IFN Working Paper No. 1067, 2015

Measuring Innovation Using Patent Data

Roger Svensson

Research Institute of Industrial Economics
P.O. Box 55665
SE-102 15 Stockholm, Sweden
info@ifn.se
www.ifn.se

Measuring Innovation Using Patent Data

Roger Svensson [*]

April 7, 2015

Keywords: Patents, commercialization, innovations, profitability, patent renewal, patent equivalents, forward citations, predicted probabilities

JEL: O31, O34

Abstract

Firms and governments spend billions of dollars on R&D every year. To increase social welfare, the results of R&D must be commercialized so that consumers can benefit from improved products and lower prices. One measure of R&D output is patents; however, most patent databases contain no information on whether patents have been commercialized, i.e., whether innovations have been introduced in the market. This paper applies a new method to identify innovations in patent databases by relating traditional patent quality indicators (patent renewal, patent equivalents and forward citations) to patent commercialization variables. For this purpose, I use a unique database on Swedish patents that includes information on whether patents are commercialized and whether the commercialization is profitable. The estimations show that commercialization is strongly positively correlated with both patent renewal and patent equivalents but only moderately positively correlated with forward citations. Further, successful innovations are most positively related to patent renewal. Based on the traditional patent quality indicators and estimated parameters in the model, probabilities of commercialization and successful innovations can be predicted. The developed parameters may be used to identify innovations across sectors and regions in other patent databases.

---

[*] Research Institute of Industrial Economics (IFN), P.O. Box 55665, SE-102 15 Stockholm, Sweden. Correspondence: roger.svensson@ifn.se.

The author wishes to thank participants at the 4th International Workshop on Entrepreneurship, Culture, Finance, and Economic Development (ECFED) in Klagenfurt for insightful comments. Financial support from Vinnova, Statistics Sweden and the Jan Wallander and Tom Hedelius Foundation is gratefully acknowledged.

1. Introduction

Several studies in economics have shown that new technology/knowledge is among the most important factors for economic growth (see Hall *et al*. 2010 for an overview). Theoretically, this finding is explained by the non-rival, non-depreciating and cumulative characteristics of technology/knowledge. However, in a free market, firms will under-invest in R&D because of spillover effects (Arrow 1962) and incomplete capital markets (Kaplan and Strömberg 2001). These market failures motivate the governments to finance R&D and stimulate innovative activities among businesses, universities and government research institutes.

Firms and governments spend billions of dollars on R&D every year, but to increase wealth, the results of R&D must be commercialized so that consumers can benefit from improved products, greater product selection and lower prices. Policy makers today have little specific information about where new technologies and ideas are commercialized; hence, they often find it difficult to identify which sectors and regions to target with public interventions in order to stimulate innovative activities.

Normally, R&D intensity and the number of patents (applied or granted) per capita are used as representative indicators for ranking sectors and countries in terms of technology intensity or innovation. R&D is a measure of the *input* of innovative activities, while patents are a measure of the *output* of innovative activities. A weakness of earlier patent databases and patent studies (especially studies using patent counts) is that they cannot identify which patents have been commercialized, i.e., which specific patents have been introduced as innovations in the market. Consequently, prior research has instead relied on several quality indicators to identify which patents are deemed valuable (see van Zeebroeck 2011 for an overview). Among these indicators are patent renewal, patent family size (patent equivalents), forward citations and oppositions. For example, many studies analyzing patent renewal patterns conclude that only a few patents have high economic value. In practice, most patents expire before the age of ten years (see Pakes and Schankerman 1984, and successive studies).

The OECD and several countries have sought to diminish if not overcome the limitations of traditional techniques for measuring innovation by distributing questionnaires to firms to obtain better information about firms' innovative activities. An exemplar of these surveys is the Community Innovation Survey (CIS), one among many similar surveys (Gault 2013). A clear disadvantage of such surveys is that firms themselves define whether and how they have created innovations. Since innovation carries a positive connotation, firms tend to overestimate their innovations in these self-reporting surveys.

This study aims to predict the probabilities of innovations and successful innovations based on patent quality indicators. The key is here to relate traditional patent quality indicators to patent commercialization variables. This has not been undertaken before. For this purpose, a unique database on Swedish patents owned by small firms and individuals is used. The database contains information on *whether*, *when* and *how* (existing firm, new firm, licensed or sold) patents were commercialized, as well as whether the commercialization was profitable. No other patent database in the world has such richly detailed information on the commercialization of patents. Here, this commercialization process refers to the introduction of an innovation in the market. The database also contains traditional patent quality variables, such as patent renewal, forward citations and patent equivalents (family size).[1]

The following methodology is adopted in this study. First, statistical correlations between the traditional patent quality variables and the commercialization variables are estimated. Next, I proceed to estimate parameters for how innovations are related to the traditional patent quality indicators by using different qualitative response models. Subsequently, based on the traditional patent value indicators and estimated parameters from the model, probabilities of innovations and successful innovations are predicted. These formulas could then be applied to other patent databases that include traditional patent quality indicators.

The estimations show that commercialization is strongly positively correlated with both patent renewal and patent equivalents but moderately positively correlated with

---

[1] Unfortunately, oppositions are not available for Swedish patents.

forward citations. Further, successful innovations are most positively related to patent renewal.

The paper is organized as follows. The database is described in section 2. In section 3, I discuss different measures of innovations and patent quality and estimate statistical correlations between commercialization and traditional patent value indicators. In section 4, the statistical methods for relating patent quality indicators to innovations are specified. The results of the estimations are then presented in section 5, and the final section summarizes the conclusions.

2. Database

I use a detailed data set of patents granted to small firms (fewer than 1,000 employees) and individual inventors based on a survey conducted in 2003−04 on Swedish patents granted in 1998. In that year, 1,082 patents were granted to Swedish small firms and individuals.[2] Information about the inventors, the applying firms and their addresses, as well as filing dates for each patent, was collected from the Swedish Patent and Registration Office (PRV). Thereafter, a questionnaire was sent out to the patent inventors, and 867 of the 1,082 inventors (80 percent response rate) completed and returned the questionnaire. Non-response in the sample is not systematically related to either International Patent Classification (IPC) or geographical region.[3]

The questionnaire asked the inventors about the type of work place where the invention was created; whether, when and how the patent had been commercialized; the profitability of the commercialization; and miscellaneous information regarding inventor characteristics. The data set was later complemented with data on patent renewal, patent equivalents, forward citations and filing routes from the Espacenet

---

[2] In 1998, 2,760 patents were granted in Sweden. Of these patents, 776 were granted to foreign firms, 902 to large Swedish firms with more than 1,000 employees, and 1,082 to Swedish individuals or firms with fewer than 1,000 employees. In a pilot survey conducted in 2002, large Swedish firms refused to provide information on individual patents. Furthermore, persuading foreign firms, which were generally large multinationals, to answer questionnaires about patents proved very difficult. The sample selection for the present data is nevertheless not problematic because the conclusions are drawn for small firms and individual inventors located in Sweden.

[3] Of the 20 percent of non-respondents, 10 percent of the inventors had outdated addresses, 5 percent had correct addresses but did not respond, and the remaining 5 percent refused to participate. The only information that I have for non-respondents is the IPC class of the patent and the region of the inventors. For these variables, there was no systematic difference between respondents and non-respondents.

(2014) website. Thus, the database includes information on several traditional patent quality indicators.

Table 1. Commercialization of patents across firm sizes, number of patents and percent.

Kind of firm where the invention was created	Commercialization		Total	Percent commercialized
	Yes	No		
Medium-sized firms (101−1000 employees)	77	39	116	66 %
Small firms (11−100 employees)	137	64	201	68 %
Micro companies (2−10 employees)	105	37	142	74 %
Inventors (1−4 inventors)	207	201	408	51 %
Total	526	341	867	61 %

In the present study, commercialization is defined to indicate that an innovation based on a patent has been introduced in the market—by the inventor, the inventing firm or an external firm that has licensed or acquired the patent. The 867 patents and the commercialization rate for the patents are described across firm groups in Table 1. As many as 408 patents (47 percent) were granted to individual inventors, and 116, 201 and 142 patents were granted to medium-sized firms (101–1,000 employees), small firms (11–100 employees) and micro companies (2–10 employees), respectively.[4] The commercialization rate for the whole sample is 61 percent. For comparison, in the few available studies that have measured the commercialization of patents, Morgan *et al.* (2001) found a commercialization rate of 47 percent for American patents, whereas Griliches (1990) found a commercialization rate of 55 percent.[5] The higher commercialization rate in the present study likely results from the focus on only patents owned by small firms and individual inventors, as large (multinational) firms have many more defensive patents than small firms. In line with this explanation, Griliches (1990) reports a commercialization rate as high as 71 percent for small firms and inventors. As shown in Table 1, the commercialization rate for firm groups is between 66 and 74 percent, whereas the commercialization rate for individuals is 51 percent. A contingency table test suggests that this difference in the

---

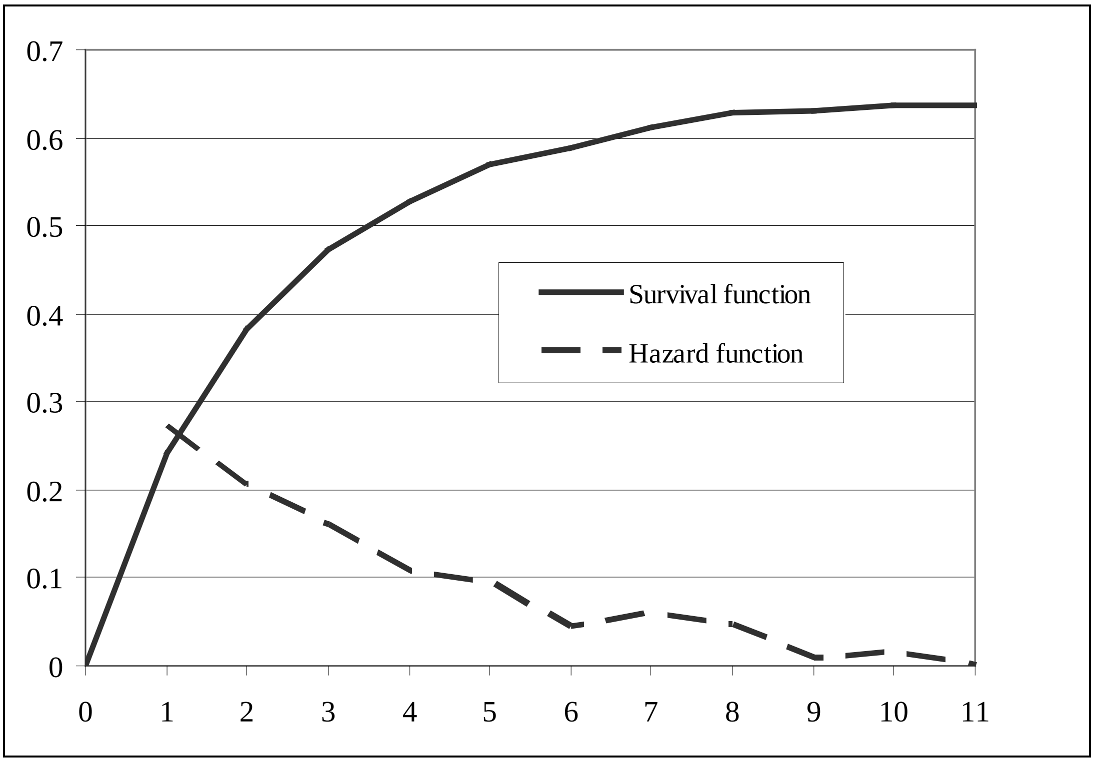
[4] The group of individual inventors includes private persons, self-employed inventors, groups of two to three inventors organized in trading companies, and private firms without employees.

[5] These other studies use a similar definition for commercialization, i.e., that the patent has been used commercially. Morgan *et al.* (2001) defines commercialization as the commercialization of a product or process or the granting of a licensing contract, whereas Griliches (1990) defines commercialization as the commercial use of a patent. In neither of these studies does the commercialization of the patent need to be profitable for the owner.

commercialization rate between firms and individuals is statistically significant at the one-percent level, with a chi-square value of 30.55 (with 3 d.f.).

Figure 1 presents the share of commercialized patents and the hazard function for the sample with the patent application year set to 0. As shown, the curve of commercialized patents increases steeply initially and reaches almost 50 percent after 3 years.[6] Thereafter, it levels out at approximately 60 percent, indicating that patent commercialization is most likely to occur within 3–4 years of application. Accordingly, the hazard function, which measures the conditional probability of a patent being commercialized, is the highest during the first three years after the application year.[7]

Figure 1. Survival and hazard functions for commercialization.



At the end point of observation period (year 2003), the inventors were asked to estimate whether the commercialized invention would yield a profit, break even or result in a loss. If they did not know, the reply was registered as a missing value

---

[6] The share of commercialized patents is estimated by using the life-table method (actuarial method) and is equal to 1 minus the survival function. The survival function shows the proportion of patents that have not been commercialized after a specific point in time (= survives).

[7] The hazard function shows the conditional probability of a patent being commercialized in a specific time period given that it has "survived" (has not been commercialized) until the beginning of this period.

(uncertain outcome).[8] In Table 2, discrete values for the outcome in terms of profit across firm groups are presented. It would have been desirable to measure the outcome in terms of money, but such information was impossible to collect.[9] As shown in the table, outcomes differ substantially across firm groups, with the group of individual inventors having the least favorable outcome.

Table 2. Performance of the commercialization across firm groups, number of patents.

Kind of firm where the invention was created	Performance				Total
	Profit	Break-even	Loss	Missing value	
Medium-sized firms	55	18	3	1	77
Small firms	97	24	15	1	137
Micro companies	60	17	27	1	105
Inventors	69	47	87	4	207
Total	281	106	132	7	526

3. Correlations with traditional patent quality indicators

Forward citations. In the patent literature, forward citations have frequently been used as a measure of patent quality or value, even though there is often skepticism about whether forward citations actually measure the private value of patents or spillover effects (Hall *et al*. 2007). Trajtenberg (1990) argues that forward citations indicate the social value of patents. For a particular patent, a higher frequency of citations by later patents is associated with a higher spillover effect and hence higher social value. A patent can be cited at any time after the application date, even after it has expired.

As an indicator of patent quality, the total number of forward citations that *a patent and its patent equivalents* have received during five years after publication is used (van Zeebroeck 2011). Self-citations are excluded from this measure, however. Forward citations are measured in two ways. First, *all forward citations* that a patent and its equivalents have received within five years of publication (as suggested by van

---

[8] For a vast majority of the patents, commercialization had reached a stage such that there was no uncertainty about the patent's performance. In 2007, importation on the profitability of commercialization was updated via phone calls to inventors who had earlier announced an uncertain outcome.

[9] Estimating profit flows is very complicated because most firms have many products in their statement of accounts, and many individual inventors do not have any statement of accounts at all.

Zeebroeck 2011), here called *Citations 1*, is used. Some patent offices, such as the United States Patent and Trademark Office (USPTO), cite patents more frequently than other patent offices. Therefore, the number of forward citations that the patent and its equivalents have received within five years of publication *from PCT applications and patents granted by the European Patent Office (EPO)*, here called *Citations 2*, is used. Table 3 shows that 318 (of 867) patents had forward citations according to the first measure but that only 209 had forward citations from the PCT or EPO. There is some evidence that a positive relationship exists between commercialization and the two measures of forward citations. The Spearman rank correlations between commercialization and the citation variables are 0.14 and 0.07, which are significant at the 1- and 5-percent levels, respectively (Table 7).

Table 3. Relationship between commercialization and forward citations.

	Citations 1, number of forward citations (from all sources)								
Commercialized	0	1	2	3−4	5−6	7−8	9−10	≥11	All
No	244	51	9	13	14	4	3	3	341
Yes	305	82	49	46	19	9	7	9	526
Total	549	133	58	59	33	13	10	12	867

	Citations 2, number of forward citations (from PCT and EPO)								
Commercialized	0	1	2	3−4	5−6	7−8	9−10	≥11	All
No	271	44	15	10	1	0	0	0	341
Yes	387	81	35	17	2	0	2	2	526
Total	658	125	50	27	3	0	2	2	867

Patent equivalents. The number of patent equivalents is an important indicator of the private value of patents (Putnam 1996). Since patent filing and enforcement are costly in many countries, only patents with a sufficiently high expected value are filed in many countries. However, once a patent is filed with any patent office, the patent owner must file patents with other offices within a year to expand the patent rights to other countries (the priority year). Maurseth and Svensson (2014) show that the probability of filing a patent abroad is highly correlated with commercialization, patent renewal and forward citations.

Turning to the filing routes, only eight of 867 patents were first filed abroad, and all of these were in the US. No patent was filed first with the EPO or WIPO and thereafter in Sweden. This pattern markedly contrasts with the filing routes undertaken by Swedish multinationals. Various explanations may account for this result, ranging from the fact that the owners in the database used in this study are individuals and small firms to the fact that the data cover patent filings in the 1990s, when it was still common to first file patents in the home country.

Table 4. Relationship between commercialization and patent equivalents.

	Number of patent equivalents								
Commercialized	0	1−2	3−4	5−6	7−8	9−10	11−12	≥13	All
No	258	34	18	12	9	2	5	3	341
Yes	274	90	45	38	25	19	10	25	526
Total	532	124	63	50	34	21	15	28	867

The 867 patents in the database together have 1,734 patent equivalents abroad, for an average of about two equivalents per patent. The frequency distribution of patent equivalents is shown in Table 4. Only 345 (40 percent) of the 867 patents have at least one equivalent. Moreover, given that a patent has at least one equivalent, the average number of equivalents per patent is 5.0. The maximum number of equivalents for a given patent is 24.

Table 4 shows that patents with many equivalents have a higher probability of being commercialized. Further, the Spearman rank correlation between commercialization and patent equivalents is 0.24, which is significant at the 1-percent level (Table 7).

However, patent equivalents should not be regarded equally, as host countries vary in market size. In the database, foreign patent filings are dominated by the large markets. Triadic patents (i.e., patents that are filed in the three largest patent offices in the world—the EPO, the USPTO and the Japanese patent office) should be especially valuable. There are 79 Triadic patents in the database, and 113 patents were filed in at least two Triadic markets. Moreover, there are 224 equivalents in the US and 141 in Japan, as well as 217 EPO patents. EPO patents must be validated in individual member countries, and EPO patents resulted in 1,104 individual patents in the EPO member countries, for an average of 5.1 individual patents per EPO patent.[10] The EPO patents in the database used in this study are filed most frequently in Germany, Great Britain and France—the large EPO countries.[11] Thus, patent equivalents are not distributed randomly across countries.[12]

Table 5. Relationship between commercialization and patent equivalents in large markets.

Commercialized	EPO patent		US patent		Patent in 2 or 3 Triadic areas		Triadic patent		Total
	No	Yes	No	Yes	No	Yes	No	Yes	
No	289	52	285	56	314	27	322	19	341
Yes	361	165	358	168	440	86	466	60	526
Total	650	217	643	224	754	113	788	79	867
Chi-square test	28.65 ***		26.00 ***		12.98 ***		8.51 ***		

Note: ***, ** and * indicate significance at the 1%-, 5%- and 10%-level, respectively. A Triadic patent means that a patent was granted at EPO, in the US and in Japan.

As shown in Table 5, commercialization is related to equivalents in the largest markets of the world. Indeed, there is a strong positive relationship between commercialization and an EPO equivalent, a US equivalent, equivalents in at least two Triadic markets and a Triadic patent (EPO, US and Japan). The chi-square tests indicate that the relationships are highly significant in all four cases; however, commercialization has a stronger relationship with EPO and US equivalents than with Triadic patents.

Patent renewal. Several previous studies have estimated the private value of patents by using the renewal scheme of patents (see, e.g., Pakes 1986, Schankerman and Pakes 1986, and successive studies).[13] Patent holders must pay an annual fee to keep their patents in force, and this fee increases over time until the maximum life span of 20 years is reached. According to Griliches (1990), rational owners will renew their

---

[10] This average number of equivalents is the same as that for EPO patents in general (van Zeebroeck 2011).

[11] Only 30 equivalents in the database were filed directly at the national patent offices in the EPO area without filing an EPO patent first.

[12] van Zeebroeck and van Pottelsberghe (2011) show that there is a strong positive correlation between market size and the probability that an EPO patent will be validated in a country. The skewed country distribution of patents above indicates that country characteristics are important for international patenting.

[13] Other studies have estimated the private value of patents by 1) asking the patent owners directly about the market value (see, e.g., Rossman and Sanders 1957, Schmookler 1966) or 2) relating firms' profit or market value to patents or innovations (see, e.g., Griliches *et al.* 1987, Hall 1993).

patents only if it is economically profitable to keep them. Further, examining the renewal pattern of patents, Schankerman and Pakes (1986) show that most patents have a low value and depreciate quickly. Only a few patents have a significant high value and last for the maximum period.

Table 6 presents the results regarding the relationship between patent renewal and commercialization for the sample. Overall, 407 patents (47 percent) expired before 10 years, whereas 460 (53 percent) were renewed for at least 10 years. As many as 237 patents (27 percent) were renewed for at least 16 years. The last group of patents (≥16 years) is right censored because some patents were applied in 1998, while the last observed renewal date is 2014. The share of commercialized patents is higher for longer lasting patents. The Spearman rank correlation between patent renewal and commercialization is 0.26, which is clearly significant at the 1-percent level (see Table 7).

Table 6. Relationship between commercialization and renewal of patents.

	Patent renewal, number of years								
Commercialized	2−3	4−5	6−7	8−9	10−11	12−13	14−15	≥16	All
No	35	56	62	49	34	27	30	48	341
Yes	20	40	84	61	44	49	39	189	526
Total	55	96	146	110	78	76	69	237	867

Oppositions. A fourth traditional patent quality indicator addresses whether oppositions have been filed against a granted patent. Oppositions by a third party signal a patent's potential value in a given market. Therefore, oppositions indicate that there is a potential market for the patent and that the patent is important enough to justify the costs and risks associated with a dispute (Lanjouw and Schankerman 1997, Harhoff *et al*. 2002, van Zeebroeck 2011). However, data on oppositions is not available for Swedish patents at Espacenet (2014).

3.1 Correlations – all patents

Table 7 presents simple Spearman rank correlations between commercialization and the traditional patent quality indicators. As shown, commercialization (indicating an innovation) is clearly more strongly correlated with equivalents and patent renewal than with forward citations. Moreover, patent renewal, patent equivalents and forward citations are all positively and significantly correlated with each other.

Table 7. Correlation matrix between commercialization and patent quality indicators, Spearman rank parameters.

Citations 1 (number, all)	0.14 ***			
Citations 2 (number, EPO + PCT)	0.07 **	0.78 ***		
Equivalents (number)	0.24 ***	0.61 ***	0.41 ***	
Renewal (years)	0.26 ***	0.23 ***	0.16 ***	0.40 ***
	Com	*Citations 1*	*Citations 2*	*Equivalents*

Note: n = 867. ***, ** and * indicate significance at the 1%-, 5%- and 10%-level, respectively.

Table 8 reports the results regarding the correlations between the profitability of patent commercialization (three levels as defined in a previous section) and the traditional patent quality indicators. Here, only commercialized patents are included in the analysis. The correlations in Table 8 between the commercialization variable and the traditional patent value indicators are somewhat weaker than those in Table 7.

Table 8. Correlation matrix between profitability of commercialization and patent quality indicators, Spearman rank parameters.

Citations 1 (number, all)	0.10 **			
Citations 2 (number, EPO + PCT)	0.07	0.75 ***		
Equivalents (number)	0.16 ***	0.57 ***	0.37 ***	
Renewal (years)	0.33 ***	0.16 ***	0.12 ***	0.34 ***
	Success	*Citations 1*	*Citations 2*	*Equivalents*

Note: n = 519. ***, ** and * indicate significance at the 1%-, 5%- and 10%-level, respectively.

3.2 Correlations – patents with EPO equivalents

Table 5 shows that 217 patents in the sample were granted an administrative EPO patent. I continue the analysis by examining the relationship between the commercialization variables and the traditional patent quality indicators for this subsample of patents. The Spearman rank correlations between commercialization and the traditional patent quality indicators for EPO patents are shown in Table 9. As the result shows, the correlations between commercialization and the patent quality indicators are clearly weaker for the EPO patent subsample than for the full sample.

Table 9. Correlation matrix between commercialization and patent quality indicators, Spearman rank parameters. EPO subsample.

Citations 1 (number, all)	−0.02			
Citations 2 (number, EPO + PCT)	−0.11 *	0.71 ***		
Equivalents (number)	0.12 *	0.05	0.05	
Renewal (years)	0.06	0.03	0.00	0.37 ***
	Com	*Citations 1*	*Citations 2*	*Equivalents*

Note: n = 217. ***, ** and * indicate significance at the 1%-, 5%- and 10%-level, respectively.

Table 10 then presents the correlations between the profitability of patent commercialization and the traditional patent quality indicators for the EPO subsample. For this analysis, only commercialized patents are included. As shown, successful innovations have a positive and significant relationship with patent renewal only. All other correlations are again weak.

Table 10. Correlation matrix between profitability of commercialization and patent quality indicators, Spearman rank parameters. EPO subsample.

Citations 1 (number, all)	−0.05			
Citations 2 (number, EPO + PCT)	−0.04	0.70 ***		
Equivalents (number)	−0.06	0.06	0.09	
Renewal (years)	0.19 **	0.03	−0.01	0.36 ***
	Success	*Citations 1*	*Citations 2*	*Equivalents*

Note: n = 163. ***, ** and * indicate significance at the 1%-, 5%- and 10%-level, respectively.

The weaker relationships between the commercialization variables and the patent quality indicators for the EPO subsample compared with the full sample may be due to systematic bias in this sample. Indeed, only inventors who think that their patent is valuable would apply for an EPO patent.

4. Estimation techniques and explanatory variables

In this section, I present the estimation techniques that I use to test the relationships between the traditional patent quality indicators and 1) the probability of an innovation (the commercialization of a patent) and 2) the probability of a successful innovation. In both cases, the purpose is to create indexes of and to forecast the probability of innovations and successful innovations, which can later be used for other patent databases.

4.1. Probability of an innovation

The dependent variable, *Com$_i$*, represents whether a patent *i* has been commercialized. It is dichotomous in nature and takes on the value of 1 if an innovation has been introduced in the market and 0 otherwise. Therefore, a probit model, based on the cumulative normal distribution function, is used to predict variation in the dependent variable. The model can be written as:

$$c_i^* = X_i\beta + \nu_i \quad , \tag{8}$$
$$c_i = 1 \;\; if \;\; c_i^* > 0 \;\; and \;\; 0 \;\; otherwise,$$

where c_i* is a latent index; c_i is the selection variable, indicating whether the patent is commercialized; $\mathbf{X_i}$ is a vector of explanatory variables, which influence the probability that the patent is commercialized; $\boldsymbol{\beta}$ is a vector of parameters to be estimated; and $\mathbf{v_i}$ ~ N(0, 1).

4.2. Probability of a successful innovation

The dependent variable, *Success*, in the empirical estimations measures the performance of the commercialization for the original owner of the patent in terms of profit. It can take on three different discrete values, denoted by index *k*:

- Profit, *k*=2;
- Break-even, *k*=1;
- Loss, *k*=0.

Since it is possible to order the three alternatives, an ordered probit model is applied.[14] A multinomial logit model fails to account for the ranking of the outcomes. By contrast, an ordinary regression would treat the outcomes 0, 1 and 2 as realizations of a continuous variable. However, treating the outcomes in this way would be inappropriate because the discrete outcomes are only ranked. The ordered probit model can be described in the following way (Greene 1997):

---

[14] There are seven observations in the database in which the owner could not specify the expected profit level of the patent commercialization. These missing values could be treated as a fourth, uncertain, outcome of *Success*. Thus, I estimate a multinomial logit model in which all four alternatives are included. Then, I perform a test for independence of irrelevant alternatives (Hausmann and McFadden 1984), and when I exclude the uncertain alternative in the multinomial logit model, this test cannot be rejected. Thus, the parameter estimates between the other outcome alternatives are almost unaffected if the uncertain alternative is excluded. Therefore, there is no problem in excluding those patents with unknown profit levels from the estimations.

$$y_i^* = X_i \alpha + \varepsilon_i \quad , \tag{6}$$

where $\mathbf{X_i}$ is a vector of patent quality indicators and technology dummies; $\boldsymbol{\alpha}$ is a vector of coefficients that indicates the influence of the independent variables on the profit level; and $\boldsymbol{\varepsilon_i}$ is a residual vector that represents the combined effects of unobserved random variables and random disturbances. The residuals are assumed to have a normal distribution, and the mean and variance are normalized to 0 and 1. The vector with the latent variable, $\mathbf{y_i^*}$, is unobserved. The model is based on the cumulative normal distribution function, $F(X\alpha)$**,** and is estimated via maximum likelihood procedures. The difference between this model and the two-response probit model is that in this model a parameter (threshold value), ω, is estimated by $\boldsymbol{\alpha}$. The probabilities $P_i(y=k)$ for the three outcomes are:

$$P_i(0) = F(-X\alpha) \quad ,$$

$$P_i(1) = F(\omega - X\alpha) - F(-X\alpha) \quad , \tag{7}$$

$$P_i(2) = 1 - F(\varpi - X\alpha) \quad ,$$

$$\textit{where} \sum_{k=0}^{2} P_i(k) = 1 \quad .$$

The threshold value, ω, must be larger than 0 for all probabilities to be positive.

An objection against the sample and the chosen statistical model may be that the patents, which are commercialized, do not comprise a random sample of patents but instead have specific characteristics that led them to be commercialized in the first place, which could result in misleading parameter estimates. An appropriate statistical model is therefore an ordered probit model with sample selectivity (Greene 2002). In the first step, a probit model estimates how different factors influence the decision to commercialize a patent:

$$d_i^* = X_i \theta + u_i \quad , \tag{8}$$
$$d_i = 1 \;\; \textit{if} \;\; d_i^* > 0 \;\; \textit{and} \;\; 0 \;\; \textit{otherwise},$$

where d_i^* is a latent index; d_i is the selection variable, indicating whether the patent is commercialized; $\mathbf{X_i}$ is a vector of explanatory variables that influence the probability that the patent is commercialized; $\boldsymbol{\theta}$ is a vector of parameters to be estimated; and $\mathbf{u_i}$ is a vector of normally distributed residuals with zero mean and a variance equal to 1.

From the probit estimates, the selection variable d_i is then used to estimate a full information maximum likelihood model of the ordered probit model (Greene 2002).[15] In addition, the first step probit model is re-estimated. The residuals [ε, u] are assumed to have a bivariate standard normal distribution and correlation ρ. There is selectivity if ρ is not equal to zero.

4.3 Explanatory variables

In all the estimations, only explanatory variables that are available in common patent databases are included, as the estimated parameters in this study should be used to predict innovations and successful innovations in other patent databases.

Most of the explanatory variables included in the estimations have already been defined in section 3. These variables are *Citations 1*, *Citations 2*, *Equivalents* and *Renewal*. In addition, additive dummies for EPO, US and Japanese equivalents are included. The traditional patent quality indicators are also squared in some of the estimations to determine whether a nonlinear relationship exists between these quality indicators and *Com*/*Success*. Definitions and descriptive statistics for the dependent and explanatory variables are shown in Appendix A, Table A1.

Since patenting and innovations are known to vary greatly between industries and technology classes (Levin *et al.* 1987), I also include additive dummies for 30 different industry classes designated by Breschi *et al.* (2004). These industry classes are based on the IPC system, and a patent may belong to several different IPC classes. However, it is not possible to determine the main IPC class because the classes are listed in alphabetical order for each patent in Espacenet (2014). Therefore, a patent in the database used in this study may belong to as many as four different industry classes. Consequently, the 30 industry dummies are not mutually exclusive. In some

---

[15] Note that this specification is not a two-step Heckman model. No lambda is computed and used in the second step.

of the estimations, I have to reduce the number of industry classes because of the limited number of observations in each class. For example, only 25 classes are included when I estimate the ordered probit model, and only 22 classes are used for the estimation with the EPO subsample.

5. Results of the estimations

5.1 All patents

The results of the probit estimations are shown in Table 11. Between 63 and 66 percent of the observations are correctly predicted with respect to commercialization (*Com*). Several variants of the model are estimated. For example, industry classes are included, forward citations are alternatively represented by *Citations 1* and *Citations 2* (all citations vs. only citations from the EPO and PCT), and equivalents are measured as the total number of equivalents or as additive dummies for the Triadic market equivalents.

Table 11. Results of the probit estimations.

Dependent variable = *Com*

Statistical model = Probit model

Explanatory variable	A	B	C	D	E	F	G	H
Citations 1	0.016	0.024	8.9 E−3	0.017				
	(0.018)	(0.020)	(0.019)	(0.020)				
Citations 2					0.015	0.034	7.6 E−3	0.022
					(0.044)	(0.049)	(0.044)	(0.047)
Equivalents	0.052***	0.061***			0.055***	0.063***		
	(0.016)	(0.017)			(0.016)	(0.017)		
Eq. EPO			0.149	0.149			0.153	0.154
			(0.135)	(0.140)			(0.135)	(0.141)
Eq. US			0.224	0.213			0.241*	0.242*
			(0.137)	(0.143)			(0.132)	(0.138)
Eq. Japan			0.107	0.235			0.106	0.227
			(0.150)	(0.162)			(0.151)	(0.162)
Renewal	0.055***	0.058***	0.058***	0.062***	0.055***	0.059***	0.058***	0.062***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Industry classes	No	Yes	No	Yes	No	Yes	No	Yes
Log likelihood	−545.3	−531.3	−546.5	−532.4	−545.7	−531.8	−546.6	−532.7
Share of correct predictions	64.0	65.5	64.8	65.9	63.6	65.5	64.7	65.9

Note: n=867. Standard errors are in parentheses. ***, ** and * indicate significance at the 1%-, 5%- and 10%-level, respectively. Parameter estimates of intercept and industry class dummies are not reported, but are available from the author on request.

Both the number of equivalents and the length of patent renewal have a strong positive relationship with the probability of commercialization, and the estimated parameters are significant at the 1-percent level (*Renewal* in all models and *Equivalents* in Models A, B, E and F). However, when the number of equivalents is substituted for dummies for EPO, US and Japanese equivalents, the parameters of these dummies are barely significant (Models C, D, G and H).[16] Notably, the estimated parameters of forward citations are never significant. The marginal effects of the explanatory variables on the probability of an innovation (calculated around the means of the X's) are shown in Table 12. As shown, if one more equivalent is filed, then the probability of an innovation increases by 2.0−2.4 percentage points. If the patent is renewed for one more year, the probability of an innovation increases by 2.1−2.4 percentage points.

Table 12. Marginal effects on the probability of an innovation.

Dependent variable = *Com*

Statistical model = Probit model

Explanatory variable	A	B	C	D	E	F	G	H
Citations 1	6.1 E-3	9.2 E-3	3.4 E-3	6.4 E-3				
	(6.9 E-3)	(7.6 E-3)	(7.1 E-3)	(7.7 E-3)				
Citations 2					5.9 E-3	0.013	2.9 E-3	8.5 E-3
					(0.017)	(0.019)	(0.017)	(0.018)
Equivalents	0.020***	0.023***			0.021***	0.024***		
	(5.9 E-3)	(6.4 E-3)			(5.9 E-3)	(6.3 E-3)		
Eq. EPO			0.056	0.056			0.058	0.058
			(0.050)	(0.052)			(0.050)	(0.052)
Eq. US			0.084	0.080			0.090*	0.090*
			(0.050)	(0.052)			(0.048)	(0.050)
Eq. Japan			0.040	0.087			0.402	0.084
			(0.056)	(0.058)			(0.056)	(0.058)
Renewal	0.021***	0.022***	0.022***	0.024***	0.021***	0.022***	0.022***	0.024***
	(4.0 E-3)	(4.2 E-3)	(4.0 E-3)	(4.2 E-3)	(4.0 E-3)	(4.2 E-3)	(4.0 E-3)	(4.2 E-3)

Note: n=867. The marginal effects are calculated around the means of the X's. Standard errors are in parentheses. ***, ** and * indicate significance at the 1%-, 5%- and 10%-level, respectively.

In Appendix B, Tables B1 and B2, I have calculated various predicted probabilities of an innovation for different values of *Equivalents* and *Renewal*. Both tables use the estimated parameters of Model B, but Table B1 is based on no forward citations, whereas Table B2 is based on five citations. As shown in Table B1, if a patent has six equivalents and expires after six years, the probability of an innovation is 58 percent,

---

[16] The model is not improved by including an additive dummy for a Triadic patent instead of the three dummies for EPO, US and Japanese equivalents.

whereas if a patent has 12 equivalents and expires after 12 years, then the probability of an innovation is 80 percent.[17]

Nonlinear relationships might exist between commercialization and some of the traditional patent quality indicators. For example, the probability of commercialization may increase with the number of equivalents, but the rate of increase may decrease for high numbers of equivalents. Estimations with squared values of the number of citations and equivalents are shown in Table 13. None of the squared variables is significantly related to commercialization.[18] Likelihood ratio tests between the estimations in Table 11 and those in Table 13 are not significant, indicating that the inclusion of the squared values does not improve the models. Further, the share of correct predictions of *Com* is not improved.

Table 13. Results of the probit estimations. Squared values.

Dependent variable = *Com*								
Statistical model = Probit model								
Explanatory variable	A	B	E	F	A	B	E	F
Citations 1	−4.8 E−3	−3.8 E−3			0.013	0.021		
	(0.041)	(0.041)			(0.018)	(0.021)		
$(Cit\ 1)^2$	1.3 E−3	2.0 E−3						
	(2.3 E−3)	(2.8 E−3)						
Citations 2			−0.081	−0.086			6.8 E−3	0.024
			(0.094)	(0.096)			(0.045)	(0.050)
$(Cit\ 2)^2$			0.019	0.025				
			(0.019)	(0.019)				
Equivalents	0.055***	0.064***	0.058***	0.067***	0.075**	0.083**	0.083**	0.091**
	(0.016)	(0.017)	(0.016)	(0.017)	(0.036)	(0.038)	(0.035)	(0.038)
$(Equiv.)^2$					−1.8 E−3	−1.7 E−3	−2.2 E−3	−2.2 E−3
					(2.5 E−3)	(2.7 E−3)	(2.4 E−3)	(2.6 E−3)
Renewal	0.055***	0.059***	0.056***	0.059***	0.054***	0.059***	0.054***	0.058***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Industry classes	No	Yes	No	Yes	No	Yes	No	Yes
Log likelihood	−545.1	−530.9	−544.9	−530.6	−545.1	−531.1	−545.3	−531.5
Share of correct predictions	63.4	65.6	64.7	65.6	64.5	65.5	64.6	65.9

Note: n=867. Standard errors are in parentheses. ***, ** and * indicate significance at the 1%-, 5%- and 10%-level, respectively. Parameter estimates of intercept and industry class dummies are not reported, but are available from the author on request.

---

[17] The 95-percent confidence interval of these predicted probabilities from Model B is on average +/− 14.3 percent. However, similar predicted probabilities from Model A yield a 95-percent confidence interval of +/− 5.2 percent on average. The smaller confidence interval in Model A depends on that fewer parameters are estimated (no industry class dummies).

[18] Similar results are also observed for squared values of renewal; however, these results are not reported in Table 13.

Table 14. Results of the ordered probit estimations.

Dep. variable: *Success*	Statistical model: Ordered probit model							
	without sample selection				with sample selection			
Explanatory variables	A	B	C	D	A	B	C	D
Citations 1	0.022	0.034	0.015	0.020	0.016	0.021	0.012	0.013
	(0.019)	(0.021)	(0.019)	(0.022)	(0.020)	(0.022)	(0.020)	(0.023)
Equivalents	−0.013	−1.5 E-3			−0.024*	−0.016		
	(0.014)	(0.015)			(0.013)	(0.013)		
Eq. EPO			0.073	0.089			0.032	0.038
			(0.149)	(0.157)			(0.15)	(0.15)
Eq. US			0.28*	0.34**			0.18	0.18
			(0.15)	(0.16)			(0.19)	(0.16)
Eq. Japan			−0.49***	−0.38**			−0.46***	−0.38**
			(0.16)	(0.18)			(0.18)	(0.17)
Renewal	0.094***	0.095***	0.091***	0.092***	0.059**	0.042***	0.062*	0.044***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.026)	(0.014)	(0.034)	(0.017)
Intercept	−0.33	−0.28	−0.33	−0.29	0.40	0.76	0.33	0.70
ω (threshold value)	0.60	0.62	0.61	0.63	0.52	0.48	0.55	0.50
Industry classes	No	Yes	No	Yes	No	Yes	No	Yes
ρ					−0.69**	−0.92***	−0.61	−0.85***
Log Likelihood	−1022.0	−997.1	−1018.5	−993.9	−1021.3	−994.2	−1018.2	−992.0
Test vs. restricted model					1.37	5.84**	0.41	3.72*

Table 14. Results of the ordered probit estimations (continued).

Dep. variable: *Success*	Statistical model: Ordered probit model							
	without sample selection				with sample selection			
Explanatory variables	E	F	G	H	E	F	G	H
Citations 2	−0.025	−0.011	−0.028	−0.020	−0.025	−0.016	−0.025	−0.019
	(0.040)	(0.044)	(0.041)	(0.045)	(0.044)	(0.050)	(0.044)	(0.050)
Equivalents	−7.3 E-3	−5.1 E-3			−0.019	−0.011		
	(0.014)	(0.015)			(0.013)	(0.013)		
Eq. EPO			0.11	0.13			0.065	0.068
			(0.15)	(0.16)			(0.15)	(0.15)
Eq. US			0.29**	0.37**			0.19	0.20
			(0.15)	(0.15)			(0.19)	(0.16)
Eq. Japan			−0.47***	−0.39**			−0.44**	−0.38**
			(0.16)	(0.18)			(0.18)	(0.17)
Renewal	0.095***	0.096***	0.091***	0.093***	0.061**	0.045***	0.062*	0.046***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.027)	(0.015)	(0.034)	(0.017)
Intercept	−0.32	−0.28	−0.32	−0.29	0.40	0.75	0.34	0.69
ω (threshold value)	0.60	0.62	0.61	0.63	0.53	0.48	0.55	0.51
Industry classes	No	Yes	No	Yes	No	Yes	No	Yes
ρ					−0.67*	−0.90***	−0.61	−0.84***
Log Likelihood	−1022.8	−998.7	−1018.7	−994.3	−1022.2	−995.8	−1018.4	−992.5
Test vs. restricted model					1.33	5.90**	0.65	3.61*

Note: n=854. Standard errors are in parentheses. ***, ** and * indicate significance at the 1%-, 5%- and 10%-level, respectively. Parameter estimates of industry class dummies are not reported, but are available from the author on request.

The results of the ordered probit estimations are presented in Table 14. The estimated parameter ρ is significant in all models where industry class dummies are included (Models B, D, F and H), indicating that there is a sample selection problem and that the model should be estimated in two steps. Regarding the explanatory variables, forward citations are not related to successful commercialization. Moreover, the number of equivalents does not have a positive relationship with *Success*; rather, the result is the opposite. However, when equivalents are measured by the dummies for Triadic market equivalents, the results show that a US equivalent is positively related to *Success*, whereas a Japanese equivalent is negatively related to *Success*. Additionally, the length of patent renewal has a strong positive relationship with successful commercialization, with a significant estimated parameter at least at the 5-percent level in all models except in Models C and G.

Table 15. Marginal effects on probability of successful innovation in the ordered probit estimations.

Dependent variable = *Success*

Statistical model: Ordered probit model with sample selection

Explanatory variables	Model B			Model H		
	P(0)	P(1)	P(2)	P(0)	P(1)	P(2)
Citations 1	−4.7 E−3	−2.4 E−3	7.1 E−3			
Citations 2				4.2 E−3	2.2 E−3	−6.4 E−3
Equivalents	−5.3 E−3	1.8 E−3	3.5 E−3			
Eq. EPO [a]				−0.015	−7.9 E-3	0.023
Eq. US [a]				−0.042	-0.023	0.066
Eq. Japan [a]				0.083**	0.036**	−0.119**
Renewal	−9.5 E-3***	−4.7 E-3***	0.014***	−0.010***	−5.3 E-3***	0.016***

Note: ***, ** and * indicate significance at the 1%-, 5%- and 10%-level, respectively. All marginal effects are calculated around the means of the X's. The sum of the marginal effects on the probabilities equals zero.

[a] Marginal effect on probabilities when dummy variable increases from 0 to 1.

The marginal effects of the explanatory variables on different commercialization outcomes (*Success*) are depicted in Table 15. If a patent is renewed one more year, then the probability of a successful innovation increases by 1.4 (Model B) or 1.6 (Model H) percentage points. Furthermore, the marginal effect of a Japanese equivalent is negative and significant. The marginal effects of the other explanatory variables are non-significant. Note here that the marginal effects of the dummy variables (*Eq. EPO*, *Eq. US* and *Eq. Japan*) are relatively large because these effects are calculated when the dummies change from 0 to 1, i.e., from the minimum to the maximum value.

In Appendix C, Tables C1 and C2, probabilities of a successful innovation (*Success*=2) given the commercialization of the patent have been calculated for various values of *Equivalents* and *Renewal*. Both tables use the estimated parameters of Model B. Moreover, Table C1 is based on *Citations 1* = 0, whereas Table C2 on *Citations 1* = 5. As shown in Table C1, if a patent has six equivalents and expires after six years, the probability of a successful innovation (given the commercialization of the patent) is 57 percent, whereas if a patent has 12 equivalents and expires after 12 years, then there is a 63 percent probability of a successful innovation, etc.

The probability of a successful innovation can also be calculated when information on whether patents have been commercialized is lacking. Such probabilities are reported in Appendix D, Tables D1 and D2. Since the probabilities in Appendix C account for the sample selection problem, the probabilities in Appendix D are simply the product of the probabilities in Appendices B and C. As shown in Table D1, if a patent has six equivalents and expires after six years, then the probability of a successful innovation is 33 percent (0.576*0.568), whereas if a patent has 12 equivalents and expires after 12 years, then there is a 49 percent (0.798*0.492) probability of a successful innovation.

5.2. Patents with EPO equivalents

As shown in Table 5, 76 percent of the patents with EPO equivalents are commercialized, whereas only 56 percent of those without EPO equivalents are commercialized. This result indicates that the subsample of patents with EPO equivalents is a biased sample with respect to innovation, i.e., a sample of potentially valuable patents. Moreover, as noted in section 2, the filing route typically entails first filing a Swedish patent application and then filing an EPO application.

In Table 16, the probit model estimates the relationships between the traditional patent quality indicators and the probability of an innovation in a manner similar to that in Table 11. As shown, only the total number of equivalents, *Equivalents*, has a positive and significant relationship with the probability of an innovation. Because of this

Table 16. Results of the probit estimations. EPO subsample.

Dependent variable = *Com*					Statistical model = Probit model			
Explanatory variable	A	B	C	D	E	F	G	H
Citations 1	5.4 E−3 (0.023)	9.0 E−3 (0.027)	4.7 E−3 (0.023)	0.011 (0.027)				
Citations 2					−0.029 (0.053)	−0.039 (0.060)	−0.028 (0.054)	−0.041 (0.060)
Equivalents	0.047* (0.024)	0.060** (0.027)			0.047* (0.025)	0.060** (0.027)		
Eq. US			0.048 (0.21)	−0.040 (0.23)			0.053 (0.21)	−0.025 (0.138)
Eq. Japan			−0.076 (0.20)	−0.021 (0.24)			−0.064 (0.20)	−0.020 (0.24)
Renewal	2.7 E−3 (0.031)	0.028 (0.035)	0.024 (0.029)	0.054 (0.033)	3.7 E−3 (0.031)	0.029 (0.035)	0.024 (0.029)	0.055 (0.033)
Industry classes	No	Yes	No	Yes	No	Yes	No	Yes
Log likelihood	−117.1	−107.8	−119.0	−110.3	−117.0	−107.6	−118.9	−110.2
Share of correct predictions	76.0 [a]	77.0	76.0 [a]	77.4	76.0 [a]	77.0	76.0 [a]	77.4

Note: n=217. Standard errors are in parentheses. ***, ** and * indicate significance at the 1%-, 5%- and 10%-level, respectively. Parameter estimates of intercept and industry class dummies are not reported, but are available from the author on request.

[a] In these models, the probit model predicts *Com*=1 for all observations.

Table 17. Results of the ordered probit estimations. EPO subsample.

Dep. variable: *Success*	Statistical model: Ordered probit model							
	without sample selection				with sample selection			
Explanatory variables	A	B	C	D	A	B	C	D
Citations 1	−3.5 E-3 (0.021)	−7.7 E-3 (0.029)	−1.3 E-3 (0.021)	−0.013 (0.029)	−4.1 E-3 (0.024)	−8.9 E-3 (0.043)	−1.3 E-3 (0.16)	−0.015 (0.048)
Equivalents	−0.045** (0.021)	−0.028 (0.023)			−0.050 (0.074)	−0.033 (0.059)		
Eq. US			9.1 E-3 (0.22)	0.013 (0.25)			9.1 E-3 (1.66)	0.013 (0.31)
Eq. Japan			−0.56*** (0.21)	−0.47* (0.27)			−0.56 (3.19)	−0.46 (0.44)
Renewal	0.109*** (0.033)	0.110*** (0.036)	0.103*** (0.031)	0.110*** (0.036)	0.106 (0.108)	0.107* (0.060)	0.103 (1.05)	0.104 (0.149)
Intercept	−0.24	−0.29	−0.25	−0.26	−0.05	−0.11	−0.25	−0.08
ω (threshold value)	0.52	0.56	0.53	0.57	0.51	0.56	0.53	0.56
Industry classes	No	Yes	No	Yes	No	Yes	No	Yes
ρ					−0.32	−0.24	−5.8 E-4	−0.25
Log Likelihood	−256.3	−241.9	−256.7	−243.7	−256.3	−241.9	−256.7	−243.7
Test vs. restricted model					0.027	0.018	0.00	0.045

Note: n=215. Standard errors are in parentheses. ***, ** and * indicate significance at the 1%-, 5%- and 10%-level, respectively. Parameter estimates of industry class dummies are not reported, but are available from the author on request.

weak overall performance, the marginal effects are not reported. Additional estimations also included squared values of *Citations 1*, *Citations 2* and *Equivalents*; however, the results did not improve.

The results of the ordered probit estimations are reported in Table 17. Only the 215 patents with EPO equivalents are used, and the EPO dummy is excluded in Models C, D, G and H. The results show that the selection parameter, ρ, is non-significant in 6 of the 8 models, indicating that an ordered probit model without sample selection is sufficient for the estimation. The selection parameter is significant in Models F and G only. The results of the log-likelihood tests are similar, indicating that there is some instability in the estimations across the models. However, the estimated parameters of the main explanatory variables are stable across estimations with and without sample selection, irrespective of whether the selection parameter is significant. Nevertheless, the standard errors of the estimated parameters increase substantially in the sample selection estimations.

Regarding the individual parameters, the results for the EPO subsample are similar to those for the full sample. Specifically, *Renewal* has a positive and strongly significant relationship with *Success* (in Models A−H without sample selection and Models F and G with sample selection). The citation variables always have a negative relationship with *Success*, but the relationship is never significant. Finally, the results regarding equivalents for the EPO subsample are similar to those for the full sample. Specifically, the estimated parameter of *Equivalents* always has a negative sign but is seldom significant. However, the dummy for a Japanese equivalent is negatively related to successful commercialization (in Models C and D without sample selection and Model G with sample selection).

The marginal effects of the estimated parameters on the probability of a successful innovation are shown in Table 18. If a patent with an EPO equivalent is renewed for one more year, then the probability of a successful innovation increases by 4.0 (Model D) or 4.2 (Model F) percentage points. Thus, I find a stronger effect for the EPO subsample than for the full sample. Furthermore, the marginal effect of a Japanese equivalent on a successful commercialization is negative and significant,[19] whereas the marginal effects of the other explanatory variables are non-significant.

Table 18. Marginal effects on probability of a successful innovation in the ordered probit estimations. EPO subsample.

Dependent variable = *Success*						
Statistical model: Ordered probit model						
Explanatory variables	Model D without sample selection			Model F with sample selection		
	P(0)	P(1)	P(2)	P(0)	P(1)	P(2)
Citations 1	3.3 E-3	1.6 E-3	−4.9 E-3			
Citations 2				4.2 E-3	2.2 E-3	−6.4 E-3
Equivalents				0.016	7.9 E-3	−0.024 E-3
Eq. US [a]	−0.033	−0.016	0.049			
Eq. Japan [a]	0.118*	0.054*	−0.172*			
Renewal	−0.027***	−0.013***	0.040***	−0.028***	−0.014***	0.042***

Note: ***, ** and * indicate significance at the 1%-, 5%- and 10%-level, respectively. All marginal effects are calculated around the means of the X's. The sum of the marginal effects on the probabilities equals zero.

[a] Marginal effect on probabilities when dummy variable increases from 0 to 1.

6. Concluding remarks

6.1 Main results

This study aims to empirically analyze whether and how strong traditional patent quality indicators are related to 1) the probability that a patent is commercialized (i.e., the probability that an innovation is introduced in the market) and 2) whether the patent commercialization is successful. To the best of my knowledge, such an analysis has never been done in the existing literature. For the analysis, a unique database of Swedish patents with information on the commercialization process of individual patents is used. Simple correlations and contingency table tests show that both patent commercialization and successful commercialization are strongly positively correlated with patent renewal and patent equivalents (family size) but only moderately positively correlated with forward citations.

---

[19] Note here that the marginal effects of the dummy variables (*Eq. US* and *Eq. Japan*) are relatively large since these effects are calculated when the dummies change from 0 to 1, i.e., from the minimum to the maximum value.

In the statistical models, patent renewal and equivalents primarily have positive relationships with the probability of an innovation, whereas the relationship between forward citations and the probability of an innovation is generally non-significant. With respect to the success of commercialization, only patent renewal is positively related to profitability. A somewhat curious result is that a Japanese equivalent is negatively related to a successful innovation. Further, inventors must decide soon after patent application in which countries to file a patent (priority year), whereas the renewal decision is undertaken every year. Therefore, it is not surprising that the probability of a successful innovation is primarily positively related to the renewal decision.

Additionally, I analyze a subsample comprising only patents with EPO equivalents. The correlations show that commercialization is positively related to equivalents, whereas the probability of a successful innovation is somewhat positively related to patent renewal. However, the strength of the relationships is weaker for the EPO subsample than for the full sample. The results further show that forward citations are not related to the commercialization variables. The statistical models confirm these relationships.

I then combine the estimated parameters from the probit and ordered probit estimations with various values of patent renewal (years), number of equivalents and number of forward citations. I so doing, I generate predicted probabilities that 1) a patent is commercialized (Appendix B); 2) an innovation is successful if the patent is commercialized (Appendix C); and 3) an innovation is successful even if information regarding whether the patent is commercialized is unavailable (Appendix D). The parameters may be applied to other patent databases in future studies.

6.2 Limitations

Since the database covers only patents owned by small firms and individuals, the estimated parameters can be used only to predict innovations and successful innovations for (Swedish or foreign) patents owned by these groups. Larger firms likely have a higher share of non-commercialized patents, which are used for defensive purposes (blocking or negotiations).

Finally, the study and method are unable to identify innovations in all sectors. In some sectors, firms traditionally prefer to protect their invented technologies by relying on secrecy and circumspection or strong lead times rather than patents. By contrast, other sectors, primarily in the large service areas, rely on other intellectual property rights (e.g., copyright) to protect artistic and literary works.

References

Arrow, K., 1962, ‘The Economic Implications of Learning by Doing’, *Review of Economic Studies*, 29(2), pp. 155–73.

Breschi, S., Lissoni, F. and Malerba, F., 2004, ‘The Empirical Assessment of Firms. Technological Coherence: Data and Methodology’. In: Cantwell, A., Gambardella, A. and Granstrand, O. (eds.), *The Economics and Management of Technological Diversification*. Routledge, London.

Espacenet, 2014, *Espacenet patent database*. Assessed at:
http://www.epo.org/searching/free/espacenet.html (January, 2014).

Gault, F., 2013, *Handbook of Innovation Indicators and Measurement*. Edward Elgar, Cheltenham, U.K. och Northampton, Ma.

Greene, W.H., 1997, *Econometric Analysis*, 3rd edition. Prentice-Hall, Upper Saddle River, NJ.

Greene, W.H., 2002, *Econometric Modelling Guide Vol. 1, Limdep Version 8.0*. Econometric software Inc. Plainview, NY.

Griliches, Z., Hall, B.H. and Pakes, A., 1987, ‘The Value of Patents as Indicators of Inventive Activity’, in Dasqupta P. and Stoneman, P. (eds.), *Economic Policy and Technological Performance*. Cambridge University Press, Cambridge, Ma, pp. 97–124.

Griliches, Z., 1990, ‘Patent Statistics as Economic Indicators: A Survey’, *Journal of Economic Literature*, 28(Dec.), pp. 1661–1707.

Hall, B.H., 1993, ‘The Stock Market Valuation of R&D Investment during the 1980s’, *American Economic Review*, 83(2), pp. 259–64.

Hall, B.H., Thoma, G. and Torrisi, S., 2007, ‘The Market Value of Patents and R&D: Evidence from European firms’, NBER Working paper No. 13426. NBER, Cambridge, Ma.

Hall, B.H., Mairesse, J. and Mohnen, P., 2010, ‘Measuring the Returns to R&D’, in Hall, B.H. and Rosenberg, N. (eds.), *Handbook of the Economics of Innovation. Vol. 2*. Elsevier-North Holland, Amsterdam, pp. 1033–82.

Harhoff, D., Scheerer, F., and Vopel, K., 2002, ‘Citations, Family Size, Opposition and Value of Patent Rights’, *Research Policy*, 32(8), pp. 1343–63.

Hausmann J. and McFadden, D., 1984, ‘A Specification Test for the Multinomial Logit Model’, *Econometrica*, 52(5), pp. 1219–40.

Kaplan, S.N. and Strömberg, P., 2001, ‘Venture Capitals as Principals: Contracting, Screening, and Monitoring’, *American Economic Review*, 91(2), pp. 426–30.

Lanjouw, J. and Schankerman, M., 1997, ‘Stylised Facts about Patent Litigation: Value, Scope and Ownership’, NBER Working paper No. 6297, NBER, Cambridge, Ma.

Levin, R.C., Klevorick, A.K., Nelson, R.R. and Winter, S.G., 1987, ‘Appropriating the Returns from Industrial Research and Development’, *Brookings Papers on Economic Activity*, 1987(3), pp. 783–831.

Maurseth, P.B. and Svensson, R., 2014, ‘Micro Evidence on International Patenting’, *Economics of Innovation and New Technology*, 23(4), pp. 398–422.

Morgan, R.P., Kruytbosch, C. and Kannankutty, N., 2001. ‘Patenting and Invention Activity of U.S. Scientists and Engineers in the Academic Sector: Comparisons with Industry’, *Journal of Technology Transfer*, 26(1–2), pp. 173–83.

Pakes, A. and Schankerman, M., 1984, ‘The Rate of Obsolescence of Patents, Research Gestation Lags and the Private Return to Research Resources’, in Griliches Z. (ed.) *R&D, Patents and Productivity*. NBER, University of Chicago Press, Chicago, pp. 73–88.

Pakes, A., 1986, ‘Patents as Options: Some Estimates of the Value of Holding European Patent Stocks’, *Econometrica*, 54(4), pp. 755–84.

Putnam, J., 1996, *The Value of International Patent Rights*. Yale University Press, Connecticut , NJ.

Rossman, J. and Sanders, B.S., 1957, ‘The Patent Utilization Study’, *Patent, Trademark and Copyright Journal of Research and Education*, 1(1), pp. 74–111.

Schankerman, M. and Pakes, A., 1986, ‘Estimates of the Value of Patent Rights in the European Countries during the Post-1950 Period’, *Economic Journal*, 96(384), pp. 1052–76.

Schmookler, J., 1966, *Invention and Economic Growth*. Harvard University Press, Cambridge, Ma.

Trajtenberg, M., 1990, ‘A Penny for Your Quotes, Patent Citations and the Value of Innovations’, *RAND Journal of Economics*, 21(1), pp. 172–87.

van Zeebroeck, N. 2011, ‘The Puzzle of Patent Value Indicators’, *Economics of Innovation and New Technology*, 20(1), 33–62.

van Zeebroeck, N. and van Pottelsberghe, B., 2011, ‘The Vulnerability of Patent Value Determinants’, *Economics of Innovation and New Technology*, 20(3), pp. 283–308.

Appendix A

Table A1. Descriptive statistics of dependent and explanatory variables.

Dependent variables	Definition	Large sample				EPO subsample			
		All observations (n=867)		Commercialized patents (n=526)		All observations (n=217)		Commercialized patents (n=165)	
		Mean	Std. dev.	Mean	Std. dev	Mean	Std. dev	Mean	Std. dev
Com	Dummy that equals 1 if commercialization, and 0 otherwise	0.61	0.49	-----	-----	0.76	0.43	-----	-----
Success	Profitability of commercialization. 2 = profit, 1 = break-even, 0 = loss	-----	-----	1.29	0.85	-----	-----	1.48	0.85
Explanatory variables									
Citations 1	Number of forward citations from all sources within five years after publishing	1.21	2.99	1.45	3.46	2.87	4.39	2.91	4.75
Citations 2	Number of forward citations from EPO and PCT within five years after publishing	0.44	1.15	0.50	1.32	0.99	1.73	0.96	1.86
Equivalents	Number of patent equivalents abroad	2.00	3.79	2.64	4.32	7.11	4.58	7.46	4.83
Eq. EPO	Dummy that equals 1 if an administrative patent at EPO, and 0 otherwise	0.25	0.43	0.31	0.46	-----	-----	-----	-----
Eq. US	Dummy that equals 1 if a US patent, and 0 otherwise	0.26	0.44	0.32	0.47	0.67	0.47	0.67	0.47
Eq. Japan	Dummy that equals 1 if a Japanese patent, and 0 otherwise	0.16	0.37	0.20	0.40	0.45	0.50	0.44	0.50
Renewal	The number of years of patent renewal (right-censored at year 16)	10.40	4.59	11.30	4.49	13.53	3.20	13.63	3.15

Appendix B

Table B1. Predicted probabilities of an innovation for different values of *Renewal* and *Equivalents*. *Citations 1* = 0. Based on Model B, Table 11.

	Renewal (years)							
No. of equivalents	2	4	6	8	10	12	14	16
0	0.367	0.409	0.452	0.496	0.539	0.583	0.625	0.666
2	0.406	0.450	0.493	0.537	0.581	0.623	0.664	0.703
4	0.447	0.491	0.535	0.578	0.621	0.662	0.701	0.738
6	0.489	0.533	0.576	0.619	0.660	0.699	0.736	0.771
8	0.530	0.573	0.616	0.658	0.697	0.734	0.769	0.801
10	0.571	0.614	0.655	0.695	0.732	0.767	0.800	0.829
12	0.612	0.653	0.693	0.730	0.766	0.798	0.827	0.854
14	0.651	0.691	0.729	0.764	0.796	0.826	0.853	0.877
16	0.689	0.727	0.762	0.795	0.824	0.851	0.875	0.897
18	0.725	0.760	0.793	0.823	0.850	0.874	0.896	0.914
20	0.758	0.791	0.821	0.849	0.873	0.895	0.913	0.929
22	0.790	0.820	0.847	0.872	0.893	0.912	0.929	0.942
24	0.818	0.846	0.871	0.892	0.911	0.928	0.942	0.953

Note: When calculating the predicted probabilities, Xβ = –0.45097 for the intercept and the industry class dummies (calculated around the means of the X’s).

Table B2. Predicted probabilities of an innovation for different values of renewal and equivalents. No. of citations = 5. Based on Model B, Table 11.

	Renewal (years)							
No. of equivalents	2	4	6	8	10	12	14	16
0	0.397	0.440	0.484	0.527	0.571	0.614	0.655	0.695
2	0.438	0.481	0.525	0.569	0.612	0.653	0.693	0.730
4	0.479	0.523	0.566	0.609	0.651	0.691	0.728	0.763
6	0.521	0.564	0.607	0.649	0.689	0.726	0.762	0.794
8	0.562	0.605	0.646	0.686	0.724	0.760	0.793	0.823
10	0.603	0.644	0.684	0.722	0.758	0.791	0.821	0.849
12	0.642	0.682	0.720	0.756	0.789	0.820	0.847	0.872
14	0.680	0.719	0.754	0.788	0.818	0.846	0.870	0.892
16	0.717	0.753	0.786	0.817	0.844	0.869	0.891	0.910
18	0.751	0.784	0.815	0.843	0.868	0.890	0.909	0.926
20	0.783	0.814	0.842	0.867	0.889	0.908	0.925	0.940
22	0.812	0.840	0.865	0.888	0.907	0.924	0.939	0.951
24	0.839	0.864	0.887	0.906	0.924	0.938	0.950	0.961

Note: When calculating the predicted probabilities, Xβ = –0.45097 for the intercept and the industry class dummies (calculated around the means of the X’s).

Appendix C

Table C1. Predicted probabilities of a successful innovation (*Success* = 2) for different values of *Renewal* and *Equivalents*, given that *Com* = 1. *Citations 1* = 0. Based on Model B, Table 13.

	Renewal (years)							
No. of equivalents	2	4	6	8	10	12	14	16
0	0.538	0.572	0.605	0.637	0.668	0.699	0.728	0.755
2	0.525	0.559	0.592	0.625	0.657	0.687	0.717	0.745
4	0.513	0.547	0.580	0.613	0.645	0.676	0.706	0.735
6	0.500	0.534	0.568	0.601	0.633	0.665	0.695	0.724
8	0.488	0.521	0.555	0.589	0.621	0.653	0.684	0.714
10	0.475	0.509	0.543	0.576	0.609	0.641	0.673	0.703
12	0.462	0.496	0.530	0.564	0.597	0.629	0.661	0.692
14	0.450	0.484	0.517	0.551	0.585	0.617	0.649	0.680
16	0.437	0.471	0.505	0.539	0.572	0.605	0.638	0.669
18	0.425	0.458	0.492	0.526	0.560	0.593	0.626	0.657
20	0.413	0.446	0.480	0.513	0.547	0.581	0.613	0.646
22	0.400	0.433	0.467	0.501	0.535	0.568	0.601	0.634
24	0.388	0.421	0.454	0.488	0.522	0.556	0.589	0.622

Note: When calculating the predicted probabilities, Xβ = 0.61174 for the intercept and the industry class dummies (calculated around the means of the X's) and ω = 0.47877.

Table C2. Predicted probabilities of a successful innovation (*Success* = 2) for different values of *Renewal* and *Equivalents*, given that *Com* = 1. *Citations 1* = 5. Based on Model B, Table 13.

	Renewal (years)							
No. of equivalents	2	4	6	8	10	12	14	16
0	0.580	0.613	0.645	0.676	0.706	0.735	0.762	0.787
2	0.568	0.601	0.633	0.665	0.695	0.724	0.752	0.778
4	0.555	0.589	0.621	0.653	0.684	0.714	0.742	0.768
6	0.543	0.576	0.609	0.641	0.673	0.703	0.731	0.759
8	0.530	0.564	0.597	0.629	0.661	0.692	0.721	0.749
10	0.518	0.551	0.585	0.617	0.649	0.680	0.710	0.738
12	0.505	0.539	0.572	0.605	0.638	0.669	0.699	0.728
14	0.492	0.526	0.560	0.593	0.626	0.657	0.688	0.717
16	0.480	0.514	0.547	0.581	0.614	0.646	0.677	0.707
18	0.467	0.501	0.535	0.568	0.601	0.634	0.665	0.696
20	0.454	0.488	0.522	0.556	0.589	0.622	0.654	0.684
22	0.442	0.476	0.510	0.543	0.577	0.610	0.642	0.673
24	0.429	0.463	0.497	0.531	0.564	0.598	0.630	0.662

Appendix D

Table D1. Predicted probabilities of a successful innovation (*Success* = 2) for different values of *Renewal* and *Equivalents*. *Citations 1* = 0. Based on Model B, Tables 11 and 13.

	Renewal (years)							
No. of equivalents	2	4	6	8	10	12	14	16
0	0.197	0.234	0.273	0.316	0.360	0.408	0.455	0.503
2	0.213	0.252	0.292	0.336	0.382	0.428	0.476	0.524
4	0.229	0.269	0.310	0.354	0.401	0.448	0.495	0.542
6	0.245	0.285	0.327	0.372	0.418	0.465	0.512	0.558
8	0.259	0.299	0.342	0.388	0.433	0.479	0.526	0.572
10	0.271	0.313	0.356	0.400	0.446	0.492	0.538	0.583
12	0.283	0.324	0.367	0.412	0.457	0.502	0.547	0.591
14	0.293	0.334	0.377	0.421	0.466	0.510	0.554	0.596
16	0.301	0.342	0.385	0.429	0.471	0.515	0.558	0.600
18	0.308	0.348	0.390	0.433	0.476	0.518	0.561	0.600
20	0.313	0.353	0.394	0.436	0.478	0.520	0.560	0.600
22	0.316	0.355	0.396	0.437	0.478	0.518	0.558	0.597
24	0.317	0.356	0.395	0.435	0.476	0.516	0.555	0.593

Table D2. Predicted probabilities of a successful innovation (*Success* = 2) for different values of *Renewal* and *Equivalents*. *Citations 1* = 5. Based on Model B, Tables 11 and 13.

	Renewal (years)							
No. of equivalents	2	4	6	8	10	12	14	16
0	0.230	0.270	0.312	0.356	0.403	0.451	0.499	0.547
2	0.249	0.289	0.332	0.378	0.425	0.473	0.521	0.568
4	0.266	0.308	0.351	0.398	0.445	0.493	0.540	0.586
6	0.283	0.325	0.370	0.416	0.464	0.510	0.557	0.603
8	0.298	0.341	0.386	0.431	0.479	0.526	0.572	0.616
10	0.312	0.355	0.400	0.445	0.492	0.538	0.583	0.627
12	0.324	0.368	0.412	0.457	0.503	0.549	0.592	0.635
14	0.335	0.378	0.422	0.467	0.512	0.556	0.599	0.640
16	0.344	0.387	0.430	0.475	0.518	0.561	0.603	0.643
18	0.351	0.393	0.436	0.479	0.522	0.564	0.604	0.644
20	0.355	0.397	0.440	0.482	0.524	0.565	0.605	0.643
22	0.359	0.400	0.441	0.482	0.523	0.564	0.603	0.640
24	0.360	0.400	0.441	0.481	0.521	0.561	0.599	0.636