APPLIED ECONOMICS
https://doi.org/10.1080/00036846.2019.1624914

Routledge
Taylor & Francis Group

Check for updates

Can patent family size and composition signal patent value?

Francois P. Kabore[a] and Walter G. Park[b]

[a]Center for Research and Action for Peace (CERAP), Abidjan, Ivory Coast; [b]Department of Economics, American University, Washington, DC, USA

ABSTRACT
Recent research has proposed a method of patent valuation based on weighting patent family size by the market size of the countries in the family. The premise is that inventors tend to seek greater international coverage for their more valuable patents. The paper presents a novel way to test the ability of market size-weighted patent families to predict patent value and compares the method against extant measures of patent valuation based on patent citations and renewal behaviour. We use forecasting techniques to show that the weighted patent family size measure outperforms other methods in terms of predicting patent life and the number of citations. An advantage of the weighted patent family size measure is that it is based on ex-ante information and is easy to construct for purposes of evaluating patent value. We demonstrate this advantage using a large, comprehensive database of international patent families.

KEYWORDS
Patent valuation; market size; patent families; citations; renewal

JEL CLASSIFICATION
O34; O33; F23; K11

I. Introduction

Patent valuation is important to firms' decision-making and strategies. For example, it shapes incentives for investments in innovation. It also plays a key role in licensing negotiations and in determining the relative bargaining strength of parties (Smith and Parr 2000). Patent litigation is also affected by perceptions of patent value. The more valuable the patent rights are, the more committed patent holders will be to monitor and enforce their rights. Patent valuation is also important to policy-makers for assessing the impact of legal provisions or innovation policies on the level and quality of new technologies.[1] Is the policy regime merely encouraging the filing of patents for low-valued inventions or generating transformative wealth-creating activities? The valuation of patents can also provide useful indicators for comparing the production of innovations across industries and countries.

However, the value of patents is not easily observed. Patent rights are not frequently or widely traded like financial instruments, such as equities. Large, international exchange institutions for providing markets for patent rights are absent.[2] Few firms publicly disclose details about their patent transactions or technology transfer agreements. Furthermore, not all patent rights have a steady cash flow, since some (if not many) patented inventions are not commercialized or are not 'worked' (put into practice). Other problems may be that some patents are invalid (not correctly granted and will be invalidated upon challenge) or that asymmetric information exists (where the seller knows the quality of the underlying invention but the buyer does not) so that contract prices may distort the true valuation of patents.[3]

Consequently, from a research standpoint, patent valuation is largely based on observing the characteristics of a patent or the behaviour of a patent holder that may *reveal* information about the underlying value. For example, previous work by Bessen (2008), Lanjouw, Pakes, and Putnam (1998), Pakes and Simpson (1989) and Schankerman and Pakes (1986) examined patent renewal behaviour. Given the costs of maintaining patent rights, the relatively most valuable patents will be selected by patent holders for renewal. Another approach has been to factor in a patent's forward citations (see Trajtenberg (1990), Harhoff et al. (1999), Jaffe and Trajtenberg

CONTACT Walter G. Park wgp@american.edu Department of Economics, American University, Washington, DC, USA

[1]See Bessen and Meurer (2009).
[2]However, patent rights are tradable at auctions, held for example by *Ocean Tomo*, an intellectual property merchant bank. Fisher and Leidinger (2014) and Mauck and Pruitt (2016) study patent values from *Ocean Tomo* auctions.
[3]See Murphy, Orcutt, and Remus (2012) for a discussion of the impact of adverse selection on patent values.

© 2019 Informa UK Limited, trading as Taylor & Francis Group

(2002), and Hall, Jaffe, and Trajtenberg (2005)). The more valuable a patent is the more likely it will find use in follow-on research or production and therefore be more often cited. Barney (2002) finds that citations and renewal are related in that patents that receive no forward citation in their first four years are less likely to be renewed than those with multiple citations. On the other hand, patent renewals and citations may not always yield the same 'signal'. A patent may lapse (for lack of market value or other) and yet continued to be cited.[4]

Another approach to gauging patent value is to examine patent family size (see Harhoff, Scherer, and Vopel 2003; Johnstone et al. 2012; Lanjouw and Schankerman 2004; Putnam 1996). Patent family size refers to the number of countries in which an invention is protected by a patent; that is, the priority filing and the subsequent filings that emanate from it.[5] Given the costs of acquiring patents in multiple jurisdictions, rights holders would reserve their more valuable technologies for international patenting. A variant of this approach is to consider triadic patents, where an invention is protected in the three major markets, namely the U.S., Japan and the European Patent Office (EPO) (see Dernis, Guellec, and van Pottelsberghe De La Potterie 2001; Kumazawa and Gomis-Porqueras 2012), or transnational patents, which are patent families with at least one EPO or Patent Cooperation Treaty (PCT) filing. But both the raw family size and the triadic variable do not take into account the potential market size of the protected areas.[6]

A key limitation of the citation and renewal methods is that patent values are assessed ex-post. As pointed out by Hall, Jaffe, and Trajtenberg (2005), a sufficient time is needed after a patent is first published to observe citations of it by later patents. Furthermore, patent citations are not distinguished by how much impact the cited patents have on an invention; for example, patent X may cite prior patents A and B, but A may be more important to X's function than B. Moreover, the reason patent X may cite A and B is to indicate that they are alternative products or methods of production, so that X is not infringing on them, rather than that patent X builds on them. Renewals are also observed later in the life of a patent. Another limitation is that renewal decisions may be made not on grounds of patent value, but on changes in the cost of renewing or maintaining patents. The limitation of patent family size as a measure of patent value is that it does not take into account the attractiveness of different market destinations.

Our contribution here is to join recent work that weights patent family size by indicators of the market potential of the countries comprising the patent family (see van Pottelsberghe De La Potterie and van Zeebroeck 2008; Frietsch et al. 2010; Ernst and Omland 2011; Neuhausler and Freitsch, 2013; Dechezleprêtre, Ménière, and Mohnen 2017).[7] The rationale for factoring in the market size is that rights holders who possess more valuable patents

[4]Moreover, renewal decisions may largely reflect private value while citations reflect some social value as well, considering the impacts of the patented technology on other inventions.

[5]There is also an extended patent family definition, as described in Hingley and Park (2003), where 'a patent family encompasses all the documents related to the patents emanating from an invention, including documents that may cross relate to other inventions as well.' See also Martinez (2011) for a thorough survey of patent family definitions and methodologies.

[6]There are other potential indicators of patent value, such as whether a patent has faced an opposition challenge (Harhoff and Reitzig 2004), the number of claims in a patent (Tong and Frame, 1994), auction prices (Nair, Mathew, and Nag 2011; Sneed and Johnson 2009), and filings strategies (van Zeebroeck and van Pottlesburg de la Potterie, 2011). Other approaches to determining patent value are event studies around court decisions or patent announcements (see Henry 2013; Austin 1993). In this paper, we focus primarily on comparing our method of valuation to the most widely used indicators of patent value thus far, namely patent renewals and citations. In this paper, we do control for patent claims which are also ex-ante information. We had examined data on opposition in preliminary work, but do not pursue them further here. While it is likely that patents with much commercial value are the ones that will be challenged, they will also be challenged if they are invalid from a legal perspective.

[7]van Pottelsberghe De La Potterie and van Zeebroeck (2008) focus on the size and age of a patent family and discuss how the market size (GDP) of destination countries is an important factor in the patent validation strategy of firms but do not construct a GDP-weighted measure of patent families. Ernst and Omland (2011) weight patent counts by market coverage in order to develop a *Patent Asset Index* for firms. Specifically, a firm's patent portfolio is evaluated based on the GDP of countries covered and its technological relevance based on citations received. Dechezleprêtre, Ménière, and Mohnen (2017) argue that the timespan between the first application date and last application date within an extended patent family can be an indicator of value, as the lag could reflect the applicant's strategy to optimize patent scope over time. Freitsch et al. (2010) incorporate export volumes and intensities to patent value indicators. Neuhausler and Freitsch (2013) use six different variables for weighting patent family sizes: the imports, GDP, population, strength of patent protection, global competitiveness, or intensity of local competition of the countries in the family. A key difference between our paper and Neuhausler and Frietsch (2013) is that we weight patent families by the level of GDP so as to capture the absolute market size of a country, whereas Neuhausler and Frietsch (2013) normalize their family size measures (so as to focus on average family sizes); for example, they use the share of a country's GDP in the world's total (or for the variables that are indexes, they use the percentage of the maximum value).

would self-select in protecting their inventions in greater and larger world markets. Our main value added to this literature is that we develop a novel way to test the extent to which patent family size, weighted by market size, is informative about patent value. Such a test and demonstration of the predictive ability of market-weighted patent family size has not been done thus far. A key advantage of the weighted patent family size as a measure of patent value is that it is ex-ante; that is, it provides more *current* information about patent value. This would be highly useful for current business decision-making or policy-making, as well as for applied research that utilizes patent data. To demonstrate this advantage, our empirical test centres on how well the weighted patent family size method can forecast patent life and forward citations, which are common measures of patent value after the fact.

The paper is organized as follows. The next section discusses the construction of our measure of patent value and the alternative measures of value. Section III discusses our dataset. Section IV discusses our methodology for testing and comparing our approach to alternative measures of value. Section V contains our results, variations and some sample trends in patent value by country and technology. Section VI concludes.

II. Valuing patents by the market size of patent families

Several reasons exist as to why patent rights are sought in large markets and why such rights are more valuable in those markets.[8] First, the value of obtaining a patent for a new innovation depends on multiple factors, such as the nature of the technology, the inventive step, the demand for that technology by consumers and potential licensees, imitation risk and the availability of alternative appropriation mechanisms. In small economies where the market for the technology and expected returns are limited, or where imitation risk is low, the incentives for obtaining a patent are relatively small or non-existent. In larger economies, markets for technology tend to be larger, which can both raise the demand for licensing as well as attract imitation and infringement given that copiers would likely target products with large mass appeal. In that vein of thought, competition is likely to be more intense in larger markets, owing to rivals that can innovate or imitate and thereby create competing technologies. Consequently, in larger economies, the incentives to procure patents and enforce them would tend to be greater.

Second, not only is it costly to file patents in the multiple countries that comprise a patent family, but the cost is higher in larger markets, such as Europe, Japan and the U.S. Figure 1 shows that the cost of patenting varies positively with the income level of the destination: the larger the market, the greater the cost per patent. Using data for 2010, the figure breaks down destinations by four quartiles, based on GDP (in PPP international dollars) and shows the average fees of obtaining a patent in each income quartile. The fees refer to official fees as well as to agent (or attorney) fees.[9] The fees exclude renewal fees, but an earlier study by Berrier (1996) confirms that the cradle-to-grave costs (i.e., during the lifespan of a patent right) tend to be greater in larger markets, such as Japan, EPO and the U.S. Furthermore, Berger (2005) finds that the cost of patenting in the EPO is compounded by requirements for translation of the patent document and validation fees (whereby an EPO patent must be activated in each individual member state that is designated in the patent). Hence, factoring in the cost and expense of international patent filing, the relatively most valuable innovations would tend to be selected for patenting in the larger markets.

To capture this, we start with a patent family and weight each country in the family by its market size, namely its gross domestic product. A patent family consists of the priority patent and its subsequent patents. Under the *Paris Convention*, a priority patent application is an initial patent application that allows an applicant to file subsequent patent applications for the same invention in other coun-

[8]The focus here is on patent *value* and not patent *quality*. Quality deals more with whether a technology should have been patented; whether it is novel and robust, or correctly awarded. Value deals with the commercial or market value of having a right. Even a minor innovation may be valuable if one can 'own' it, particularly if it is overly broad and generic (e.g. algorithm for sorting, where it would indeed be quite profitable to charge other people for the right to use). See Kappos and Graham (2012) for a discussion of measures to improve patent quality.

[9]The data are from Park (2010) and sources cited therein.

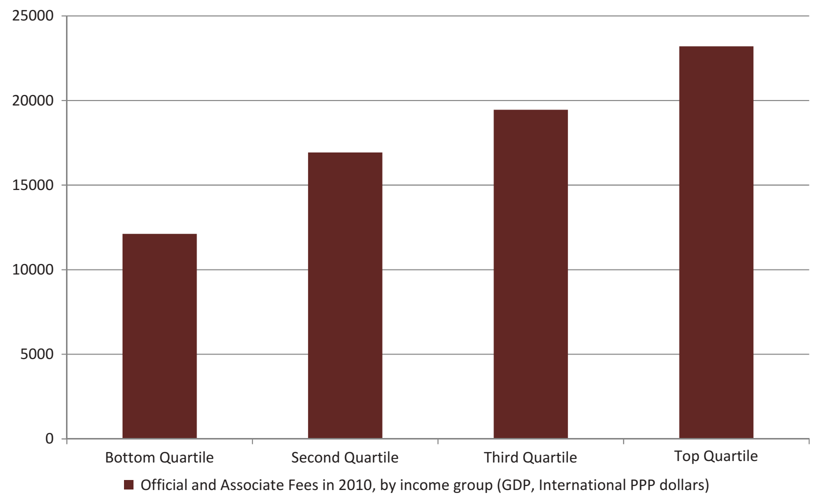

Figure 1. Relationship between patenting cost and market size.

tries within a year. Within each patent family, we focus on those countries in which the patent has been granted and validated.

Suppose that V_j^* is the unobserved value of a priority patent j. Let V_j be an observable indicator of this value; that is, $V_j^* = V_j^*(V_j)$ is some function of the observable. For example, V_j could be the number of citations received by priority patent j within the first T years of its life. Alternatively, it could be the number of countries in the patent family; namely, the country of priority patenting and countries of subsequent patenting. Let the raw family size (count) be

$$V_j = \sum_{n=1}^{N} I_n \tag{1}$$

where I_n is an indicator function which equals one if priority patent j is patented in country n and zero otherwise, and N is the total number of countries (including the country of priority filing). Equation (1) is what de Rassenfosse, Dernis, and Boedt (2014) call 'geographic' family size. The more countries covered by the priority patent, the more valuable it may be. However, the raw family size in (1) does not take into account the market sizes of countries. Larger markets are likely to be not only more attractive for exploiting an invention but also more competitive for acquiring patent rights and costlier to enforce them. Hence, we modify (1) to take into account market size:

$$V_j = \sum_{n=1}^{N} \omega_n I_n \tag{2}$$

where $\omega_n = GDP_n$ (real gross domestic product of country n). We refer to (2) as the GDP-weighted family size of patent j, which takes into account the composition of the international patent family in terms of the market size of countries.

More specifically, we sample the GDP-weighted family size of priority patent j at the date of its application (time t):

$$V_{jt} = \sum_{n=1}^{N} \omega_{nt} I_{nt} \tag{3}$$

That is, we evaluate the market sizes of all the countries in the patent family *as of time t*. Thus, once we evaluate the global market size of the patent family at time t, we hold that value *fixed* in our subsequent ex-ante analyses of patent value. We do not update the GDPs of the countries in the patent family after time t, as our goal is to evaluate the ex-ante ability of the GDP-weighted patent family size to predict the patent value of the priority patent.

We can then assess this patent valuation ex-post by judging how long the patent lived on the premise that patent holders will maintain their more valuable patents. Hence, priority patent j is renewed in country n at some future time $\tau > t$ if the benefit of having a patent right over not having it exceeds the cost of maintaining and enforcing the property right:

$$R^j_{n\tau}(V^*) - NR^j_{n\tau}(V^*) \geq \kappa_{n\tau} \quad (4)$$

and not renewed otherwise, where $R_{n\tau}$ is the benefit of renewing the patent in country n at time τ and $NR_{n\tau}$ the benefit of exploiting the technology without the protection of a patent in country n at time τ, and $\kappa_{n\tau}$ denotes the cost of renewing the patent right in country n at time τ.[10] In the empirical section, the duration of patent j is defined as the period between the date of filing and the date when the patent is no longer renewed in any of its family countries.[11]

III. Data sources

We utilize a dataset of patent families, each of which is composed of the priority patent and subsequent patents that have been granted. Data on this come from a worldwide patent statistical database, *PATSTAT*, a database gathered by the European Patent Office (EPO) on behalf of the OECD Taskforce on Patent Statistics. We employed the autumn 2018 version of *PATSTAT* Global which contains bibliographical data on several millions of patent documents from around the world, including the legal events involving patents (such as payment of renewal fees and cessation of patent rights).

We focus on granted patents so that we can study their renewal and lapses, the key aspects that determine their value. Our empirical analysis covers the cohorts of priority patents associated with patent application years 1982 to 2012 (inclusive).[12] We use the application date of the priority filing as the application date for the whole family. While we collected data on patent grants data up to 2017, we end the sample in 2012 to enable us to adequately observe the outcomes of various metrics of patent valuation, such as patent citations received in the first few years of patent life beyond 2012. Our sample also requires that priority patents have expired, or lapsed, during the sample period so that we can know the ultimate lifespan of the priority patents in the sample.

In our empirical analysis, we also restricted the sample to priority patent grants that have a family size of at least two countries. There are two reasons for doing so. The first is to test whether international market coverage reveals the underlying value of a patent, which we would not capture well if we included domestic-only patents (or singletons). The second is that a significant number of these domestic-only patents are where the priority filing is in the U.S. The U.S. is a very large market – its GDP is typically greater than the sum of that of the other four nations – and so domestic U.S. patents would still be valuable on that basis, even if their family sizes are just one. Thus, a potential problem is that some patents with no subsequent filings may send mixed messages: they could represent patents of modest value or high value if they cover a very large single market, like the U.S.[13] For these reasons, we focus our analyses on those patent families whose family size exceeds one.[14]

Construction of the dataset

We initially extracted just over eight million patent priority filings from five major inventor (source) countries – the USA, Japan, Germany, France and the United Kingdom – and four major applicant types – companies, individuals, universities and governments and non-profits – from 1982–2017. Again, we selected granted patents and patents of inventions. We considered all destination countries in which there could be subsequent filings, including the five source countries. Our data include first filings at the EPO and their subsequent validation in member states. After matching this first draw of the data to the available data on citations and lapses, our dataset

[10] See Schankerman and Pakes (1986) for more detailed analyses of the patenting decision. In the empirics, we do not explicitly model the costs of filing and renewing patent rights since we rely on revealed preference; namely, that the costs did not exceed the benefit of obtaining and maintaining rights for those patents still in force.

[11] For example, suppose a patent is filed in 1980 and has a patent family of three countries. In the first country, the patent is not renewed in 1985, in the second, it is not renewed in 1990, and in the third, it is not renewed in 1995. The patent's duration is then 15 years (1995 minus 1980). Van Zeebroeck (2007) calls this the single renewal approach. One advantage of this definition is that it is much less influenced by the maintenance fees of an individual country.

[12] We begin our sample after 1981 to allow for the shift in U.S. practice of introducing maintenance fees in December 1980.

[13] Our results still hold qualitatively if we include 'singletons', but the results are sharper and stronger without them.

[14] On the other hand, Dechezleprêtre, Ménière, and Mohnen (2017) provide an analysis that exploits information about the multiple patents within each country that can be part of an international patent family. This allows for the valuation of patents that are never patented abroad.

diminishes to just over two million observations. We culled further by matching to the available data on patent claims and arrived at a dataset of 1,354,926 observations (or priority filings).

Our construction of a patent family follows what Martinez (2011) calls the *single first filing based* families; each first filing forms a family with the subsequent filings that claim it as a priority. The number of subsequent filings plus the priority filing gives the count of the family. It is the geographic family size, or number of countries covered in the patent family.[15] The composition of countries in the patent family is identified by the application authorities associated with the filings. The GDP-weighted patent family size is obtained by summing the real gross domestic product of the countries in the patent family; the GDP is as of the year of application and are converted into constant 2005 U.S. dollars.[16] For EPO grants, we summed the GDP of the jurisdictions in which the patent was validated. We determined the countries in which an EPO grant was validated by examining the member states to which the payments of renewal fees were made.

The duration of a patent is computed as the length of time between the date of its earliest application and the date when it lapsed in all of the countries in the patent family. Figure 2 shows the distribution of patents in our sample by their duration. The mode of the distribution is between 11 and 12 years of patent life.

Information about citations comes from published patent documents. The citations-based indicator used here is the forward citations (including self-citations) received by a priority patent during the first 5 years after its earliest date of publication, which is our point of reference for counting citations. As Hall, Jaffe, and Trajtenberg (2001) discuss, the longer the time passes, the more opportunity a patent has to be cited, and so as to minimize this bias we consider citations over this fixed length (first 5 years).[17] In any case, citations tend to drop substantially after 6–8 years of publication. Details on citations are in *PATSTAT's* published patent documents. We identified the cited patent publication number and counted the number of times it was cited by other published patents within five years of the earliest publication year of the cited patent. We do not include citations in the non-patent literature.

The number of inventors associated with a priority patent is readily provided in *PATSTAT*. Patent portfolio size is the count of priority patents granted to the patent holder – namely, the person name or entity associated with the priority patent – by application year.[18] We use patent portfolio size as

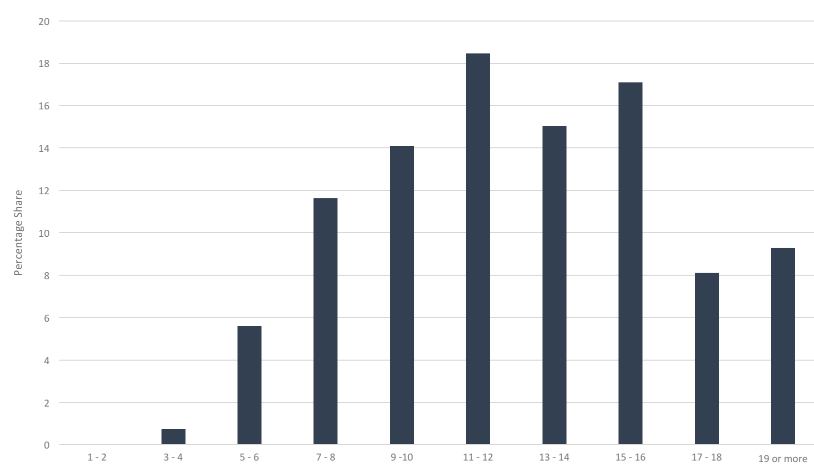

Figure 2. Percentage distribution of patents by duration.

[15]*PATSTAT* Global also readily provides the size of simple families via the DOCDB_family_id identifier.
[16]GDP data are from the World Bank's *World Development Indicators*. In preliminary analyses, we also worked with private GDP – netting out government expenditures – and found the results to be very similar.
[17]In preliminary analyses, we also worked with citations received during the first 8 years after publication and found the results to be highly similar.
[18]For example, if person X has priority (granted) patents A, B, and C indexed by application year T, the portfolio size equals three for person X at time T. Each patent A, B, and C is then associated with a portfolio size of three. The person name is based on a harmonized or standardized name. From the *PATSTAT* database, we aggregated priority patent grants by person identification number and application year.

a proxy for firm (or entity) size. The number of claims in a patent is also available in the published patent documents but is only available for certain U.S. and EPO patent publications. Hence, the sample size falls when we utilize data on patent claims.[19]

Lastly, we identify the technological field of the priority patent based on the World Intellectual Property Organization (WIPO)'s technology concordance with the International Patent Classification system. Our dataset spans 35 technology classes, from Electrical Machinery to Civil Engineering.[20]

Summary statistics

Table 1 presents some descriptive statistics. The top part shows the means and standard deviations of the variables of interest broken down by inventor country, and the bottom part by applicant type. The priority patent grant is the unit of analysis in this table. Hence, the entries show the statistics *per* patent. The sample consists mostly of the U.S. and Japanese patents (about 35% each). About three-quarters of the sample are company patents and about one-fifth are individual inventor patents.

The average patent duration is just under 12 years, compatible with Figure 2. Japan has the highest (at 13) while Germany has relatively the lowest (at 12.3). the U.S. patent grants have the most citations in the first 5 years, garnering just above 22 cites. The U.K. has the largest mean family size. However, once family sizes are weighted by GDP, the average Japanese patent covers the largest market size in terms of real GDP (i.e., $11.7 trillion in real 2005 U.S. dollars). By applicant type, university patents have the longest duration, most citations received and largest family sizes (weighted and unweighted).

Other indicators in the table will be used as control variables. Patent portfolio size is the number of patent grants held by the inventor (associated with the priority patent grant). It is intended to capture whether an inventor (or firm) is large or small. As Table 1 shows, the average portfolio size is about 69 patents, but this varies considerably across source countries and sectors. Companies typically have the largest patent portfolios and individual inventors the smallest. Japanese inventors have on average the largest patent portfolios and inventors from the U.K. the smallest among the

Table 1. Patent sample statistics, by source country and applicant type.

Group (% group share)	Duration (Years)	Citations first 5 yrs	Family Size	Weighted Family Size	Patent Portfolio Size	No. of Inventors	Patent Claims	Triadic Patents
France	12.8	11.3	7.6	11.0	13.1	2.2	12.6	0.6
(6.9%)	*(4.5)*	*(21.2)*	*(4.8)*	*(5.1)*	*(26.6)*	*(1.4)*	*(9.5)*	*(0.5)*
Germany	12.3	10.8	5.9	10.6	41.6	2.5	12.2	0.5
(18.7%)	*(4.4)*	*(17.9)*	*(4.4)*	*(4.7)*	*(91.1)*	*(1.8)*	*(8.3)*	*(0.5)*
Japan	13.0	15.0	4.8	11.7	106.7	2.8	11.8	0.3
(35.5%)	*(3.8)*	*(25.7)*	*(3.8)*	*(3.9)*	*(220.6)*	*(1.9)*	*(9.7)*	*(0.5)*
U.K.	12.8	15.6	8.4	10.5	7.3	2.1	13.7	0.4
(3.8%)	*(4.4)*	*(28.6)*	*(6.9)*	*(4.3)*	*(16.4)*	*(1.5)*	*(9.6)*	*(0.5)*
U.S.A.	12.4	22.6	6.3	10.2	62.7	2.4	16.7	0.2
(35.1%)	*(4.0)*	*(40.9)*	*(7.0)*	*(3.4)*	*(193.3)*	*(1.7)*	*(13.4)*	*(0.4)*
Company	12.8	16.1	6.0	10.90	90.7	2.5	13.4	0.4
(75.1%)	*(4.2)*	*(29.7)*	*(5.5)*	*(4.4)*	*(204.9)*	*(1.7)*	*(10.9)*	*(0.5)*
Gov Non-Profit	12.8	16.3	6.0	10.69	12.3	2.8	13.4	0.4
(1.5%)	*(4.1)*	*(39.0)*	*(4.6)*	*(4.2)*	*(15.5)*	*(1.7)*	*(10.1)*	*(0.5)*
Individual	12.2	18.1	5.4	10.86	1.6	2.6	14.5	0.1
(22.0%)	*(3.6)*	*(33.7)*	*(5.7)*	*(2.7)*	*(1.7)*	*(2.0)*	*(11.7)*	*(0.2)*
University	13.3	25.6	6.2	10.92	11.0	2.7	17.8	0.2
(1.4%)	*(3.8)*	*(46.1)*	*(5.5)*	*(3.5)*	*(22.8)*	*(1.6)*	*(14.9)*	*(0.4)*
Total	12.6	16.7	5.9	10.9	68.8	2.5	13.7	0.3
N = 1,354,926	*(4.1)*	*(31.1)*	*(5.5)*	*(4.1)*	*(181.7)*	*(1.8)*	*(11.1)*	*(0.5)*

The unit of analysis is a granted patent. The table shows the means, with standard deviations in parentheses. Sample period is 1982–2012. Duration of patent life is measured from the date of application to the date of last. non-renewal. Family size is the count of countries associated with a priority patent. Weighted family size is the count, weighted by the country's GDP in trillions of real 2005 U.S. dollars. Citations are counts of forward citations by other patents during the first 5 years of a patent's life. Patent portfolio size is the number of patents granted to the inventor associated with the priority patent. The number of inventors is the number of such persons associated with the granted patent. Patent claims are the number reported in the published patent. Triadic indicates the fraction of patent families comprising the EPO, USPTO, and JPO.

[19]We have performed the analysis in a much larger dataset without patent claims data. The results are qualitatively similar and are available upon request. To avoid showing too many tables and figures, we only present the results with claims data.

[20]See http://www.wipo.int/ipstats/en/statistics/patents/pdf/wipo_ipc_technology.pdf for details.

countries we have considered. We also consider the number of inventors associated with each patent grant; the average is close to three across countries or across applicant types. the U.S. patent grants contain the most claims, while Japanese ones contain the least. Among applicant types, university patents contain relatively the most claims. Lastly, about 30% of patents in the sample are Triadic, having been filed in the USPTO, EPO and JPO. Percentage-wise, triadic patents are relatively more common in France, Germany and the U.K. (given their membership in the EPO), than in Japan and the U.S. They are also more common among company and government (and non-profit) patents than among individual and university patents.

IV. Methodology: forecasting exercise

Our task is to compare among the various indicators of patent value and conduct a kind of 'horse race' to see which one might contain the better signal of patent value. First, we compare the GDP-weighted patent family size measure against patent citations obtained within the first 5 years to predict the *duration of patents*. Second, we compare the GDP-weighted patent family size measure against the duration of patents to predict *patent citations* received within the first five years.

To carry out our test of predictive properties, we need an estimation period and an out-of-sample forecasting period. To obtain these, we randomly split our sample of priority patents into two balanced groups. The sampling is done to ensure that the two resulting subsets are representative of the overall population. The two resulting groups are balanced in terms of the representation of source country, technological fields, application years, family size, patent duration, portfolio size, number of inventors, patent claims and forward citations (in the first five years).

Let us call one sample A and the other sample B. We then have an equation for patent duration that contains one of the patent value indicators as one of the independent variables. The methodology here is to estimate this model using sample A, and then to use the estimated model (or fitted model) to predict the duration in sample B, and finally to compare the actual duration of the B sample against the forecasted duration of the B sample. The model, with whichever indicator of patent value, that has the best forecast accuracy (for example, the lowest root mean square percentage error) is deemed to be the best at predicting duration. For robustness, we perform the reverse: we estimate the model using sample B and use it to predict duration in sample A. This methodology is intended to mimic how the GDP-weighted patent family size measure might be used in practice; namely, to predict at time t (the present), how long the patent will live (an ex-ante perspective). The reason we do not perform this exercise (i.e., to forecast duration and then compare it to actual duration) with the original (whole) sample is that we would be using the same patent data to make the predictions that we used to estimate the parameters. By splitting the sample into two subsamples (A and B), we use an estimated model from one dataset (one realization of the world) to forecast duration in another draw of the data. In other words, we create separate environments in which to estimate and test the model, thereby creating conditions similar to out-of-sample forecasting.

To recap, we estimate the following equations on sample A:

$$T_i^{jnst} = \alpha_j + \alpha_n + \alpha_s + \alpha_t + \beta C_{it} + \gamma X_{it} + \varepsilon_i^{jnst} \tag{5}$$

$$T_i^{jnst} = \alpha_j + \alpha_n + \alpha_s + \alpha_t + \beta F_{it} + \gamma X_{it} + \varepsilon_i^{jnst} \tag{6}$$

$$T_i^{jnst} = \alpha_j + \alpha_n + \alpha_s + \alpha_t + \beta V_{it} + \gamma X_{it} + \varepsilon_i^{jnst} \tag{7}$$

where i indexes the (granted) priority patent, our unit of analysis, and T denotes the duration of the patent, and the ε error term. In each model, we control for fixed effects for the technological field j, the inventor country n, the applicant type s and the year of application t.[21] Equation (5) uses C_{it}, the number of citations by the fifth year, along with a vector of control variables, X, such as the number of inventors in the patent, the patent portfolio size of the assignee associated with the patent, and the number of claims in the patent, to predict duration T. Equation (6) uses F_{it}, the raw family size (i.e., simple counts of countries in the

[21]Applicant types are company, individual, government and non-profit, and universities.

family), instead of C, to help predict T, and Equation (7) uses V_{it}, the GDP-weighted family size measure, to help predict T.

We then obtain the following 'fitted' equations:

$$\hat{T_i^{jnst}} = \hat{\alpha}_j + \hat{\alpha}_n + \hat{\alpha}_s + \hat{\alpha}_t + \hat{\beta}C_{it} + \hat{\gamma}X_{it} \quad (8)$$

$$\hat{T_i^{jnst}} = \hat{\alpha}_j + \hat{\alpha}_n + \hat{\alpha}_s + \hat{\alpha}_t + \hat{\beta}F_{it} + \hat{\gamma}X_{it} \quad (9)$$

$$\hat{T_i^{jnst}} = \hat{\alpha}_j + \hat{\alpha}_n + \hat{\alpha}_s + \hat{\alpha}_t + \hat{\beta}V_{it} + \hat{\gamma}X_{it} \quad (10)$$

and apply them to the dataset, sample B. That is, using the estimated α's, β's and γ's from sample A, we plug in the data for C, F, V and X from sample B to generate predicted values $\hat{T}$ and compare them to the actual values of duration T in sample B.

As measures of forecast accuracy, we use two kinds. First, the *Root Mean Square Percentage Error*:

$$RMSPE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\hat{T}_i - T_i}{T_i}\right)^2}$$

where N is the sample size of sample B. The i^{th} subscript refers to the i^{th} patent in sample B (with other subscripts suppressed to avoid clutter). This gives us the average forecast errors $(\hat{T}_i - T_i)$ as a percentage of the actual value.

Another measure of forecast accuracy is *Theil's inequality coefficient* (the U1 version):

$$U = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{T}_i - T_i\right)^2}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{T}_i\right)^2} + \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(T_i\right)^2}}$$

where U = 0 if there is a perfect fit. Again, for robustness, we will repeat the entire exercise by reversing sub-samples: using sample A for estimation and sample B for prediction. For further tests of robustness, we will change the dependent variable of interest from duration to forward citations, and determine how well the ex-ante GDP-weighted family size measure of a priority patent predicts the citations that a patent will receive in five years' time and compare that to the ability of a patent's duration to predict the number of citations it receives in the first five years.

V. Results

Our first 'horse race' test involves predicting the duration of patents, and the second involves predicting the citations received. To control for the many unobservable factors, our estimation equations include a full set of source country fixed effects, year fixed effects, applicant type fixed effects, and technological field fixed effects. These fixed effects should capture differences in technologies, organizations, policy regimes, patenting costs and research productivity shifts over time, among others.

Forecasting

Table 2 reports on the fitted equations for predicting the duration of patents. The first four models were estimated using the data in sample A, and the last four models using sample B. Recall that we use the estimated equations from one dataset to perform out-of-sample forecasting on the other dataset.

In each table, columns 1 and 5 show the model without any of the key patent value indicators of interest, such as family size or GDP-weighted family size. We refer to this as the base case. The other columns include a patent value indicator of interest, one by one. If it were the case that the base case models predict the best, this would cast doubt on the usefulness of our indicators of patent value. As Table 2 shows, the coefficient estimates of the patent indicators – citations received in the first five years, family size and GDP-weighted family size – are all statistically significant at the 1% level. They indicate that patents that are relatively more heavily cited and have wider geographic coverage, particularly in large markets, tend to be more valuable in terms of being longer-lived (or not letting them lapse too soon by not renewing their patent rights). The control variables are generally positive influences. For example, patents with more claims tend also to be longer-lived and thus revealed to be more valuable. Patents with longer duration are also associated with more inventors and larger patent portfolios.

Table 3 contains the results of the 'horse race' tests. It summarizes the performances of the various patent value indicators at forecasting patent duration using the two measures of forecast accuracy discussed earlier: the root mean square percentage error (RMSPE) and Theil's U. Panel (i) reports on

Table 2. Regression model for predicting duration.

Dependent Variable:ln (Duration)								
	Sample A				Sample B			
Split Sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pat Portfolio Size	0.0323***	0.0319***	0.0319***	0.0033***	0.0315***	0.0311***	0.0309***	0.0027***
	(0.0003)	(0.0003)	(0.0003)	(0.0002)	(0.0003)	(0.0003)	(0.0003)	(0.0002)
No. of Inventors	0.0701***	0.0688***	0.0417***	0.0318***	0.0707***	0.0695***	0.0422***	0.0325***
	(0.0008)	(0.0008)	(0.0008)	(0.0006)	(0.0008)	(0.0008)	(0.0008)	(0.0006)
No. of Claims	0.1235***	0.1196***	0.1046***	0.0070***	0.1244***	0.1207***	0.1054***	0.0076***
	(0.0009)	(0.0009)	(0.0008)	(0.0005)	(0.0009)	(0.0009)	(0.0008)	(0.0005)
Citations, 1st 5 yrs.		0.0143***				0.0139***		
		(0.0005)				(0.0005)		
Family Size			0.2216***				0.2217***	
			(0.0010)				(0.0010)	
Weighted Fam				0.0836***				0.0836***
				(0.0001)				(0.0001)
Observations	677,247	677,247	677,247	677,247	677,676	677,676	677,676	677,676
Adj R-squared	0.971	0.972	0.974	0.984	0.971	0.972	0.974	0.984

Robust standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1.
Source country, year, sector, and technological field fixed effects are included.

Table 3. Predicting duration: measures of forecast accuracy.

	(1)	(2)	(3)	(4)
Additional RHS Variables:	Base Case (No Value Indicators)	Citations 1st 5 years	Family Size	Weighted Family Size
i. Using estimates from sample A to predict duration in sample B				
RMSPE	0.188	0.188	0.180	0.153
– Ratio to Weighted Fam	1.226	1.225	1.174	1.000
THEIL U	0.085	0.085	0.081	0.064
– Ratio to Weighted Fam	1.330	1.328	1.264	1.000
ii. Using estimates from sample B to predict duration in sample A				
	(5)	(6)	(7)	(8)
RMSPE	0.188	0.188	0.180	0.154
– Ratio to Weighted Fam	1.220	1.219	1.168	1.000
THEIL U	0.085	0.085	0.081	0.064
– Ratio to Weighted Fam	1.330	1.328	1.264	1.000

RMPSE denotes Root Mean Squared Percentage Error and THEIL U is Theil's measure of forecast accuracy. Each column should correspond to the column in Table 2 showing the estimated model. Columns 1, 5 show the forecast accuracy associated with not using any of the patent value indicators. Columns 2–4 and 6–8 show the accuracies associated with adding *one* of the indicators shown.

the use of the estimated model from sample A to predict duration in sample B, and panel (ii) shows the reverse. For each of these panels, the two different forecast accuracy values are reported, and these values are also normalized with respect to the forecast accuracy value associated with the GDP-weighted family size measure, which is shown in column 4 and column 8. This helps us better see how the other patent value indicators perform relative to our proposed GDP-weighted family size measure. Essentially, column 1 of panel (i) corresponds to the model in column 1 of Table 2, and so on for columns 2–4 in the same panel. Column 5 of panel (ii) corresponds to the model in column 5 of Table 2, and so on for columns 5–8 in this panel.

Based on both forecast accuracy criteria, the model which uses the GDP-weighted family size variable predicts the best in terms of yielding the lowest forecast errors. In columns 1–4 of panel (i), the RMSPE criterion shows that the base case model has a forecast error of 18.8%; that is, the deviation between actual and predicted is 18.8% of the actual duration value. Meanwhile the model which utilizes the GDP-weighted family size has a forecast error of 15.3%. Thus, the model which incorporates GDP-weighted family size performs almost 22.6% better than the base case model (see the ratio of the RMSPE of the base case to that of the weighted family size in column 4). It also performs 22.5% better than a model which utilizes forward citations, and 17.4% better than one using the simple (unweighted) family size. The Theil U criterion is in agreement. The GDP weighted family size variable predicts patent duration more accurately than forward citations by 32.8% (see column 2). The family size count is only marginally more accurate than forward citations at predicting duration (see column 3). The weighted family count measure improves forecast performance (according to the Theil U) by 26.4% over the simple count measure. While the advantage of patent family size is its simplicity – namely, it is easy to observe and apply – it ignores the market sizes of the countries in the patent family and thus does not have the predictive power of the weighted measure. Throughout panel (i), the weakest performance at predicting duration is the base case model which does not use any of three patent value indicators of interest.

The results in panel (ii) which come from reversing the roles of the sub-samples (i.e., using estimates from sample B to predict duration in sample A) are practically identical (to the first three decimal points). Hence, this split sample serves as a useful robustness check. Of course, the forecasting models could all be improved by combining the various patent indicators, such as citations and GDP-weighted patent family size, and entering them all in the forecasting equations. But our objective was to conduct a 'horse race' test and so it was more appropriate to demonstrate the usefulness by entering these patent indicators one at a time. The more important reason, however, is that in practice, citations are not observed ex-ante. At the time of a patent priority application, we do not observe reliable citation counts until some years later. We do, nevertheless, observe the family size and the GDP levels of countries in the patent family around the time of priority filing. Hence, a chief advantage of the GDP-weighted patent family size measure is that it utilizes information that is available relatively earlier, whether at the time of application or grant.

The horse race exercise we conducted focused on patent duration as the benchmark for patent value. We now turn to patent citations as the benchmark of value and see how the GDP-weighted family size measure performs against duration at predicting citations received by a patent in the first five years. The significance is to show that the GDP-weighted family size measure has broader influences on patent value beyond patent renewal behaviour, since duration is also an imperfect measure of value; for example, duration can be longer if the costs of renewing patents are lower, holding changes in the benefits or value of patents constant.

When forecasting citations, we need to account for the potential endogeneity between duration and citations during the time the patent is in force. Even though a patent can be cited long after the right expires or lapses, duration and citations may be correlated during the lifetime of a patent, since the more time passes, the more citations the patent may receive. For instance, more citations are likely to have been received in a patent's fifth year of life than in its first year (and the same patent may keep on receiving citations long after it expires). Our citation counts are based on a fixed duration. We have effectively stopped the citations 'clock' at the fifth year. Thus, to avoid possible simultaneity between citations and duration, we focus on the sample of patents that lived longer than five years. Hence, we measure the ability of the duration of patents that lived longer than five years to explain citations received in the first five years. This way, we determine if the eventual duration of the patent signals high patent value in terms of the number of cites it receives in the first five years, and compare the predictive ability of duration against the forecast performance of family size and GDP-weighted family size.

Table 4 shows estimates of the models for predicting forward citations. Note the reduction in the sample size as a result of dropping patents that only lived for five years or less. The first four columns show estimates using sample A and the last four using sample B. The RHS variables are statistically significant determinants; however, family size has a negative association with forward citations. Compared to the models for predicting duration, the models for predicting forward citations have a lower goodness of fit. Apparently, a lot of noise in patent citations exist that is hard to capture. With that in mind, we summarize the measures of forecast accuracy for the citations models in Table 5. As anticipated the forecast

Table 4. Regression model for predicting citations received in first 5 years.

Dependent Variable: ln (Citations)								
	Sample A				Sample B			
Split Sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pat Portfolio Size	0.0291***	0.0258***	0.0293***	0.0127***	0.0282***	0.0250***	0.0284***	0.0121***
	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)
No. of Inventors	0.0912***	0.0841***	0.1006***	0.0695***	0.0920***	0.0849***	0.1017***	0.0705***
	(0.0023)	(0.0024)	(0.0024)	(0.0023)	(0.0023)	(0.0024)	(0.0024)	(0.0023)
No. of Claims	0.2690***	0.2561***	0.2752***	0.2038***	0.2693***	0.2564***	0.2757***	0.2045***
	(0.0019)	(0.0019)	(0.0019)	(0.0019)	(0.0019)	(0.0019)	(0.0019)	(0.0019)
Duration		0.1047***				0.1032***		
		(0.0035)				(0.0035)		
Family Size			−0.0729***				−0.0753***	
			(0.0025)				(0.0025)	
Weighted Fam				0.0467***				0.0463***
				(0.0004)				(0.0004)
Observations	661,627	661,627	661,627	661,627	662,089	662,089	662,089	662,089
Adj. R-squared	0.777	0.777	0.777	0.781	0.778	0.778	0.778	0.782

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Source country, year, sector, and technological field fixed effects are included.

Table 5. Predicting citations: measures of forecast accuracy.

	(1)	(2)	(3)	(4)
Additional RHS Variables:	Base Case (No Value Indicators)	Duration	Family Size	Weighted Family Size
i. Using estimates from sample A to predict duration in sample B				
RMSPE	0.672	0.671	0.672	0.666
– Ratio to Weighted Fam	1.009	1.008	1.009	1.000
THEIL U	0.298	0.298	0.298	0.293
– Ratio to Weighted Fam	1.019	1.018	1.016	1.000
ii. Using estimates from sample B to predict duration in sample A				
	(5)	(6)	(7)	(8)
RMSPE	0.673	0.672	0.673	0.667
– Ratio to Weighted Fam	1.009	1.008	1.009	1.000
THEIL U	0.299	0.298	0.298	0.293
– Ratio to Weighted Fam	1.019	1.018	1.016	1.000

RMPSE denotes Root Mean Squared Percentage Error and THEIL U is Theil's measure of forecast accuracy. Each column should correspond to the column in Table 4 showing the estimated model. Columns 1, 5 show the forecast accuracy associated with not using any of the patent value indicators. Columns 2–4 and 6–8 show the accuracies associated with adding *one* of the indicators shown.

errors are larger, according to both the RMSPE and Theil U criteria. Nevertheless, the GDP-weighted family size variable performs relatively best at predicting future citations, including better than the duration of patents at predicting citations received.

Next, as a robustness check, we evaluate the use of a triadic patent indicator in a forecasting model. The triadic variable equals one if the patent included filings in the U.S. Patent and Trademark Office (USPTO), European Patent Office (EPO) and Japanese Patent Office (JPO), and zero otherwise. Right away, there are some issues with this indicator. Like the simple family count, it does not take into account market size. For example, triadic patents vary in terms of the other countries (or offices) in which the patent was filed; some might only have been filed in those three offices (EPO, JPO and USPTO) while others might have also been filed in relatively large markets like Canada, Australia, S. Korea, China and Brazil, and others only in relatively smaller markets like Ghana, Egypt and the Philippines. Furthermore, there might be differences in where the patent might have been validated in the EPO member states. Yet, another consideration is that the GDP value of a triadic patent (even holding the composition of countries constant) can vary over time, as GDP varies over time. Thus, the Triadic dummy variable does not capture all the market nuances.

Table 6 shows some investigations of the use of the Triadic dummy variable. Part (i) shows some sample statistics between patents that are part of a triadic patent family and those that are not. Triadic patents

Table 6. Triadic patents.

Patent Families	Duration	Citations 1st 5 yrs.	Family Size	GDP-weighted Fam Size	No. of Inventors	Patent Portfolio Size	No. of Claims
(i) Sample Means							
Triadic	13.8	9.0	7.2	12.1	2.6	46.6	12.2
Non-Triadic	12.1	20.4	5.2	10.3	2.5	79.8	14.5
(ii) Model Estimates							
		Dep. Var →	Duration		Citations 1st 5 yrs.		
			Sample A	Sample B	Sample A	Sample B	
		Pat Portfolio Size	0.0343***	0.0240***	0.0335***	0.0227***	
			(0.0003)	(0.0008)	(0.0003)	(0.0008)	
		No. of Inventors	0.0627***	0.1092***	0.0634***	0.1103***	
			(0.0008)	(0.0022)	(0.0008)	(0.0022)	
		No. of Claims	0.1181***	0.2826***	0.1191***	0.2829***	
			(0.0009)	(0.0018)	(0.0009)	(0.0018)	
		Triadic Patent	0.2449***	−0.6523***	0.2422***	−0.6584***	
			(0.0013)	(0.0032)	(0.0013)	(0.0032)	
		No. Observ.	677,247	677,247	677,676	677,676	
		Adj. R-sq.	0.973	0.790	0.973	0.791	
		Forecast Accuracy of Model relative to GDP-weighted Family Size:					
		in →	Sample B	Sample A	Sample B	Sample A	
		RMPSE	1.205	1.197	1.005	1.004	
		THEIL U	1.289	1.288	1.001	1.003	

Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Source country, year, sector, and technological field fixed effects are included.

have longer durations, larger family sizes (simple or weighted by GDP), but non-triadic patents have on average received far more citations, are associated with larger patent portfolios, and contain more claims. Thus, triadic patents do not dominate non-triadic patents across indicators of patent value. Part (ii) shows the regression results and the forecast accuracy statistics (relative to the GDP-weighted family size measure). What is notable here is that in one split sample, triadic patents are positively associated with duration and forward citations, but in the other sample, they are negatively associated with those variables. Thus, the Triadic variable is not a consistent determinant in the forecasting model. In all cases, the root mean square percentage errors and the Theil U values are higher when the forecasting models include the triadic variable than when they include the GDP-weighted family size variable. For predicting duration, even the simple family size measure would outperform the triadic patent variable (which can be seen by comparing the relative forecast accuracies in Table 3).

Select samples

This section examines how the forecasting results hold up if we drill down to further sub-samples of the dataset. First, we re-estimated the models (as shown earlier in Tables 2 and 4) by different types of applicants. We then repeated the forecasting exercises. The results for predicting duration are shown in Table 7. To conserve space, we only report the forecast accuracy measures. The results at the disaggregated level are remarkably similar to those for the pooled sample. For example, the models with the GDP-weighted patent family size variable produce the lowest RMSPE; that is, the deviation of predicted duration from actual duration is between 15% and 16% of the actual duration, whether we perform the out-of-sample forecasting in sample A or B. In the pooled sample (recall Table 3), the forecast error was 15.3% or 15.4%. The other models, using either raw family size or forward citations, and the base case model all perform worse; by the RMPSE criterion, the other models produce forecast errors that are more than 15% greater than the errors generated by the model with the weighted family size variable, and the base case model produces errors at least 20% greater. By the Theil U criterion, the other models produce forecast errors that are at least 25% greater than those of the model with the weighted family size variable. Overall, the GDP-weighted family size measure dominates the forecasting outcomes for all applicant types, but it performs relatively best for company patents.

Table 8 shows the results for predicting citations by applicant sector. Again, it appears to be

Table 7. Predicting duration: forecast accuracy by applicant type.

		Sample A estimates to Predict in sample B				Sample B estimates to Predict in sample A			
		(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Sector	Forecast Accuracy	Base Case	Citations First 5 years	Family Size	Weighted Family Size	Base Case	Citations First 5 years	Family Size	Weighted Family Size
Company	RMSPE	0.189	0.189	0.181	0.155	0.189	0.189	0.181	0.156
	Ratio to Col. 5	1.219	1.217	1.165	1.000	1.210	1.208	1.157	1.000
	THEIL U	0.086	0.085	0.082	0.065	0.086	0.085	0.082	0.065
	Ratio to Col. 5	1.317	1.314	1.262	1.000	1.315	1.312	1.260	1.000
Individual	RMSPE	0.196	0.195	0.188	0.157	0.197	0.195	0.188	0.157
	Ratio to Col. 5	1.248	1.242	1.197	1.000	1.249	1.240	1.194	1.000
	THEIL U	0.086	0.086	0.082	0.066	0.086	0.086	0.082	0.066
	Ratio to Col. 5	1.312	1.311	1.253	1.000	1.315	1.312	1.254	1.000
Gov Non-Profit	RMSPE	0.194	0.193	0.189	0.156	0.194	0.193	0.189	0.158
	Ratio to Col. 5	1.239	1.237	1.208	1.000	1.230	1.227	1.199	1.000
	THEIL U	0.089	0.089	0.084	0.067	0.088	0.088	0.084	0.067
	Ratio to Col. 5	1.324	1.324	1.259	1.000	1.323	1.323	1.258	1.000
Universities	RMSPE	0.196	0.196	0.188	0.157	0.196	0.196	0.188	0.156
	Ratio to Col. 5	1.249	1.249	1.203	1.000	1.252	1.252	1.202	1.000
	THEIL U	0.097	0.097	0.091	0.069	0.097	0.097	0.092	0.069
	Ratio to Col. 5	1.406	1.407	1.328	1.000	1.402	1.403	1.320	1.000

Number of observations: Company (1,018,094), Individual (20,809), Government & Non-Profits (297,668), and Universities (18,355)

Table 8. Predicting citations first 5 years: forecast accuracy by applicant type.

		Sample A estimates to Predict in sample B				Sample B estimates to Predict in sample A			
		(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Sector	Forecast Accuracy	Base Case	Duration	Family Size	Weighted Family Size	Base Case	Duration	Family Size	Weighted Family Size
Company	RMSPE	0.641	0.641	0.643	0.634	0.641	0.641	0.643	0.634
	Ratio to Col. 5	1.010	1.010	1.013	1.000	1.011	1.010	1.013	1.000
	THEIL U	0.310	0.309	0.308	0.303	0.310	0.309	0.308	0.303
	Ratio to Col. 5	1.023	1.020	1.017	1.000	1.023	1.020	1.017	1.000
Individual	RMSPE	0.931	0.928	0.933	0.902	0.927	0.924	0.930	0.900
	Ratio to Col. 5	1.032	1.029	1.035	1.000	1.030	1.027	1.034	1.000
	THEIL U	0.295	0.294	0.295	0.288	0.295	0.294	0.295	0.288
	Ratio to Col. 5	1.024	1.021	1.023	1.000	1.023	1.020	1.023	1.000
Gov Non-Profit	RMSPE	0.688	0.694	0.692	0.674	0.709	0.713	0.711	0.699
	Ratio to Col. 5	1.021	1.030	1.027	1.000	1.014	1.020	1.017	1.000
	THEIL U	0.305	0.303	0.305	0.298	0.302	0.301	0.302	0.296
	Ratio to Col. 5	1.021	1.017	1.021	1.000	1.019	1.017	1.020	1.000
Universities	RMSPE	0.898	0.903	0.872	0.946	0.898	0.901	0.868	0.951
	Ratio to Col. 5	0.949	0.954	0.921	1.000	0.944	0.947	0.913	1.000
	THEIL U	0.301	0.300	0.299	0.298	0.302	0.302	0.300	0.299
	Ratio to Col. 5	1.009	1.009	1.004	1.000	1.007	1.007	1.000	1.000

Number of observations: Company (1,018,094), Individual (20,809), Government & Non-Profits (297,668), and Universities (18,355)

harder to forecast the number of citations a patent will receive within the first 5 years. The RMSPE and Theil U take on relatively high values. The GDP-weighted family size measure performs relatively best at predicting citations except among university patents. Here, the simple family size measure performs relatively best and the GDP-weighted family size measure relatively worst. To speculate, it might be that the GDP-weighted family size variable is more suitable for predicting the citations of commercially oriented innovations rather than those of the academic-oriented ones. The results are qualitatively the same whether we use sample A or B for the out-of-sample forecasting.

Next, we break the sample down by technological fields. We tested the predictive ability of the GDP-weighted family size to forecast the patent duration of different technologies and found that this variable helps produce the lowest RMSPE and Theil U values uniformly for all 35 technological fields. In Table 9, we report just a select few fields, such as Electrical Machinery, Telecommunications, Computer Technology, Biotechnology, Pharmaceuticals, Medical Technology, Chemical Engineering and Transportation. These fields represent diverse industrial innovative areas, from the life sciences, such as biotech and pharm, to capital intensive sectors such as transportation and telecommunications. According to the results, the simple family size measure predicts duration less accurately than the weighted measure but more accurately than forward citations. The differences in forecast accuracy between the model with forward citations and the base case model are marginal. These results hold in both subsamples of the data. Furthermore, the relative forecast performances by technological fields are similar to those we found for the pooled sample (recall Table 3).

The results, however, are mixed for predicting citations, as shown in Table 10. To avoid clutter, Table 10 merely reports which included indicator is associated with the lowest forecast errors. Thus, for each technological field, there are four outcomes, depending on the two forecast accuracy criteria – RMSPE or Theil U – and the two subsamples – A or B. For 14 of the 35 technological fields, forecasting models that include the GDP-weighted patent family size measure yield uniformly the highest forecast accuracy, i.e., in all four outcomes. These fields include Chemical Engineering, Surface Technology, Electrical Machinery and Environmental Technology. For other technological fields, one forecast criterion might favour a model that includes family size or duration, while the other criterion favours a model with GDP-weighted family size (e.g., Semiconductors, Macromolecular Chemistry and Polymers, the Analysis of Biological Materials, Pharmaceuticals and Biotechnology). For Medical technology and Microstructures and Nanotechnology, there are instances where the base case model dominates in terms of forecasting citations received in the first five years. But for the most part, the GDP-weighted family size is the variable that appears most frequently as contributing to the lowest forecast errors. Thus, while the evidence is not as strong as it was for predicting duration, the GDP-weighted family size measure stands out as the leading predictor of forward citations among the different technological fields.

Applications

This paper focused on constructing GDP-weighted patent family size and testing their ability to predict which patents will become valuable. The next step is to apply these patent valuation measures in empirical research. As alluded to in the introduction, the GDP-weighted patent family size measure could, for example, be used to assess the productivity of R&D, the outcomes of patent reform, or the quality of technology transfers. We leave these kinds of applications for future scholarly work. In this section, we use our constructed

Table 9. By select technological field: predicting duration.

		Using sample A's estimates to predict in sample B				*Using sample B's estimates to predict in sample A*			
	Forecast	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Tech Field	Accuracy (Ratio to col. 4)	Base Case	Citations First 5 years	Family Size	Weighted Family Size	Base Case	Citations First 5 years	Family Size	Weighted Family Size
Electrical	RMSPE	1.269	1.267	1.220	1.000	1.268	1.266	1.219	1.000
Mach.	THEIL U	1.380	1.378	1.299	1.000	1.380	1.378	1.300	1.000
Telecomm	RMSPE	1.209	1.207	1.162	1.000	1.213	1.211	1.167	1.000
	THEIL U	1.385	1.383	1.304	1.000	1.386	1.386	1.305	1.000
Computer	RMSPE	1.211	1.211	1.170	1.000	1.210	1.211	1.170	1.000
Tech	THEIL U	1.406	1.406	1.323	1.000	1.406	1.406	1.325	1.000
Biotech	RMSPE	1.207	1.209	1.118	1.000	1.210	1.212	1.120	1.000
	THEIL U	1.368	1.368	1.301	1.000	1.364	1.363	1.295	1.000
Pharma.	RMSPE	1.218	1.218	1.119	1.000	1.214	1.214	1.120	1.000
	THEIL U	1.364	1.364	1.332	1.000	1.361	1.361	1.334	1.000
Medical	RMSPE	1.257	1.261	1.184	1.000	1.258	1.262	1.186	1.000
Tech.	THEIL U	1.381	1.383	1.306	1.000	1.384	1.386	1.310	1.000
Chemical	RMSPE	1.283	1.282	1.205	1.000	1.281	1.281	1.204	1.000
Engineer.	THEIL U	1.378	1.377	1.312	1.000	1.382	1.379	1.314	1.000
Transp.	RMSPE	1.286	1.283	1.246	1.000	1.287	1.284	1.244	1.000
	THEIL U	1.354	1.354	1.274	1.000	1.362	1.361	1.277	1.000

The combined sample sizes of samples A and B are as follows: Electrical Machinery 86,806; Telecommunications 42,537; Comp. Tech 79,125; Biotechnology 25,011; Pharmaceuticals 38,866; Chemical Engineering 37,642; Medical Technology 52,777; and Transportation 68,795.

Table 10. By select technological field: predicting forward citations.

	Technological Field	From Sample A to Predict in B RMSPE	From Sample A to Predict in B THEIL U	From Sample B to Predict in A RMSPE	From Sample B to Predict in A THEIL U	Sample Size
1	Electrical Mach.	Weighted Fam	Weighted Fam	Weighted Fam	Weighted Fam	86,806
2	Audio Visual	Family Size	Weighted Fam	Family Size	Weighted Fam	75,683
3	Telecomm	Family Size	Weighted Fam	Family Size	Weighted Fam	42,537
4	Digital Com.	Family Size	Weighted Fam	Family Size	Weighted Fam	26,017
5	Basic Comm Proc.	Family Size	Weighted Fam	Family Size	Weighted Fam	21,587
6	Computer Tech	Family Size	Weighted Fam	Family Size	Weighted Fam	79,125
7	IT Management	Family Size	Weighted Fam	Family Size	Weighted Fam	3,941
8	Semi-conductors	Family Size	Weighted Fam	Family Size	Weighted Fam	54,628
9	Optics	Family Size	Weighted Fam	Family Size	Weighted Fam	68,955
10	Measurement	Family Size	Weighted Fam	Family Size	Weighted Fam	63,227
11	Analy. Bio Mat	Duration	Weighted Fam	Duration	Weighted Fam	9,227
12	Control	Family Size	Weighted Fam	Family Size	Weighted Fam	20,692
13	Medical Tech	Weighted Fam	Weighted Fam	Base Case	Weighted Fam	52,777
14	Organic Chem.	Family Size	Weighted Fam	Family Size	Weighted Fam	59,890
15	Biotech	Family Size	Weighted Fam	Duration	Weighted Fam	25,011
16	Pharmaceuticals	Duration	Weighted Fam	Family Size	Weighted Fam	38,866
17	Macro Chem.	Family Size	Family Size	Family Size	Family Size	40,448
18	Food Chem.	Weighted Fam	Family Size	Weighted Fam	Family Size	8,025
19	Basic Mat. Chem.	Weighted Fam	Weighted Fam	Weighted Fam	Family Size	40,885
20	Metallurgy	Weighted Fam	Weighted Fam	Weighted Fam	Weighted Fam	28,789
21	Surface Tech.	Weighted Fam	Weighted Fam	Weighted Fam	Weighted Fam	26,326
22	Micro. Nanotech	Base Case	Weighted Fam	Family Size	Weighted Fam	1,048
23	Chem. Eng.	Weighted Fam	Weighted Fam	Weighted Fam	Weighted Fam	37,642
24	Environment	Weighted Fam	Weighted Fam	Weighted Fam	Weighted Fam	16,561
25	Handling	Weighted Fam	Weighted Fam	Weighted Fam	Weighted Fam	40,033
26	Mach. Tools	Weighted Fam	Weighted Fam	Weighted Fam	Weighted Fam	37,316
27	Engines ...	Weighted Fam	Weighted Fam	Weighted Fam	Weighted Fam	48,511
28	Textiles ...	Weighted Fam	Weighted Fam	Weighted Fam	Weighted Fam	44,587
29	Oth. Special Mach.	Weighted Fam	Weighted Fam	Weighted Fam	Family Size	43,235
30	Thermal Proc.	Weighted Fam	Weighted Fam	Weighted Fam	Weighted Fam	15,963
31	Mechan. Elem.	Weighted Fam	Weighted Fam	Family Size	Weighted Fam	50,588
32	Transportation	Weighted Fam	Weighted Fam	Weighted Fam	Weighted Fam	68,795
33	Furniture, Games	Weighted Fam	Weighted Fam	Weighted Fam	Weighted Fam	23,168
34	Other Consumer	Weighted Fam	Weighted Fam	Weighted Fam	Weighted Fam	22,544
35	Civil Eng.	Weighted Fam	Weighted Fam	Weighted Fam	Weighted Fam	31,493

The table reports the indicator with the best forecast accuracy under the given sub-sample and accuracy measure. Weighted Fam is short for weighted family size. The sample size is the sum of samples A and B.

measures to provide some trends in patent value over time and across countries and technological fields. This is of interest in and of itself as scholars debate the optimal design of patent systems (see Bessen and Maurer, 2009; Baker, Jayadev, and Stiglitz 2017).

Figures 3 and 4 present trends in the mean patent value by inventor country and select technological fields, respectively. For these figures, we did not restrict the sample to patents that lapsed, which we needed to do for purposes of our empirical testing. If we had plotted these figures with only lapsed patents, the graphs would show a steep decline in mean value during the 2010–-2016 period. This is because those patents that lapsed during that period had short lives and were deemed less valuable. Here, we bring back all patent grants, lapsed or not, and calculated the market value (in terms of GDP) of the patent family at the time of application.

We observe a steady rise in patent value, as measured by our approach, from the early 1980s to around 2008. At that point, the average value for the U.S. falls slightly but continues to grow. The growth in average patent value for Japan slows after 2008 but does not exhibit a significant decline. However, for the EPO member states like Germany, France and the U.K., the average patent value falls, especially in France where its 2014 value is about what it was in 2002. However, average patent value appears to rebound in the U.K. and France in 2015, but not in Germany.

Turning to the technological fields, we observe steady growth in the average value of electrical machinery, telecommunication and computer technology patents. In the life sciences fields, like pharmaceuticals and biotechnology, the mean patent value grows more quickly during the 2000s, but late in that decade, the mean values

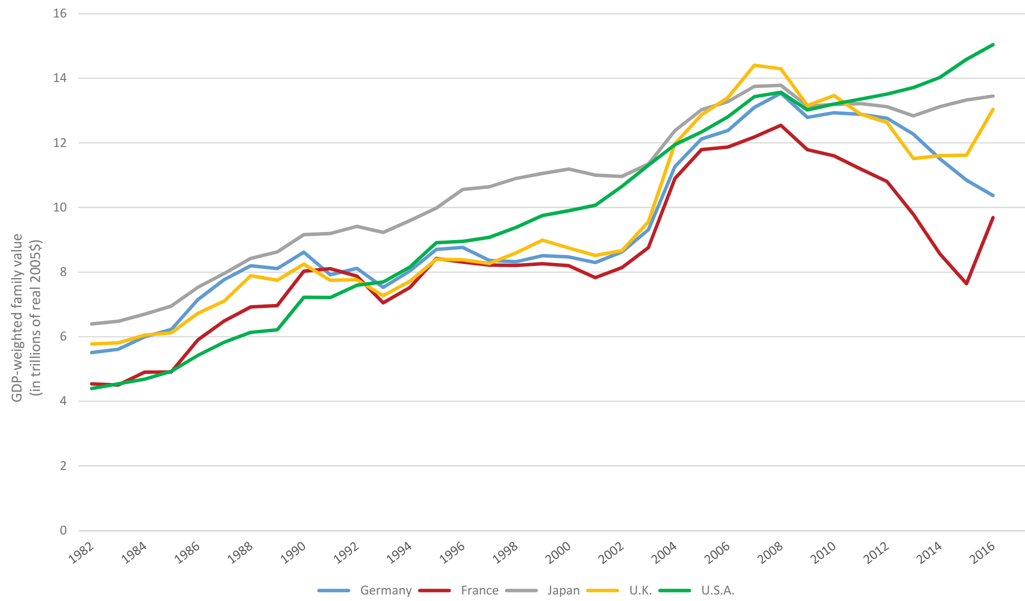

Figure 3. Trends in GDP-weighted family value, by inventor country.

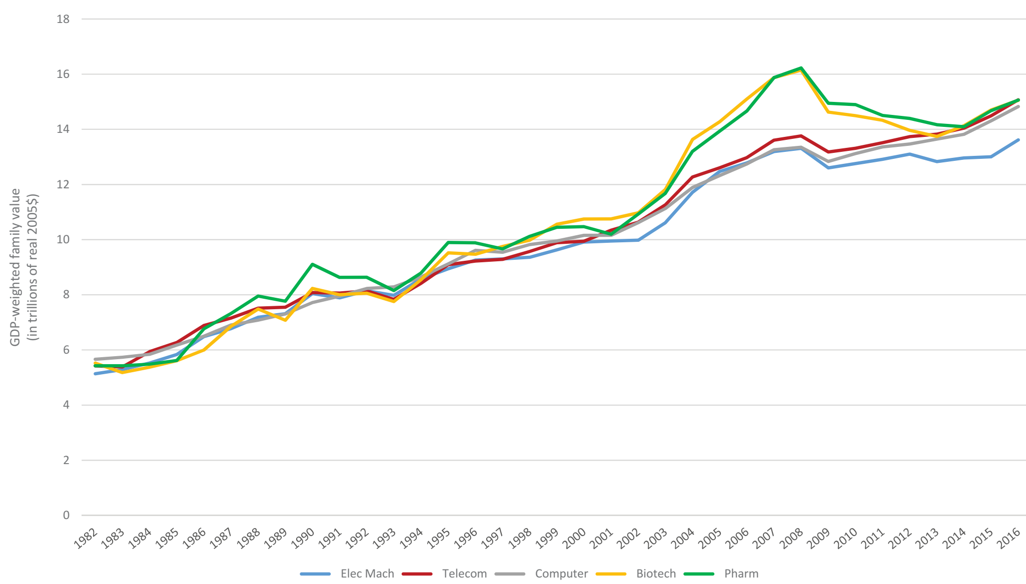

Figure 4. Trends in GDP-weighted family value, by select technology fields.

begin a declining trajectory till about 2013. It would be useful to study the determinants of these differing growth paths of patent value by technological field, whether they are driven by international trade agreements, technology policies, or shifts in innovation potential.

Of course, other developments may be occurring as well, such as shifts in the geographic coverage of countries in a patent family (due to trends in FDI) or in the costs of filing or maintaining a patent right. Shifts in geographic coverage can be accounted for when measuring trends in patent value; for example, we can construct family values per country in the patent family; that is, $\frac{\text{GDP-weighted family size}}{\text{Family Size}}$, which is essentially the ratio of Equation (3) to Equation (1), and can be readily constructed using the data we have. This approach can be useful if the sample consists only of lapsed patents. Figures 5 and 6 show the path of the ratio during the sample period. The interpretation of the ratio is the average market size of the

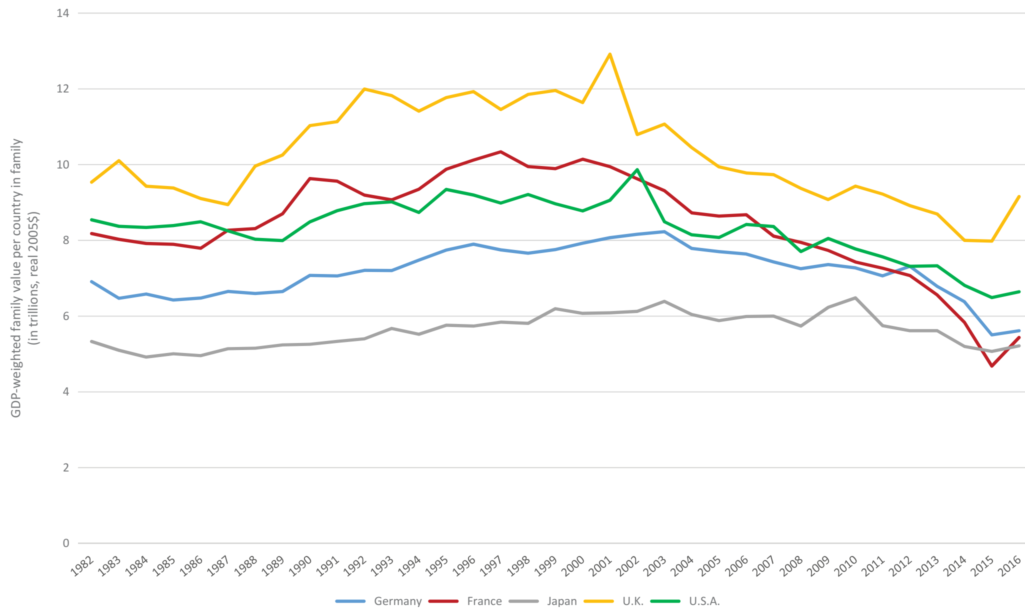

Figure 5. Trends in family value relative to family size, by inventor country.

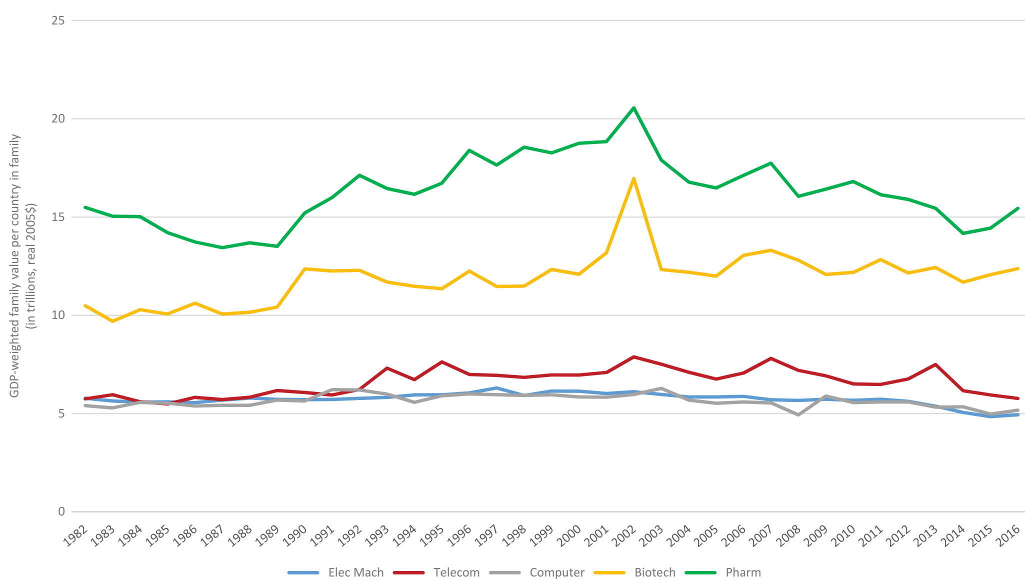

Figure 6. Trends in family value relative to family size, by select technology fields.

patent family, or the average GDP of the countries in the family.

Figure 5 shows that the rapid growth in mean patent values is tempered. The average market size of patent families appears stable until the 2000s, when they mostly decline. Figure 6 shows that for life science patents, the average market size experiences a temporary spike during the early 2000s, possibly reflecting a productivity boost that was initially concentrated in only a few markets. Again, these and other developments warrant further research, and could contribute to normative debates on the welfare effects of patent protection.

VI. Conclusion

In this paper, we constructed an indicator of patent value based on a patent's family size and composition, weighting each country in the family by its market size (namely GDP). This approach overcomes the disadvantage of existing measures of patent valuation which yield information about patent value after the fact; for example, after a patent's record of citations has been observed over a period of time, or after a patent right has been observed to lapse or be renewed at a later point in time. To the extent that our GDP-weighted patent family measure is a good *ex-ante* indicator of patent value, then *ex-post*, patents with greater GDP-weighted patent family values should live longer and have greater citations (over some specified period of time). This was the motivation for conducting the forecasting exercises; namely, to see whether our measure of interest can predict the valuable patents that we would observe ex post. To do this, we had split the sample into two groups, estimated a forecasting model over one group, and then used the other group, treating it as another draw of data, to conduct out-of-sample forecasts. We then reversed the situation, estimating the model over the latter group and conducting forecasting exercises with the first group. Our main result is that our measure of patent valuation out-performs the citations method at predicting patent duration and renewal, and out-performs patent duration at predicting citations, at least for the pooled sample. And it improves upon models that do not use any indicator of patent value or models that use a close substitute like an indicator for triadic patent families.

We also repeated the analysis by different technological fields and different applicant types. For predicting duration, models that incorporate the GDP-weighted patent family variable performs best across all fields and applicant sectors. For predicting forward citations, the models with the GDP-weighted patent family variable performs quite well in many different fields of technology, but not for all fields, such as patents for Macromolecular Chemistry and Polymers, nor for university patents. The evidence overall demonstrates the usefulness of the GDP-weighted family size measure as an ex-ante method of patent valuation.

Of course, this is not to say that ex-post valuation matters less. In many cases, especially for policy evaluation, it is useful to know whether patents granted in earlier periods turned out to be valuable and to provide feedback on the policies that gave rise to them (or failed to). However, where anticipating 'value' is crucial for making current resource allocation decisions, the valuation method discussed here serves to help minimize forecast errors.

But how well the method serves as an ex-ante indicator depends on the timeliness of the information needed to construct the measure. For our approach, we primarily need to know the countries covered by the patent family and the real value of the GDP of those countries at the time of application of the priority filing. Under the *Paris Convention*, after a first application is filed in one of the member states, the subsequent filings for the same invention can be made in other member states within a year and be regarded as if they were filed at the same time as the first application. In this case, there would be a 1-year lag before the country composition of the patent family can be fully known. But with more complex-structured patent families, the timeliness of information on country coverage would be a bigger issue as the lags between the first application and later filings in other member states would be much longer.

As extensions to this paper, the GDP-weighted patent family size measure can be used to create indicators of patent values by industry and/or country, and they in turn can be applied to studies on productivity, R&D, or technology diffusion. We have already provided some preliminary applications of the method for purposes of analysing the trends in patent value by country and technological field. It will be useful in future work to study the effects of changes in the propensity to patent, which need not remain stable over time. The GDP-weighted family size measure may potentially fluctuate over time because of shifts in the patenting strategies of firms. Thus, the measure may be more useful if smoothing techniques were applied to it (like a moving average adjustment or exponential smoothing). Another exercise would be to test the valuation method by relating patent licensing fees to the market value of the patent's family size. Future work could also try to compute the market size of patent families under a more extensive family definition.[22] As Martinez (2011) points out, about

[22]An extended definition seeks to consolidate all the direct and indirect priority linkages among the patent applications within a family.

25% of extended priority patent families do not overlap with the single first filing families, leading to different counts of patent families as well as different sizes of patent families.

Acknowledgments

We thank Jonas Anderson, Mafini Dosso, Robert Feinberg, Peter Hingley, Keun Lee, Alan Marco, Kara Reynolds, Nate Sawadogo, Wonkyu Shin and seminar participants at STEPI and the U.S. Patent and Trademark Office (USPTO) for feedback and discussions. We are also grateful to the referee for thorough comments and suggestions, and to the Institute for Humane Studies (IHS) for financial support.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was partially supported by the Institute for Humane Studies, George Mason University.

References

Austin, D. H. 1993. "An Event-Study Approach to Measuring Innovation Output: The Case of Biotechnology." *American Economic Review* 83 (2): 253–258.

Baker, D., A. Jayadev, and J. Stiglitz. 2017. *Innovation, Intellectual Property, and Development: A Better Set of Approaches for the 21st Century*. Washington, DC: Center for Economic and Policy Research.

Barney, J. 2002. "A Study of Patent Mortality Rates: Using Statistical Survival Analysis to Rate and Value Patent Assets." *AIPLA Quarterly Journal* 30 (3): 317–352.

Berger, R. 2005. *A Study on the Cost of Patenting, Prepared on Behalf of the European Patent Office*. Munich, Germany: Roland Berger Market Research.

Berrier, E. F., Jr. 1996. "Global Patent Costs Must Be Reduced." *IDEA: the Journal of Law and Technology* 36 (4): 473–511.

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

Bessen, J., and M. Meurer. 2009. *Patent Failure: How Judges, Bureaucrats, and Lawyers Put Innovators at Risk*. Princeton, NJ: Princeton University Press.

de Rassenfosse, G., H. Dernis, and G. Boedt. 2014. "An Introduction to the PATSTAT Database with Example Queries." *Australian Economic Review* 47 (3): 395–408.

Dechezleprêtre, A., Y. Ménière, and M. Mohnen. 2017. "International Patent Families: From Application Strategies to Statistical Indicators." *Scientometrics* 111: 793–828.

Dernis, H., D. Guellec, and B. van Pottelsberghe De La Potterie. 2001. "Using Patent Counts for Cross-Country Comparisons of Technology Output, OECD." *Science Technology Industry Review* 27: 128–146.

Ernst, H., and N. Omland. 2011. "The Patent Asset Index - A New Approach to Benchmark Patent Portfolios." *World Patent Information* 33: 34–41.

Fischer, T., and J. Leidinger. 2014. "Testing Patent Value Indicators on Directly Observed Patent Value - an Empirical Analysis of Ocean Tomo Patent Auctions." *Research Policy* 43: 519–529.

Frietsch, R., U. Schmoch, B. van Looy, J. P. Walsh, R. Devroede, M. Du Plessis, T. Jung, et al. 2010. *The Value and Indicator Function of Patents, Studien Zum Deutschen Innovationssystem Nr. 15-2010*. Berlin: Expertenkommission Forschung und Innovation (EFI).

Hall, B., A. Jaffe, and M. Trajtenberg. 2001. "The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools." *NBER Working Paper No. 8498*.

Hall, B., A. Jaffe, and M. Trajtenberg. 2005. "Market Value and Patent Citations." *RAND Journal of Economics* 36 (1): 16–38.

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

Harhoff, D., F. M. Scherer, and K. Vopel. 2003. "Citations, Family Size, Opposition and the Value of Patent Rights." *Research Policy* 32 (8): 1343–1363.

Harhoff, D., and M. Reitzig. 2004. "Determinants of Opposition against EPO Patent Grants the Case of Biotechnology and Pharmaceuticals." *International Journal of Industrial Organization* 22 (4): 443–480.

Henry, M. 2013. "The Market Effects of Patent Litigation." *Technology and Investment* 4 (1): 57–68.

Hingley, P., and W. G. Park. 2003. "Patent Family Data and Statistics at the European Patent Office." *WIPO-OECD Workshop on Statistics in the Patent Field*.

Jaffe, A., and M. Trajtenberg. 2002. *Patents, Citations and Innovations: A Window on the Knowledge Economy*. Cambridge, MA: MIT Press.

Johnstone, N., I. Hai, J. Poirier, M. Hemar, and C. Michel. 2012. "Environmental Policy Stringency and Technological Innovation: Evidence from Survey Data and Patent Counts." *Applied Economics* 44 (17): 2157–2170.

Kappos, D., and S. Graham. 2012. "The Case for Standard Measures of Patent Quality." *MIT Sloan Management Review*. Spring. http://sloanreview.mit.edu/article/the-case-for-standard-measures-of-patent-quality

Kumazawa, R., and P. Gomis-Porqueras. 2012. "An Empirical Analysis of Patents Flows and R&D Flows around the World." *Applied Economics* 44 (36): 4755–4763.

Lanjouw, J., A. Pakes, and J. Putnam. 1998. "How to Count Patents and Value Intellectual Property: The Uses of

Patent Renewal and Application Data." *Journal of Industrial Economics* 46 (4): 405–432.

Lanjouw, J., and M. Schankerman. 2004. "Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators." *Economic Journal* 114 (495): 441–465.

Martinez, C. 2011. "Patent Families: When Do Different Definitions Really Matter?" *Scientometrics* 86 (1): 39–63.

Mauck, N., and S. W. Pruitt. 2016. "The Valuation of Patents Using Third-Party Data: The Ocean Tomo 300 Patent Index." *Applied Economics* 48 (42): 3995–3998.

Murphy, W., J. Orcutt, and P. Remus. 2012. *Patent Valuation: Improving Decision-Making Through Analysis*. New Jersey: Wiley Finance.

Nair, S., M. Mathew, and D. Nag. 2011. "Dynamics between Patent Latent Variables and Patent Price." *Technovation* 31 (12): 648–654.

Neuhausler, P., and R. Frietsch. 2013. "Patent Families as Macro Level Patent Value Indicators: Applying Weights to Account for Market Differences." *Scientometrics* 96: 27–49.

Pakes, A., and M. Simpson. 1989. "Patent Renewal Data." *Brookings Papers on Economic Activity: Microeconomics* 1989: 331–410.

Park, W. G. 2010. "On Patenting Costs." *The WIPO Journal: Analysis of Intellectual Property Issues* 2 (1): 38–48.

Putnam, J. 1996. "The Value of International Patent Rights." Ph.D dissertation, Yale University, New Haven, CT.

Schankerman, M., and A. Pakes. 1986. "Estimates of the Value of Patent Rights in European Countries during the Post-1950 Period." *Economic Journal* 96 (384): 1052–1077.

Smith, G., and R. Parr. 2000. *Valuation of Intellectual Property and Intangible Assets*. 3rd ed. New York, NY: John Wiley and Sons Inc.

Sneed, K., and D. Johnson. 2009. "Selling Ideas: The Determinants of Patent Value in an Auction Environment." *R&D Management* 39 (1): 87–94.

Tong, X., and F. J. Davidson. 1994. "Measuring National Technological Performance with Patent Claims Data." *Research Policy* 23: 133–141.

Trajtenberg, M. 1990. "A Penny for Your Quotes: Patent Citations and the Value of Innovations." *Rand Journal of Economics* 21: 172–187.

van Pottelsberghe De La Potterie, B., and N. van Zeebroeck. 2008. "A Brief History of Space and Time: The Scope-Year Index as A Patent Value Indicator Based on Families and Renewals." *Scientometrics* 75 (2): 319–338.

van Zeebroeck, N., and B. Van Pottelsberghe de la Potterie. 2011. "Filing Strategies and Patent Value." *Economics of Innovation and New Technology* 20 (6): 539–562.

van Zeebroeek, N. 2007. "Patents Only Live Twice: A Patent Survival Analysis in Europe." *Universit Libre de Bruxelles Working Paper*.