DISCUSSION PAPER SERIES

No. 9264

THE ECONOMIC VALUE OF PATENT PORTFOLIOS

Alfonso Gambardella, Dietmar Harhoff and Bart Verspagen

INDUSTRIAL ORGANIZATION

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP9264.asp**

ISSN 0265-8003

THE ECONOMIC VALUE OF PATENT PORTFOLIOS

Alfonso Gambardella, Bocconi University
Dietmar Harhoff, LMU Munich and CEPR
Bart Verspagen, University of Maastricht

Discussion Paper No. 9264
December 2012

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 9264

December 2012

ABSTRACT

The Economic Value of Patent Portfolios

Patent holders may choose to protect innovations with single patents or to develop portfolios of multiple, related inventions. We propose a simple decision-making model in which patent-holders may allocate resources to either expanding the number of related patents or investing in higher value of patents in the portfolio. We estimate the derived value equation using portfolio value data from an inventor survey. We find that investments in individual inventions exhibit diminishing returns, and that much of the value of a portfolio depends on adding new inventions. These effects are less pronounced in high-techology industries, when the inventions rely on external information, and when the inventor holds a doctorate. We also find higher returns to an increase of the number of inventions when firms perceive patent protection to be strong. Thus, a higher number of inventions in a portfolio may reflect both genuine creation of value or stronger appropriability via patents.

JEL Classification: L20, O31, O33 and O34
Keywords: intellectual property rights, inventors, patents and technical change

Alfonso Gambardella
Department of Management & Technology
Università Bocconi
Via Roentgen 1
20136 Milan
ITALY

Email: alfonso.gambardella@unibocconi.it

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=157918

Dietmar Harhoff
LMU Munich
Kaulbachstr. 45
D-80539 Munich
GERMANY

Email: harhoff@lmu.de

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=125340

Bart Verspagen
University of Maastricht
Keizer Karelplein 19
6211 TC Maastricht
THE NETHERLANDS

Email: b.verspagen@maastrichtuniversity.nl

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=164985

Submitted 08 December 2012

1. INTRODUCTION

Much of the literature investigating the role of patents in innovation and value creation has focused on individual patents. However, it is well-known that firms frequently pursue the development of patent portfolios whose patented inventions cover multiple, yet related aspects of an innovation (Knight 2001, Harhoff and Reitzig 2001). To the best of our knowledge, the decision-making tradeoffs in constructing such portfolios of related patents have not been investigated. It is unclear under which conditions firms will be content with single patents covering an innovation, and when they will opt for portfolios with multiple, related inventions. Much of the literature has been descriptive, e.g., by using categories such as "discrete" vs. "complex" technologies (Cohen et al. 2000). Nor has empirical evidence been produced that sheds light on the question how strongly the value of such patent portfolios is determined by the number of patents contained in them or by the quality of the individual inventions represented in the patent portfolio. Our paper seeks to provide a first theoretical and empirical investigation of these issues.

Pioneering work by Rosenberg (1982) and by Teece (1986) has shown that the size and quality of downstream assets explain much of the value created during the development, production, and commercialization of innovations. However, a better understanding of the factors that affect economic value in *earlier* stages of the innovation process is critical as well, as implied by the growing importance of technology licensing and more generally the exchange of technologies disembodied from products (e.g., Arora et al., 2001).

In most industries, patented inventions represent important outcomes of the upstream research process. Not all inventions are patented, but patents cover a relevant share of invention output (Giuri et al. 2007) and are important assets of the firm (Lev 2001, Hall et al. 2005). Moreover, patents may be the object of a license, or an important component of it, and on many occasions, the assessment of the value of relevant patents is central for defining the terms of a technological alliance or in standard-setting technology pools that involve many firms (Zuniga and Guellec 2009). We consider a portfolio of related patents which are linked from a technical point of view or with respect to their impact on the value of an innovation.[1] Combining related patents is increasingly common as firms have started to apply for different patents to cover the different components of their complex inventions or to create patent "fences" that strengthen their protection (e.g., Ziedonis, 2004; Knight 2001). Accordingly, the value of a portfolio of patents might increase because the number of inventions increases or because the average value of the inventions in the portfolio increases. In our theoretical model, we study the extent to which the overall value of the portfolio is determined by the former or the latter.

[1] The term "patent portfolio" is used in different connotations in practice. Practitioners occasionally use the term to describe the set of patents which cover the same invention in different jurisdictions. In this paper, we focus on one jurisdiction, but analyze patents that are related technologically or in terms of the strategic impact that they have regarding a particular innovation or new product.

We use the insights from our theoretical analysis to estimate the returns associated with single inventions as well as patent portfolios using survey data from an inventor survey. We find evidence of diminishing returns to the work-months invested in an individual patented invention, as well as complementarity across inventions. Thus, much of the value of a portfolio of inventions depends on the addition of new inventions, rather than persistent investments devoted to one invention. We further find that this effect is less pronounced i) in high-technology industries, ii) when external information (from science, users, competitors or suppliers) guides the invention process and iii) when the inventor holds a doctoral degree. We interpret these results as evidence that technological opportunities, external information, and science-based education produce less complex inventions in which the average value of the inventions in the portfolio is relatively more important than their number. Finally, we find that the returns from increasing the number of inventions in the portfolio increase as patenting is associated more strongly with the goal of preventing imitation.

Our study contributes to the growing body of workon the value of patents, but our focus on the economic logic of patent portfolios is relatively new in the economic literature.[2] We emphasize that filling this research gap is tantamount to understanding real-world patenting behavior and the recent rise in patent filings in many patent systems. Our paper also offers additional contributions to the study of the relationship between resource utilization and patenting. Following Hausman et al. (1984), a vast literature reviews the relationship between R&D (or other measures of investments) and the number of patents at the firm level. However, little research considers the relationships between innovation or patent output and the resources invested in the project that leads to an invention. Although the literature on indicators provides insights into the distribution of patent value, it does not address the question as to what determines that value causally. Similarly, the main goal of the literature investigating the link between market value and patent counts or citations is to estimate the market's evaluation of a firm's patents; it does not seek to identify the mechanisms or factors that affect this value. The current study is also one of the few to consider, along with firm and technological factors, the impact of inventor characteristics on the value of patents (Gittelman and Kogut 2003; Toivanen and Väänänen 2012). Moreover, we are the first to study this relationship at the level of empirically identified portfolios of related patents.

[2] There are a large number of contributions studying the value of either single patents or the overall patent portfolio of a firm. For example, the use of patent renewal fees, as pioneered by Schankerman and Pakes (1986), only looks at the former, such as in the studies on the transfer of patent rights (Serrano, 2010). By contrast, citations (Trajtenberg, 1990), or other indicators (e.g., like the international coverage of a patent, or whether a patent is opposed or litigated; e.g., Harhoff et al., 2003; Lanjouw and Schankerman, 2004), are more likely to be correlated with the value of the invention. The estimated impacts of the patent or citation stocks of the firms and their market value can cover both the value of the inventions and the patent rights (e.g., Arora et al. 2008; Bessen 2008, 2009; Hall et al., 2005; Bessen, 2009; see also Arora et al., 2008, and Bessen, 2008). More recently, Hsu and Ziedonis (2012) use the valuation of venture capitalists to assess the value of patents as protection vs quality signals. (See also Greenberg, 2012.)

In the next section, we develop our model before presenting the data in section 3. Section 4 presents the empirical results, and in section 5 we conclude with an in-depth discussion.

2. THEORY

2.1 A Model of the Patent Portfolio Decision

In our model a firm allocates resources (e.g., inventors' time) to inventive activity. Rather than investing all resources in one particular patented invention, the firm has to decide how many related inventions it wants to develop and how much to invest in each of them. Specifically, the firm invests time H_i, $i = 1, 2, \ldots N$, in N related inventions that it plans to patent. The expected value of the portfolio is $V = K\left(\sum_{i=1}^{N} H_i^{\mu}\right)^{\frac{\alpha}{\mu}} P^{g(N)}$, where $\frac{\alpha}{\mu}$ measures complementarity across investments in the single inventions, μ measures the returns from investing in each inventions when $N > 1$, K accounts for any other factor known to, but not under the control of the firm that affects V, and $P^{g(N)}$ is the probability that the value of the portfolio is realized. Typically, a more complex portfolio, which is composed of a larger number of related inventions, is less likely to be technically successful because more individual inventions need to be jointly successful. We capture this idea by assuming that the probability that the portfolio is technically successful is $P^{g(N)}$, where $0 \leq P \leq 1$, $0 \leq g(N) \leq N$, and $g(\cdot)$, the number of core inventions in the portfolio, is an increasing function of N. Each core invention is technically successful with probability P, and the portfolio is technically successful only if all the core inventions are technically successful. We assume further that the success of each core invention is independent of the success of the other core inventions, and if the core inventions are technically successful, all the other (non-core) inventions are technically successful with probability 1. With this setting, we allow for the possibility that the technical successes of the individual patented inventions are not all independent from one another. As an alternative, we can think of $P^{g(N)}$ as an approximation of the probability of success of the portfolio as a whole, ranging from P (the success of the portfolio depends on one key idea) to P^N, which means that it depends on the independent successes of all the N patented inventions. Finally, we assume that $P = P(Y, G)$, where Y is a choice variable of the firm that can then optimize the probability of success P, and G is a set of exogenous factors that affect this probability.

This structure makes a few simplifying assumptions. First, we do not identify some specific core inventions, but simply posit that there are $g(N)$ core inventions, and we focus on the fact that a portfolio composed of more inventions (higher N) requires more inventions to be jointly successful from a technical point of view. Second, the productivity μ associated with each invention is the same, whether the invention is core or not, and so is the unit cost W. As we shall see, this implies that the firm invests the same amount of time H_i on core or non-core inventions. While a simplifying assumption (with not much additional insight if we released it), it helps to emphasize our distinction between the technical success of the invention, which depends on the probability $P^{g(N)}$, and its economic value, which depends on the summation across all the N patented inventions that defines V. Finally, we assume that the relationship across inventions, captured by the parameter $\frac{\alpha}{\mu}$, is common across all the invention pairs. In principle, different pairs may exhibit different relationships. We assume this possibility away, and only estimate the average relationships across the inventions in our portfolios.

In summary, a firm chooses Y, N, and H_i, i = 1, 2, … N to maximize the expected value of the portfolio of inventions. The cost of developing the portfolio is $W\sum_{i=1}^{N} H_i$, where W is the unit cost of H. The cost of raising the probability P is $C(Y)$, where, using subscripts to denote derivatives, C_Y, $C_{YY} > 0$. The choice of N resolves the trade-off highlighted earlier: The larger the number of inventions, the higher the value of the portfolio, yet the lower the probability that the portfolio as a whole will work. The choice of H_i, i = 1, 2, … N, aims to increase the value of each invention, and of the portfolio as a whole, if it works. Higher H_i improve the features of the inventions, making the whole portfolio more effective or more appealing to the market. The optimization problem of the firms then is

$$\underset{Y;N;\{H_i\}}{Max}\; K\left(\sum_{i=1}^{N} H_i^{\mu}\right)^{\frac{\alpha}{\mu}} P(Y,G)^{g(N)} - W\sum_{i=1}^{N} H_i - C(Y) \;.$$

The first-order condition (*foc*) with respect to the generic H_i is

$$\alpha K\left(\sum_{i=1}^{N} H_i^{\mu}\right)^{\frac{\alpha}{\mu}-1} H_i^{\mu-1} P^{g(N)} = W \;. \tag{1}$$

If we take the ratio of any two of these *focs* $H_i = H$, $\forall\ i = 1, 2, \ldots N$. Therefore,

$$\alpha K N^{\frac{\alpha}{\mu}-1} H^{\alpha-1} P^{g(N)} = W \;, \tag{2}$$

and

$$H = \left(\frac{\alpha K}{W}\right)^{\frac{1}{1-\alpha}} N^{\theta-1} P^{\frac{g(N)}{1-\alpha}} \;, \tag{3}$$

where $\theta \equiv \dfrac{\alpha(1-\mu)}{\mu(1-\alpha)}$. From Equation (2), we know $W\sum_{i=1}^{N} H_i = WNH = \alpha K N^{\frac{\alpha}{\mu}} H^{\alpha} P^{g(N)}$. Similarly, $V = KN^{\frac{\alpha}{\mu}} H^{\alpha} P^{g(N)}$. Now we replace H in both expressions with Equation (3), and the optimal problem becomes

$$\underset{N;Y}{Max}(1-\alpha)K\left(\frac{\alpha K}{W}\right)^{\frac{\alpha}{1-\alpha}}N^{\theta}P(Y,G)^{\frac{g(N)}{1-\alpha}}-C(Y)\,,$$

whose *foc* with respect to N is

$$\theta N^{\theta-1}P^{\frac{g(N)}{1-\alpha}}-N^{\theta}\frac{P^{\frac{g(N)}{1-\alpha}}}{1-\alpha}\frac{\partial g(N)}{\partial N}logP=0\,.$$

Define $g(N) = N^{\gamma}$, where $0 \leq \gamma \leq 1$. Thus, $\gamma = 0$ denotes the case in which the innovation associated to the portfolio as a whole is technically successful when one core invention is technically successful; the other extreme is $\gamma = 1$, that is, all the inventions need to be technically successful, and the success of each invention is independent of the others. We assume that each invention is technically successful if an exponentially distributed latent variable is greater than a threshold G, and Y is the expected value of the latent variable. This means that $P(Y, G) = e^{-\frac{G}{Y}}$, and the optimal N is

$$N=\left[\frac{\theta(1-\alpha)}{\gamma}\frac{Y}{G}\right]^{\frac{1}{\gamma}}, \tag{4}$$

where a higher G means that the probability of technical success of the inventions is smaller, and Y accounts for the fact that the firm draws inventions from a better distribution.

Finally, the firm chooses Y, i.e., the distribution from which inventions are drawn. Define $C(Y) = Y^{\beta}\cdot X$, where $\beta > \frac{\theta}{\gamma}$, and X is a set of variables that affect this cost. If we replace the optimal N in $P^{\frac{g(N)}{1-\alpha}}$, we obtain $P^{\frac{\theta Y}{\gamma G}}$. Using the fact that $\frac{Y}{G} = -(logP)^{-1}$, then $logP^{\frac{\theta Y}{\gamma G}} = -\frac{\theta}{\gamma}$, which implies $P^{\frac{g(N)}{1-\alpha}} = e^{-\frac{\theta}{\gamma}}$. The optimization problem then becomes

$$\underset{Y}{Max}(1-\alpha)K\left(\frac{\alpha K}{W}\right)^{\frac{\alpha}{1-\alpha}}\left(\tilde{\theta}\frac{Y}{G}\right)^{\frac{\theta}{\gamma}}-Y^{\beta}X\ ,$$

where $\tilde{\theta} \equiv \frac{\theta(1-\alpha)}{e\gamma}$. The *foc* is now

$$\frac{\theta}{\gamma}(1-\alpha)K\left(\frac{\alpha K}{W}\right)^{\frac{\alpha}{1-\alpha}}\left(\frac{\tilde{\theta}}{G}\right)^{\frac{\theta}{\gamma}}Y^{\frac{\theta}{\gamma}-1}=\beta Y^{\beta-1}X\,, \tag{5}$$

which yields an optimal expression for Y as a function of the exogenous variables of the model.

We also define R to be the total work-months invested, encompassing both the total effort $W\sum_{i=1}^{N} H_i$ spent on each invention and the resources committed to the realization of P. Thus, we can write $C(Y) = Y^{\beta} \cdot X = WR - WNH$. From Equation (5), and using the earlier expression for WNH, we know $\psi WNH = WR - WNH$, where $\psi \equiv \frac{1-\mu}{\gamma\beta\mu}$ or $H = \frac{R}{(1+\psi)N}$. Replacing it in the expression for $V = KN^{\frac{\alpha}{\mu}} H^{\alpha} P^{g(N)}$ yields

$$V = const \cdot KR^{\alpha} N^{\frac{\alpha(1-\mu)}{\mu}}, \tag{6}$$

where given the optimal N, $P^{g(N)} = e^{-\frac{\theta(1-\alpha)}{\gamma}}$, and it is captured by the constant term. Equation (6) is the reduced form of the expected value of the portfolio once the optimal choices have been accounted for, and it is the value equation that we estimate.

2.2 Estimated Value Equation

In our empirical analysis, we estimate the log-log version of (6) after instrumenting for N and R. Our empirical measure of value is the expected value of the portfolio when the patent is granted. As we will see in Section 3.2, we take care of expectation automatically because we use a survey-based measure where the respondent is asked to predict value by using all the available information at the moment of the survey. A potential concern is that, to simplify the analysis, our model treats N as a continuous variable despite of its integer nature. This raises two questions. The first one is whether $N = 1$ is a reasonable solution. Not only is it natural to allow for the possibility that portfolios are composed by singletons, but as we will see in our discussion of the data below, about two-thirds of the portfolios in our sample are composed of one patent. The second question is whether the comparative static of our continuous problem is a good approximation, especially at low levels of N, and particularly between $N = 1$ and $N = 2$.

As far as the first question is concerned, take (4), our optimal expression for N. We want to make sure that an optimal $N = 1$ does not correspond to unreasonable parameter values. From (4) it is not difficult to see that $N = 1$ if $\frac{Y}{G} = \frac{\gamma\mu}{\alpha(1-\mu)}$. Since in our empirical analysis we will not be able to identify γ, we focus on α and μ. Since $\frac{Y}{G} > 0$, we need $\alpha > 0$ and $0 < \mu < 1$. As we will see in our empirical sections, our estimated values of α and μ are well within these boundaries. As an additional check, our estimated value of $\frac{\alpha(1-\mu)}{\mu}$ is 1.2. For the extreme case of $\gamma = 1$, this corresponds to $\frac{Y}{G} = \frac{5}{6}$ and P equal to about 30%. This is a reasonable probability of technical success.

To address the second question, the firm's optimum is reached when choosing a particular N and resource allocations conditional on N dominate all other possible choices of N. We compare the optimal V when $N = 1$ and $N = 2$, and derive the conditions under which V is higher when the optimal N is equal to 2 with respect to 1. These conditions should be approximately the same as the comparative statics from the continuous model. Use the optimal V after optimizing for H and P. To streamline the analysis, set K, W, and $\gamma = 1$, which is inconsequential for this discussion. This yields $V = \alpha^{\frac{\alpha}{1-\alpha}} N^{\theta} P^{\frac{N}{1-\alpha}}$, where we replaced the optimal expression for H. Take the difference of this expression when $N = 2$ and $N = 1$. This means that we value V at $N = 2$ when this is the optimal choice for N, and similarly for $N = 1$. As noted at the end of the previous section, P adjusts to change in N in such a way that $P^{\frac{N}{1-\alpha}}$ does not change as the optimal N changes. Thus, using the expression above, the difference between $V(N{=}2)$ and $V(N{=}1)$ is simply proportional to $2^{\theta} - 1$, which is positive for $\theta > 0$. In the continuous case, if we take the derivative of this expression with respect to N, it is proportional to $\theta N^{\theta-1}$, which is also positive for $\theta > 0$. Moreover, anticipating our empirical results, we estimate that θ is about 2. This implies $2^2 - 1 = 3$. At the same time, if we set $N = 1.5$ (mid-point between 1 and 2), then $\theta = 2$ implies $2 \cdot (1.5)^{2-1} = 3$. Our estimates imply that the change in V as N increases from 1 and 2 yields similar results in the discrete and continuous models.

Finally, our model guides the search for the excluded instruments in (6) that identify the impacts of N and R. The variable V does not depend directly on G or X, which instead affect the optimal Y. The optimal Y affects N, as implied by (4). Similarly, R is proportional to $Y^{\beta} \cdot X$, and therefore, it also depends on the factors that affect Y, particularly G and X. We can use, as instruments of N and R, proxies for the probability of technical success of the portfolio, or else proxies for the costs of the research that lowers P. Our model also provides a natural interpretation of the parameters that we estimate. The parameter α represents the elasticity of V with respect to R. The elasticity of V with respect to N is affected by μ, which measures the extent of the diminishing returns associated with the individual inventions. Moreover, we can identify both α and μ, and thus the degree of complementarity, $\frac{\alpha}{\mu}$.

2.3 Propositions

In our empirical analysis, we first study an implication of the diminishing returns of effort on individual inventions. Typically, early inventions are solutions to technical problems. As Rosenberg (1982) put it, the economic value of an invention depends on several factors that go beyond its technical success, including demand, network externalities, and complementarity with other goods, which are rarely under the control of the inventive team. The inventive team might talk to managers, the marketing department, or users to capture some such factors, yet their main task remains finding technical solutions. As a result, increasing effort on a particular invention produces diminishing economic returns, because the growing technical perfection of an individual invention likely does not raise its economic value much if it is not accompanied by some nontechnical elements and offerings.

Consider the meaning of patents. A single inventive step that gives rise to a patent tends to focus on a particular invention. Moreover, the inventive step that makes an invention patentable is a technical matter, as implied by the patentability requirement that looks at technical rather than economic features of any invention. For example, Astebro and Dahlin (2005) show that technical feasibility determines the decision to patent an invention but not the decision to commercialize it. As a result, for work put into improving a specific inventive step, the returns of additional efforts are likely to fade away rapidly. At the same time, inventors can increase the economic value of the patent portfolio by adding and combining different technical features or inventions. In our model, this situation suggests that μ is small; from Equation (6), the elasticity of *V* with respect to *N* is higher than the elasticity of *V* with respect to *R*.

Proposition 1. *Strong diminishing returns on the individual inventions (small* μ*) make the returns of N higher than the returns of R — that is,* $\frac{\alpha(1-\mu)}{\mu} > \alpha$.

Proposition 1 does not imply complementarity. The individual inventive steps may simply add to the value of the invention because it is more beneficial to invest additional resources in a different patented invention rather than insisting on the same one—which is the essential outcome of diminishing returns. Yet the inventions may be complementary, in the sense that increasing effort on one of them raises the value of the other. Therefore, $\frac{\alpha}{\mu} > 1$, which can raise the impact of *N* on *V* beyond the effect of diminishing returns, as implied by the finding that a larger $\frac{\alpha}{\mu}$ increases $\frac{\alpha(1-\mu)}{\mu}$.

Proposition 2. *Complementarity (higher* $\frac{\alpha}{\mu}$*) raises the returns of N,* $\frac{\alpha(1-\mu)}{\mu}$.

We also consider effects on the extent of diminishing returns on single inventions. As suggested by Propositions 1 and 2, they have implications for the returns of *N*, both relative to *R* and in absolute terms. The diminishing returns to *N* result because the individual inventions, even if technically valid, might not be valuable from an economic point of view if they are not accompanied by complementary economic factors. The addition of new inventions may address this issue somewhat, if they target complementary technical aspects and make the portfolio as a whole economically more appealing.

Other conditions may make the diminishing returns less pronounced. First, in high-tech industries, new technological opportunities raise the returns to technical improvements, because in these industries, the returns on an effort aimed at technical improvements are higher. By definition, high-tech industries produce greater technological opportunities, and the unit effort spent improving a technology is productive. Although these traits may not fully satisfy the economic conditions for obtaining a valuable invention, higher productivity in the generation of technical improvements can only raise the returns on the efforts of a single invention. Second, a similar rationale applies to two other cases: availability of information and education. Information from users, competitors or suppliers focus research on more valuable inventions. For example, information about user needs raises the odds that there will be demand for the invention in the market. Similarly, competitors or suppliers provide information that reduce the uncertainty about the value of the invention. Scientific information plays a similar role. As noted by Arora and Gambardella (1994) or Fleming and Sorenson (2004), science guides the search for inventions, and therefore projects that rely to a greater extent on scientific knowledge are more likely to identify inventions in which effort is technically, and possibly economically, more productive. If so, these projects will exhibit less pronounced diminishing returns. Education has a similar effect. More educated inventors (typically Ph.D.s) hold more general knowledge and thus are more likely to focus on more productive invention opportunities.

Proposition 3. *The returns to an increase of N (portfolio size) are lower in high-tech industries, when inventors have better information about the inventors, and in the case of scientifically trained inventors..*

Finally, a higher number of patents in a portfolio may strengthen the protection of a key idea, because the firm can patent different aspects of it and make it harder to invent around (von Graevenitz et al., 2011; Ziedonis, 2004). More patents then can mirror a set of technical connections among inventions, which cover different facets of an innovation and make it more valuable, and mere connections in value, which raise the value of the patent protection.

Proposition 4. *The returns to an increase of N (portfolio size) are higher when patents are effective as protection against imitation.*

3. SAMPLE AND VARIABLES

3.1 Survey Data

With a survey, we successfully collected data on 9,550 patents after asking 28,470 European inventors for participation. Inventor names were sampled from patent applications filed and granted at the European Patent Office (EPO) with priority years between 1993 and 1997. The address of the first inventor listed in the patent was in Denmark, France, Germany, Hungary, Italy, the Netherlands, Spain, or the United Kingdom. The original 28,470 patents reflected several characteristics, available from the patent document, of the universe of patents with priority dates 1993–1997 in these countries. We slightly oversampled "important" patents (i.e., those with at least one citation and that were opposed by third parties, as allowed for by the European Patent Convention). All our analyses employ sample weights to account for this oversampling.

Details of the survey and an analysis of the variables that describe the context of invention processes in Europe appear in Giuri *et al.* (2007).[3] Here, we focus on the private value of the patents held by firms. Thus, the sample that we use excludes patents held by universities, individuals, or nonprofit institutions. We consolidated firms according to their ultimate parents, using *Who Owns Whom* (several editions).

3.2 Dependent Variables

Following Harhoff et al. (2003), our survey asked inventors: "Suppose that on the day on which this patent was granted, the applicant had all the information about the value of the patent that is available today. In case a potential competitor of the applicant was interested in buying the patent, what would be the minimum price (in Euro) the applicant should demand?" The survey offered 10 interval responses: less than €30K; 30–100K; 100–300K; 300K–1M; 1–3M; 3–10M; 10–30M; 30–100M; 100–300M; and more than 300M. Gambardella et al. (2008) show that the distribution of this measure is skewed to the left, and in general, it conforms well with other distributions of the value of patents or inventions (Astebro, 2003; Harhoff et al., 1999; Scherer and Harhoff, 2000). It is also strongly correlated with some standard indirect indicators of patent value, and for a subsample of 354 patents, it is not different from the distribution of responses given by company managers (Gambardella et al., 2008).

The inventors indicated whether the patent was part of a group (portfolio) and, if so, whether the number of patents in the portfolio fell into the following intervals: 2, 3–5, 6–10, 11–20, or >20. The exact definition that we used was whether the focal patent was part of "a group of patents which crucially depend on each other in terms of their value, or in a technical way." The survey then asked for the value of the overall portfolio, using the same wording as described before. In this case, the menu consisted of 12 interval responses: The first nine intervals were the same as for single patents, but the final intervals became 300M–1B, 1–3B, and more than 3B.

The proxy for V that we used in our regressions was the geometric mean of the boundaries of the intervals for the value of the whole patent portfolio. The distribution of this variable also was skewed to the left, similar to the one for the individual patent values, though with a longer tail. The proxy for N was equal to 1 for all patents that the respondent declared not part of a group. For the patents declared to be part of a group, N equaled the geometric mean of the boundaries of the other intervals (N = 20–40 for the last class). Approximately two-thirds of the patents in our survey were singletons, N = 1. We show in Appendix 1 that the average value of the portfolio — that is, the ratio of our proxies V and N — are

[3] Giuri et al. (2007) only report patents in six countries. Data about Denmark and Hungary were collected later.

highly correlated with several indirect indicators of patent value.

A potential concern with survey-based measures, such as our proxies for V and N, is that they may be affected by shocks associated with some characteristics of the respondents that do not have much to do with the relevant characteristics of the variables for our purposes. For example, inventors who are particularly excited about their inventions may boost the evaluation of their patents or the extent of the related patents. We are not particularly concerned about these problems for two reasons. First, we instrument for N in (6), which removes potential biases due to a common shock affecting both N and V. Second, the problem of a respondent shock would be relevant if we looked at the predicted values or number of patents, but because our analysis focuses on the impacts (elasticities) of the covariates on the dependent variables, it is a less important concern.

Finally, we could have employed alternative measures of the number of technically related patents derived directly from patent data. We did not follow this route because technically connected patents coming out of a given research project are not easy to identify. We could have looked at the patents of the same inventor, in a given technological area, and around the same time. However, it is not clear that technically related patents are produced by the same inventor, or in similar technological classes, or around the same time. More generally, this approach would have required assumptions about the search algorithm, and it is not obvious that these assumptions would produce fewer errors than asking inventors.

3.3 Covariates

Table 1 defines all the variables that we employ in our analysis, along with descriptive statistics. In estimating (6), we control for the endogeneity of N and the total work months invested in the project leading to the portfolio of patents (R). Note that controlling for the endogeneity of R tackles the same problem mentioned previously about potential common respondent shocks across surveyed variables. The next section discusses the instruments for these endogenous variables.

TABLE 1 ABOUT HERE

We control for four sets of factors in our regressions: (1) characteristics of the technological projects, (2) characteristics of the firms, (3) characteristics of the inventor, and (4) patent premium. In our survey, targeting a specific patent, the controls refer to each patent, which is not a problem for the singleton patents in our regressions, whereas for the remaining patents, the surveyed patent is a random patent in the portfolio. Because the patents in the portfolio are related, we expect their characteristics to be correlated with those of the focal patent.

We control for the characteristics of the innovation project by using 30 technology class dummies, dummies for the priority year of the patent (1993–97), and dummies for the countries in which the patent was invented.[4] In addition, GOVFUND is a dummy value equal to 1 if the funding of the research leading to this patent came from government research programs or related funds. Subsidies to R&D are common in Europe, both from the European Union and from national or even regional governments. This variable accounts for the costs of the project, correlated with the wage W or parameters of the cost function $C(\cdot)$ in our model. The variable INFO_SCIENCE accounts for the importance of science as a source of knowledge for the patented invention, while INFO_OTHER accounts for information received from customers, competitors or suppliers..

We controlled for firm characteristics using their SALES in 1995, consolidated at the level of the ultimate parent firm. The variable RD is consolidated R&D expenditures in 1995. However, R&D and sales figures were not available for some firms in our sample. Because we estimate log-log regressions, we set SALES and RD equal to 1 when their value is missing (logs equal to 0), and we introduced two dummies for missing RD and for missing SALES, equal to 1 when the corresponding variable is missing and 0 otherwise. The reason why we cared about missing observations for RD and SALES is that the missing values are concentrated on smaller firms.[5] We also controlled for the relative specialization of the firm in the technology class of the patent (RTA). Because our survey mirrors the distribution of patents in the population, the use of surveyed patents to compute RTA is unlikely to introduce any particular bias.

Finally, we controlled for the age, education, and gender of the inventor (see the dummies in Table 1). With two variables, we controlled for the value associated with the patent protection of the invention, that is, for the "patent premium." The RACE dummy variable equaled 1 if the inventor declared that the invention had to be patented quickly because other firms or researchers were working on the same idea. The second control, IPC4_COMP, represented the complement to 1 of the Herfindahl index of the number of applicants in our sample in the same four-digit IPC class of the patent. The scale of our survey, along with the sample stratification, which reflected the distribution of the population of patents, ensured that this variable provided a good proxy of the Herfindahl index we would have obtained had we employed the full population. Moreover, following our consolidation of the applicants in the sample, the Herfindahl index was computed by considering applicants as different only if they belonged to unaffiliated organizations.[6] Both RACE and IPC4_COMP indicate that a patent right is more valuable if there is actual or potential competition.

---

[4] We retrieved the technology classes from the ISI-INIPI-OST concordance classification between patent IPC classes and the 30 technology classes elaborated by the German Fraunhofer Institute of Systems and Innovation Research (ISI), the French Patent Office (INIPI), and the Observatoire des Sciences and des Techniques (OST). See Giuri *et al*. (2007) for details.

[5] As we shall see in Section 4.2, we performed a robustness check that rules out that missing observations for other covariates in our analysis influenced our results.

[6] An alternative to IPC4_COMP would be a measure of the product market competition of the firm, but it is very hard to find information on product market competitors associated with the specific technological class of the patent.

3.4 Instruments

We employed two sets of instruments in our analyses. The first set is composed of the three instruments SERENDIPITY, INTRATECY, and INVENTOR_BREADTH. The second set uses IPC3 – the dummies for IPC patent classes at 3-digits – in lieu of INVENTOR_BREADTH. The SERENDIPITY variable measured whether the invention was the outcome of a "stroke of genius" rather than a well-planned research project. Moreover, as implied by its definition in Table 1, serendipitous inventions are not furthered in a formal project, and they do not entail additional R&D costs before being patented. This definition rules out later investments in the project, after receiving news about its value, and thus can be viewed as exogenous shocks affecting R and N. The variable INTRATECY measured the business cycle in the country. As a macro variable, it is exogenous with respect to decisions taken at the level of firm projects. The interest rate affects investments in innovation and thus N and R.

We also conducted an extensive data collection to retrieve all the European Patent Office (EPO) patents by inventors in our sample before the date of the surveyed patent. The variable INVENTOR_BREADTH measured the technological specialization of the inventor's past inventions. It is the complement to 1 of the Herfindahl index of the stock of the inventor's past patents one year before the focal patent across four-digit IPC classes. An inventor whose past patents span several technological classes is more likely to hold some general competence that enables her or him to develop technologies in different domains. Higher values of INVENTOR_BREADTH thus suggest that the inventor holds a general rather than specialized expertise, implying that the inventor team can take advantage of this general capability. Even if only the focal inventor has such expertise, it is likely to be an asset available to the entire research process leading to the innovation. At the same time, some inventors with low INVENTOR_BREADTH may team up with more generalist inventors, and thus our measure underestimates the actual breadth of knowledge inside the team. Although the lack of information about other inventors is a limitation, our survey is directed to the first inventor or, if he or she does not answer, to the subsequent inventors in the list. Therefore, our respondent inventor is either an important or a random inventor in the team. If the team has no general capability, INVENTOR_BREADTH is small; if it does, INVENTOR_BREADTH captures it at least sometimes.

If the inventor has a general capability, the successes of the individual inventions in a portfolio are unlikely to be independent, because an inventor, or an inventor team, with a general capability is likely to produce inventions that hinge on one or a few general ideas from which they generate different inventions. The probability success of an innovation then depends largely on the probability of success of its core ideas, and not of all inventions that can stem from them. As a result, INVENTOR_BREADTH is likely to be negatively correlated with G in our model. Generalist inventors may also be more capable and exhibit a higher P or lower costs of producing it (i.e., lower X). All these effects raise N or R.

In principle, INVENTOR_BREADTH may affect V directly, especially if interpreted as a measure of ability. The low value of the Hansen J statistics that we report in our estimation seems to rule out this possibility. Because INVENTOR_BREADTH is a Herfindahl index, it may be correlated with the number of past patents of the inventor. Some technically related patents that are part of the portfolio of our focal patent also may be past patents that enter the computation of INVENTOR_BREADTH. We tried alternative instruments that assess the average or maximum number of citations of the past patents of the inventor at least three years before the focal patent. They are both strong and relevant instruments, and the Hansen J statistics rules out their direct correlation with V. The returns of N on V that they produce are even higher than what we show using INVENTOR_BREADTH. Finally, our results do not change if we use the Herfindahl across three- rather than four-digit IPC classes.

In addition, we report the results from using the IPC3 dummies as an alternative instrument to INVENTOR_BREADTH. Another factor that likely affects G, and thus N, is the degree of modularity of the technology, or more generally the nature of the technology itself. Some technologies rely on a larger set of differentiated modules, while others tend to rely on fewer basic elements. In our value Equation (6) we control for 30 ISI-INIPI-OST technological classes, which we list under Table 1. These classes are obtained from aggregations of the IPC classes to define more meaningful classes from an industrial point of view. In contrast, the IPC classes at three-digits, IPC3, are strictly technological in nature. To further note the distinction between the two classifications, the former aggregates, under the same class, IPC classes with four-digits or more that belong to different IPC classes at three-digits. The rationale for this instrument is that a classification of patents with a stronger industrial bent is more likely to have a direct effect on the value of patent portfolios. It captures factors related to demand or the underlying markets. In contrast, technological modularity or related matters are more inherently technical factors, and thus they affect N without directly affecting V.

To be sure, IPC3 is a rather coarse and noisy as an instrument. In the first place, while technological dummies may well capture factors like the degree of modularity or other features of technology, they also capture other factors. This may affect the strength of IPC3 as an instrument. Nonetheless, we employed it as an alternative to INVENTOR_BREADTH to check whether they produce consistent results. We also tried other alternative instruments that may resonate with the logic of capturing factors that affect modularity or more generally G in our model. For example, we used the number of four-digit IPC classes in each three-digit IPC class, as a proxy for the granularity of the IPC three-digit classes. We also employed the average N by four-digit IPC classes in our sample. They both produce consistent results, even though the former is a weaker instrument.

4. EMPIRICAL RESULTS

4.1 GMM Estimation and Results

In Table 2 we present our generalized method of moments (GMM) estimation of the log-log version of (6). We distinguish between Specification A and B. In the former, we employ as instruments SERENDIPITY, INTRATECY and INVENTOR_BREADTH, while in the latter we replace the IPC3 dummies for INVENTOR_BREADTH. Table 3 shows that the instruments of Specification A are strong and relevant. The Hansen J-statistic indicates that our identification depends on exclusion restrictions that are not statistically significant. The F-statistics that test for the significance of the three excluded instruments of our GMM regressions are well above 10, the threshold most commonly invoked in these cases. The instruments for Specification B are relevant (as implied by the underidentification test), and the Hansen J-statistic indicates that the exclusion restrictions are not statistically significant. The F-statistic for the *R*-equation is well above 10, while the F-statistics for the *N*-equation is statistically significant, though smaller than 10. Unlike the previous case, this second set of instruments does not pass the weak identification test. All in all, the instruments of Specification A appear to be stronger.

TABLES 2 AND 3 ABOUT HERE

Table 4 reports the first-stage regressions, which confirm the significance of our instruments. Moreover, SERENDIPITY, INTRATECY, and INVENTOR_BREADTH have the expected sign: SERENDIPITY reduces *R* and *N* because the project is not scaled up after the initial shock; INTRATECY reduces *R* and *N*, while INVENTOR_BREADTH increases them, as suggested in our discussion in the previous section. The IPC3 dummies are, overall, statistically significant, as implied by the F-tests reported below the table. The first-stage regressions produce other reasonable results. As expected, R&D increases both *R* and *N*; working in areas in which the firm does not have a revealed technological advantage raises costs (higher *R*), which is likely to stem from the lack of expertise in the field. Older inventors are more likely to manage larger teams, with implied higher *R*, and government funds raise *R* because of the lower costs of research. More educated inventors are more productive (higher *R*), and they may have more general capabilities, which helps them produce more inventions *N*. Gender does not matter. Finally, a competitive race increases investment *R* because once the company has invested in the innovation, it seeks to beat the competitors; the firm also increases *N* to strengthen protection by creating fences around the core inventions (von Graevenitz et al., 2011; Ziedonis, 2004). More potential competitors, as implied by a higher Herfindahl index of applicants in the four-digit IPC class, reduces *R* and *N*, as expected, because a higher number of potential competitors reduces the prospective returns and dampens investment.

TABLE 4 ABOUT HERE

Turning to our key parameters in Table 2, the estimated elasticity of V with respect to N, $\frac{\alpha(1-\mu)}{\mu}$, is higher than that of V with respect to R, α, which corroborates Proposition 1. In Specification A we estimate $\frac{\alpha(1-\mu)}{\mu}$ = 1.201 and α = 0.423. The estimated μ = 0.261 (p = 0.000) denotes sharp diminishing returns, while the estimated $\frac{\alpha}{\mu}$ = 1.625 (p = 0.000) implies a fairly pronounced complementarity among the technologies in the portfolio (Proposition 2). The results in Table 2 confirm that a competitive race has a positive impact on V, which likely reflects the importance of the patent rights when there is tight competition. The impact of IPC4_COMP is positive too. As noted, this variable reduces investment because of the lower returns when there are more potential competitors. However, the patent right is more valuable, because stronger competition raises the relative returns from holding a patent. Finally, we find evidence of a direct effect of age-based experience on value. The results obtained with Specification B are practically identical.

To test our Proposition 3, we used the five macro-areas in which our 30 technology class dummies are classified (see Table 1). We ran our log-log GMM regression of Equation (6) separately for high-tech (electrical engineering, instruments, chemicals & pharmaceuticals) versus other industries (process engineering, mechanical engineering). Similarly, we split our sample for INFO_SCIENCE or INFO_OTHER = 1 and INFO_SCIENCE = INFO_OTHER = 0 and for PHD = 1 or PHD = 0. The former accounts for technologies that rely on scientific or other sources; the latter accounts for inventors holding a Ph.D.. For high-tech versus other industries and info vs. no info, the sample is almost split in half, whereas inventors holding a Ph.D. constitute approximately one-quarter of the sample.

TABLE 5 ABOUT HERE

Table 5 corroborates our predictions in Proposition 3. The returns of N are systematically lower in high-tech industries, when the invention hinges on relevant information, and when the inventors hold a Ph.D. The results holds under Specification A and B. All the parameters are well measured except the returns of N in the case of Ph.D.s in Specification A. This is not a serious concern though, because of the lower number of observations of Ph.D.s compared with the other cases. The lower returns of N suggest that, other things being equal, in these three cases patent portfolios tend to be more compact. Notice from Table 5 that in all three cases $\frac{\alpha}{\mu}$, which is the sum of the elasticities with respect to R and N, is smaller than in the alternative samples. This means that the higher elasticity of N in non-high tech, no-information or no-PhD samples is produced by a lower μ or stronger diminshing returns. As predicted by our logic in

Section 2, in these cases the lack of technological opportunities, information, or a broader technical or scientific vision on the part of the inventors, entails that the returns produced by persisting investments on individual inventions fade away rapidly. Recall that μ measures the returns to the economic not technical value of the portfolio. Therefore the lower μ likely measures that the individual inventions are not perfectly focused on the goals of the market, and value is produced by the addition of related inventions rather than by attempts to make a specific invention more valuable. As a matter of fact, the higher $\frac{\alpha}{\mu}$ means that in these cases the links between the inventions in the portfolio is tighter, in the sense that, as implied by stronger complementarity, the additional value produced by investing more resources in one invention of the portfolio increases when the inventive team has invested more resources in other inventions of the bundle.

At the same time, in high-tech industries the returns of R are higher than in other industries. Following our simple logic in Section 2.1, this means that in high-tech industries increases in R raise the average value of the portfolio (i.e., keeping N constant) more than in other sectors. In turn, this suggests that, other things being equal, in high-tech industries we tend to observe smaller portfolios of higher average value. We find the same difference when the invention process is informed by science, customers, suppliers or competitors, even though the result holds only for Specification A. In contrast, we find that increases in R lower the average value of the portfolio of inventors holding a doctorate. This may be because these inventors conduct more exploratory research.

Finally, we test Proposition 4 by splitting our sample according to the extent to which firms seek patent protection. We employ two questions in our survey. The first asks the inventor to rank (0–5) whether a key reason for patenting the invention was to protect it from imitation by competitors. The second uses the same 0–5 scale and asks whether the patent aimed to block rivals from developing competing inventions. We distinguish between the sample in which the answer to at least one of these questions is 4 or 5 versus 0–3 and the sample in which they are both 0-3. The logic is that answering 4 or 5 to either one or both questions reflects the importance of protection in the context of this specific patent. We then run our log-log GMM regression for (6) on the two subsamples for each question. We report the corresponding estimated returns of R and N in Table 6.

TABLE 6 ABOUT HERE

As expected there are many more observations in which firms seek patent protection than observations in which the respondents rank prevention from imitation or blocking rivals as 0-3. Nonetheless, we have enough observations in the latter case to provide a meaningful comparison. As a robustness check, in the last two columns of Table 6, we split our sample according to whether the answers to our two questions are 3-5 versus 0-2, and consider only Specification A because the large number of IPC3 dummies combined with the small number of observations prevented the convergence of the estimation. In general, we find that the returns of *N* are higher when patent protection or strategic patenting is more important.

We can also compare the extent to which the returns of *N* depend on protection versus other more genuine technical interdependencies among inventions. If we stick to Specification A, the differences between the returns of *N* in Table 5 are larger than in Table 6. For example, being in high-tech versus other industries changes the returns of *N* from 66.6% to 167.3%, while the largest difference for Specification A in Table 6 is from 85.4% to 118.2%. Differences between the returns of *N* due to protection versus other reasons are less pronounced according to Specification B. However, in other regressions not shown here, the returns of *N* when protection matters are less clear. While we cannot rule out that strategic patenting may have become a more serious concern in recent years, in our sample patent portfolios are not driven mainly by patent protection. Real interdependencies across inventions appear to be an equally if not more important reason.

4.2 Robustness checks

We performed several robustness checks. First, we estimated our key regression in (6) using different controls. The results did not change. Second, as discussed in Section 3.4, we tried alternative instruments for INVENTOR_BREADTH or IPC3. Third, we ran our main GMM regression using the same controls and the number of forward citations up to five years since the patent's grant as dependent variable. As well known, this is a commonly used indirect indicator of value. This variable refers to the focal patent in our survey. Thus, it corresponds to the value of the individual patent, which is the marginal value of the portfolio. The impact of *N* in turn corresponds to the elasticity of the value of the portfolio minus 1. Table 7 reports the estimated returns of *R* and *N* using the sample composed of all observations for which patent citations are available, and then the same sample of the main GMM regression in Table 2. The elasticity of R is similar across all equations. The elasticity of *N* is roughly 0 or slightly higher, which corresponds to an elasticity of 1 or slightly higher than 1 in the portfolio. This is consistent with the results that we obtained using *V*. Proposition 3 predicts that the returns of *N* are higher in non–high-tech industries. As Table 7 shows, this prediction is confirmed when we use citations and Specification A, while the difference between high-tech and non-high tech industries is less pronounced when we use Specification B. We find the same pattern for the no-information and no-PhD samples: Proposition 3 is corroborated for Specification A, and the results are less clear for Specification B. In summary, the effects that we theorized in Proposition 3 are not falsified by the use of a more partial and noisy measure of value.

TABLE 7 ABOUT HERE

Fourth, to address the potential concern that we only employed 5439 observations, while our survey covers 8515 patents held by firms, we ran a probit regression whose dependent variable takes the value 1 if the observation is in our sample of 5439 patents and 0 if it is one of the other patents in the survey. The regressors are the indirect indicators of value employed in Appendix 1, along with the dummies for countries, technological sectors, and priority years. The goal of this robustness check is to assess whether the sample that we used correlated with our indirect indicators of value. None of the indicators was statistically significant, apart from forward citations, which correlated negatively (and significantly at 5%). We then extracted a random sample of 4639 observations, such that the same probit regression run on the 8515 observations exhibited no correlation with any of the indirect indicators of value, including past citations. When we ran our main GMM regression on the sample of 4639 data points, the key results did not change.

5. Discussion and Conclusions

The main contribution of this study has been to distinguish the technical value of the individual inventions from the economic value of a portfolio of related inventions. Typically, increased efforts on an invention raise its technical value; inventions are perfected and made more and more exact from a technical point of view. From an economic perspective, these efforts may not yield—at least after some threshold of technical performance—a significantly higher economic value. Value may arise instead from the combination of inventions that represent different components of a more complex product or process. There are two issues here. First, this finding implies that the resources invested in individual inventions exhibit diminishing returns. After exerting some effort on an invention, firms should redirect their efforts toward other inventions, which in turn suggests that value increases by portfolio size, that is, by raising the number of inventions in the portfolio. Second, this situation does not imply complementarity, which arises instead if effort exerted toward one invention increases the value of the other inventions in the portfolio—the typical situation when inventions as a whole are worth more than their sum. Complementarity further encourages an increase in the number of patents in the portfolio. Both theoretical predictions are confirmed in our empirical study.

Our theoretical and empirical analysis focuses on predictions at the level of patent portfolios related to a particular innovation. While similar issues have been studied for a firm's overall patent portfolio in work by von Graevenitz et al. (2012), we concentrate on determinants of portfolio value and portfolio size and use data at the disaggregate level of the innovation rather than the firm level. We find that the rise in value through the number of patents, as opposed to their average quality, is sizable. In turn, the former effect is composed of a proportionality effect (i.e., the portfolio includes an additional patent), and by a complementarity effect (i.e., greater average value of the portfolio stemming from the synergies among patents). These results suggest that in the invention process, it is not always a good strategy to concentrate resources and to invest in one specific invention. Spreading investments across technically related inventions instead may generate higher valuae.

In addition, the invention process often seems uncertain and subject to many vagaries, and investments in resources matter. In their pioneering study, Hausman et al. (1984) found that R&D is a good predictor of the number of patents produced by companies. We confirm and extend this result. Specifically, along with the sheer investment in resources, the value of a portfolio of inventions can increase if firms spread a given amount of resources across a greater number of related patented inventions. Simply put, while perfecting the technical features of an invention may have a strong technical value, from an economic point of view, it soon leads to sharply diminishing returns. In this case, shifting toward a related but different invention may yield higher economic returns.

We also find that the returns through numbers are less pronounced in high-tech industries, when the inventor team can rely on information from users, competitors, suppliers or science, or when the inventor holds a Ph.D. In all these cases, the effort exerted on a given idea is more productive. High-tech industries entail more technological opportunities, so there are more technical returns to exploit from investing in a given idea. Similarly, the availability of information directs research toward more fertile domains, and Ph.D.s control a broader knowledge base that facilitates their search for fruitful opportunities and recombination of knowledge.[7]. One way of thinking about our results holds that in high-tech industries, when information is available, or when inventors hold a doctorate, there are greater chances to increase returns on individual inventions because of the advantages associated with technological opportunities, information, or education. However, when these factors are missing, firms, and their inventors, can partially compensate for their effect by increasing the number of patents in the portfolio. In short, portfolio size can partly offset the lack of opportunities from depth. Our analysis is silent about the relationships among patent portfolios within the same firm though, so we cannot say whether or to what extent technological opportunities, information, or education provide better options to obtain more valuable portfolios (i.e., firms' "portfolios of portfolios"). Instead, we can say that if the firm cannot lever these factors, it can limit its relative disadvantage by increasing the size of the individual portfolios by creating related, differentiated inventions.

The returns of portfolio size increase if the firm seeks patent protection. This finding suggests that increases in portfolio size are also associated with increases in the value of patent rights. However, we also find that, at least within the framework of our sample, the concern that this may be the leading cause

[7] See Gruber et al. (2012) for an empirical study of an invention's recombination breadth. The study shows that scientists are more likely to generate broad inventions than engineers, and that a doctoral degree is associated with increased recombination breadth for all groups of inventors.

of larger patent portfolios may be exaggerated. While things may have changed in more recent years, our estimations suggest that the factors discussed earlier—technological opportunities, availability of information, education of the inventor—have a stronger effect on the size of patent portfolios than protection. A precise comparison between these two reasons is behind the scope and the possibility (especially in terms of available data) of this paper, and it is clearly a good topic for future research. However, we can fairly conclude that the generation of technically related inventions appears to be at least as important for creating economic value from patent portfolios as the search for protection.

This study also suffers several limitations. As we noted previously, we could have developed alternative measures, rather than the self-reported number of technically related patents, though our measure also may not be any worse or even better than metrics derived from patent register data. A potentially more serious limitation is that we lack information about the characteristics of the other patents (and inventors) in the portfolio. Additional research should collect specific information about actual patents that constitute the portfolio. Nor do we have any measures of capital assets for producing patents; so that we have to rely on company characteristics for this measure. It is therefore not clear that our estimates of the elasticities of work months refer specifically to work months, or if instead they combine the effects of other assets needed to produce patented inventions. An additional, obvious limitation is that we rely on a cross-section. Thus, our identification strategy rests on the validity of our instruments and the implied *ad hoc* exclusion restrictions. Although we test the validity of these instruments, additional research could reassess our results using panel data and fixed effects. Finally, we assume that all the patents in the portfolio are either complementary or substitutes, as implied by their interrelations through the same parameter $\frac{\alpha}{\mu}$. In contrast, some patents in the portfolio may be substitutes, while others are complements; generally, $\frac{\alpha}{\mu}$ could differ across the subset of patents.

The aforementioned limitations arise largely because our contribution seeks to cast light on a firm's decision-making regarding the composition and value of patent portfolios at the level of innovations. We are not aware of any other study that has attempted to do that. In that regard, this study should be seen as a complement to analyses at the firm level and to descriptive accounts which have used classifications such as "discrete" vs. "complex" technologies to highlight important differences in patenting behavior between firms (Cohen et al. 2000). We look forward to additional contributions that will resolve the concerns we could not addressed here and extend our knowledge of patent portfolios.

References

Arora, A., Ceccagnoli, M., and W. Cohen (2008). "R&D and the Patent Premium," International Journal of Industrial Organization 26 (5), 1153-1179

Arora, A. and A. Gambardella (1994). "Evaluating Technological Information and Utilizing it," *Journal of Economic Behavior and Organization* 24, 91-114.

Arora, A., Fosfuri, A., and A. Gambardella (2001). Markets for Technology. MIT Press, Cambridge MA

Astebro, T. (2003). "The Return to Independent Invention: Evidence of Unrealistic Optimism, Risk Seeking or Skewness Loving?" *Economic Journal* 113 (484), 226-239.

Astebro, T and K. Dahlin (2005). "Opportunity Knocks," *Research Policy* 34 (9), 1404-1418.

Bessen, J. (2008). "The Value of US Patents by Owner and Patent Characteristics", *Research Policy* 37 (5), 932–945.

Bessen, J. (2009). "Estimates of Patent Rents from Firm Market Value," *Research Policy* 38(10), 1604-1616.

Cohen, W.M., R. R. Nelson, and J. P. Walsh (2000). "Protecting their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)," NBER Working Paper 7552, Cambridge, Mass.

Fleming, L. and O. Sorenson (2004). "Science as a Map in Technological Search," *Strategic Management Journal* 25 (8-9), 909-928

Gambardella, A., Harhoff, D., and B. Verspagen (2008). "The Value of European Patents," *European Management Review* 5(2), 69-84.

Gittelman, M. and B. Kogut (2003). "Does Good Science Lead to Valuable Knowledge? Biotechnology Firms and the Evolutionary Logic of Citation Patterns," *Management Science* 49 (4), 366-382.

Giuri, P., Mariani, M., Brusoni, S., Crespi, G., Francoz, D., Gambardella, A., Garcia-Fontes, W., Geuna, A., Gonzales, R., Harhoff, D., Hoisl, K., Lebas, C., Luzzi, A., Magazzini, L., Nesta, L., Nomaler, O., Palomeras, N., Patel, P., Romanelli, M., and B. Verspagen (2007). "Inventor and Invention Processes in Europe. Results from the PatVal-EU Survey," *Research Policy* 36 (8), 1107-1127.

Greenberg, G. (2012). "Small Firms, Big Patents? Estimating Patent Value Using Data on Israeli Start-Ups' Financing Rounds," draft, Bocconi University

Gruber, M., D. Harhoff and K. Hoisl (2012). "Knowledge Recombination across Technological Boundaries: Scientists vs. Engineers", *Management Science, forthcoming.*

Hall, B., Jaffe, A. and M. Trajtenberg (2005). "Market Value and Patent Citations," *RAND Journal of Economics* 36, 16-38.

Harhoff, D., Narin, F, Scherer, F., and K. Vopel (1999). "Citation Frequency and the Value of Patented Innovation," *Review of Economics and Statistics*, 81 (3), 511-515.

Harhoff, D. und M. Reitzig (2001): Strategien zur Gewinnmaximierung bei der Anmeldung von Patenten, *Zeitschrift für Betriebswirtschaft*, 5, 509-529.

Harhoff, D., Scherer, F., and K. Vopel (2003). "Citations, Family Size, Opposition and the Value of Patent Rights - Evidence from Germany," *Research Policy*, 32, 1343-1363.

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

Hsu, D. and R. Ziedonis (2012). "Resources as Dual Sources of Advantage: Implications for Valuing

Entrepreneurial-Firm Patents." *Strategic Management Journal*, forthcoming

Knight, H.J. (2001): *Patent Strategy for Researchers and Research Managers*, 2. Aufl., Chichester, Wiley.

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

Lev, B. (2001): *Intangibles: management, measurement and reporting*, Washington D.C., Brookings Institution.

Rosenberg, N. (1982). *Inside the Black Box: Technology and Economics*, Cambridge University Press, Cambridge UK.

Schankerman, M. and A. Pakes (1986). "Estimates of the Value of Patent Rights in European Countries during the Post-1950 Period," *Economic Journal*, 97, 1-25.

Scherer, F. and D. Harhoff (2000). "Policy Implications for a World with Skew-Distributed Returns to Innovation," *Research Policy,* 29, 559-566.

Serrano, C. (2010). "The dynamics of the transfer and renewal of patents," *RAND Journal of Economics*, 41(4), 686-708.

Teece, D. (1986). "Profiting from Technological Innovation," *Research Policy* 15 (6), 285-305

Toivanen, O. and L. Väänänen (2012). "Returns to Inventors," *Review of Economics and Statistics*, forthcoming.

Trajtenberg, M. (1990). "A Penny for Your Quotes: Patent Citations and the Value of Inventions," *RAND Journal of Economics*, 21, 172-187.

von Graevenitz, G., Wagner, S. and Harhoff, D. (2011). "How to Measure Patent Thickets: A Novel Approach," *Economic Letters* 111 (1), 6-9.

von Graevenitz, G., S. Wagner and D. Harhoff (2012). "Incidence and Growth of Patent Thickets - The Impact of Technological Opportunities and Complexity," *Journal of Industrial Economics* (forthcoming).

Ziedonis, R.H. (2004). "Don't Fence Me In: Fragmented Markets for Technology and the Patent Acquisition Strategies of Firms," *Management Science* 50(6), 804-820.

Zuniga, M. and Guellec, D. (2009). "Who licenses out patents and why? Lessons from a business survey". Organisation for Economic Co-ordination and Development, STI Working Paper 2009/5, Paris.

Table 1: Definition of variables and descriptive statistics (^)

Variable, Definition	*Mean*	*SD*	*Min*	*Median*	*Max*
Endogenous Variables					
V: Value of the patent portfolio. Geometric mean of the boundaries of the PatVal-EU classes: ≤€30K; 30-100K; 100-300K; 300K-1M; 1-3M; 3-10M; 10-30M; 30-100M; 100-300M; 300M-1B; 1-3B; ≥3B (3-10B to compute mean of the last class) (in 000 Euros)	67758.1	$4.84 \cdot 10^{5}$	5.477	547.7	$5.48 \cdot 10^{6}$
N: Number of patents that "crucially depend on each other in terms of their value, or in a technical way." N = 1 if patents do not belong to a portfolio, otherwise = geometric mean of the boundaries of the PatVal-EU classes: 2; 3-5; 6-10; 11-20; ≥ 20. (20-40 to compute the mean of the last class)	3.237	5.236	1	1	28.284
R: Total work months invested in the portfolio of patents. Geometric mean of the boundaries of the PatVal-EU classes: ≤1; 1-3; 4-6; 7-12; 13-24; 24-48; 48-72; 72-96; 96-120; ≥120 (120-144 to compute mean of the last class)	25.146	37.237	1	8.485	131.453
Project Characteristics					
COUNTRIES: Dummies for the country in which the first inventor listed in the patent is located (DE, DK, ES, FR, HU, IT, NL, UK)			--		
TECHNOLOGICAL_CLASSES: Dummies for the 30 ISI-INIPI-OST technology classes in which the patents are classified (+)			--		
YEAR: Dummies for the priority year of the patent (1993–97)			--		
GOVFUND: Dummy = 1 if the funding of the research leading to this patent came from Government research programs or related government funds	0.065	0.247	0	0	1
INFO_SCIENCE: Dummy = 1 if the PatVal-EU inventor ranked 4 or 5 (on a 0-5 scale) the importance of any of the following sources of knowledge for the patented invention: university, other non-profit labs, technical conferences, scientific literature	0.484	0.500	0	0	1
INFO_OTHER: Dummy = 1 if the PatVal-EU inventor ranked 4 or 5 (on a 0-5 scale) the importance of customers, suppliers or competitors as sources of knowledge for the patented invention	0.436	0.496	0	0	1
Firm Characteristics					
RD: R&D expenditures of the firm in 1995 (000 Euros)	1591.3	1618.9	0.751	1257.0	8387.9
SALES: Sales of the firm in 1995 (000 Euros)	23174.8	24402.1	3.060	14106.3	$1.65 \cdot 10^{5}$
RTA: Revealed technological advantage of the ultimate parent of the applicant in the technological class of the patent (share of firm patents in class over total patents in class on total patents). Obtained from PatVal-EU data.	11.927	14.350	0.020	8.064	185.143

		Inventor Characteristics			
AGE30: Dummy = 1 if inventor's age < 30	0.046	0.210	0	0	1
AGE30-40: Dummy = 1 if inventor's age is 30-40	0.334	0.472	0	0	1
AGE40-50: Dummy = 1 if inventor's age is 40-50	0.313	0.464	0	0	1
AGE50-60: Dummy = 1 if inventor's age is 50-60	0.258	0.437	0	0	1
AGE60: Dummy = 1 if inventor's age > 60	0.049	0.215	0	0	1
PRECOLLEGE: Dummy = 1 if the inventor has a high-school or lower degree	0.176	0.381	0	0	1
BA/MASTER: Dummy = 1 if the inventor has a BA or Master	0.554	0.497	0	1	1
PHD: Dummy = 1 if the inventor has a Ph.D.	0.270	0.444	0	0	1
MALE: Dummy = 1 if the inventor is a man	0.981	0.137	0	1	1
		Patent Premium Controls			
RACE: Dummy = 1 if inventor declared that the invention had to be patented quickly because other firms or researchers were working on the same idea	0.293	0.455	0	0	1
IPC4_COMP: 1 – HerfindahlHerfindahl index of the share of different applicants in the IPC4 class of the patent	0.935	0.072	0	0.956	0.995
		Excluded Instruments			
SERENDIPITY: Dummy = 1 if, as stated in the formulation of the PatVal-EU question, "the idea for the invention came from pure inspiration or creativity or from your normal job (which is not inventing), and was not further developed in a (research or development) project (and it was patented without further research or development costs)."	0.116	0.320	0	0	1
INTRATECY: Three-year moving average of the interest rate of the country before the priority date of the patent	5.721	2.368	1	5.802	11.729
INVENTOR_BREADTH: 1 – Herfindahl index of the IPC3 classes of the inventor's past patents up to 1 year before the focal patent	0.244	0.328	0	0	0.981
IPC3: Dummy for 71 IPC3 technological classes of the patent (used as an alternative instrument to INVENTOR_BREADTH)			--		

(*) 5439 observations, but RD (2338 observations) and SALES (3248 observations).

(+) The 30 technology classes are divided in 5 macro-classes: (1) Electrical engineering: electrical devices, electrical engineering, electrical energy; audio-visual technology; telecommunications; information technology; semiconductors. (2) Instruments: optics; analysis, measurement, control technology; medical technology; nuclear engineering. (3) Chemicals and pharmaceuticals: organic fine chemistry; macromolecular chemistry, polymers; pharmaceuticals, cosmetics; biotechnology; agriculture, food chemistry; chemical and petrol industry, basic materials chemistry. (4) Process engineering: materials, metallurgy; chemical engineering; surface technology, coating; materials processing, textiles, paper; environmental technology; handling, printing; agricultural and food processing, machinery and apparatus. (5) Mechanical engineering: thermal processes and apparatus; machine tools; engines, pumps, turbines; mechanical elements; transport; space technology, weapons; consumer goods and equipment; civil engineering, building, mining.

Table 2: GMM Estimation of Equation (6), Value, Dependent Variable *V*

Variables	*Specification A (excludes INVENTOR_BREADTH)*	*Specification B (excludes IPC3)*
R	0.423***	0.391***
	(0.042)	(0.035)
N	1.201***	1.172***
	(0.272)	(0.204)
RD	-0.044	-0.015
	(0.053)	(0.045)
SALES	0.060	0.051
	(0.054)	(0.047)
RTA	0.042	0.047*
	(0.032)	(0.026)
AGE30-40	0.257	0.284**
	(0.157)	(0.142)
AGE40-50	0.307*	0.417***
	(0.162)	(0.140)
AGE50-60	0.451***	0.463***
	(0.156)	(0.143)
AGE60	0.407*	0.596***
	(0.231)	(0.191)
BA/MASTER	0.055	0.086
	(0.095)	(0.082)
PHD	-0.055	0.065
	(0.136)	(0.114)
MALE	0.022	0.140
	(0.231)	(0.207)
INFO_SCIENCE	0.036	0.063
	(0.075)	(0.068)
INFO_OTHER	0.048	0.029
	(0.074)	(0.062)
GOVFUND	-0.038	0.018
	(0.139)	(0.121)
RACE	0.300***	0.292***
	(0.101)	(0.082)
IPC4_COMP	1.505*	1.804***
	(0.881)	(0.649)
Constant	3.808***	3.436***
	(0.754)	(0.604)
# obs.	5439	5439

Notes: Log-log functional form: all variables, but dummies, in logs; log(1+variable) if variable can take value 0; robust standard errors in parenthesis, clustered by ultimate parent company (1801 clusters); equations include country dummies, 30 dummies for technological sectors, dummies for patent priority years, and weights to adjust the oversampling of important patents in the survey. Weights = inverse of the relative shares of important patents cited at least once or opposed in the sample and in the population of EPO patents with the same priority dates as the surveyed patents ("important" patents in the survey = cited at least once or opposed). *R*, *N* endogenous. Excluded instruments: SERENDIPITY, INTRATECY, and INVENTOR_BREADTH (in Specification A) or IPC3 (in Specification B). Missing SALES or RD coded as 1 (logs=0), and regressions include two dummies = 1 when SALES or RD are missing, and 0 otherwise. ***$p < 0.001$, **$p < 0.05$, *$p < 0.10$

Table 3: Tests for strong and relevant instrument, Specifications A and B in Table 2

	Specification A	***Specification B***
UNDERIDENTIFICATION TEST		
Null hypothesis: matrix of reduced form coefficients has rank = # endogenous regressors – 1 *(equation is underidentified)*		
Kleibergen-Paap rk LM statistic	89.78	119.0
p-value (§)	0.000	0.000
WEAK IDENTIFICATION TEST		
Null hypothesis: equation is weakly identified		
Kleibergen-Paap Wald rk F statistic	31.12	2.489
Stock-Yogo weak ID test critical values: 10% maximal IV size (+)	13.43	134.32
OVERIDENTIFICATION TEST OF ALL INSTRUMENTS		
Hansen J statistic	0.271	77.30
p-value (*)	0.603	0.257
F-STATISTICS FOR THE IMPACT OF THE EXCLUDED INSTRUMENTS IN THE FIRST STAGE REGRESSIONS (TABLE 4) (^)		
R-equation	376.49	41.51
N-equation	43.83	3.46

(§) Under the null hypothesis the Kleibergen-Paap statistic is distributed as $\chi^2_{(2)}$ (equation A) or $\chi^2_{(71)}$ (equation B)

(+) Critical values are for Cragg-Donald F statistic and i.i.d. errors

(*) Under the null hypothesis the Hansen J statistic is distributed as $\chi^2_{(1)}$ (equation A) or $\chi^2_{(70)}$ (equation B)

(^) Degrees of freedom based on 1801 clusters. F(3, 1800) (Specification A), F(72, 1800) (Specification B)

Table 4: First-stage regressions, Specifications A and B in Table 2

Variables	*R, specification A*	*N, specification A*	*R, specification B*	*N, specification B*
RD	0.058*	0.049**	0.058*	0.057**
	(0.033)	(0.022)	(0.033)	(0.023)
SALES	-0.040	-0.020	-0.039	-0.028
	(0.026)	(0.020)	(0.028)	(0.021)
RTA	-0.037**	-0.000	-0.040**	-0.009
	(0.017)	(0.014)	(0.017)	(0.015)
AGE30-40	0.128	0.010	0.152*	0.047
	(0.078)	(0.057)	(0.078)	(0.057)
AGE40-50	0.144*	0.042	0.201**	0.116*
	(0.085)	(0.063)	(0.084)	(0.063)
AGE50-60	0.191**	0.062	0.261***	0.156**
	(0.084)	(0.065)	(0.082)	(0.066)
AGE60	0.263**	0.097	0.355***	0.207**
	(0.114)	(0.090)	(0.113)	(0.090)
BA/MASTER	0.202***	0.054*	0.227***	0.077**
	(0.046)	(0.031)	(0.046)	(0.031)
PHD	0.283***	0.198***	0.332***	0.242***
	(0.061)	(0.043)	(0.060)	(0.044)
MALE	0.024	0.086	0.045	0.140*
	(0.116)	(0.081)	(0.117)	(0.079)
INFO_SCIENCE	0.187***	0.032	0.189***	0.041
	(0.039)	(0.027)	(0.039)	(0.027)
INFO_OTHER	0.036	-0.013	0.033	-0.019
	(0.037)	(0.030)	(0.038)	(0.030)
GOVFUND	0.321***	0.073	0.312***	0.076
	(0.076)	(0.057)	(0.078)	(0.060)
RACE	0.244***	0.233***	0.251***	0.238***
	(0.044)	(0.036)	(0.044)	(0.037)
IPC4_COMP	-0.921**	-0.720*	-1.419***	-0.955**
	(0.446)	(0.427)	(0.490)	(0.418)
SERENDIPITY	-0.284***	-0.159***	-0.300***	-0.178***
	(0.055)	(0.041)	(0.054)	(0.042)
INTRATECY	-4.846***	-0.252***	-4.800***	-0.233***
	(0.155)	(0.082)	(0.160)	(0.084)
INVENTOR_BREADTH	0.500***	0.582***	--	--
	(0.089)	(0.060)		
IPC3	NO	NO	YES(^)	YES(^)
Constant	11.283***	0.959**		
	(0.549)	(0.374)		
# obs.	5439	5439	5439	5439

Notes. See Table 2. (^) F-tests for exclusion of the IPC3 dummies (p-values in parenthesis): R-eq. $F(71, 1800)=2.01$ (0.000); N-eq. $F(71, 1800) = 2.46$ (0.000). ****p* < 0.001, ***p* < 0.05, **p* < 0.10

Table 5: GMM Estimation of Equation (6), Dependent Variable *V*, Alternative Samples, Specifications A and B in Table 2

	High-Tech Industries	*Other Industries*	*Info*	*No Info*	*PhD*	*No PhD*
			Specification A			
R	0.488***	0.367***	0.457***	0.388***	0.358***	0.447***
	(0.054)	(0.065)	(0.058)	(0.074)	(0.089)	(0.055)
N	0.666*	1.673***	0.935***	1.587***	0.482	1.512***
	(0.350)	(0.416)	(0.353)	(0.428)	(0.391)	(0.372)
			Specification B			
M	0.465***	0.375***	0.435***	0.437***	0.378***	0.466***
	(0.054)	(0.051)	(0.038)	(0.059)	(0.043)	(0.047)
N	0.770***	1.310***	0.833***	1.199***	0.790***	1.086***
	(0.289)	(0.295)	(0.148)	(0.258)	(0.131)	(0.254)
# obs.	2530	2909	2632	2807	1469	3970

Notes: Same log-log regressions as in Table 2. Robust standard errors in parentheses, clustered by ultimate parent company (1801 clusters). *High-tech industries*: Electrical Engineering, Instruments, Chemicals & Pharmaceuticals; *Other industries*: Process Engineering, Mechanical Engineering. *Info/No Info*: Sample with INFO_SCIENCE or INFO_OTHER = 1 vs. INFO_SCIENCE=INFO_OTHER=0. *PhD*, *No PhD*: Sample with PHD = 1 vs. 0. ***p < 0.001, **p < 0.05, *p < 0.10.

Table 6: GMM Estimation of Equation (6), Dependent Variable *V*, Protection

	Specification A		*Specification B*		*Specification A*	
Variables	*Protection*	*Other Patents*	*Protection*	*Other patents*	*Protection*	*Other Patents*
R	0.401***	0.501***	0.370***	0.535***	0.415***	0.412***
	(0.047)	(0.113)	(0.038)	(0.069)	(0.043)	(0.140)
N	1.126***	1.052	0.907***	0.486**	1.182***	0.854
	(0.277)	(1.214)	(0.182)	(0.191)	(0.272)	(1.353)
# obs.	4494	945	4494	945	4946	493

Notes: Same log-log regressions as in Table 2. Robust standard errors in parenthesis, clustered by ultimate parent company (1801 clusters). *Protection* = respondent checked 4–5 (versus 0–3) to *Prevention from Imitation* or *Blocking Rivals* as reasons for patenting the invention. In the final two columns *Protection* defined as checking 3-5 rather than 4-5 to the survey questions. Specification B not available in this case because the high number of IPC3 dummies relatively to the number of observations prevented the estimation from converging. ***p < 0.001, **p < 0.05, *p < 0.10.

Table 7: GMM Estimation of Equation (6) using Forward Citations as Dependent Variable

	All available obs.		***GMM sample***	
	All industries	*Non High-Tech*	*All industries*	*Non High-Tech*
		Specification A		
R	0.037***	0.026*	0.032***	0.029*
	(0.011)	(0.015)	(0.012)	(0.015)
N	0.071	0.219**	0.044	0.137
	(0.069)	(0.100)	(0.072)	(0.100)
# obs.	6648	3556	5439	2909
		Specification B		
R	0.034***	0.026**	0.032***	0.029**
	(0.010)	(0.013)	(0.010)	(0.014)
N	0.208***	0.233***	0.173***	0.173**
	(0.053)	(0.066)	(0.062)	(0.084)
# obs.	6648	3556	5439	2909

Notes: Same log-log regressions as in Table 2. Robust standard errors in parenthesis, clustered by ultimate parent company (1801 clusters). The first set uses as the dependent variable the number of forward citations of the patent up to five years after the priority date; the second set uses the number of equivalent patents in different patent offices. GMM sample = sample from Table 2. *Non High-Tech* = Process Engineering and Mechanical Engineering. ***$p < 0.001$, **$p < 0.05$, *$p < 0.10$.

APPENDIX 1: Correlation of the average value of the patent portfolio and indirect indicators

Variable	Estimates	Variable	Estimates	Variable	Estimates
CITES	0.274***	OP	0.259***	CONSTANT	5.147***
	(0.046)		(0.096)		(0.202)
CLAIMS	0.149***	ACCEX	0.196*		
	(0.006)		(0.108)	Statistics	
STATES	0.220***	PCT	0.234***	# obs.	6850
	(0.061)		(0.072)	Adjusted R^2	0.102
EQUIVALENTS	0.271***	OBS3PARTY	0.896**		
	(0.051)		(0.381)		

Notes: Log-log functional form: All variables, but dummies, in logs; log(1+variable) if variable can take value 0. Dependent variable: *V/N*. Robust standard errors in parenthesis; equation includes country dummies, 30 dummies for technological sectors, dummies for patent priority years, and weights to adjust the oversampling of important patents in PatVal-EU. (Weights, see Table 2.) CITES = number of forward citations up to five years after the priority date; CLAIMS = number of claims at grant; STATES = number of EPO countries in which the patent has been applied for; EQUIVALENTS = number of equivalent patents; OP, ACCEX, PCT, OBS3PARTY = dummies = 1 if the patent was opposed, the applicant requested an accelerated examination procedure, the patent is a PCT, the patent was subject to observations by third parties before the grant (according to art. 115 of the EPC). All these indicators refer to the focal PatVal-EU patent. Regression employs all the available observations (6850). Results are similar if we use the 5439 observations in Table 2. ***$p < 0.001$, **$p < 0.05$, *$p < 0.10$;