPO and US citations include all citations to their equivalents. Probably because EPO citations are more parsimonious in general, the EPO citation-patent stock ratio has a mean and standard deviation of 2.9 and 3.3, much smaller than the US citation-patent stock ratio, with 12.6 and 15.3 respectively. The elasticity of market value with respect to the EPO (with US equivalents) cite per patent ratio is 7.0 per cent, as compared to 2.8 per cent for the US with European equivalents, which suggests that these cites are even more informative about value than would be suggested by the 4 to 1 ratio in which they are received.

[Table 6 about here]

The second pair of columns in Table 6 report similar results using the patent quality index based on 3-year forward citations, family size, and number of IPC classes instead of forward citations alone. The other coefficients in the regression are little affected by the change in quality indicator. However, the elasticity of market value with respect to the index is greater than that with respect to forward citations, suggesting that it is a somewhat better proxy for the average quality of a firm's patented inventions. For EPO patents with US equivalents a one standard deviation increase in average patent quality is associated with an increase in the market value of the firm equal to 5.0 per cent. The same calculations for US patents yields a 7.4 per cent increase. The corresponding numbers for the forward citation measure are 11.1 per cent for average cites to EPO patents with US equivalents and 6.2 per cent for average cites to US patents with EPO equivalents. Thus it appears that citations to EPO patents and their equivalents are a somewhat stronger value indicator than the constructed index, while for US patents both are about the same.

Table 7 reports the results of similar estimations that include software patents. The only significant result for software patents is the positive premium for patents with equivalents, as before. There is no premium for higher "quality" software patents, at least not using our measures of quality; in fact the elasticity of market value with respect to the quality indicators for software patents is almost exactly the same as that for ordinary patents. This result indicates that the financial market does not recognize any additional premium from the "intrinsic" value of software-related inventions.

[Table 7 about here]

Various robustness checks of the above results have been done using regressions that excluded extreme values of R&D stocks, patent stocks, the composite 'quality' index and software citation stocks. The qualitative results are very similar. However, these estimations do not account for bias due to unobserved firm-specific heterogeneity. We defer this to future research using panel data estimation.

5. Discussion and conclusions

This paper reports some new estimates of the economic value of patents in a sample of European firms. The main novelty of the paper consists in the use of both EPO and USPTO patents and quality-adjusted patents in the market value equation. In addition, we explored the question of whether software-related patents in Europe are valued differently from other patents. This exercise was motivated by the growing number of software patents in the EPO, the debate over the patentability of Computer Implemented Inventions and the supposedly poor quality of 'software-related' patents due to their strategic nature.

As far as total patents are concerned, our results demonstrate clearly that the financial markets primarily value those patented inventions for which patents are obtained in both European and US jurisdictions. Although EPO patents held by European firms are valued somewhat more highly than USPTO patents held by the same firms, this result is entirely accounted for by the fact that USPTO patents are slightly more numerous, so that the elasticity of market value with respect to patenting of either type is the same. Compared to USPTO patents held by US firms, patents of either type held by European firms have a slightly greater impact on value than those held by U.S. ICT firms during a similar time period (an elasticity of 0.035 versus about 0.016 reported by Hall and MacGarvie 2006) or those held by all US firms during the 1980s (0.02 reported by Hall, Jaffe, and Trajtenberg 2005).

Although quality adjusting these patents is significant, using either forward citations or an index based on forward citations, family size, and number of IPC classes, it adds only about 0.1 per cent to the explanatory power of the regression. It is also noteworthy that forward citations do almost as well as the 3-component factor index for EPO patents. One reason for this may be that by including patents with equivalents separately in the regression we have already captured much of the information associated with family size. That is, the taking out of a patent at both the EPO and the USPTO is a good enough indicator that it is more valuable than other patents. However, for US patents, the factor index provides more information than citation weights alone.

The insignificant share of software firms in software patenting suggests that most software firms in Europe are not using patents to protect their inventions. It is also true that the very small number of EPO patents we obtain when using a definition designed to capture "pure" software patents suggests that the EPO has been successful in excluding such patenting. Nevertheless, patents identified as software-related in general are more valuable that other patents, whether taken out at the EPO or at the USPTO. More interestingly, the quality-weighted software patents are no more valuable than other patents, suggesting that the value of these patents derives from their numbers rather than the quality of the inventions that they cover.

The present paper is a first investigation of the EPO patent dataset based only on European firms from a large set of countries. In future research we will try to correct for some limitations of the dataset. First, we want to extend the analysis to firms of non-European countries such as the United States. Second, we will control for differences between citations to patents held by other firms and self-citations. Although we have included self-citations, we do not expect significant changes in our results from this exploration. Previous work on US data by Hall *et al.* (2005) and Hall and MacGarvie (2006) have found that removing self-citations yields real but limited changes in the impact of citation-adjusted patents on the firm's market value.

Finally, we will control for changes in corporate structure. The results presented in this paper rely on the pooled ownership links of the firms mainly in the period from 1998 to 2005, which was used to match the name of patent assignees in the EPO database with that of companies in Amadeus. Therefore in earlier years our patent variables may include more or fewer patents than are actually owned by the firm, which introduces an unknown source of bias. Moreover, we do not take into account the date of a merger, acquisition or the entry of a subsidiary firm. We recognize that this is a potential source of bias because expectations about future firms' performance (our dependent variable) may be correlated with future acquisitions of patents, implying that the patent variable proxies for growth expectations in some cases.

Bibliography

Arundel, A. (2001), "The Relative Effectiveness of Patents and Secrecy for Appropriation." *Research Policy* 30: 611-624.

Arundel, A. (2003), *Patents in the Knowledge-Based Economy*, Report of the KNOW Survey, MERIT, University of Maastricht.

Bacchiocchi, E. and F. Montobbio (2004), "EPO vs. USPTO Citations Lags." in *CESPRI Working Papers*.

Bergstra, J. A., and P. Klint (2007), "How to find a software patent?" Informatics Institute, University of Amsterdam: manuscript.

Bessen, J and Hunt, R. M. (2004), "An Empirical Look at Software Patents," Boston University, Working Paper: http://www.researchoninnovation.org/swpat.pdf

Bloom, N., and J. Van Reenen (2002), "Patents, real options and firm performance." *Economic Journal* 112: C97-C116.

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 (1995), "Market Share, Market Value, and Innovation in a Panel of British Manufacturing Firms." *University College London Discussion Paper 95/19*, London, October.

Bureau van Dijk (2005) Amadeus – Company Directory Database, available also at www.amadeus.bvdep.com

DaDalt, P.J., Donaldson, J.R. and Garner J.L. (2003) "Will Any q Do?", *Journal of Financial Research*, 4: 535-551.

Datastream (2006) Thomson Financial, www.datastream.com.

Gambardella, A., D. Harhoff, and B. Verspagen (2005), "The Value of Patents." Universita Bocconi, Ludwig-Maximiliens Universitaet, and Eindhoven University, Working Paper: http://www.creiweb.org/activities/sc_conferences/23/papers/gambardella.pdf

Gold, B. (1977), *Research, Technological Change and Economic Analysis*, Lexington Books, Lexington, Mass.

Graham, S. J. H. and D. C. Mowery (2003), "Intellectual Property Protection in the U. S. Software Industry," in W. M. Cohen and S. A. Merrill (eds.), *Patents in the Knowledge-Based Economy*, Washington, DC: National Academies Press.

Greenhalgh, C., and M. Rogers (2006), "The value of innovation: The interaction of competition, R&D and IP." *Research Policy* 35: 562-580.

Griliches, Z. (1979), "Issues in Assessing the Contribution of Research and Development to Productivity Growth." *Bell Journal of Economics* 10(1): 92-116.

Griliches, Z. (1980), "Returns to Research and Development Expenditures in the Private Sector." In Kendrick, J. W., and B. N. Vaccara (eds.), *New Developments in Productivity Measurement and Analysis*, Chicago: University of Chicago Press, 419-454.

Griliches, Z. (1981), “Market Value, R&D and Patents.” *Economic Letters* 7: 183-87.

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

Griliches, Z., Hall, H. B. and Pakes, A. (1991), “R&D, Patents. And Market Value Revisited: Is There a Second (Technological Opportunity) Factor?.” *Economics of Innovation and New Technology* 1: 183-202.

Guellec, D. and B. Van Pottelsberghe de la Potterie (2000), “Applications, Grants, and the Value of Patent,” *Economic Letters* 69 (1): 109-114.

Hall, B. H. (1993), The Stock Market Valuation of R&D Investment during the 1980s.” *American Economic Review* 83: 259-64.

Hall, B. H. (2000), “Innovation and Market Value.” In R. Barrell, G. Mason, and M. O’Mahoney (eds.), *Productivity, Innovation and Economic Performance*, Cambridge: Cambridge University Press.

Hall, B. H. (2006), “R&D, Productivity, and Market Value,” revised version of paper presented at the International Conference in memory of Zvi Griliches, Paris, August 2003. UC Berkeley and NBER.

Hall, B. H., and M. MacGarvie (2006), “The Private Value of Software Patents,” Cambridge, Mass.: National Bureau of Economic Research Working Paper 12195 (April).

Hall, B. H., and Mairesse, J. (1995), “Exploring the Relationship between R&D and Productivity in French Manufacturing Firms,” *Journal of Econometrics* 65: 263-94.

Hall, B. H., A. Jaffe, and Trajtenberg, M. (1998), “The Economic Significance of Patent Citations.” UC Berkeley, Nuffield College, Brandeis University, University of Tel Aviv, mimeo.

Hall, B. H., A. Jaffe, and M. Trajtenberg (2001), “The NBER Patent Citations Data File: Lessons, Insights, and Methodological Tools.” In A. Jaffe and M. Trajtenberg (eds.), *Patents, Citations and Innovations,* Cambridge, MA: The MIT Press. Also Cambridge, Mass.: National Bureau of Economic Research Working Paper 8498 (October).

Hall B. H., A. Jaffe, and M. Trajtenberg (2005), “Market Value and Patent Citations,” *Rand Journal of Economics* 36: 16-38.

Hall, B. H., and R. Oriani (2006), “Does the market value R&D investment by European firms? Evidence from a panel of manufacturing firms in France, Germany.” *International Journal of Industrial Organization* 24: 971-993.

Hansen, L. P. (1982), “Large Sample Properties of Generalised Methods of Moments Estimators.” *Econometrica* 50: 1029-1054.

Harhoff, D., K. Hoisl, and C. Webb (2006), “European patent citations – How to count and how to interpret them?” University of Munich and CEPR (London), University of Munich, and OECD.

Harhoff, D., F. Narin, and K. Vopel (1999), “Citation Frequency and the Value of Patented Inventions.” *Review of Economics and Statistics* 81(3): 511-15.

Hoover’s (2005), *Hoover’s Company Directo*ry, www.hoovers.com.

IMF (2006), *Global Financial Stability Report. Market Development and Issues*, International Monetary Fund, Washington DC, September www.imf.org / External / Pubs/FT/GFSR/2006/02/index.htm.

IMF (2007), *Global Financial Stability Report. Market Development and Issues*, International Monetary Fund, Washington DC, September www.imf.org / External / Pubs/FT/GFSR/2007/01/index.htm.

Jaffe, A., Trajtenberg, M. and Fogarty, M. (2000), "The Meaning of Patent Citations: Report of the NBER/Case Western Reserve Survey of Patentees." Working Paper No. 7631, NBER. Reprinted as Chapter 12 in Jaffe and Trajtenberg (2002), *Patents, Citations and Innovations: A Window on the Knowledge Economy*, Cambridge: MIT Press,.

Jolliffe, I T. (2002), *Principal Component Analysis*, Second Edition. New York: Springer-Verlag.

Lanjouw, J. O., and M. Schankerman (2004), "Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators." *Economic Journal* 114: 441-465.

Lanjouw, J. O., A. Pakes, and J. Putnam (1998), "How to Count Patents and Value Intellectual Property: The Uses of Patent Renewal and Application Data," *Journal of Industrial Economics* 46 (4): 405-432.

Layne-Farrar, A. (2005). "Defining Software Patents: A Research Field Guide," LECG, Chicago: manuscript.

Lerner, J. (1994), "The Importance of Patent Scope: An Empirical Analysis," *Rand Journal of Economics* 25 (2): 319-33.

Levine, R. and Zervos, S. (1998), "Stock Markets, Banks, and Economic Growth." *American Economic Review* 88(3): 537-58.

Lindenberg, E.B. and S. A. Ross (1981), "Tobin's q Ratio and Industrial Organization." *Journal of Business* 54: 1-32.

Mairesse, J. and M. Sassenou (1991), "R&D and productivity: a survey of econometric studies at the firm level." *STI Review* (OECD) 8: 9-46.

Mansfield, E. (1968), *Industrial Research and Technological Innovation. An Econometric Analysis*. Norton, New York.

Montgomery, C.A. and B. Wernerfelt (1988), "Diversification, Ricardian Rents, and Tobin's q." *Rand Journal of Economics* 19(4): 623-32.

Nelson, R.R. (1959), "The Simple Economics of Basic Scientific Research." *Journal of Political Economy* 67(June): 297-306.

Norsworthy, J. R. and S. L. Jang (1992), *Empirical Measurement and Analysis of Productivity and Technological Change. Applications in High-Technology and Services Industries*, North Holland, Amsterdam.

PATSTAT (2006), *EPO Worldwide Patent Statistical Database* – Available under license from OECD-EPO Task Force on Patent Statistics.

Perfect, S. and Wiles, K. (1994), "Alternative construction of Tobin's q: an empirical comparison," *Journal of Empirical Finance*, 1: 313-341.

Quillen, C. D., O. H. Webster, and R. Eichmann (2002), "Continuing Patent Applications and Performance of the U.S. Patent and Trademark Office – Extended," *The Federal Circuit Bar Journal* 12(1): 35-55

Rajan, R. G. and Zingales, L. (2003) "Banks and Markets: The Changing Character of European Finance", *CEPR discussion paper series financial economics*,3865, May, www.cepr.org/pubs/dps/DP3865.asp and http://ssrn.com/abstract=419440.

Schumpeter, J. (1934), *The Theory of Economic Development,* Harvard University Press, Cambridge, Mass.

Teece, D.J. (1986), "Profiting from Technological Innovation: Implications for Integration, Collaboration, Licensing and Public Policy." *Research Policy* 15: 285–305.

Thoma G. and Torrisi S. (2005), "Evolution of Software Industry in Europe," *Study of the effects of allowing patents claims for computer-implemented inventions*, Brussels: EU DG Commission Research Report.

Toivanen, O., P. Stoneman, and D. Bosworth (2002), "Innovation and the market value of UK firms, 1989-1995." *Oxford Bulletin Of Economics And Statistics* 64: 39-___.

Wildasin, D. E. (1984), "The q Theory of Investment with Many Capital Goods." *American Economic Review* 74(1): 203-10.

WIPO (2007), *Patent statistics database*, available at http://www.wipo.int/ipstats/en/statistics/patents/

Appendix A – Sample description

[Table A.1 about here]

[Table A.2 about here]

[Table A.3 about here]

[Table A.4 about here]

[Table A.5 about here]

Appendix B – Correcting for citation truncation

The HJT method to identify the random process generating citations is based on the estimation of a semi-structural model which is made of two equations. With the first equation the citation frequency is modelled as a multiplicative function of cited-year effects (s), citing-year (t) effects, technology field (k) effects and citation lag effects (Hall *et al.*, 2001). The equation can be written as follows:

$$C_{kst} / P_{ks} = \alpha_0 \alpha_s \alpha_t \alpha_k \exp[f_k(L)]$$

where C_{kst} is the total number of citations received by patents with application date s and in technology k from patents with application date t. P_{ks} is the number of patents in technology k, year s. C_{kst} / P_{ks} is then the average number of citations received by patents k-s by all patents in year t. The parameters α_s, α_t, α_k measure the effect of, respectively, cited-year, citing-year and technology on the probability of citations. The function $f_k(L)$ describes the shape of the citation-lag (L=t-s) distribution, which is allowed to vary across fields. The multiplicative form of the citation frequency relies on the assumption of proportionality, i.e., the shape of the lag distribution is assumed to be independent of the number of citations received.

The α parameters are normalized so that each parameter measures the proportional difference in the citation propensity with respect to the base category. For instance, an estimated coefficient α_k= (k=chemicals field) = 2 implies that the expected citation rate of patents in the chemical field is twice the citation rate of patents in the base field.

The second equation in the model is the following:

$$f_k(L) = \exp(-\beta_{1k} L)(1 - \exp(-\beta_{2k} L)$$

where the parameters β_{1k} and β_{2k} measure the depreciation or obsolescence of the knowledge protected by patents in field k and the diffusion effect, respectively.

Following Hall et al (2001), we estimated this model by non linear least squares. Estimated α parameters can be used to remove cited-patent, citing-patent and technology field effects. Since we are primarily interested in truncation, we used the estimates of β parameters to calculate the expected distribution lags. Table B.1 reports the cumulative citation lag distributions in the seven technological groups defined by Fraunhofer-ISI and the Observatoire des Sciences et des Techniques over the cited period 1978-2004.[33] We used these proportions to correct the observed citation counts. Consider, for example, a chemical patent in year 2002 which has received 5 citations until 2005. Table B.1 shows that the typical chemical patent in year 2002 receives about 48.2% of citations after three years from its application. To correct for truncation we have to 'deflate' the observed citations by 0.48183 obtaining 10.38 citations.

The weights reported in Table B.1 and Table B.2. are obtained by using all citations respectively to EPO and USPTO by year of cited patents, year of citing patents, citation lag and technological field of the cited patent. The source of data is PATSTAT (2006), which reports citations received by EPO and the USPTO patents from the main world patent offices, including the USPTO, the JPTO and the WIPO. The weights reported above have been estimated separately for EP and US patents without conditioning on the patent office from which they originate from, and we have used the same weights to correct all citations received by the patents in our sample, assuming that the shape of the simulated cumulative lag distributions does not vary with the citing patent's office.

Unfortunately, the EPO system does not require examiners to indicate the 'main' technological field. The PATSTAT Data Catalogue_3_22 states that '…For other authorities, like the EPO, there is in general no meaning in the position – classes may be quoted in alphabetical order for instance …" p. 50). The problem is serious since many patents are classified in two or more 2-digit IPC fields. In this case, we used the arithmetic mean of the citation lag distribution weighted by the patent's own IPC distribution (e.g., if it has 3 chemical classes and one drug, we used 3/4 the chemical cite lag and 1/4 the drug cite lag; given the

[33] http://www.obs-ost.fr

similarity of the lag distributions, this procedure is not likely to introduce much error into the measure.

[Table.B.1 about here]

[Table.B.2 about here]

Appendix C – A Composite Patent Quality Indicator

The construction of the multidimensional measure of patent quality relies on factor analysis. In factor models each series of data (quality indicator in our case) is decomposed into a common component and an idiosyncratic component. The common component is only driven by a few common shocks, denoted by V< N, where N is the number of indicators. In a static factor model, the common shocks affect the indicators only contemporaneously. The basic model is given by $X = UB + E = K + E$, where X is the $(T \times N)$ matrix of observations on N series (indicators) of length T. The series are normalized to have mean 0 and variance 1. U is the $(T \times V)$ matrix of V common shocks and B is the $(V \times N)$ matrix of factor loadings, which determines the impact of common shock v on series n. The common shocks and the factor loadings together make up the common component K. After the influence of common shocks has been removed, only the idiosyncratic component (E) remains. To estimate the common component we have to find a linear combination of the indicators in X that explains as much as possible the total variance of each indicator, minimizing the idiosyncratic component (for a technical discussion of factor models see Jolliffe (2002).

The parallel with least squares estimation is clear from this formulation, but the fact that the common shocks are unobserved complicates the problem. The standard way to extract the common component in the static case is to use principal component analysis. In principal component analysis the first V eigenvalues and eigenvectors are calculated from the variance-covariance matrix of the dataset X. The common component is then defined as $K = XVV'$, with $V = [p_1,\ldots,p_V]$ and where p_i is the eigenvector corresponding to the ith largest $(i = 1 \ldots Q)$ eigenvalue of the covariance matrix of X. This method does not guarantee a unique solution. A further problem is that *ex ante* it is not known how many common shocks V affect the series in X. Following the approach suggested by Lanjouw and Schankerman (2004), we use a multiple-indicator model with an unobserved common factor:

$$y_{k\,i} = \lambda_k v_i + \beta' X + e_{ki}$$

where y_{ki} indicates the value of the kth patent indicator for the ith patent; v is the common factor with factor loadings λ_k and normally distributed, while X is a set of controls. The main underlining assumption is that the variability of each patent indicator in the sample may be generated by the variability of a common factor across all the indicators and an idiosyncratic component $e_{k\sim} N(0,\sigma^2{}_k)$ which is not related to other 'quality' indicators.

In our setting, the common factor is the unobserved characteristic of a patent that influences positively three 'quality' indicators: family size, forward citations, and the number of 8-digit IPC technology fields. The analysis is based on the total number of EPO patents granted between 1978 and 2002 (around 785,740 observations) and of US patents granted between 1978 and 2002 (around 2,756,353 observations).

More precisely, to estimate v we followed a two step estimation procedure. In the first step we regressed the three patent 'quality' indicators against two observable patent characteristics, the year of application and the main technology class of the patent (out of 30 macro-technological classes) using three stage least squares. Estimation of the common quality index v is then based on information extrapolated from the covariance matrix of three observable indicators conditional on year and technology class. In the second step we used maximum likelihood to estimate a factor model using the residuals from the first step under the assumption that $v \sim N(0, \sigma^2)$. We found evidence of the existence of a single common factor which we used as our multidimensional measure of patent 'quality' in the market value estimations. Factor analysis in the second step yields the following factor loadings:

Variable	**EPO patents**	**USPTO patents**
Forward citations	0.289	0.173
Family size	0.301	0.106
Number of IPC classes	0.170	0.334

Table 1
Empirical studies of the market value of innovation using European data

Paper	R&D	Innovation output	Patent citations	Sample size	Geographical coverage	Time period
Blundell *et al.* (1999)	NO	USPTO patents, SPRU innovation counts	NO	340	UK	1972-1982
Bloom and Van Reenen (2002)	NO	USPTO patents	5-year cite stock	404	UK	1968-1996
Toivanen *et al.* (2002)	YES	NO	NO	1519	UK	1988-1995
Greenhalgh and Rogers (2006)	YES	UK and EPO patents	NO	3227	UK	1989-2002
Hall and Oriani (2006)	YES	NO	NO	2156	US, UK, FR, IT, DE	1989-1998
Our study	YES	USPT and EPO patents	Yes	7168	21 European countries	1991-2002

Table 2
Country-size distribution of R&D-reporting firms in our sample

Average sales (euros in 2000)	*< 10M*	*10M-100M*	*100M-1B*	*1B-10B*	*> 10B*	*Total*	*Market cap/GDP**
Austria	1	4	6	3	0	14	0.41
Belgium & Luxembourg	0	9	9	5	1	24	0.77
Switzerland	1	11	39	18	3	72	NA
Germany	8	73	67	35	11	194	0.43
Denmark	2	9	9	4	0	24	0.62
Eastern Europe	0	2	7	1	0	10	0.11
Spain & Greece	0	8	21	3	0	32	0.82
Finland	1	26	21	13	2	63	0.94
France	9	43	44	23	13	132	0.73
UK	51	123	100	55	14	343	1.73
Ireland	2	2	6	2	0	12	0.45
Italy	0	0	1	1	1	3	0.45
Netherlands	2	10	10	9	4	35	0.81
Norway	2	8	8	3	2	23	NA
Sweden	13	28	21	15	3	80	0.97
Totals	92	356	369	190	54	1061	

This variable is the total stock market capitalization for the country over GDP (source: IMF 2006)

Table 3a
Descriptive Statistics

5312 observations, 1061 firms, 15 country/regions, 1991-2002

	Number**	Mean	S.D.	Median	1Q	3Q	Min	Max
Sales*	5312	3749.8	11996.8	306.3	66.9	1950.6	0.0	194,724
Tobin's q	5312	2.99	3.54	1.71	1.14	3.18	0.10	24.85
Employment	4729	16864	47119	1423	298	9600	1	477,100
R&D expenditures*	5312	129.32	485.91	8.11	1.87	36.62	0.000	6,787
R&D stock*	5312	637.45	2396.47	35.16	8.44	183.46	0.01	33,127
R&D stock/assets	5312	0.51	0.74	0.25	0.09	0.59	0.000	4.99

*In millions of current euros

**The number of good observations.

Table 3b
Descriptive statistics for patent variables

5312 observations, 1061 firms, 15 country/regions, 1991-2002

	N nonzero	Mean	S.D.	Median	1Q	3Q
Granted patents by application date						
EPO	3980	13.22	64.04	0	0	4
EPO with US equivalents	3758	9.94	48.40	0	0	3
EPO only	3309	3.28	18.46	0	0	1
USPTO	4253	25.80	134.46	1	0	7
USPTO with European equivalents	4020	12.96	70.41	0	0	3
USPTO only	3383	12.85	81.91	0	0	3
Granted software patents by application date						
EPO	277	0.05	0.58	0	0	0
EPO with US equivalents	205	0.04	0.46	0	0	0
EPO only	150	0.01	0.21	0	0	0
USPTO	2393	1.73	13.35	0	0	0
USPTO with European equivalents	1925	0.68	5.90	0	0	0
USPTO only	1732	1.05	9.30	0	0	0
Stock of granted patents						
EPO	3980	81.67	365.23	3.53	0.00	26.02
EPO with US equivalents	3758	61.75	267.71	2.44	0.00	17.29
USPTO	4253	132.40	607.71	5.89	0.38	39.30
USPTO with European equivalents	4020	74.54	361.60	3.10	0.07	21.21
Stock of granted software patents						
EPO	277	0.25	2.38	0.00	0.00	0.00
EPO with US equivalents	205	0.19	1.72	0.00	0.00	0.00
USPTO	2393	9.56	57.84	0.00	0.00	1.73
USPTO with European equivalents	1925	4.28	30.19	0.00	0.00	0.72
Patent-R&D stock ratios						
EPO	3980	0.30	0.87	0.08	0.00	0.28
EPO with US equivalents	3758	0.21	0.65	0.05	0.00	0.20
USPTO	4253	0.44	1.36	0.15	0.02	0.46
USPTO with European equivalents	4020	0.26	0.75	0.07	0.00	0.25
Patent-R&D stock ratios - software patents						
EPO	277	0.001	0.016	0.000	0.000	0.000
EPO with US equivalents	205	0.000	0.015	0.000	0.000	0.000
USPTO	2393	0.023	0.084	0.000	0.000	0.012
USPTO with European equivalents	1925	0.010	0.047	0.000	0.000	0.003
Citation-patent stock ratios						
EPO	3809	2.92	3.33	2.47	0.00	4.06
EPO with US equivalents	3599	3.10	3.48	2.64	0.00	4.49
USPTO	4141	12.57	15.25	10.00	2.91	16.17
USPTO with European equivalents	3909	14.59	23.94	9.50	0.00	16.69
Citation-patent stock ratios - software patents						
EPO	234	0.21	1.31	0.00	0.00	0.00
EPO with US equivalents	194	0.19	1.30	0.00	0.00	0.00
USPTO	2298	12.80	31.91	0.00	0.00	16.92
USPTO with European equivalents	1861	13.94	49.98	0.00	0.00	12.84

*In millions of current euros

Table 4
Market value regressions with patent stocks

5312 observations for the 1991-2002 period. Dependent variable = log Tobin's Q

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
R&D stock-assets ratio	**0.675 (0.061)**	**0.728 (0.067)**	**0.782 (0.071)**	**0.763 (0.072)**	**0.753 (0.070)**	**0.748 (0.071)**	**0.788 (0.074)**
Patent stock-R&D ratios:							
European		**0.157 (0.034)**					
European only				0.014 (0.030)		0.026 (0.032)	0.007 (0.034)
US			**0.177 (0.031)**				
US only					**0.121 (0.045)**	**0.100 (0.047)**	**0.124 (0.046)**
European with US equivalents				**0.330 (0.066)**		**0.274 (0.065)**	
US with European equivalents					**0.191 (0.042)**		**0.211 (0.049)**
Dummies for zero patent stocks							
European		-0.016 (0.037)					
European only				**0.188 (0.044)**		**0.202 (0.043)**	**0.145 (0.043)**
US			0.062 (0.041)				
US only					**-0.132 (0.034)**	**-0.114 (0.037)**	**-0.161 (0.035)**
European with US equivalents				**-0.101 (0.043)**		-0.047 (0.047)	
US with European equivalents					**0.160 (0.042)**		**0.112 (0.048)**
Log sales (millions of euros)	0.016 (0.005)	0.020 (0.006)	0.024 (0.006)	0.028 (0.006)	0.019 (0.006)	0.022 (0.007)	0.024 (0.007)
R-squared (s.e.)	0.255 (0.729)	0.263 (0.725)	0.266 (0.723)	0.269 (0.722)	0.269 (0.722)	0.271 (0.721)	0.271 (0.721)
Average elasticity (standard deviation)							
R&D stock-assets ratio	**0.194 (0.011)**	**0.201 (0.012)**	**0.205 (0.012)**	**0.199 (0.012)**	**0.203 (0.012)**	**0.198 (0.012)**	**0.205 (0.012)**
Patent stock-R&D ratios:							
European		**0.032 (0.006)**					
European only				0.001 (0.002)		0.002 (0.002)	0.001 (0.002)
US			**0.049 (0.007)**				
US only					**0.014 (0.005)**	**0.011 (0.005)**	**0.013 (0.005)**
European with US equivalents				**0.041 (0.007)**		**0.035 (0.007)**	
US with European equivalents					**0.033 (0.006)**		**0.035 (0.007)**

These regressions include 15 country dummies, 24 industry dummies, and 12 year dummies, as well as dummies for obs with zero patent stocks.

Nonlinear least squares with robust standard errors.

Table 5
Market value regressions with software patent stocks

5312 observations for the 1991-2002 period. Dependent variable = log Tobin's Q

Variable	(8)	(9)	(10)	(11)
R&D stock-assets ratio	**0.722 (0.068)**	**0.790 (0.072)**	**0.710 (0.068)**	**0.746 (0.070)**
Patent stock-R&D ratios:				
US only			**0.100 (0.052)**	**0.125 (0.052)**
European with US equivalents	**0.293 (0.059)**		**0.243 (0.057)**	
US with European equivalents		**0.206 (0.047)**		**0.158 (0.041)**
US only software			-0.003 (0.313)	-0.101 (0.298)
EP software with US equivalents	**2.59 (1.34)**		**2.54 (1.18)**	
US software with European equiv.		**1.57 (0.61)**		**1.50 (0.57)**
Log sales (millions of euros)	**0.020 (0.006)**	**0.028 (0.006)**	**0.015 (0.006)**	**0.018 (0.006)**
R-squared (s.e.)	0.266 (0.723)	0.265 (0.724)	0.268 (0.722)	0.270 (0.722)
	Average elasticity (standard deviation)			
K/A	**0.198 (0.012)**	**0.205 (0.012)**	**0.198 (0.012)**	**0.201 (0.012)**
Patent stock-R&D ratios:				
US only			**0.011 (0.006)**	**0.014 (0.005)**
European with US equivalents	**0.038 (0.006)**		**0.032 (0.006)**	
US with European equivalents		**0.034 (.007)**		**0.027 (0.006)**
US only software			0.0009 (0.0025)	0.0002 (0.0023)
EP software with US equivalents	**0.0004 (0.0001)**		**0.0004 (0.0001)**	
US software with European equiv.		**0.0096 (0.0029)**		**0.0092 (0.0028)**

These regressions include 15 country dummies, 24 industry dummies, and 12 year dummies, as well as dummies for obs with zero patent stocks.
Nonlinear least squares with robust standard errors.

Table 6
Market value regressions with patent stocks and patent value indicators
5312 observations for the 1991-2002 period. Dependent variable = log Tobin's Q

Variable	(12)	(13)	(14)	(15)
R&D stock-assets ratio	**0.806 (0.083)**	**0.807 (0.078)**	**0.854 (0.116)**	**0.852 (0.099)**
Patent stock-R&D ratios:				
European only	0.029 (0.037)	-0.002 (0.033)	0.034 (0.037)	0.000 (0.036)
US only	**0.109 (0.053)**	**0.105 (0.047)**	**0.113 (0.056)**	**0.081 (0.046)**
European with US equiv.	**0.300 (0.073)**		**0.308 (0.081)**	
US with European equiv.		**0.237 (0.053)**		**0.258 (0.060)**
Value indicator stock-patent ratios:	*Forward citations*		*Index*	
European only	-0.002 (0.007)	0.001 (0.007)	-0.088 (0.060)	-0.062 (0.056)
US only	0.000 (0.001)	0.000 (0.001)	-0.005 (0.050)	
European with US equiv.	**0.0320 (0.0080)**		**0.200 (0.084)**	
US with European equiv.		**0.0026 (0.0008)**		**0.155 (0.040)**
Log sales (millions of euros)	0.019 (0.007)	0.022 (0.007)	0.022 (0.007)	0.023 (0.007)
R-squared (s.e.)	0.274 (0.719)	0.272 (0.721)	0.272 (0.721)	0.273 (0.720)
	Average elasticity (standard deviation)			
R&D stock-assets ratio	**0.191 (0.012)**	**0.202 (0.012)**	**0.197 (0.012)**	**0.203 (0.012)**
Patent stock-R&D ratios:				
European only	0.002 (0.002)	0.000 (0.002)	0.002 (0.002)	0.000 (0.002)
US only	**0.011 (0.005**	**0.011 (0.005)**	**0.011 (0.005)**	**0.008 (0.004)**
European with US equiv.	**0.034 (0.007)**		**0.034 (0.007)**	
US with European equiv.		**0.038 (0.007)**		**0.039 (0.007)**
Value indicator stock-patent ratios:	*Forward citations*		*Index*	
European only	-0.002 (0.009)	0.002 (0.009)	-0.047 (0.034)	-0.035 (0.033)
US only	0.003 (0.005)	-0.004 (0.005)	-0.003 (0.030)	-0.007 (0.031)
European with US equiv.	**0.070 (0.015)**		**0.140 (0.048)**	
US with European equiv.		**0.028 (0.008)**		**0.102 (0.023)**

These regressions include 15 country dummies, 24 industry dummies, and 12 year dummies, as well as dummies for obs with zero patent stocks.

Nonlinear least squares with robust standard errors.

Table 7
Market value regressions
with patent stocks, software patent stocks, and patent value indicators

5312 observations for the 1991-2002 period. Dependent variable = log Tobin's Q

Variable	(17)	(19)
R&D stock-assets ratio	**0.774 (0.074)**	**0.799 (0.093)**
Patent stock-R&D ratios:		
US only	**0.113 (0.053)**	**0.094 (0.052)**
US with European equiv.	**0.185 (0.045)**	**0.196 (0.049)**
SW: US only	-0.149 (0.309)	-0.177 (0.309)
SW: US with European equiv.	**1.47 (0.59)**	**1.55 (0.64)**
Value indicator-patent ratios:	*Forward citations*	*Index*
US with European equiv.	**0.0026 (0.0008)**	**0.157 (0.038)**
SW: US with European equiv.	0.0000 (0.0002)	-0.040 (0.028)
Log sales (millions of euros)	**0.017 (0.006)**	**0.017 (0.007)**
R-squared (s.e.)	0.272 (0.721)	0.272 (0.720)
	Average elasticity (standard deviation)	
R&D stock-assets ratio	**0.199 (0.012)**	**0.199 (0.013)**
Patent stock-R&D ratios:		
US only	**0.012 (0.005)**	**0.010 (0.005)**
US with European equiv.	**0.030 (0.006)**	**0.031 (0.006)**
SW: US only	-0.0003 (0.0027)	-0.0007 (0.0023)
SW: US with European equiv.	**0.0087 (0.0028)**	**0.0090 (0.0029)**
Value indicator-patent ratios:	*Forward citations*	*Index*
US with European equiv.	**0.029 (0.009)**	**0.108 (0.023)**
SW: US with European equiv.	**0.030 (0.008)**	**0.091 (0.031)**

These regressions include 15 country dummies, 24 industry dummies, and 12 year dummies, as well as dummies for obs with zero patent stocks.
Nonlinear least squares with robust standard errors.

Table A.1. Distribution by year of incorporation and market capitalisation

Year of incorporation	Market Capitalisation (million mil EUR - latest year available)									
	<100		100-1000		1000-5000		> 5000		All	
before 1970	142	27.3%	135	42.2%	71	54.6%	49	53.8%	397	37.4%
1971-1980	41	7.9%	27	8.4%	10	7.7%	5	5.5%	83	7.8%
1981-1990	140	26.9%	63	19.7%	25	19.2%	9	9.9%	237	22.3%
1991-2000	183	35.2%	89	27.8%	20	15.4%	25	27.5%	317	29.9%
After 2000	11	2.1%	5	1.6%	1	0.8%	2	2.2%	19	1.8%
N.A.	3	0.6%	1	0.3%	3	2.3%	1	1.1%	8	0.8%
All	**520**	**49.0%**	**320**	**30.2%**	**130**	**12.3%**	**91**	**8.6%**	**1061**	**100.0%**

Table A.2. Distribution by stock market listing

Main exchange	Companies	Share (%)
Athens Stock Exchange	31	2.9%
Australian Stock Exchange	1	0.1%
Budapest Stock Exchange	5	0.5%
Dusseldorf Stock Exchange	1	0.1%
Euronext Amsterdam	24	2.3%
Euronext Brussels	22	2.1%
Euronext Paris	136	12.8%
Frankfurt Stock Exchange	93	8.8%
Hamburg Stock Exchange	1	0.1%
Helsinki Stock Exchange	1	0.1%
Irish Stock Exchange	10	0.9%
Italian Continuous Market	3	0.3%
London Stock Exchange (SEAQ)	154	14.5%
London Stock Exchange (SETS)	182	17.2%
Madrid Stock Exchange	1	0.1%
NASDAQ National Market	3	0.3%
NASDAQ OTC Bulletin Board	1	0.1%
New York Stock Exchange	3	0.3%
Not available	2	0.2%
OFEX	1	0.1%
OMX - Copenhagen Stock Exchange	23	2.2%
OMX - Helsinki Stock Exchange	62	5.8%
OMX - Stockholm Stock Exchange	80	7.5%
OMX - Tallinn Stock Exchange	1	0.1%
Oslo Stock Exchange	24	2.3%
Stuttgart Stock Exchange	3	0.3%
Swiss Electronic Stock Exchange	13	1.2%
Swiss Exchange	57	5.4%
Vienna Stock Exchange	13	1.2%
Warsaw Stock Exchange	2	0.2%
XETRA	106	10.0%
Zagreb Stock Exchange	2	0.2%
Total	1061	100.0%

Table A.3. Distribution of companies by industry – 2.5 digit industry class

2.5 digit industry class	with R&D		with EP pats		with EP software pats		with US pats		with US software pats	
	firms	%	firms	%	firms	%	firms	%	firms	%
01 Food & tobacco	39	3.7	31	4.5	0	0.0	28	3.8	18	4.6
02 Textiles, apparel & footwear	20	1.9	12	1.7	1	2.3	11	1.5	4	1.0
03 Lumber & wood products	7	0.7	3	0.4	0	0.0	4	0.5	2	0.5
04 Furniture	10	0.9	8	1.2	1	2.3	8	1.1	4	1.0
05 Paper & paper products	17	1.6	13	1.9	0	0.0	15	2.0	6	1.5
06 Printing & publishing	14	1.3	6	0.9	1	2.3	7	0.9	3	0.8
07 Chemical products	46	4.3	40	5.8	1	2.3	41	5.5	21	5.4
08 Petroleum refining & prods	20	1.9	16	2.3	0	0.0	17	2.3	12	3.1
09 Plastics & rubber prods	17	1.6	12	1.7	1	2.3	13	1.8	7	1.8
10 Stone, clay & glass	22	2.1	18	2.6	0	0.0	16	2.2	8	2.1
11 Primary metal products	24	2.3	15	2.2	0	0.0	15	2.0	5	1.3
12 Fabricated metal products	28	2.6	21	3.0	1	2.3	22	3.0	10	2.6
13 Machinery & engines	89	8.4	76	10.9	0	0.0	75	10.1	44	11.3
14 Computers & comp, equip,	29	2.7	20	2.9	3	6.8	22	3.0	16	4.1
15 Electrical machinery	39	3.7	30	4.3	3	6.8	32	4.3	17	4.4
16 Electronic inst, & comm, eq,	127	12.0	82	11.8	7	15.9	90	12.1	50	12.8
17 Transportation equipment	10	0.9	9	1.3	1	2.3	9	1.2	8	2.1
18 Motor vehicles	25	2.4	22	3.2	3	6.8	22	3.0	12	3.1
19 Optical & medical instruments	41	3.9	32	4.6	2	4.5	34	4.6	17	4.4
20 Pharmaceuticals	61	5.8	47	6.8	2	4.5	49	6.6	26	6.7
21 Misc, manufacturing	23	2.2	15	2.2	0	0.0	17	2.3	8	2.1
22 Soap & toiletries	11	1.0	10	1.4	0	0.0	11	1.5	6	1.5
24 Computing software	159	15.0	42	6.0	8	18.2	65	8.8	30	7.7
25 Telecommunications	21	2.0	9	1.3	5	11.4	9	1.2	7	1.8
26 Wholesale trade	26	2.5	14	2.0	0	0.0	16	2.2	5	1.3
27 Business services	16	1.5	9	1.3	1	2.3	10	1.3	6	1.5
29 Mining	13	1.2	11	1.6	0	0.0	11	1.5	6	1.5
30 Construction	19	1.8	12	1.7	0	0.0	11	1.5	4	1.0
31 Transportation services	6	0.6	4	0.6	1	2.3	4	0.5	3	0.8
32 Utilities	21	2.0	20	2.9	1	2.3	18	2.4	11	2.8
33 Trade	7	0.7	2	0.3	0	0.0	2	0.3	0	0.0
34 Fire, Insurance, Real Estate	2	0.2	0	0.0	0	0.0	0	0.0	0	0.0
35 Health services	4	0.4	2	0.3	0	0.0	2	0.3	2	0.5
36 Engineering services	38	3.6	26	3.7	1	2.3	28	3.8	9	2.3
37 Other services	10	0.9	6	0.9	0	0.0	7	0.9	3	0.8
Total	1061	100.0	695	100.0	44	100.0	741	100.0	390	100.0

Table A.4. Distribution of R&D, patents and software patents by industry

2.5 digit industry classes (1060 firms)

2.5 digit industry class	R&D		EP patents		EP software patents		US patents		US software patents	
	Mil EUR	%	n	%	n	%	n	%	n	%
01 Food & tobacco	24875	3.5	1752	2.5	0	0.0	3534	2.6	103	1.1
02 Textiles, apparel & footwear	658	0.1	103	0.1	0	0.0	175	0.1	4	0.0
03 Lumber & wood products	52	0.0	1	0.0	0	0.0	3	0.0	0	0.0
04 Furniture	2402	0.3	162	0.2	1	0.4	264	0.2	2	0.0
05 Paper & paper products	2170	0.3	444	0.6	0	0.0	477	0.3	12	0.1
06 Printing & publishing	1243	0.2	4	0.0	0	0.0	18	0.0	9	0.1
07 Chemical products	73977	10.5	10964	15.6	0	0.0	15701	11.5	262	2.8
08 Petroleum refining & prods	25109	3.6	1822	2.6	0	0.0	3610	2.6	165	1.8
09 Plastics & rubber prods	5515	0.8	1291	1.8	1	0.4	1115	0.8	54	0.6
10 Stone, clay & glass	5215	0.7	1627	2.3	0	0.0	2137	1.6	79	0.9
11 Primary metal products	2594	0.4	349	0.5	0	0.0	765	0.6	13	0.1
12 Fabricated metal products	2134	0.3	790	1.1	0	0.0	1824	1.3	44	0.5
13 Machinery & engines	16343	2.3	2984	4.2	0	0.0	4707	3.4	137	1.5
14 Computers & comp, equip,	3185	0.5	171	0.2	1	0.4	700	0.5	181	2.0
15 Electrical machinery	93255	13.2	19372	27.6	139	51.3	35529	25.9	2930	31.8
16 Electronic inst, & comm, eq,	93435	13.3	6605	9.4	33	12.2	26183	19.1	2657	28.8
17 Transportation equipment	22424	3.2	409	0.6	1	0.4	579	0.4	33	0.4
18 Motor vehicles	145932	20.7	8922	12.7	3	1.1	17272	12.6	1130	12.3
19 Optical & medical instruments	5580	0.8	670	1.0	17	6.3	1316	1.0	185	2.0
20 Pharmaceuticals	116961	16.6	4852	6.9	2	0.7	11501	8.4	404	4.4
21 Misc, manufacturing	1503	0.2	108	0.2	0	0.0	275	0.2	4	0.0
22 Soap & toiletries	8972	1.3	2532	3.6	0	0.0	2722	2.0	42	0.5
24 Computing software	9645	1.4	205	0.3	7	2.6	474	0.3	183	2.0
25 Telecommunications	16885	2.4	1089	1.6	65	24.0	1524	1.1	337	3.7
26 Wholesale trade	493	0.1	16	0.0	0	0.0	21	0.0	3	0.0
27 Business services	3701	0.5	28	0.0	1	0.4	94	0.1	18	0.2
29 Mining	1894	0.3	375	0.5	0	0.0	1012	0.7	14	0.2
30 Construction	2645	0.4	116	0.2	0	0.0	140	0.1	10	0.1
31 Transportation services	3697	0.5	1473	2.1	0	0.0	1690	1.2	46	0.5
32 Utilities	8445	1.2	719	1.0	0	0.0	1273	0.9	125	1.4
33 Trade	178	0.0	0	0.0	0	0.0	0	0.0	0	0.0
34 Fire, Insurance, Real Estate	10	0.0	0	0.0	0	0.0	0	0.0	0	0.0
35 Health services	217	0.0	54	0.1	0	0.0	60	0.0	2	0.0
36 Engineering services	2192	0.3	181	0.3	0	0.0	289	0.2	12	0.1
37 Other services	287	0.0	27	0.0	0	0.0	88	0.1	13	0.1
Overall	703823	100.0	70217	100.0	271	100.0	1E+05	100.0	9213	100.0

*This is the total over all years of the sample, in constant year 2000 euros.

Table A.5. Distribution of R&D expenditures by country and sector

		R&D expenditure in millions of euros						As a share of total expenditure				HTT sample relative to	
Country	**Year**	**Business Sector**	**Govt Sector**	**HEI Sector**	**Other**	**Total R&D**	**HTT sample**	**Business Sector**	**Govt Sector**	**HEI Sector**	**Other**	**Business sector**	**Total R&D**
Austria	2002	3131	266	1266	21	4684	65.1	66.8%	5.7%	27.0%	0.4%	2.1%	1.4%
Belgium	2000	3589	312	1005	58	4964	907.7	72.3%	6.3%	20.2%	1.2%	25.3%	18.3%
Bulgaria	2000	15	49	7	0	71	0.0	21.4%	68.6%	9.8%	0.2%	0.0%	0.0%
Switzerland	2000	5065	90	1566	132	6852	8794.5	73.9%	1.3%	22.9%	1.9%	173.7%	128.3%
Cyprus	2000	5	11	6	2	25	0.0	21.3%	46.6%	24.8%	7.3%	0.0%	0.0%
Czech Rep.	2000	446	188	106	4	744	0.0	60.0%	25.3%	14.2%	0.5%	0.0%	0.0%
Germany	2000	35600	6873	8146	0	50619	28094.8	70.3%	13.6%	16.1%	0.0%	78.9%	55.5%
Denmark	2000	2596	492	770	34	3892	979.8	66.7%	12.6%	19.8%	0.9%	37.7%	25.2%
Estonia	2000	8	9	19	1	37	1.0	22.5%	23.1%	52.4%	1.9%	11.4%	2.6%
Spain	2000	3069	905	1694	51	5719	0.0	53.7%	15.8%	29.6%	0.9%	0.0%	0.0%
Finland	2000	3136	468	789	30	4423	731.3	70.9%	10.6%	17.8%	0.7%	23.3%	16.5%
France	2000	19348	5361	5804	439	30954	14557.8	62.5%	17.3%	18.8%	1.4%	75.2%	47.0%
Greece	2001	278	188	383	3	852	55.1	32.7%	22.1%	44.9%	0.4%	19.8%	6.5%
Croatia	2002	115	60	95	0	271	50.5	42.7%	22.2%	35.1%	0.0%	43.8%	18.7%
Hungary	2000	180	106	97	23	405	35.4	44.3%	26.1%	24.0%	5.6%	19.7%	8.7%
Ireland	2000	842	96	238	0	1176	403.8	71.6%	8.1%	20.2%	0.0%	47.9%	34.3%
Iceland	2000	142	64	41	5	251	0.0	56.4%	25.5%	16.2%	1.9%	0.0%	0.0%
Italy	2000	6239	2356	3865	0	12460	37.0	50.1%	18.9%	31.0%	0.0%	0.6%	0.3%
Lithuania	2000	16	31	27	0	73	0.0	21.5%	41.9%	36.5%	0.0%	0.0%	0.0%
Luxembourg	2000	337	26	1	0	364	1.7	92.6%	7.1%	0.2%	0.0%	0.5%	0.5%
Latvia	2000	15	8	14	0	38	0.0	40.3%	22.1%	37.6%	0.0%	0.0%	0.0%
Malta	2002	3	2	7	0	12	0.0	24.7%	16.4%	58.8%	0.1%	0.0%	0.0%
Netherlands	2000	4458	974	2120	75	7626	8370.4	58.5%	12.8%	27.8%	1.0%	187.8%	109.8%
Norway	2001	1814	444	780	0	3037	321.3	59.7%	14.6%	25.7%	0.0%	17.7%	10.6%
Poland	2000	432	386	377	2	1197	10.1	36.1%	32.2%	31.5%	0.1%	2.3%	0.8%
Portugal	2000	258	222	348	100	927	0.0	27.8%	23.9%	37.5%	10.8%	0.0%	0.0%
Romania	2000	103	28	18	0	149	0.0	69.4%	18.8%	11.8%	0.0%	0.0%	0.0%
Russia	2000	2087	721	134	7	2948	0.0	70.8%	24.4%	4.5%	0.2%	0.0%	0.0%
Sweden	2001	8118	297	2085	10	10511	7119.9	77.2%	2.8%	19.8%	0.1%	87.7%	67.7%
Slovenia	2000	167	77	49	3	297	0.0	56.3%	25.9%	16.6%	1.2%	0.0%	0.0%
Slovakia	2000	94	35	14	0	143	0.0	65.8%	24.7%	9.5%	0.0%	0.0%	0.0%
Turkey	2000	464	86	839	0	1389	0.0	33.4%	6.2%	60.4%	0.0%	0.0%	0.0%
UK	2000	18884	3672	5985	529	29070	17214.8	65.0%	12.6%	20.6%	1.8%	91.2%	59.2%
Europe	2000	121054	24902	38694	1528	186177	87752.0	65.0%	13.4%	20.8%	0.8%	72.5%	47.1%
EU15	2000	109883	22508	34499	1351	168239	78539.2	65.3%	13.4%	20.5%	0.8%	71.5%	46.7%
EU25	2000	111365	23436	35233	1385	171417	78585.7	65.0%	13.7%	20.6%	0.8%	70.6%	45.8%
US	2000	216552	29926	33221	10218	289917	0.0	74.7%	10.3%	11.5%	3.5%	0.0%	0.0%
Japan	2000	109181	15217	22354	7108	153860	0.0	71.0%	9.9%	14.5%	4.6%	0.0%	0.0%

(Source: Eurostat and Stan, OECD, 2007)

Table B.1. Weights implied by estimated cumulative lag distributions for EPO patents

cited year	lag	CHEM	DRUG	ELEC	IND	MECH	OTHR	INST
2004	1	25.743	25.656	26.754	27.755	27.581	26.453	27.786
2003	2	12.613	12.574	12.629	12.828	12.723	12.374	12.970
2002	3	8.367	8.343	8.255	8.300	8.226	8.058	8.429
2001	4	6.272	6.255	6.146	6.148	6.092	5.992	6.255
2000	5	5.023	5.010	4.906	4.894	4.849	4.783	4.983
1999	6	4.194	4.184	4.090	4.073	4.037	3.989	4.148
1998	7	3.604	3.596	3.513	3.495	3.465	3.428	3.558
1997	8	3.162	3.155	3.082	3.065	3.039	3.011	3.118
1996	9	2.819	2.814	2.749	2.733	2.711	2.688	2.779
1995	10	2.545	2.540	2.484	2.469	2.450	2.431	2.509
1994	11	2.322	2.317	2.268	2.254	2.238	2.222	2.288
1993	12	2.135	2.131	2.088	2.075	2.061	2.048	2.105
1992	13	1.977	1.974	1.936	1.925	1.913	1.902	1.951
1991	14	1.842	1.840	1.806	1.796	1.786	1.776	1.819
1990	15	1.726	1.723	1.694	1.685	1.676	1.668	1.705
1989	16	1.623	1.621	1.596	1.588	1.580	1.573	1.605
1988	17	1.533	1.531	1.509	1.502	1.496	1.490	1.517
1987	18	1.453	1.452	1.433	1.427	1.421	1.416	1.439
1986	19	1.381	1.380	1.364	1.359	1.354	1.350	1.370
1985	20	1.317	1.316	1.302	1.298	1.294	1.291	1.307
1984	21	1.259	1.258	1.247	1.243	1.240	1.237	1.250
1983	22	1.206	1.205	1.196	1.193	1.191	1.188	1.199
1982	23	1.157	1.157	1.150	1.148	1.146	1.144	1.152
1981	24	1.113	1.113	1.108	1.106	1.105	1.103	1.109
1980	25	1.072	1.072	1.069	1.068	1.067	1.066	1.070
1979	26	1.035	1.035	1.033	1.033	1.032	1.032	1.034
1978	27	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table B.2. Weights implied by estimated cumulative lag distribution for US patents

cited year	lag	CHEM	DRUG	ELEC	IND	MECH	OTHR	INST
2004	1	31.692	36.501	28.857	33.062	27.891	33.167	33.167
2003	2	13.617	15.548	13.224	13.977	12.151	14.284	14.284
2002	3	8.454	9.580	8.516	8.584	7.623	8.874	8.874
2001	4	6.114	6.882	6.288	6.165	5.558	6.415	6.415
2000	5	4.799	5.368	4.995	4.816	4.392	5.030	5.030
1999	6	3.961	4.404	4.150	3.963	3.647	4.145	4.145
1998	7	3.381	3.739	3.555	3.376	3.130	3.533	3.533
1997	8	2.957	3.252	3.114	2.949	2.751	3.084	3.084
1996	9	2.634	2.880	2.774	2.624	2.461	2.741	2.741
1995	10	2.379	2.587	2.503	2.368	2.233	2.470	2.470
1994	11	2.172	2.350	2.283	2.162	2.048	2.251	2.251
1993	12	2.002	2.155	2.100	1.993	1.896	2.070	2.070
1992	13	1.860	1.991	1.946	1.851	1.768	1.918	1.918
1991	14	1.738	1.852	1.814	1.730	1.659	1.789	1.789
1990	15	1.634	1.732	1.701	1.626	1.565	1.678	1.678
1989	16	1.543	1.627	1.601	1.536	1.484	1.581	1.581
1988	17	1.463	1.535	1.514	1.457	1.413	1.496	1.496
1987	18	1.393	1.454	1.436	1.387	1.349	1.420	1.420
1986	19	1.330	1.382	1.367	1.325	1.293	1.353	1.353
1985	20	1.273	1.317	1.305	1.270	1.243	1.293	1.293
1984	21	1.223	1.258	1.249	1.219	1.198	1.239	1.239
1983	22	1.177	1.205	1.198	1.174	1.157	1.190	1.190
1982	23	1.135	1.157	1.151	1.133	1.120	1.145	1.145
1981	24	1.097	1.112	1.109	1.095	1.086	1.104	1.104
1980	25	1.062	1.072	1.069	1.061	1.055	1.066	1.066
1979	26	1.030	1.035	1.033	1.029	1.026	1.032	1.032
1978	27	1.000	1.000	1.000	1.000	1.000	1.000	1.000