Paper to be presented at the DRUID Academy Conference 2017 at University of Southern Denmark, Odense, Denmark on January 18-20, 2017

Measuring Patent Value: a Critical Reappraisal of the Most Popular Indicators

Giovanna Capponi
Sant'Anna School of Advanced Studies
Institute of Economics
g.capponi@sssup.it

Arianna Martinelli
Sant'Anna School of Advanced Studies
Institute of Economics
a.martinelli@sssup.it

Alessandro Nuvolari
Sant'Anna School of Advanced Studies
Institute of Economics
a.nuvolari@sssup.it

Abstract

Measuring Patent Value: a Critical Reappraisal of the Most Popular Indicators Authors: G. Capponi, A. Martinelli, A. Nuvolari Email: g.capponi@sssup.it Affiliation: Sant'Anna School of Advanced Studies Ph.D status: enrolled in the 3rd year, expected final date October 2018 The increasing number of patent applications filed in the past few decades calls for more reliable estimates of patent value to understand their actual impact on the economy (Gambardella et al, 2008). Notably, the nature and the relevance of inventions vary considerably, ranging from the technologies which reshape the landscape to those which do not reach the commercialization phase (Arts et al, 2013). So far, a number of patent based indicators have been used as proxies for value: forward citations count, number of renewals and family size to name a few. At the same time, a constant effort has been devoted to the development of an effective composite index able to capture the value-related variation of the single measures (Squicciarini et al, 2013 and Lanjouw and Schankerman, 2004). Still, these approaches present some limitations: above all, being indirect measures, the actual relationship between these indicators and patent value remains widely untested; in addition, patent based indicators may suffer from the so called "patent premium" bias (Gambardella et al. 2008). Against this background, we seek to assess which patent based indicator(s) best captures the actual relevance of the underlying invention. More in details, we expect this paper to contribute to the growing literature on the value of patents in two main ways: 1) providing a

direct and exogenous measure of patent value and 2) comparing the performance of indirect indicators in predicting this measure, disentangling the variation related to technological importance or to market opportunities. The external indicator we propose for validation is based on the Queen Award for Innovation (QAI), one of the most prestigious accolade for technological achievements in the UK. The outline of the scheme specifies that the Award looks at “outstanding innovations reaching commercial success”, so, it is important to stress that while it certainly recognizes market performance, the definition of “outstanding” covers a range which is broader than technical relevance (Queen's Award Committee, 2013). We collect data on this prize between 1976 and 2015, matching the innovations Awarded to patents. From a total of 1234 technological achievements, we identify 401 as patented. For the purpose of this study, within the treated sample we keep only patent applications filed in the US which find correspondence within the OECD Quality database, which lists around 7 million USPTO patents filed since 1976, along with the relevant indicators of patent quality (Squicciarini et al, 2013). From the same database, we retain as control patents only those having the same filing year and first IPC-CLASS as the treated group, dropping the observations with missing values in the variables of interest. As a first approach, to have a balance sample, we randomly select one control patent per each treated one, ending up with 642 total observations. The paper we present tests the performance of patent indicators as proxies of value by regressing them against a binary variable taking value of 1 if the patent refers to an Award winning innovation, 0 otherwise. The choice of the indicators falls on the typical measures used in the literature on patent quality: number of forward citation within five years after publication, number of claims, family size and number of renewals. As a first approach, we construct logit models to compare the performance of the different indicators in predicting the binary outcome, controlling for the filing year and the technical field (Arts et al, 2013). In light of the results at single indicators level, we specify a confirmatory factor analysis with a generalized structural equation model, in order to extract two latent variables capturing separately patents' technological and commercial value (Lanjouw and Schankerman, 2004). At the single measures level, the result show that the best predictor of an “outstanding innovations reaching commercial success” is patents’ family size. In line with this, looking at the latent variables we observe that the market-related factor best captures patent value and it is mostly correlated to family size and number of renewals. The technology-related factor is a significant predictor of the external validation measure but its impact is weaker compared to the market-related factor. Also, we observe a positive correlation between the two latent variables, highlighting that patents with high technological relevance, are characterized by a greater market value as well. References: Arts, Sam, Francesco Paolo Appio, and Bart Van Looy. "Inventions shaping technological trajectories: do existing patent indicators provide a comprehensive picture?." Scientometrics 97.2 (2013): 397-419. Gambardella, Alfonso, Dietmar Harhoff, and Bart Verspagen. "The value of European patents." European Management Review 5.2 (2008): 69-84. Lanjouw, Jean O., and Mark Schankerman. "Patent quality and research productivity: Measuring innovation with multiple indicators." The Economic Journal 114.495 (2004): 441-465. Queen’s Award Committee, Report (2013) Squicciarini, Mariagrazia, Hélène Dernis, and Chiara Criscuolo. "Measuring Patent Quality." (2013).

Jelcodes:O31,O34

Measuring Patent Value: a Critical Reappraisal of the Most Popular Indicators

Giovanna Capponi,* Arianna Martinelli,† and Alessandro Nuvolari‡
Scuola Superiore Sant'Anna

This paper investigates the performance of five popular patent–based indicators in measuring patent value. We introduce a unique measure of patent value based on the Queen's Award for Innovation, a prestigious British Award that recognizes businesses' outstanding achievements in innovation. Then, we test the effectiveness of indicators such as renewals, forward citations, family size, number of claims and a composite index estimated as a latent factor in discriminating the most from the least valuable patents. We find that the composite indicator better captures the value related variation of patents, however the magnitude of the impact is very small. Further investigation shows that the relationship between patent–based indicators and patent value is complex and that the exact informative content of each measure is still to be defined.

I. Introduction

The increasing number of patent applications filed in the past few decades calls for more reliable estimates of patent value to understand their economic and technical significance (Gambardella et al. 2008). Notably, the nature and the relevance of inventions vary considerably, ranging from the technologies which reshape the landscape to those which do not reach the commercialization phase (Arts et al. 2013). As a consequence, patents are highly heterogeneous and their value distribution is very skewed, so that a simple patent count does not constitute a good indicator of innovative output by itself (Harhoff et al. 1999).
In this respect, a growing strand of literature investigates the behaviour of selected patents' features – renewals, forward citations and family size to name a few – to see whether and to which extent they could be used as a proxy of patent value. In principle, patent–based indicators are an attractive solution to this measurement problem: one the one hand, they are standardized, widely and easily available, on the other hand, the studies performed so far show that they are empirically correlated to patent value. However, being indirect measures, they fail to tackle a substantial share of the variation in patent value (Gambardella et al. 2008). Given these premises, further investigation in this direction is required to better understand which indicators are most suitable, what does each of them tell us, and how can we fully exploit their informative power.
As part of the research effort addressing this issue, the paper we present focuses on the relationship between the private value of patents and their characteristics. In particular we expect to contribute to the literature in two main ways: first, we provide a unique indicator of patent value based on 243 innovations winning a Queen's Award for Innovation (hereafter QAI), a British Award which acknowledges successful products or processes few years after commercialization. Through a matching procedure, we retrieved the original patents referring to the awarded innovations and we set them as a benchmark of valuable patents. Second, we compare the effectiveness of five commonly used indicators in discriminating the benchmark group from other control patents selected within the portfolios of QAI winning firms. In particular, the value measure we adopt is a binary variable signalling whether the patents led to an Award winning innovation or not. As predictors, we include renewals,

* g.capponi@sssup.it
† a.martinelli@sssup.it
‡ a.nuvolari@sssup.it

forward citations, family size, claims and an adjusted version of a composite indicator developed by Lanjouw and Schankerman (2004) which estimates a value variable as a latent factor among the other four indicators. Our results show that the composite index outperforms the single indicators in terms of discriminatory power, suggesting that patent features do share a value component which can most effectively be estimated as an underlying latent construct. More in general, we find that all the indicators we tested have a systematic but very small impact on patent value.
The reminder of this paper is structured as follows. Section II provides an overview of the literature on patent value estimation, highlighting the different approaches. Section III presents the Queen's Award prize scheme and the process behind the dataset construction. Section IV goes through the empirical analysis, describing the framework and the results. In the last part we discuss the findings and we provide some final remarks.

II. Literature review

A. Patent value: what is it and why is it important?

The concept of 'patent value' can be interpreted in two main ways: first, it can refer to the economic private value to the patent owner, that is the discounted stream of revenues that he or she expects to generate from a patent (OECD 2009). As part of these returns, we include also the intrinsic value of patents, the so called 'patent premium', which can be defined as the difference between the value of the patented invention and the technological contribution of the invention to the state of the art (Arora et al. 2008). Second, there is a social component of value that is linked to the patent's contribution to the society's stock of technology (OECD 2009).
To be able to estimate the value of patents is becoming increasingly important especially because of the unprecedented surge in the number of applications filed worldwide. There is consensus among practitioners and academics that the recent changes in the patent legal framework fueled this trend by directly strengthening patent protection and by indirectly lowering patent examination standards. These premises also paved the way to a growing use of patents for strategic reasons, generating the so called 'patent paradox', which refers to a situation in which firms and industries do not rely on patents to profit from innovations, but are still patenting aggressively (Hall and Ziedonis 2001). Not surprisingly, it's long been established that simple patent count as a measure on inventive output is far too naive and that it should be weighted by more discriminatory value indicators in order to be informative (e.g. Schankerman and Pakes (1986)).

B. Measuring patent value

Most of the empirical studies in the field focused on the measurement of patents private value, developing and cross-testing different indicators to reach a reliable estimate. Following van Zeebroeck and van Pottelsberghe de la Potterie (2011), we can classify the most popular measures into two main categories: the 'market–based indicators' and the 'patent–based indicators'. The formers enable researchers to estimate patents' value in a direct way. In particular, the line of literature following this approach investigates the value of patents and the implications of their variability in terms of number by using data on the stock market

valuation of firms (Griliches (1981), Pakes (1985), Hall (1993), Hall et al. (2007)).
Instead, patent–based indicators are considered a more indirect way of measuring value. The stream of studies on this topic analyzes the relationship between patents' characteristics and the technological and economic importance of the inventions. The main patent related variables considered in the literature are forward and backward citations ((Carpenter et al. 1981), Trajtenberg (1990), Harhoff et al. (2003)) along with derived indexes of generality, originality and breakthrough (Trajtenberg et al. (1997), Ahuja and Morris Lampert (2001) Hall et al. (2001)), the number of claims (Tong and Frame 1994), the renewals data (Schankerman and Pakes (1986), Bessen (2008)), the family size (Putnam 1996) and opposition and litigation occurrence (Lanjouw and Schankerman (2001), Harhoff et al. (2003)). In general, there is evidence of positive correlation between the indicators and the value of the patents. However, the greatest portion of this relationship remains largely unexplained and it is necessary to gain new insights to better define the potential and the limitations of the indicators. To tackle this issue, we consider five commonly used measures:

1. Renewals

Early contributions adopted renewal decisions, viewed as the economic response of patentees, to infer the value of patents (Schankerman and Pakes (1986), Pakes (1986), Lanjouw et al. (1998)). These studies exploit the fact that each year, a patent holder should pay a renewal fee to keep a patent in force. As it is reasonable to assume that agents decide to renew a patent depending on the expected value it carries, it follows that renewals and renewal fee schedules embed relevant information on patents' value (Schankerman and Pakes 1986). Also, the number of renewals as a measure, allows researchers to obtain monetary estimates of patent rents (Bessen 2008). Still, this indicator alone presents some drawbacks: above all, it does not reach the extreme right tail of the distribution because it captures decisions taken within the limit of a patent statutory lifetime, setting a lower rather than an upper bound of patent value (Harhoff et al. (2003), Gambardella et al. (2008), Bessen (2009)). In addition, this approach assumes the value of patents to decrease monotonically, while the actual distribution may be more skewed (Gambardella et al. 2008).

2. Forward citations

Patents' forward citations as a measure of value were first introduced by Trajtenberg (1990) who suggested that patent count weighted by citations would in fact provide a measure of patent value. The link between the number of times that a patent is cited and its technological significance emerges from the following procedure: during a patent examination process, the examiner decides which other patents representing the prior art should be cited. Specifically, the new patent could protect an improvement of the prior technology, or an alternative way to apply it, or a different method to produce it (Jaffe et al. 2000). In principle, the existence of later patents building on cited patents, signals their role in establishing a successful line of innovation. Moreover, a certain continuity along the same line reflects a costly research effort which is presumably justified by a high expected economic value (Trajtenberg 1990).

3. *Claims*

Another variable related to patents value is the number of claims, to the extent that 'An inventor's invention is embodied in his or her claims' (Tong and Frame 1994). More in details, each patent contains a number of claims which disentangle the invention in different contributions, which represent an intrinsic invention themselves. Assuming that the technological effort is related to both scientific and economic activities, then the number of claims appear to be a good indicator of patent value (Tong and Frame 1994).

4. *Family size*

A patent family size refers to the number of countries in which a patent is filed and enforced. Since the geographical extension of a patent coverage is costly, only the inventions with a sufficiently high expected value on the market will be protected abroad (Lanjouw et al. 1998). Moreover, by comparing the patents within a family it is possible to single out their economic relevance as the difference in value across countries: given that they protect the same technology, most of the variation in quality is imputable to market size (Putnam 1996).

5. *Composite indicator*

Based on this evidence, researchers built also composite indicators either to obtain a comprehensive measure of value or to disentangle the variables embedding economic value from those carrying more technology-related information. In this study we test the performance of a composite indicator resembling the one developed by Lanjouw and Schankerman (2004). Relating forward and backward citations, number of claims and family size through a confirmatory factor analysis, the authors show that the use of multiple indicators leads to a greater variance reduction in quality estimates. For the purpose of our paper, we construct the index including renewals instead of the number of backward citations.

C. Comparing patent-based indicators

As we mentioned before, patents' features are an indirect measure of patent value. The main limitation to their validity lies in the fact that patents contain selective and therefore, incomplete information, especially because of the lag between their filing and the product commercialization (Hsieh 2013). For this reason, it is necessary to test the performance of indicators against external evidence of patent value. So far, survey based studies addressed this issue providing a benchmark to assess the economic relevance of patents (Harhoff et al. (2003), Gambardella et al. (2008)). For example, Gambardella et al. (2008) grounded their analysis on the Pat–Val EU survey, a comprehensive investigation conducted on EPO patents with priority date in 1993–1997. In particular, they asked inventors the price at which the patent would be sold at the moment of grant. This methodology provides a continuous estimate of the economic share of value of a patent, however, it doesn't control for the patent premium.
Following the same line of research, the objective of this study is to assess and to compare

the discriminatory power of selected patent–based indicators. The value measure we propose is based on whether the patent refers to a QAI winning innovation or not. In contrast with previous contributions, the main advantage of our approach is related to the *timing*: while existing value measures rely on the expected potential of an invention, the QAI indicator depend on the (at least partially) realized potential of an innovation. Moreover, the Award winning event is a direct and exogenous way of classifying a patent as valuable, it therefore overcomes the issue of including the patent premium in the evaluation. Lastly, the benchmark and the control patents are selected within a close set of firms which are all based in the UK, so that the 'noise' affecting the patents' value estimation is reduced.
To be clear on this points, in the next section we describe the Queen's Award prize scheme and we explain in details the dataset construction steps.

III. Data description

A. The Queen's Award for Enterprise: background

The scheme of the Queen's Awards was initially announced on February the 4th 1965 in the House of Commons by the Prime Minister Harold Wilson, following the recommendations of a Committee chaired by HRH The Duke of Edinburgh. In a later speech in the House of Common Wilson specified that: 'The purpose of this new scheme is twofold: to reward and to stimulate. I hope that the Award will encourage industry in its efforts to achieve the improvements in exports and the technological advance on which our national future so much depends.'
To keep the scheme's parameters updated, the first draft released in 1965 announced that 'the working of the Award Scheme as presented in our Report should be reviewed after five years so that any modifications which practical experience of its operation had shown to be desirable could be introduced'. In practice, apart from the first presentation draft, they made only three other official reports: in 1970, 1975 and 1999. In each review, the Committee in charge collected some feedback from applicants to assess whether and to which extent to amend the scheme. Yet, over its lifetime the nature of the prize remained largely unaffected and, still today, it represents the most prestigious distinction for individuals and businesses in the UK (Groom 2015). In fact, every year major British newspapers such as the Financial Times and The Guardian publish the list of winners along with specific articles devoted to a selection of them.
At present, what is known as 'The Queen's Awards for Enterprise' recognizes achievements in three separate fields: export, innovation and sustainability (Report 1999). In general, there are no restrictions in terms of sectors nor predetermined pattern of regional allocation of the prize: all large, medium or small organizations which regularly operate as a 'business unit' in the UK, are eligible to apply. Through the years, another issue commonly discussed was whether the grant should be associated to tangible rewards, for example special tax reliefs, however these suggestions have never been endorsed as 'their inclusion would detract from the dignity of the Honour' (Report 1970).
For the purpose of this study, we only focus on the Queen's Award for Innovation; having explained the general framework, in the next part we provide some details on the past and current eligibility criteria to give a precise idea of the parameters used to classify an innovation as successful.

B. Eligibility and selection criteria for the QAI

At the very beginning, what was known as the Queen's Award for Technological Achievement was rewarding 'A significant advance, leading to increased efficiency, in the application of an advanced technology to a production or development process in British industry or the production for sale of goods which incorporate new and advanced technological qualities'. Already in the first report in 1970, the Scheme Review Committee underlined the applied connotation of the definition by removing the word 'advanced' before 'technology', and emphasizing that 'the timely application of established technology, as against advanced technology, may well be equally important and deserving of recognition, particularly in the less sophisticated sectors of industry'. In the same spirit, in 1999 the Review Committee in charge drafted the final and most substantial reformulation of the criteria for the Award. Above all, the name changed from 'Queen's Award for Technological Achievement' to 'Queen's Award for Innovation', with the aim to broaden the eligible set of subject matters to innovation in services and, in general, innovations which are not technology-driven. Moreover, the eligibility criteria were grouped under two main heading:

- Outstanding innovation, resulting in substantial improvement in business performance and commercial success, sustained over not less than two years, to levels which are outstanding for the goods or services concerned and for the size of the applicant's operations, and arising in the fields listed below. Or:
- Continuous innovation and development, resulting in substantial improvement in business performance and commercial success, sustained over not less than five years, to levels which are outstanding for the goods or services concerned and for the size of the applicant's operations, and arising in the fields listed below.

> Achievements under either criterion may be assessed in any of the following fields: the invention, design, production (in respect of goods), performance (in respect of services, including advice), marketing, distribution, after sale support, of goods or services (Report 2014).

Along with a detailed description of their achievement, applicants should demonstrate the commercial success of their innovation by providing the relevant financial figures covering two to five years before the submission. In particular, they should support their application with the patterns of growth of earnings imputable to the innovation, using as a reference an indicator such as profitability, market share, compliance with a target, and others. Finally, the applications are screened by a group of contracted technical assessors and then judged on the basis of merit by a Panel of Judges chaired by the Permanent Secretary of the Department for Business, Innovation and Skills (Report 2014).

To sum up, the innovations are awarded after few years of demonstrated commercial success on the market and, within the context of this analysis, the patents protecting them can be effectively used as a benchmark of economically valuable patents. In the next section we describe the steps we followed to build the dataset, the patent matching procedure and the final selection of the benchmark and control patents for the patent–based indicators comparison.

C. Database construction

1. *Data collection*

Every year the full list of the QAI winners is published with the name of the company, the location and a brief description of the innovation, for a total of 1420 prizes awarded to relevant innovations between 1966 and 2015. For example, a typical entry taken from 1993 looks as follow:

> Clinical Reagents Division of Amersham International Plc, Little Chalfont, Buckinghamshire Amerlite laboratory diagnostic system (jointly with the Wolfson Research Laboratories of the Department of Clinical Medicine of the University of Birmingham.)

Private, public and non for profit units are eligible for a QAI provided that the innovation contributes to industrial efficiency. In case of joint development, both entities are awarded and if there is a specific business division involved, it is mentioned. Since 2009 The Queen's Award Office provides a more detailed description, often specifying if the innovation is patented or not. For instance, from the report in 2011:

> Checkmate Lifting & Safety LLP, New Road, Sheerness, Kent ME12 1PZ, Website: www.checkmateuk.com, Employees: 41, Managing Director: Mr O Auston, Ultimate Parent: N/A, Contact for Press enquiries: Mr Oliver Auston, Tel: 01795 662590; 07770 395919 (Mobile), E-mail: oa@checkmateuk.com. An Innovation Award is made to Checkmate Lifting Safety LLP for the design, development and successful sales of it's Xcalibre Fall Arrest Blocks. The device is designed for use when working at height and arrests a worker should they fall and reduces the peak forces to the body and structure it is attached to. The company's patented technology including the SBM (Sealed Brake Module) and injection moulded one piece drum were both inspired by the automotive industry and use the very latest aerospace materials and automated manufacturing techniques. These unique features allow the device to function in more arduous environments and reduces both the weight and manufacturing costs. In a conservative market the company has forged successful sales at home and overseas.

Starting from this listing, we first matched the names of the winners to a unique code identifying them in AMADEUS – Bureau van Dijk (BvD), a database containing comprehensive information on 21 million European companies. AMADEUS covers all the companies having more than ten employees over a moving window of ten years, leaving out those which ran out of business at each update. Also, it keeps track of MA history and it is possible to find an enterprise even after a structural operation or a change of ownership. We could perform additional checks on the actual correspondence between our list and AMADEUS by looking at the addresses and at the main sector of operations of the companies. Eventually we retrieved 1129 BvD codes, which enable us to access all the information stored in AMADEUS, for each firm matched.

2. *Patent matching*

As a next step, we looked for possible matches between each innovation winning a QAI and one or more patents. For this purpose, we didn't impose any restrictions in terms of application authorities, however we limited the research to a certain time interval which varies depending on the rules governing the prize scheme in every period. Also, we looked only at applications filed in English and we kept the patent family as a matching unit to make the procedure less expensive. Starting from the BvD codes, we saved all the patents publications numbers assigned to a firm in the QAI dataset within AMADEUS. Then, we retrieved the patents' family codes, application titles and filing dates from PATSTAT, the EPO Worldwide Patent Statistical Database.
At this point, we applied time filters: to win an Award, applicants should provide from two to five years of sales figures as a prove of commercial success of their technology; for this reason, there should be a certain lag between the patent filing and the prize. To define the correct interval, we first approached the most recent Awards which were mentioned to be patented. We considered as relevant for commercial success those patents granted up to one year before the first year of figures to be submitted. For example, an innovation awarded in 1995 with three years of sales required, could be patented in 1991 the latest, any later filing is considered an improvement. We set the lower bound at nine years before the Award, assuming that any prior filing would be no longer innovative enough to receive a prize. From this pool, we singled out the relevant patents manually, assigning them a score depending on the quality of the matching. As for the companies which are not in AMADEUS, or do not have patents assigned in AMADEUS, we searched directly in PATSTAT, by matching the name of the innovator to the assegnee, and retrieving the variables mentioned above using the PERSON ID code in PATSTAT. The subsequent steps were the same as for the other group.
In this phase, we considered only prizes awarded from 1976 till 2015, mostly because the first precise reference we have in terms of years of sales figures to be submitted, comes from the 1975 Review Committee. With very few exceptions, all the innovations were patentable and on a total of 1234, we found that 32% of them were in fact protected by patents at the time of the Award. In particular, we singled out 1466 patent families referring to 401 innovations in the QAI dataset.

3. *Selecting the benchmark and the control group*

To assess the performance of the selected indicator in predicting the Award winning cases, we focus on the applications filed at the USPTO. Table I reports how many matched patent families have at least one application filed in a major Patent Office, and we can observe that within the Queen's Award sample, almost half of them has a filing in US. Since the winning companies are all based in the UK, we can assume that only the most valuable inventions get protection oversea. To control for confounding factors and to eventually have a full spectrum of patents reflecting also strategic actions, we decide to build the control group as the set of all USPTO applications filed by the QAI winning companies since 1976. Specifically, this sampling choice allows us to compare inventions generated in the same context as the awarded one, i.e. leveraging on the same set of resources and capabilities.

To obtain the final dataset for our analysis, we proceeded as follow:

Table I: Main filing authorities within the QAI dataset

Authority	No. of families	Share
GB	1,179	0.80
USPTO	720	0.49
EPO	558	0.38
JPO	465	0.32
USPTO, EPO, JPO	267	0.18

1. We matched the list of QAI related patent applications filed at the USPTO with the OECD Quality database, which covers around 7 million USPTO patents filed since 1976, along with their score on the most common patent–based indicators of patent value (Squicciarini et al. 2013).
2. Using the BvD codes as a refence, we retrieved all the patent applications filed at the USPTO by the companies having at least one QAI related patent filed in the same office; in other words, we included in the control group all the inventions with the same assegnee. Also in this case, we looked for matching observations in the OECD quality database.
3. To get a more balance dataset composition, reaching a share of QAI patents of at least 10%, we dropped control observations filed more than three years before (after) the oldest (most recent) QAI patent within each company portfolio.

The final dataset consists of 5,706 patent applications filed at the USPTO by 180 UK based companies. Within the set we have 596 QAI related patents, referring to 243 innovations. The patent features we include in our analysis have been retrieved from the OECD quality database and, within our sample there are no missing values. In the next section we present the descriptive statistics and the results of the empirical analysis.

IV. Horse race: comparing indicators

A. Descriptive evidence

We start describing the four variables of interest, both over the whole dataset and separately between the awarded and the control group. Table II provides the summary statics of the value indicators. As expected, we observe that the mean tends to be lower than the median, signalling a skewed distribution for forward citations, family size and claims. The only exception is renewals, which is obviously not affected by outliers as it only goes over the 20 years of statutory life of the patents.

By looking at the two subsets separately, we observe that the well known property of skewness emerges also graphically. Figure 1 depicts the kernel density for the four variables, grouping the results for the awarded patents (solid line) and for the control patents (dot line). A visual inspection suggests that awarded patents scores on average higher on forward citations, family size and claims, with lower peaks on the zero/low values and a slower fall

Table II: Summary statistics of the single indicators

Indicators	n	Mean	S.D.	Min	.25	Mdn	.75	Max
Renewal	5706	8.65	4.53	0.00	5.00	9.00	13.00	20.00
Forward citations	5706	3.64	5.22	0.00	1.00	2.00	5.00	71.00
Family size	5706	7.51	5.91	1.00	4.00	6.00	9.00	50.00
Claims	5706	13.58	12.63	1.00	7.00	10.00	17.00	244.00

on the right end of the distribution. Analogously, in the case of renewals, the awarded patents do not display peaks over the first ten years of the patent life but they show a higher concentration between the 10th and the 15th year. Accordingly, the Fligner–Policello test rejects the null hypothesis of equal population for the awarded and the control patents on all four indicators. Still, even though the differences between the two subsets are persistent across the graphs, they are not substantial as the curves are very close to each other.

We conclude this descriptive assessment with the correlation matrix. Table III lists the Spearman's rank correlation coefficients with significance levels. In line with other studies we find that the correlation among the variables is positive, steady and basically low (e.g. Lanjouw and Schankerman (2004)). The most correlated variable is the number of forward citations, which reports relatively higher coefficients on claims (0.18), renewals (0.14) and family size (0.12). Despite the small magnitude, the overall significance of the correlation coefficients allows for the specification of a common factor based on the unobservable value related characteristics that influence all four indicators.

Table III: Spearman correlation between the single indicators

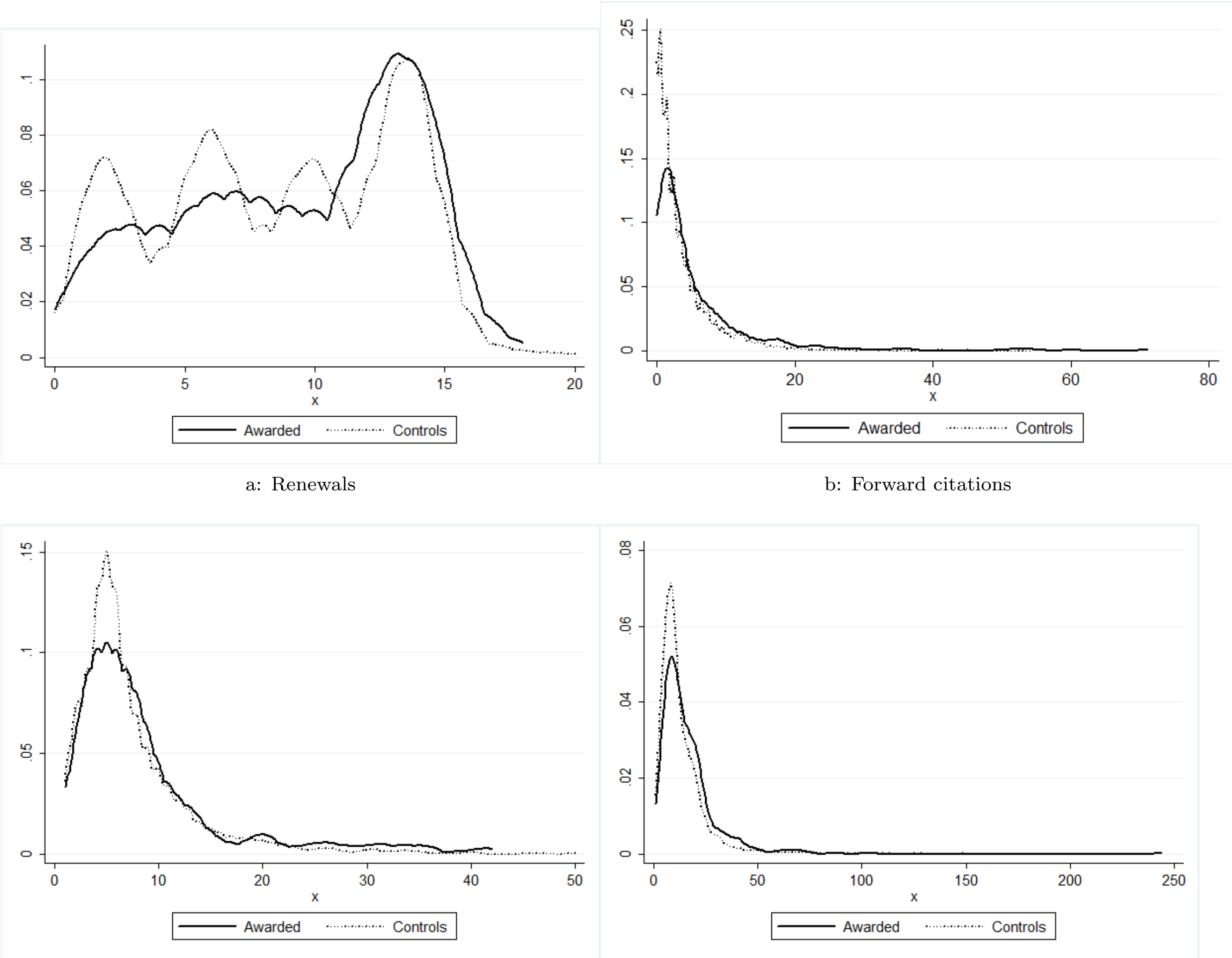
	Renewals	Fwd cit.	Fam. size	Claims
Renewals	1.0000	–	–	–
Forward Citations	0.1408***	1.0000	–	–
Family size	0.0473***	0.1188***	1.0000	–
Claims	0.0446***	0.1815***	0.0858***	1.0000

1. *** (**, *) indicate a significance level of 1% (5%, 10%).

B. Estimating the composite indicator

We now proceed with the estimation of the composite index developed by (Lanjouw and Schankerman 2004). As we mentioned before, we introduce a modified version by (1) including the number of renewals instead of the number of backward citations, and (2) by using a non linear latent variable model which better resembles the variables' distribution. In particular, we estimate the indicator through a confirmatory factor analysis using a Generalized Structural Equation Model, specified as follows:

$$E[Y_k|V^*] = G(\alpha_k + V^*\beta_k) \tag{1}$$



a: Renewals

b: Forward citations

c: Family size

d: Claims

Figure 1: Kernel density of the single patent–based indicators

where $Y_k(K = 1, ..., 4)$ is the vector of values for the K single patent–based indicators, $G(.)$ is the link function which varies depending on the distributional features of the variables, α is the constant and V^* is the vector of latent patent private value with factor loading β_k. Note that V^* is a latent exogenous variable and, to allow for identification, we should anchor it to an observed variable. For this reason we constrain the path coefficient of family size to 1. Once estimated the factor loadings, we predict the latent variable 'Value', which is the adjusted version of the composite patent–based indicator developed by Lanjouw and Schankerman (2004). Table IV shows the correlation coefficients between the estimated variable and the single indicators. We notice that the number of renewals is basically not captured by the composite measure, which is mostly correlated to forward citations and claims.

Figure 2 shows the kernel density curves of the composite measure for the awarded (solid line) and the control (dot line) patents. Compared to the graph of the single indicators, in this case we observe a greater difference between the two distributions. In other words,

Table IV: Spearman correlation between 'Value' and the single indicators

	Value	Renewals	Fwd cit.	Fam. size	Claims
Value	1.0000	0.0898***	0.6223***	0.4864***	0.6985***

1. *** (**, *) indicate a significance level of 1% (5%, 10%).

assuming that the common variation of the four single features is related to patent value, the specification of a latent factor capturing this variation better discriminates the most from the least valuable patents.

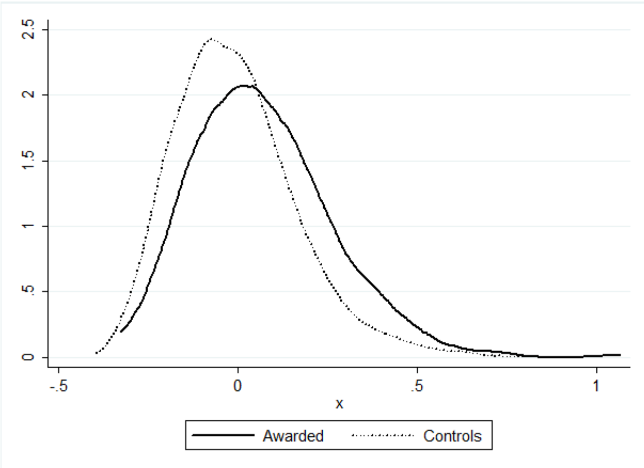


Figure 2: Kernel density of the composite indicator 'Value'

C. Regression results

Now we empirically assess the discriminatory power of the indicators described before. Given the binary nature of our dependent variable QAI, we construct five separate logit models, each of them having a patent–based indicator as single predictor. Note that, to control for exposure, we limit the forward citations count to the five years after the publication date. In the case of renewals, this is less of a problem because we control for the application filing year. In this respect, we have 38 year dummies covering the period from 1976 to 2013 as well as a set of 35 dummies for the technical fields based on Schmoch (2008).

To evaluate the relative performance of each indicator we look at the odds ratios and at the percentage of the cases correctly classified by the logit models. Since the awarded patents are only the 10% of the total, we impose a cutoff of 0.1 as a rule, i.e. the patents are classified as valuable if the predicted $Pr(QAI) >= 0.1$. Note that the purpose of this adjustment is to rescale the outcome in order to facilitate the comparison as, in practice, it does not affect the relative performance of the indicators.
Table VI shows the regression results: M(0) refers to the baseline model with dummies alone, M(1) to M(4) use the four single indicators as independent variables and M(5) reports the

Table V: Variables employed in the analysis

Variable	Description
QAI	Binary variable equal to 1 if the patent refers to an awarded innovation, 0 otherwise
Renewals	# of years in which a granted patent has been kept alive
Fwd cit. 5y	# of citations that a patent receives over a period of five years after the publication date
Fam. size	# of patent offices at which an invention has been protected
Claims	# of claims per patent
Value	Continuous variable obtained as a latent factor from Renewals, Fwd cit. 5y, Fam. size and claims
Year dummies	38 dummies for the patents' filing year
Tech. dummies	35 dummies for the patents' technical field

coefficients for the estimated factor 'Value'. We immediately notice that all the predictors are positive and strongly significant. Also, looking at the odds ratios, there is a large difference between the magnitude of the single measures and the one of the composite index. More precisely, the odds ratio for renewals, forward citations, family size and number of claims ranges from 8.4% to 1.2%, instead the result for 'Value' is over 900%, meaning that a unit increase in 'Value' increases the odds of being a QAI patent nine folds. While the odds ratios give an idea of the effect that a change in the predictor can have on the dependent variable, they cannot be a meaningful ground of comparison as the relevance of a unitary shift varies depending on the distributional features of the indicators. As a discriminatory criterion we look at the percentage of correctly classified observations, which consistently highlights the composite measure as the most effective, with 63% correctly classified cases.
So far, the message emerging from our results is the following: in general, patent–based indicator are positive and persistent predictors of patent value, in particular, the effects are greater if we construct a factor variable leveraging on the correlation between each indicator. Yet, the overall magnitude of their impact is extremely small. Analysing the Pat–Val survey, Gambardella et al. (2008) found that the share of variance explained by a model having four measures as independent variables (forward and backward citations, number of claims and designated European countries) is only 2.7% higher than a baseline model with only dummies. Our outcome is not different: looking at the pseudo R^2, the difference in variance explained between M(0) and M(5) is only 1.7%, which is even lower because we test one indicator in each model.

D. The relationship between patent value, renewals, forward citations, family size and claims

Given the results we just described, in this section we attempt to investigate further the relationship between patent characteristics and patent value. Starting from the inference that a latent construct is most suitable in capturing the value related share of the variables, we perform an exploratory factor analysis to see how the single indicators relate to each other if we do not impose a specific path.
Using a principal component factor method, we obtained two factors with eigenvalues greater than 1. Table VII reports the loadings of the indicators on the rotated factors and the

Table VI: Comparing indicators performance

	M(0) OR(se)	M(1) OR(se)	M(2) OR(se)	M(3) OR(se)	M(4) OR(se)	M(5) OR(se)	M(6) OR(se)
Renewal	–	1.085*** (0.020)	–	–	–	–	–
Fwd cit	–	–	1.046*** (0.007)	–	–	–	–
Fam. size	–	–	–	1.058*** (0.008)	–	–	–
Claims	–	–	–	–	1.012*** (0.003)	–	–
Value	–	–	–	–	–	9.023*** (2.414)	–
F1	–	–	–	–	–	–	1.389*** (0.059)
F2	–	–	–	–	–	–	1.279*** (0.087)
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tech. dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cons	0.093***	0.079***	0.075***	0.059***	0.074***	0.833***	0.114***
No. of Obs.	5706	5706	5706	5706	5706	5706	5706
Pseudo R2	0.078	0.083	0.088	0.091	0.081	0.095	0.099
Correctly classified	60.55%	61,58%	61.46%	62.58%	60.94%	62.99%	63,69%

1. *** (**, *) indicate a significance level of 1% (5%, 10%).
2. The cutoff for classification is set at 0.1

uniqueness levels. First of all, we observe that the uniqueness level is quite high, confirming the results of Van Zeebroeck (2011) and Bessen (2008) among the others, who found that the relationship between these patent features and their economic value is affected by a lot of noise and there seems to be many unexplained factors playing a role in it. Looking at the factor score, F1 resembles the composite index we introduced before with high and positive loading on forward citations, claims and family size. Instead, F2 is dominated by renewals, which appears to be quite orthogonal to the other measures. Also, on F2 we have a negative loading on claims of 35% and a positive score of 27% for forward citations, which is the indicator with the highest commonality.

In trying to give an interpretation to the two factors, it's important to keep in mind the composition of our dataset. As we described before, by selecting both the awarded and the control patents from the same group of companies, we are able to at least partially reconstruct winning firms' patent portfolio. Presumably, this sampling choice has two effects: on the one hand, it grants a certain degree of homogeneity of the patentees, on the other hand, it allows us to cover a representative spectrum of the selected firms' patenting activity. As a consequence, we would expect our analysis to include patents protecting a technically and/or economically valuable invention, and patents filed for strategic reasons. Note that,

Table VII: Rotated factor loadings matrix

Variable	F1	F2	Uniqueness
Renewal	0.0365	0.9227	0.1472
Forward citations	0.7262	0.2717	0.3987
Family size	0.5819	0.0520	0.6587
Claims	0.6520	–0.3483	0.4535

by labelling a patent as 'strategic', we do not mean to attach any connotation to the actual value of the invention. Rather, we simply argue that the main reason why a patent is filed is not necessarily related to the patent effectiveness in protecting the invention.
Having said that, a tentative explanation is that, given the high scores on forward citations, number of claims and family size, F1 captures both the technical and the economic value of a patent. This interpretation also validate Lanjouw and Schankerman (2004) composite indicator as a comprehensive measure of patent's private value. Instead, F2 could be associated to both the economic and the strategic value of a patent. The strong orthogonality of renewals may indicate either patents which, regardless of the technical merits of the inventions, effectively bring positive returns to the firms, or patents that, regardless of the technical merits of the invention, are kept alive for strategic reasons. In fact, the number of renewals is a value indicator which is affected by the 'patent premium' bias (Arora et al. 2008).
As a final step of this exploratory section, we regress our binary measure of value against F1 and F2, the results are reported in Table VI, model M(6). Both factors are positive and significant, with odds ratio of approximately 40% and 30%. Following our previous reasoning and acknowledging that the economic component of patent value is captured by both factors, patents scoring high on F1 and F2 are more likely to be QAI patents. Also, we observe that a latent construct with two factors performs better the previous indicators both in terms of correctly classified cases and variance explained. However, as for the comparative analysis of the other measures, the improvement is only marginal.

V. Conclusion

In this paper we introduce a unique dataset which gives us the possibility to estimate patent value from a different perspective compared to previous studies. By retrieving a sample of valuable patents from successfully commercialized innovations, we are able to directly tie intellectual property rights to realized economic value. Moreover, working with the portfolio of a limited set of companies, we can investigate the behaviour of patent features both on an effective and on a strategic ground. Still, a major shortcoming of our setting is the binary nature of the measure we adopt, which may over simplify the issue of patent value estimation.
In line with earlier contributions, our results show that the relationship between patent–based indicators and actual patent value is systematic and non random. Moreover, we find that composite indicators better discriminate the awarded from the control group in our dataset, suggesting that the selected patent features share an unobservable private value

component. To proceed further in this direction, we perform an exploratory factor analysis in order to have a clearer view on the information content of each indicator and on their interaction. This exercise highlights that, with a disaggregation level leading to two factors instead of one, patent features are likely to define different value dimensions. However, in all cases patent–based indicators report a relevant share of unique variance which confines their explanatory power to a little portion of what makes a patent valuable. To conclude, our analysis emphasizes two main points: on the one hand, the untapped potential of patent–based indicators as a measure of patent value is still remarkable and additional research is needed to fully exploit it. Specifically, the combination and recombination of selected patent features can disclose interesting insights on the information content of single indicators. On the other hand, the magnitude of their impact on patent value is very small and substantial research effort is required to give a decisive contribution in this sense. In general, given the recent trends in patent filing, the understanding of patent value is an increasingly important issue that shall be further investigated.

Ahuja, G. and C. Morris Lampert (2001). Entrepreneurship in the large corporation: A longitudinal study of how established firms create breakthrough inventions. *Strategic Management Journal 22*(6-7), 521–543.

Arora, A., M. Ceccagnoli, and W. M. Cohen (2008). R&d and the patent premium. *International journal of industrial organization 26*(5), 1153–1179.

Arts, S., F. P. Appio, and B. Van Looy (2013). Inventions shaping technological trajectories: do existing patent indicators provide a comprehensive picture? *Scientometrics 97*(2), 397–419.

Bessen, J. (2008). The value of us patents by owner and patent characteristics. *Research Policy 37*(5), 932–945.

Bessen, J. (2009). Estimates of patent rents from firm market value. *Research Policy 38*(10), 1604–1616.

Carpenter, M. P., F. Narin, and P. Woolf (1981). Citation rates to technologically important patents. *World Patent Information 3*(4), 160–163.

Gambardella, A., D. Harhoff, and B. Verspagen (2008). The value of european patents. *European Management Review 5*(2), 69–84.

Griliches, Z. (1981). Market value, r&d, and patents. *Economics letters 7*(2), 183–187.

Groom, B. (2015). Half a century of the Queen ' s Awards. *Financial Times*.

Hall, B. H. (1993). The stock market's valuation of r&d investment during the 1980's. *The American Economic Review 83*(2), 259–264.

Hall, B. H., A. B. Jaffe, and M. Trajtenberg (2001). The nber patent citation data file: Lessons, insights and methodological tools. Technical report, National Bureau of Economic Research.

Hall, B. H., G. Thoma, and S. Torrisi (2007). The market value of patents and r&d: Evidence from european firms. In *Academy of Management Proceedings*, Volume 2007, pp. 1–6. Academy of Management.

Hall, B. H. and R. H. Ziedonis (2001). The patent paradox revisited: an empirical study of patenting in the us semiconductor industry, 1979-1995. *RAND Journal of Economics*, 101–128.

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.

Hsieh, C.-H. (2013). Patent value assessment and commercialization strategy. *Technological Forecasting and Social Change 80*(2), 307–319.

Jaffe, A. B., M. Trajtenberg, and M. S. Fogarty (2000). The meaning of patent citations: Report on the nber/case-western reserve survey of patentees. Technical report, National bureau of economic research.

Lanjouw, J. O., A. Pakes, and J. Putnam (1998). How to count patents and value intellectual property: The uses of patent renewal and application data. *The Journal of Industrial Economics 46*(4), 405–432.

Lanjouw, J. O. and M. Schankerman (2001). Characteristics of patent litigation: a window on competition. *RAND journal of economics*, 129–151.

Lanjouw, J. O. and M. Schankerman (2004). Patent quality and research productivity: Measuring innovation with multiple indicators. *The Economic Journal 114*(495), 441–465.

OECD (2009). OECD Patent Statistics Manual. Technical report, OECD.

Pakes, A. (1985). On patents, r&d, and the stock market rate of return. *Journal of Political Economy 93*(2), 390–409.

Pakes, A. (1986). Patents as options: some estimates of the value of holding european patent stock. *Econometrica 54*(4), 755–784.

Putnam, J. (1996). *The Value of International Patent Rights*. Ph. D. thesis, Yale University.

Report (1970). The Queen's Award to Industry. Technical report, 1970 Review Committee.

Report (1999). The Queen's Award for Enterprise. Technical report, 1999 Review Committee.

Report (2014). The Queen's Award for Enterprise. Technical report, 2014 Review Committee.

Schankerman, M. and A. Pakes (1986). Estimates of the value of patent rights in european countries during the post-1950 period. *The economic journal 96*(384), 1052–1076.

Schmoch, U. (2008). Concept of a technology classification for country comparisons. *Final report to the world intellectual property organisation (wipo), WIPO*.

Squicciarini, M., H. Dernis, C. Criscuolo, et al. (2013). Measuring patent quality: Indicators of technological and economic value. Technical report, OECD Publishing.

Tong, X. and J. D. Frame (1994). Measuring national technological performance with patent claims data. *Research Policy 23*(2), 133–141.

Trajtenberg, M. (1990). A penny for your quotes: patent citations and the value of innovations. *The Rand Journal of Economics*, 172–187.

Trajtenberg, M., R. Henderson, and A. Jaffe (1997). University versus corporate patents: A window on the basicness of invention. *Economics of Innovation and new technology 5*(1), 19–50.

Van Zeebroeck, N. (2011). The puzzle of patent value indicators. *Economics of Innovation and New Technology 20*(1), 33–62.

van Zeebroeck, N. and B. van Pottelsberghe de la Potterie (2011). The vulnerability of patent value determinants. *Economics of innovation and new technology 20*(3), 283–308.