KATHOLIEKE UNIVERSITEIT
LEUVEN

Faculty of Business and Economics

"Wacky" patents meet economic indicators

Dirk Czarnitzki, Katrin Hussinger and Cédric Schneider

DEPARTMENT OF MANAGERIAL ECONOMICS, STRATEGY AND INNOVATION (MSI)

OR 1109

Non-technical Summary

The empirical literature on the economics of patents employs a large array of patent-based measures to proxy economic characteristics of patents such as the “science-base”, “importance” or “value” of patented inventions. Although scholars have attempted to assess the validity of these proxies by relating them to e.g. inventor surveys, the appropriateness of these indicators is currently heavily debated.

We contribute to this discussion by relating a group of obviously “wacky” patents to two control groups. If the patent-based indicators are appropriate, they should unambiguously identify the “wacky” patents.

We present descriptive statistics and run probit regressions to evaluate the performance of commonly used patent-based measures. Our findings show that forward citations are good predictors of importance. However, the “wacky” patents have higher originality, generality and average citation lags than the controls, which suggests that these indicators should be interpreted carefully. At best, scholars should provide an external validation for their interpretation of patent-based measures.

“Wacky” patents meet economic indicators

A revised version of this paper is forthcoming in *Economic Letters*.

Dirk Czarnitzki [a,b,c], Katrin Hussinger [d,a,c] and Cédric Schneider [e*]

[a] *K.U. Leuven, Dept. of Managerial Economics, Strategy and Innovation (Belgium)*
[b] *Center for R&D Monitoring (ECOOM) at K.U. Leuven (Belgium)*
[c] *Centre for European Economic Research (ZEW), Mannheim (Germany)*
[d] *Maastricht University, Dept. of Organization and Strategy (The Netherland)*
[e] *Copenhagen Business School, Dept. of Economics (Denmark)*

May 2011

Abstract

This study investigates whether standard patent measures for the importance and basicness of patents are able to distinguish between “wacky” patents and a control group of randomly drawn patents. Our findings show that forward citations are good predictors of importance. However, the “wacky” patents have higher originality, generality and average citation lags than the controls, which suggests that these indicators should be interpreted carefully.

Keywords: patents, patent indicators, quality measurement

JEL-Classification: O34

[*] Corresponding author. Address: Copenhagen Business School, Dept. of Economics, Porcelaenshaven 16a, 2000 Frederiksberg, Denmark. Email: csc.eco@cbs.dk. Tel: +45 38152316. Fax: +45 38152576

1 Introduction

The empirical literature on the economics of patents employs a large array of patent-based measures to proxy economic characteristics such as "science-base", "importance" or "value" of patented inventions. Although scholars have attempted to assess the validity of these proxies by relating them to e.g. inventor surveys, the appropriateness of these indicators is currently heavily debated (e.g. Gambardella et al., 2008, Gittelman, 2008, Alcacer and Gittelman, 2006, Harhoff et al., 2003). We contribute to this discussion by relating a group of "wacky" patents to two control groups. If the patent-based indicators are appropriate, they should unambiguously identify the "wacky" patents. We present descriptive statistics and run probit regressions to evaluate the performance of commonly used patent-based measures.

2 Data and Variables

2.1 Data and sample selection

The analysis is based on patents downloaded from www.patentoftheweek.com. This website provides a list of "wacky" patents collected by an employee of the World Intellectual Property Organization (WIPO). Patents classified as "wacky" were selected by the site author for their futile nature, as they do not involve a high inventive step or only marginally satisfy the "non-obviousness" criterion. One example is patent US4866863 for a religious device named "empty tomb". Another example is patent US5078642 for a toy bar soap slide that can be attached to a bathtub for kids' entertainment.[1]

In total, 188 U.S. granted patents are listed on the webpage that have been applied for between 1974 and 2002. We construct two control groups. The first control group consists of five randomly drawn patents in the same application year and three-digit technology class for each "wacky" patent. The second control group contains five randomly drawn patents in the same six-digit technology class for each "wacky" patent. The second control group is constructed because the use of the three-digit technology classes may generate spurious matches (Thompson and Fox-Kean, 2005).

[1] We invite the reader to look at the other patents on the website. For reasons of brevity we do not provide more examples of "wacky" patents.

As there are not always enough patents filed in the same six-digit technology class in the same application year, control patents were drawn from a three-year window around the application year of the focal patent. For five “wacky” patents, control patents had to be taken from an even broarder time window. In these cases, the patents that were closest to the application year of the focal patent were chosen. Finally, we dropped one “wacky” patent because there were not enough control patents available in the same six-digit technology class over all years. All patent related variables were drawn from the 2006 edition of the NBER Patent and Citation Database (Hall et al., 2001).

2.2 Variables

We use the most commonly applied patent measures in the empirical innovation literature to test whether they are able to distinguish the “wacky” patents from the controls.

Forward citations are defined as the number of all U.S. citations received by a focal patent from subsequent patents. This measure is typically interpreted as the “importance”, the “quality” or the “significance” of a patented invention. Previous studies have shown that forward citations are highly correlated with the social value (Trajtenberg, 1990) and the private value of the patented invention (Harhoff et al., 1999, Hall et al., 2005). Furthermore, forward citations reflect the economic and technological “importance” as perceived by the inventors themselves (Jaffe et al., 2000) and knowledgeable peers in the technology domain (Albert et al., 1991).

Backward citations determine the legal boundaries of an invention by defining a related body of prior art. Empirical evidence shows a positive relationship between the number of backward citations and the patent value (Harhoff et al., 2003). The number of cited patent reflects the extent of patenting in a given technological area and therefore the profitability of the inventions in that domain.

The citation lag: Patents covering more basic or fundamental technologies are often argued to be cited later than applied patents because it takes longer for basic inventions to be understood and used by others (Sampat et al., 2003). We use the average citation lag to test for this.[2]

Non-patent references (NPR): While the meaning of NPRs is not unambiguous, there is some recognition of their use as an indicator of science-technology linkages (Callaert et al., 2004, Meyer, 2000). Therefore, patents citing NPRs may reflect inventions resulting from fundamental research and are thus further away from market applications.

Originality and Generality (Trajtenberg et al., 1997): Originality is defined as one minus the Hirschman-Herfindahl index of the cited 3-digit technology classes. Patents drawing from many different technology areas are presumably more original and more complex. Generality is defined as one minus the Hirschman-Herfindahl index of the citing 3-digit technology classes and is typically interpreted as a measure for the basicness of a technology: the more inventions in different fields a patented technology triggers the more basic it is. Both measures are adjusted for small number bias following Hall (2005).

Technological distance is defined as suggested by Trajtenberg et al. (1997). They assign the values 0/.33/.66 if the citing and cited patents are in the same 3-/2-/1-digit technology class, respectively. The value one is assigned if citing and cited patent are in different 1-digit technology classes. The measure of technological distance is the average distance of all citing patents to the focal patent. This can be interpreted as a measure of basicness: the further away the follow-up patents the more basic and fundamental is the focal patent.

We also control for 6 broad technology fields based on the classification of the “wacky” patents on the patent-of-the week website. We also realized that the “wacky” patents are frequently filed by individual applicants. Consequently, we generated a dummy variable indicating whether the applicant is an individual or not after drawing the control group.

---

[2] The empirical findings do not change if we use the time until a patent receives the first citation as an alternative measure.

3 Empirical Results

3.1 Descriptive statistics

Table 1 shows descriptive statistics for the sample of "wacky" patents and both control groups.

A first interesting observation is that "wacky" patents receive forward citations. Further, most "wacky" patents are owned by individual applicants rather than corporations.[3] The subsequent regression analysis includes a dummy controlling for this difference. Regarding the patent measures discussed in section 2.2, "wacky" patents score higher in terms of generality than the patents in both control groups. There is a significant difference in originality between "wacky" patents and the first control group. Section 3.2 tests whether this holds in a multivariate framework.

[3] Note that also in the control groups the share of patents owned by individual applicants is high.

Table 1: Descriptive Statistics

	"Wacky" patents		Control group I			Control group II		
	187		935			935		
	mean	std. dev.	mean	std. dev.	sig. mean diff. t-test	mean	std. dev.	sig. mean diff. t-test
forward citations	5.22	5.95	10.10	15.30	***	8.44	13.28	***
backward citations	4.45	4.61	5.67	7.32	**	7.37	1.69	***
generality[A]	0.45	0.37	0.36	0.34	***	0.38	0.36	***
originality[A]	0.40	0.39	0.30	0.35	***	0.36	0.36	
average forward citation lag	6.46	5.28	5.76	4.57	*	6.25	4.76	
NPRs	0.42	1.28	0.61	3.44		0.88	3.50	*
technological distance	0.64	0.26	0.67	0.21		0.64	0.25	
no forward citations	0.12	0.33	0.08	0.27		0.11	0.32	
individual applicant	0.84	0.36	0.49	0.50	***	0.49	0.50	***
share of forward self-citations	0.02	0.09	0.06	0.16	***	0.03	0.14	

[A] Generality and originality are adjusted for small numbers bias. It is set to zero for patents that receive less than two forward citations or have less than two backward citations respectively.

3.2 Probit regressions

Table 2 features probit models where the dependent variable takes the value one if the patent is "wacky". We show different estimation results: models I and II show the basic regressions for the two different control groups. Model III and IV show the same specifications, but self-citations are excluded from the forward citation based measures. The share of forward self-citations is used as an additional regressor in models III and IV.[4] Table 3 shows the marginal effects.

The regression results support the standard interpretation of forward citations. "Wacky" patents receive fewer citations indicating that they are less "important" than the controls. Note that the average probability of being a wacky patents amounts to 1/6 (= 16.7%) in our sample. An additional forward citation decreases the likelihood of a patent being "wacky" by 0.4-0.7 percentage points, keeping all other variables at their mean. Increasing the number of forward citations by one standard deviation (i.e. 14 citations), decreases the likelihood of a "wacky" patent by 5.6-9.8 percentage points.

However, "wacky" patents score higher on generality, originality and receive their citations later. If the generality of a patent increases by 0.10 the likelihood to be a "wacky" patent increases by 0.8-1.6 percentage points at the means of all other variables. A decrease in generality by one standard deviation (0.35) increases the likelihood of being "wacky" by 28-56 percentage points. The effects for originality have a similar magnitude. The usual interpretation to these measures suggests that "wacky" patents are more basic, fundamental and complex.

The standard interpretation of the citation lag might be misleading as well. While this measure is usually interpreted as the degree of basicness, our results for the comparison of "wacky" patents and the first control group suggest that longer citation lags might simply identify older and/or slower-moving technologies (Hall and Trajtenberg, 2006). The results for control patents drawn from the same six-digit

---

[4] All regressions include two dummy variables indicating if a patent receives less than two forward or backward citations respectively in order to control for the fact that the bias-adjusted measures for originality and generality are not defined for patents with less than two backward/forward citations.

technology class do not show any significant differences between “wacky” patents and controls with regard to the citation lag.

The originality and generality measures rely heavily on the USPTO’s classification system and treat each patent class as roughly comparable in size and importance (Hall and Trajtenberg, 2006). This is unlikely to hold because some subclasses are more important than others and some subclasses refer to closely related technologies whereas others refer to more distant technologies. Control group I drawn at the three-digit technology class level could be subject to such biases. Originality and generality for the “wacky” patents might be overestimated (relative to control group 1) as the citing and cited patents might come from closely related and/or less important technology classes.

Control group II that contains control patents drawn from the same six-digit technology class is not subject to such biases. Still, differences between “wacky” patents and controls exist with regard to generality and originality. A likely explanation is that “wacky” patents combine distant technologies that should not be joined because the combination is trivial or useless and that “wacky” patents receive citations by other patents that propose combinations of technologies that have been rarely combined before.[5] An example for a “wacky” patent that combines existing distant technologies is patent US6385796 a self-flushing urinal with an integrated gaming and reward system. This patent receives three citations by patents using similar technology combinations. Similarly, the patent US4866863 for a religious shrine receives citations by patents for other religious devices.

An alternative explanation for the difference between “wacky” patents and controls with regard to generality and originality could be that the “wacky” patent applications are based on a sloppy search for prior art, and that “wacky” patents receive citations by patents filed with little efforts in prior art search.

[5] Moir (2008) argues, for instance, that it is less likely that the patent office can reject a patent application that combines old ideas.

Table 2: Probit regression for being a "wacky" patent

	I	II	III	IV
			w/t self citations[A]	w/t self citations[A]
	control group I	control group II	control group I	control group II
	coeff. (s.e.)	coeff. (s.e.)	coeff. (s.e.)	coeff. (s.e.)
forward citations	-0.03***	-0.02***	-0.04***	-0.02***
	(0.01)	(0.01)	(0.01)	(0.01)
backward citations	-0.01	-0.02*	-0.01	-0.02*
	(0.01)	(0.01)	(0.01)	(0.01)
generality	0.71***	0.43**	0.85***	0.39**
	(0.18)	(0.17)	(0.24)	(0.17)
originality	0.67***	0.39**	0.69***	0.38**
	(0.17)	(0.17)	(0.17)	(0.17)
average fwd. cit. lag	0.08***	0.02	0.07***	0.02
	(0.02)	(0.02)	(0.02)	(0.02)
techn. distance	-0.28	-0.07	-0.19	0.03
	(0.29)	(0.27)	(0.28)	(0.27)
NPRs	0.01	-0.01	0.01	-0.02
	(0.02)	(0.02)	(0.02)	(0.02)
individual applicant	0.87***	0.86***	0.86***	0.89***
	(0.12)	(0.11)	(0.12)	(0.11)
share of fwd self-cit			-0.79*	0.24
			(0.47)	(0.42)
less than 2 bwd cit	0.35**	0.30*	0.36**	0.29*
	(0.17)	(0.17)	(0.18)	(0.17)
less than 2 fwd cit	0.60***	0.12	0.54***	0.16
	(0.19)	(0.18)	(0.18)	(0.17)
constant	-2.44***	-1.96***	-2.36***	-2.05***
	(0.50)	(0.48)	(0.49)	(0.48)
#	1,122	1,122	1,122	1,122
Log-likelihood	-417.43	-417.81	-418.78	-441.26

6 application period dummies (each covering 4 years) and 5 technology field dummies are included in all regressions.

[A] Self-citations are excluded for all measures based on fwd citations. Those are fwd citations, generality, average forward citation lag, technological distance, and time to first citation.

*,**,*** indicates significance at the 10%, 5%, 1% level.

Table 3: Marginal effects

	I	II	III	IV
			w/t self citations[A]	w/t self citations[A]
	control group I	control group II	control group I	control group II
	coeff. (s.e.)	coeff. (s.e.)	coeff. (s.e.)	coeff. (s.e.)
forward citations	-0.006***	-0.004***	-0.007***	-0.005***
	(0.001)	(0.002)	(0.002)	(0.01)
backward citations	-0.002	-0.004*	-0.002	-0.004**
	(0.002)	(0.002)	(0.002)	(0.002)
generality	0.136***	0.089**	0.161***	0.080**
	(0.034)	(0.036)	(0.046)	(0.035)
originality	0.129***	0.080**	0.131***	0.079**
	(0.032)	(0.034)	(0.032)	(0.035)
average fwd. cit. lag	0.015***	0.005	0.014***	0.005
	(0.003)	(0.004)	(0.004)	(0.004)
techn. distance	-0.054	-0.015	-0.035	0.007
	(0055)	(0.056)	(0.002)	(0.056)
NPRs	0.002	-0.003	0.001	-0.003
	(0.003)	(0.004)	(0.004)	(0.004)
individual applicant	0.160***	0.173***	0.157***	0.177***
	(0.020)	(0.021)	(0.020)	(0.021)
share of fwd self-cit			-0.150	0.050
			(0.088)	(0.086)
less than 2 bwd cit	0.074*	0.067	0.076*	0.065
	(0.041)	(0.041)	(0.041)	(0.041)
less than 2 fwd cit	0.143***	0.026	0.124**	0.036
	(0.054)	(0.039)	(0.049)	(0.040)
#	1,122	1,122	1,122	1,122

6 application period dummies (each covering 4 years) and 5 technology field dummies are included in all regressions.

[A] Self citations are excluded for all measures based on fwd citations. Those are fwd citations, generality, average forward citation lag, technological distance, and time to first citation.

*,**,*** indicates significance at the 10%, 5%, 1% level.

4 Conclusion

This paper tests whether the standard indicators for the importance and basicness of patents are able to distinguish between "wacky" patents and a control group. Our results show that if we would interpret the measures originality, generality and average citation lag as is common in the empirical literature we would have to conclude that our "wacky" patents are more basic, fundamental and complex. This shows that patent indicators should be interpreted with care. Alternative interpretations than those provided by prior research should be taken into account. At best, scholars should provide an external validation for their interpretation of patent-based measures.

Unfortunately, our results are not based on an exhaustive list of wacky patents so that it is not possible to make conclusion about the frequency of "wacky" patents, their distribution over technology classes or to identify the technology subclasses that contain most "wacky" patents.

Acknowledgements

We are grateful to an anonymous referee for helpful comments, and to Tom Magerman for providing the data on non-patent references. Czarnitzki gratefully acknowledges financial support from K.U.Leuven's basic research fund (grant: OT/07/012)

References

Albert, M.B., Avery, D., Narin, F., and McAllister, P. (1991). Direct Validation of Citation Counts as Indicators of Industrially Important Patents. *Research Policy* 20: 251-259.

Alcacer, J. and M. Gittelman (2006). Patent Citations as a Measure for Knowledge Flows: The Influence of Examiner Citations. *The Review of Economics and Statistics* 88(4): 774-779.

Callaert, J., Van Looy, B., Verbeek, A., Debackere, K. Thijs, B. (2004). Traces of Prior Art: An Analysis of Non-patent References Found in Patent Documents. *Scientometrics* 69(18): 3-20.

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

Gittelman, M. (2008). The Value of European Patents. *European Management Review* 5: 85-89.

Hall B.H., Jaffe, A.B., Trajtenberg, M. (2001). *The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools*. NBER Working Paper 8498.

Hall B.H. (2005). A Note on the Bias in Herfindahl-type Measures Based on Count Data. *Revue d'Economie Industrielle* 110: 149-156.

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

Hall, B.H. and Trajtenberg, M. (2006). Uncovering GPTs using Patent data. In: Antonelli, C., Foray, D., Hall, B.H. and Steinmuller, W.E. (eds). "New Frontiers in the Economics of Innovation and New Technology – Essays in honor of Paul A. David" Edward Elgar Publishing.

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

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

Jaffe, A., Fogarty, M.S., Trajtenberg, M. (2000). Knowledge Spillovers and Patent Citations: Evidence from a Survey of Inventors. *American Economic Review* 90: 215-218.

Meyer, M. (2000). Does Science Push Technology? Patents Citing Scientific Literature. *Research Policy* 29: 409–434.

Moir, H. (2008). A Question of Balance. Submission to the ACIP Review of Patentable Subject Matter. Australian National University. Canberra, Australia.

Sampat, B., Mowery, D., Ziedonis, A. (2003). Changes in University Patent Quality after the Bayh-Dole Act: A Re-examination. *International Journal of Industrial Organization* 21(3): 1371-1390.

Thompson, P. and Fox-Kean, M. (2005). Patent Citations and the Geography of Knowledge Spillovers: A Reassessment. *American Economic Review* 95(1): 450-460.

Trajtenberg, M. (1990). A Penny for your Quotes: Patent Citations and the Value of Innovations. *The RAND Journal of* Economics 21(1): 172-287.

Trajtenberg, M., Henderson, R., Jaffe, A.B. (1997). University versus Corporate Patents: A Window on the Basicness of Invention. *Economics of Innovation and New Technologies* 5(19), 19-50.