NBER WORKING PAPER SERIES

PATENT VALUE AND CITATIONS:
CREATIVE DESTRUCTION OR STRATEGIC DISRUPTION?

David S. Abrams
Ufuk Akcigit
Jillian Grennan

Working Paper 19647
http://www.nber.org/papers/w19647

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2013, Revised December 2018

We thank seminar and conference participants at the NBER Summer Institute IO and Intellectual Property Policy and Innovation group meetings, University of Pennsylvania Law School, Wharton, IP Scholars Conference (New York), American Law and Economics Association Conference (Nashville), Understanding Entrepreneurship Conference (Israel), Patent Conference (Chicago), and Works in Progress Intellectual Property Conference (Newark). Salomé Baslandze provided excellent research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Patent Value and Citations: Creative Destruction or Strategic Disruption?
David S. Abrams, Ufuk Akcigit, and Jillian Grennan
NBER Working Paper No. 19647
November 2013, Revised December 2018
JEL No. K1,L2,O3

ABSTRACT

Prior work suggests that more valuable patents are cited more. Using novel revenue data for tens of thousands of patents held by non-practicing entities (NPEs), we find that the relationship between citations and value forms an inverted-U, with fewer citations at the high end of value than in the middle. We explain the inverted-U with a model of innovation that has productive and strategic patents. Empirically, we observe more strategic patents where the model predicts: among inventors in fields of rapid development and where divisional applications are employed. These findings have important implications for our understanding of growth, innovation, intellectual property policy, and patent valuation.

David S. Abrams
University of Pennsylvania
Law School
3501 Sansom Street
Philadelphia, PA 19104
dabrams@law.upenn.edu

Ufuk Akcigit
Department of Economics
University of Chicago
1126 East 59th Street
Saieh Hall, Office 403
Chicago, IL 60637
and NBER
uakcigit@uchicago.edu

Jillian Grennan
Duke University
Fuqua School of Business
100 Fuqua Drive
Durham, NC 27708
jillian.grennan@duke.edu

One of the core questions of economics, both at the micro and macro level, is what leads to productivity gains. In order to understand what policies impact innovative activity and ultimately productivity, it is crucial to start with a good metric to value innovation.[1] The importance of such a metric has long been recognized (Scherer (1967), Griliches (1981), Aghion and Jaravel (2015)), yet few studies have the necessary data to provide external validation of existing proxies for valuing innovation (Schankerman and Pakes (1986), Hall and Harhoff (2012)).

Over the last several decades, a number of pioneering efforts were made to overcome the challenges inherent in measuring the value of innovation. Given that patent records contain a wealth of information on each patented invention as well as citations to previous patents, patent counts and citation-weighted patent counts have become popular proxies for the value of innovation.[2] The intuition is straightforward: ceteris paribus, fields with greater innovative activity will have more value to protect and will do so by applying for more patents.[3] Weighting patent counts by forward citations[4] is a natural augmentation to simple patent counts, given the well-known fact that patents vary tremendously in value.[5]

The use of this measure implicitly assumes that a larger number of citations corresponds to higher value. Testing this assumption is challenging, given that there are few sources of patent-level value. Of course, if there were, it would be unnecessary to proxy for it. Several problems have held back this inquiry: the reluctance of companies to share proprietary patent data, the lack of generality and sufficient observations from any single patent portfolio, the fact that only a small portion of the patents that are exchanged have patent-specific transaction values, and the fact that stock market measures of value reflect expectations instead of realized value.

We add to this literature by using novel licensing data from a set of non-practicing entities (NPEs).[6] By using a proprietary data set with licensing revenue for tens of thousands of patents, we are able to rigorously examine the relationship between patent value and

[1]Throughout the paper we generally use the term "value" to mean private value, but we do explore social value as well in Appendix D and discuss it in Section 2.3.1. See Abrams and Sampat (2018) for detailed discussion about the meaning of value in the innovation context.

[2]Lerner and Seru (2017) document that academic researchers increased their use of patent citation data as a proxy for innovation between 2005 and 2017.

[3]Of course, there is substantial variation in patenting across fields that is unrelated to the value of the underlying innovation as documented most famously by Levin et al. (1987).

[4]Forward citations is the number of citations received by a particular patent by subsequent patents.

[5]Fewer than 10 percent of patents are worth the money spent to secure them (Allison et al. (2009)), but the most valuable ones are thought to be worth hundreds of millions of dollars (Hall, Jaffe and Trajtenberg (2005)).

[6]An NPE is a company that earns money by licensing, selling, or enforcing patents, but not primarily by producing products.

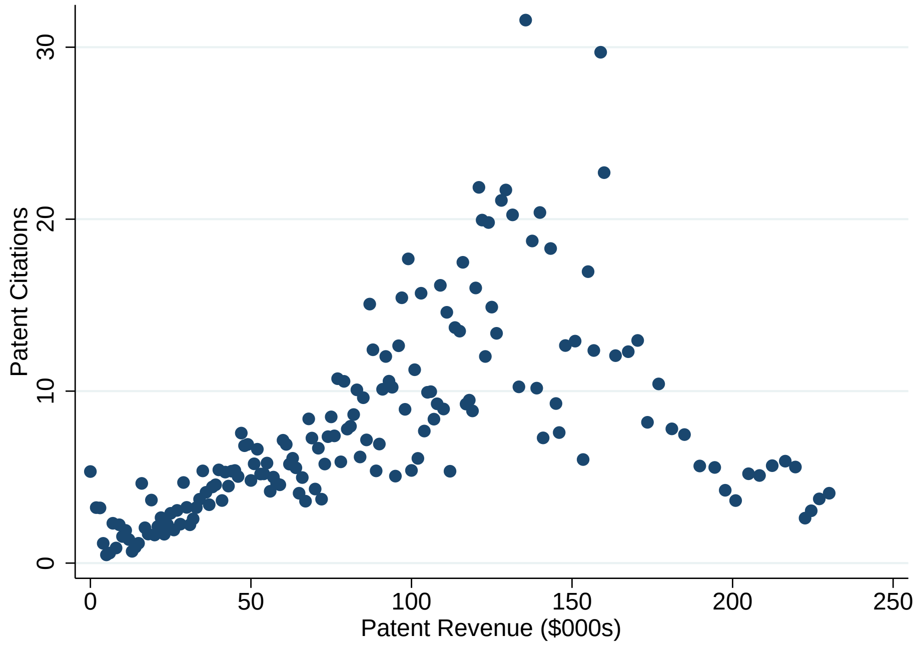


FIGURE 1: FORWARD CITATIONS AND PATENT VALUE

NOTE. FORWARD CITATIONS ARE DEFINED AS THE NUMBER OF CITATIONS A PATENT RECEIVES WHILE HELD BY A NON-PRACTICING ENTITY (NPE). PATENT VALUE IS THE SUM OF CONSUMER PRICE INDEX (CPI) DEFLATED ANNUAL REVENUES DERIVED FROM THE NPEs' LICENSING AGREEMENTS. THE REVENUE DATA ARE NORMALIZED TO A MEAN ANNUAL REVENUE OF $30,000 FOR CONFIDENTIALITY PURPOSES.

citations. This data set has the major advantage that it allows us to compute patent-level licensing revenue. But the generality of the findings is limited by the technological fields represented by the portfolio and the particular selection process used to assemble it.[7] While the patents analyzed are held by NPEs, almost all were originally granted to individual inventors or those employed by practicing firms.

This new data suggest that the relationship between forward citations and patent value is not only non-linear, it is not even monotonic. Figure 1 displays this relationship, computed from tens of thousands of observations.[8] There is still an overall positive correlation between citations and value, similar to what has been found in prior work, but it comes primarily from lower private value patents. The full citation-value pattern is more complex. Regression results are consistent with the evidence in this figure. We perform a large number of specification checks, use different functional forms, examine different subsets of the data, include an array of controls, all of which support the basic inverted-U relationship.

[7]We discuss details of the data set in Section 1. Confidentiality agreements limit our ability to disclose actual revenue numbers or number of observations.

[8]Further details on the normalization and other aspects of the production of this figure are discussed in Section 3.

To better understand the economic forces behind the result, we introduce a theoretical model that predicts an inverted-U relation between citations and value due to competing effects in the generation of citations. The model includes two types of innovative effort, which we characterize as *productive* and *strategic*. Productive innovative effort leads to the traditional increasing relationship between patent value and citations; strategic innovative effort, however, leads to a negative relationship. Strategic innovation is aimed at protecting high-value innovations. This can be achieved through aggressive action in a product market or strategies such as fencing patents, which seek to expand the area of protection available to previously granted patents. The intuition is that high value innovations create a greater incentive to protect them, which deters downstream innovation and leads to a negative relationship between citations and value above a threshold.[9] A key prediction from the model is that strategic innovative efforts generate the downslope. To test this prediction, we examine the citation-value relationship using characteristics that proxy for strategic efforts. First, we focus on parent patents, which are defined as patents that subsequently have follow-on patent applications via continuation or division. Then, we examine patents in technology classes with rapid growth as they are likely to generate greater expected profits. Using both of these proxies for strategic patenting we find evidence that is consistent with the prediction of the model.

This paper makes two major contributions to the innovation literature. Using detailed patent-level revenue data, we empirically demonstrate a non-monotonic relationship between forward citations and value for a large number of patents. We develop a model to explain this relationship whereby innovators engage in two types of effort, productive and strategic. We then test additional predictions of the model and find support for it. The empirical and theoretical insights may help future work on innovation by providing a more complex and hopefully more accurate description of the innovative process. For example, our evidence suggesting that innovators strategically defending against follow-on competitors helps to explain the high value-low citation patents provide insights for those studying economic growth, innovation, intellectual property policy, and patent valuation.

The rest of the paper proceeds as follows. In Section 1 we provide substantial detail about the extensive literature we build on, incentives to patent and cite, the business models of NPEs and further description of the data. Section 2 introduces our model which we believe captures some of the key elements of innovation and the patenting and citing processes. In Section 3 we present the main empirical results and a discussion of them. Section 4 concludes

[9] In Section 2.3.2, we discuss an additional model focused on strategic choices that may be more common among NPEs. That model also predicts a non-monotonic relationship between forward citations and value and may be found in Appendix E.

and makes the point that the goal of this work is not to undermine the large body of work on innovation that has relied on widely-held assumptions about the patent value-citations relationship. Rather, we hope that this will help build a more robust literature that informs some of the central economic issues of our time. Finally, Appendix A contains additional data descriptions and Appendices C-F contain additional theoretical proofs and derivations.

1 Background and Data Description

1.1 Prior Literature

By examining the relationships between patent value and citations through the lens of theory and data, we make several contributions to the existing literature. First, this paper illustrates that the heterogeneous motives individuals and firms have for seeking patent protection can render the traditional relationship between patent value and citations unrecognizable. While research indicates that frictions among inventors (Jaravel, Petkova and Bell, 2018; Jaravel, 2018) and in the patent system (Bloom, Schankerman and Van Reenen, 2013; Williams, 2017) can have meaningful economic consequences, our study quantifies how meaningful these frictions are for a large number of patents. In particular, we document a key tension: socially productive investments in innovation that merit high private values are also the ones that encourage their owners to misuse the patent system by investing in innovation and applying for patent protection in a strategic manner. To formalize this tension, we provide a model for how citations accrue to patents. Unlike prior citation models which suggest citations are a function of obsolescence (Caballero and Jaffe, 1993), our model distinguishes between productive and strategic investments in innovation. Thus, our model incorporates real-world aspects of patenting such as sequential innovation (Green and Scotchmer, 1995; O'Donoghue, Scotchmer and Thisse, 1998) and the precautions taken to avoid imitation (Gallini, 1992; Lerner, 1995) while recognizing the intrinsic value from breakthrough (Kerr, 2010) and general purpose technologies (Bresnahan and Trajtenberg, 1995).

Given that our findings support the notion that not all patents have a productive role in technological progress, we contribute to research exploring distortions to effective cumulative innovation (Williams, 2013; Galasso and Schankerman, 2015; Budish, Roin and Williams, 2015; Sampat and Williams, 2018) as well as theoretical work on strategic patenting (Farrell and Shapiro, 2008; Noel and Schankerman, 2013) and empirical work (Hall and Ziedonis, 2001; Ziedonis, 2004; Hegde, Mowery and Graham, 2009; Galasso and Schankerman, 2010; Cockburn and MacGarvie, 2011; Von Graevenitz, Wagner and Harhoff, 2013).[10] Generating

[10] See also Nicholas (2013) for a survey on this topic.

a strategic patent by assimilating and exploiting existing information from previous R&D investments also relates to research on innovation and learning (Cohen and Levinthal, 1989; Geroski, Reenen and Walters, 1997). Strategic patenting has also received a great deal of attention in the legal literature, which we do not review here. Some have argued that we have arrived at a point where the patent system is actually detrimental to innovation (Bessen and Meurer, 2008; Boldrin and Levine, 2013). We capture these observations and intuitions by modeling strategic patents as ones which do not lead to substantial further work in a field and in fact may stifle it. Thus, there may be extremely valuable strategic patents that receive very few citations, leading to a null or negative relationship between forward citations and revenue.

By providing a novel test of the efficacy of the relationship between patent value and citations, we contribute to a large literature on the value of innovation. Our primary finding is that the citation-value relationship has an inverted-U shape instead of the monotonic relationship previously assumed. Trajtenberg (1990) is the first paper to examine the citation-value relationship. His paper focuses on a relatively small number of patents in the computed tomography (CT) field, with values imputed from a structural model of the CT device market and finds an approximately linear relationship between citations and value. Harhoff et al. (1999) use survey data to test the citation-value relationship and find a positive but noisy correlation. Their evidence comes from categorical measures of patent value for 772 German patents with a single priority date and all of which renew to full term. Several studies examine the patent value distribution using the patent renewal decision to infer value (Pakes (1986), Schankerman and Pakes (1986), Bessen and Meurer (2008)). These papers find a linear relationship, but the range of patent values they evaluate is on the low-end of the distribution because of the renewal fee bound.

Finally, some studies examines patents issued to public firms to estimate patent value from the firm's overall valuation. Given that public firms follow significantly different innovation strategies than private firms (Bernstein, 2015), such findings are unlikely to generalize. As discussed in Hall, Jaffe and Trajtenberg (2001), the public firm approach originally showed that patent counts do not add to firm value after R&D is included in a Tobin's q regression, but a significant relation exists between citation-weighted patents and firm value. The market premium associated with citations stems from the high valuation of the upper tail of cited patents as opposed to a smoother increase in value as citation intensity increases. More recently, Kogan et al. (2017) employ an event study framework to estimate the expected value of a granted patent. Their study documents a positive correlation between the grant-day stock market reaction and citation-weighted patent counts, yet they also show considerable differences in relation to subsequent realizations of firm growth.

1.2 NPE Business Model

Since a major impediment to greater understanding of patent value has been the lack of available data on patent revenues, it is worth discussing the data sources and characteristics in some detail. The data in this paper was provided by a set of NPEs. The patents held by NPEs primarily cover the technology sector.[11] Our NPEs follow a portfolio business model, meaning their revenue primarily derives not from producing products based on patented technology, but from licensing the patents that they acquire (FTC, 2016).[12] The incentives associated with patents within the technology sector encourages licensing (Lerner and Tirole, 2002, 2005), and the number of patents held by NPEs continues to grow rapidly (Shapiro, 2012)).

This is fortunate for those interested in learning about innovation as NPEs function as an excellent data source in many ways, and when compared to traditional patent holding firms, NPE-derived data sets have several advantages. Their portfolios can be substantially larger than practicing firms, since their capital is almost exclusively employed in assembly and licensing, rather than production. NPEs are more diversified than practicing firms as well, since it is often easier to acquire the breadth of expertise necessary to acquire and license patents in a large array of fields, rather than to practice them. The data available from NPEs are also likely to be substantially more useful for researchers, as they compute patent-specific revenues. This is something that few practicing firms do, unless licensing is a major part of their business. This should come as little surprise, since ultimately most firms care about overall profit from innovation, not specifically from which patent the profit derives.

Our NPEs often acquire patents hundreds at a time through portfolio purchases. As such, the vast majority of patents represented in the data were not targeted for acquisition, but "came along for the ride." Given that these non-random patents are likely the most valuable patents, when conducting our empirical tests we assess the extent to which the top 1% of top 5% of patents in terms of realized value alter the observed inverted-U relation. Moreover, reviewing the data from our NPEs, we see no evidence that those patents they target tend to be particularly different than other patents. Rather their strategy is to accumulate patents in a broad range of technology areas that they believe may be important, so often the targeted patents represented a technology area they wanted to expand into. To understand the extent of selection in our data, we further compare characteristics of the data with the universe of

[11] See Abrams et al. (2018) for an investigation into the broader impacts of NPEs.

[12] Other NPEs follow a litigation business model in which infringement lawsuits are filed as a negotiating tactic for a settlement. Several recent papers explore the implications of the litigation NPE model (Kesan and Schwartz, 2014; Bessen and Meurer, 2014; Cohen, Gurun and Kominers, 2018; Feng and Jaravel, 2018).

patents in Section 3.5.

One concern about the data source may be that our NPEs use patents differently than practicing entities and as such the revenue data we use may not reflect the value of patents. In particular, there may be concern that the NPEs behave opportunistically and hold up individuals and firms that may be infringing by threatening litigation. We discuss and model this possibility in more detail in Section 2.3.2 and Appendix E. Here it is important to note that all revenue data we use is derived from licensing, not litigation. In addition, the NPEs tend to license each patent to multiple customers. Thus, it is far more likely that the licensing fees paid to the NPEs reflect real value for the licensees. Even if the value that the NPEs receive may be higher than the value a practicing entity would receive, this may not be because of holdup. An alternative view is that NPEs as intermediaries reduce search frictions for knowledge and thus reduce the rate of private obsolescence of knowledge. In this sense, they are monetizing value that was already there but not being realized because of search frictions. Additional details about the revenue data may be found in Appendix A.

1.3 Data Description

By characterizing the existing literature on patent value and the business model of the set of NPEs, the advantageous aspects of our data are clear. The detailed contracts for many licensing agreements across multiple customers enables us to use the realized value for individual patents in our empirical analyses. Such data help to accurately characterize the full distribution of patent value rather than relying on truncated values as with inferences from renewal decisions or relying on the implicit assumption that stock market investors rationally price patents at issuance. Moreover, since we have data from a set of NPEs, our sample size of tens of thousands of patents is much larger than most previous studies and allows us to include patents issued to government entities, individual inventors, private, and public firms.

Before presenting summary statistics, it is important to note additional distinctive characteristics of the data. At the request of the portfolio owners, we have agreed to not report the exact number of observations beyond noting that there are tens of thousands of patents in the data set. In the calculation of lifetime patent value (see Appendix A) we have also normalized the data such that mean annual revenue is \$30,000.[13] Thus throughout the paper, all dollar values are subject to this normalization. While absolute values are not accurate, relative values are and this normalization does not impact our ability to examine the for-

[13] While the dollar sign on the normalized values is superfluous, we keep it as a reminder to the reader that the original variables were denominated in dollars.

ward citation-value relationship or other correlates of value. Appendix A also discusses the normalization procedure for comparing forward citations across patents of different ages.

With these points in mind, we present summary statistics for the primary patent and assignee characteristics in Table 1. We restrict the data to U.S. utility patents, and exclude design and plant patents. We obtain annual licensing revenues from 2006–2015 for each patent and calculate lifetime value from this data. Some of the patents expire during this time period, and some are granted after 2008, but most are active for the full period. If a patent is never active during this period, it is excluded.

Table 1: Summary Statistics for NPE and Comparable USPTO Patents

	Panel 1: NPE Patents			**Panel 2: PTO Patents**		
Variables	**Mean**	**Median**	**Sd**	**Mean**	**Median**	**Sd**
Firm Size	4.82	4.87	2.99	6.68	7.37	2.99
Individual Inventor	0.05	0	0.21	0.01	0	0.12
Lifetime Forward Citations	32.7	11.2	65.0	21.2	7.6	43.5
Backward Citations	24.6	9.0	69.6	14.1	7	33.4
Parent Patent	0.22	0	0.42	0.14	0	0.34
Claims	12.3	9	10.3	12.4	10	9.9
Frac Bkwd Cites in Past 3 years	31.3	25.0	29.6	30.1	25	29.6
Application Year	2001	2001	5.1	2002	2002	4.5
Revenue (Thousands of 2010 dollars)	236	48	2099	n/a	n/a	n/a
Sale Indicator	1	1	0	0.015	0	0.12

Note. Panel 1: Patent application data from 1987–2014 for all U.S. utility patents granted in the NPE dataset (2014 application year only has two observations). Revenue data spans 2006–2015. Panel 2: Patent application data from 1987–2014 for all U.S. utility patents granted and weighted by subcategory. The post-normalization by patent subcategories was initially introduced in Hall, Jaffe and Trajtenberg (2001). Individual inventor is one if there is a single listed inventor and no assignee. Please see the appendix for additional variable definitions.

We define patent value as the sum of the normalized annual revenues realized by a patent while it is held by an NPE. The mean patent value is $235,723 (all figures are 2010 dollars). Note that the standard deviation of $2.09 million is almost 9 times the mean and more than 40 times the median value of $47,955.

The high level of dispersion (and skewness) is consistent with prior studies of patent value. Bessen (2008) uses the patents as options methodology and finds that U.S. patents issued in 1991 have a mean value of $121,000 and a median of $11,000. A closer comparison to the current study may be made by focusing on technology categories. Bessen finds a mean-to-median value ratio of 5.7 for Electrical and Electronic patents and 2.1 for the Computers and Communications category. The data set under study has a mean-to-median ratio of approximately 5, in between these two figures. Serrano (2010) determined the average private value of a patent right to be $90,799 and the median $19,184, which exhibits a similar mean-to-median ratio as our data.

One variable of interest, lifetime forward citations, also has a skewed distribution with a mean of 32.7, standard deviation of 65.0 and median of 11.2. The degree of skewness in the distribution of forward citations, the very wide range of forward citations, and the concentration of patents with 1 or fewer citations replicate familiar patterns such as those reported in Trajtenberg (1990), Harhoff et al. (1999), and Hall, Jaffe and Trajtenberg (2005). We also compare the raw median number of forward citations (not adjusted for patent age) in our data with that of the entire universe restricted to the same PTO technology categories and find them to be very similar: 8.75 for our data and 8.0 for the universe. Backward citations are also skewed, with a mean of 24.6, median of 9.0 and standard deviation of 69.6. About 30% of backward citations are for patents issued within the prior 3 years. We use this measure as an indicator of how active or hot a field is.

Most (67%) patents are original applications and the remainder are divisionals or continuations. Under U.S. law, inventors may file continuations or divisionals for their patent applications to cover new improvements to their inventions or to cover different aspects of their inventions (see Hegde, Mowery and Graham (2009) for more on the use of continuation patents). The difference between a divisional and continuation patent is that divisional applications make a distinct, new independent claim not in the parent application. The median application year is 2001, meaning the median patent had about 7 years of protection left by the end of our revenue data. Noncorporate inventors account for 14% of the patents, which is similar to that reported in Bessen (2008).

Table 2 shows that value and forward citations vary substantially by technology class.[14] The most valuable patents are found in the Information Storage category with a mean value of $623,000 but with only an average of 6.6 citations. At the low end are Optics and Electrical Devices which average 92 and 74 thousand respectively.

2 Theory of Patent Valuations and Citations

In this section, we offer a new model of innovation, patents, and citations. Our purpose is to develop a better understanding of the underlying reasons for the observed inverted-U relationship between citations and patent value. We embed intuitive assumptions into a structural model, and show that the model fits the observed pattern well.

[14] These classifications are subcategories from Hall, Jaffe and Trajtenberg (2001) and have been adopted by the USPTO

Table 2: Revenue and citation data for prominent subcategories of NPE patents

Subcategory	**Mean Revenue (Thousands of 2010 dollars)**	**Citations while held by NPE**
Communications	216	6.08
Hardware/Software	282	9.72
Computer Peripherals	278	6.73
Information Storage	623	6.64
Information Processing	236	13.96
Electrical Devices	74	3.12
Power Systems	227	4.39
Semiconductors	152	3.01
Misc-Elec.	117	3.98
Optics	92	3.04

Note. Patent revenue and citation data for U.S. utility patents granted in NPE dataset with at least 500 patents in the technology subcategory initially introduced in Hall, Jaffe and Trajtenberg (2001).

In our model, we rely on the *Schumpeterian theory of creative destruction* (see the survey by Aghion, Akcigit and Howitt (2014) for more on this topic), where each new innovation builds on previous technologies, but also makes them obsolete by introducing a better one. This tension between the incumbent technology owner's wish to defend its monopoly power and the future innovator's wish to utilize the spillovers generated by the current incumbent helps us rationalize the non-monotonic relationship between patent value and subsequent entry, identified by forward citations.[15] Our model emphasizes the decision to innovate productively or strategically.

Our model features two distinct types of innovation efforts – *productive* and *strategic*. The intuition for productive innovation follows the traditional economic view that patents are offered as a contract between society and the inventor. In return for a limited period of exclusivity, the inventor agrees to make his invention public rather than keeping it secret. This institutional arrangement promotes the diffusion of ideas (spillover) and economic growth. New big ideas generate a higher profit for the original inventor and also generate more spillovers for subsequent innovations. Hence a positive relationship occurs between patent value and subsequent entry (forward citations).

However, this is likely not the full story. Therefore, we also introduce the notion of the strategic innovation, a type of destructive creation. This idea seeks to capture the fact that

[15] Relatedly, Farrell and Shapiro (2008) emphasize the ability of patent holders, even of weak or less productive patents, to hold up firms through the threat of infringement. Arora, Ceccagnoli and Cohen (2008) show conditions under which firm's with equivalent R&D efforts patent differently; Bloom, Schankerman and Van Reenen (2013) model negative and positive technology spillovers based on a firms position in technology and product market spaces.

when firms and individuals are endowed with an exclusionary right, they may use it strategically to defend their existing market share in ways that do not serve the original intent of the legislation that created the right in the first place. Hence a valuable strategic innovation is one that prevents subsequent entry. This structure generates a negative relationship between patent value and subsequent entry (forward citations).

In order to highlight the distinct features and impacts of productive and strategic innovations, we introduce the model in two steps: In Section 2.1, we first introduce a model with productive innovations only. In this version of the model, we abstract from incumbent innovations and focus only on entrants' innovations. This assumption is relaxed in the subsequent model in Section 2.2 where we allow incumbent firms to create strategic innovations, which protect their valuable productive patents and market share. For reasons that we explain formally below, our model predicts that the link between patent value and citations is positive for productive innovation efforts and negative for strategic innovation efforts.

2.1 The Case of Productive Innovations

In this section, we introduce a continuous-time model with a representative household. The household consumes a basket of goods, each of which is produced by a different incumbent monopolist. The household's intertemporal consumption/saving decision, which does not impact the innovation dynamics in this economy, is provided in Appendix C for the interested readers.[16]

The economy features a large number of outside entrepreneurs who invest in *productive innovations*. Through these productive innovations, outside entrepreneurs replace existing incumbents and obtain some market share. The key feature of the productive innovation model that relates to citations is how new innovations arrive. Specifically, we assume that new innovations and innovative efforts arrive in clusters and that each new patent cites the prior art within the same technology cluster. Intuitively, this corresponds to the fact that certain markets become hot and attract top talent to invest their innovative efforts in that market. This simple logic leads to clustering of innovations by technology sector over time. Although this is an assumption of the model, it is one that has empirical support (Jaffe and Lerner (2004)). In terms of the model, what follows from this logic is an endogenous-citation dynamic.

The positive link between citations and patent value comes from the fact that more novel innovations will have larger mark-ups due to their originality, denoted by the step size of a

[16]Household's saving decision pins down the equilibrium interest rate in this model and provided for completeness.

new innovation, which corresponds to larger patent values. At the same time, more novel innovations generate larger spillovers for subsequent innovations, which will encourage new investment in innovation by outside entrepreneurs. With more entrepreneurs entering the market, a natural cluster of innovative effort over time by technology is created. Since a new innovation must cite the previous related patents upon which it builds, more novel patents receive more citations on average. Thus, the first simple model of productive innovation effort leads to the traditional conclusion of a positive correlation between citations and patent value. Given this intuition for the model of productive innovation, we now turn to the details.

Basic Environment Consider the following continuous time economy that admits a representative household. A unique final good, Y_t , is produced using a continuum of varieties indexed by $j \in [0, 1]$ as follows:

$$Y_t = \exp \int_0^1 \ln y_{jt} dj. \tag{1}$$

In this expression, y_{jt} is the quantity of variety j at time t. We normalize the price of the final good Y_t to be 1 in every period without loss of generality. The final good is produced in a perfectly competitive market.

Each variety j is produced by a monopolist who owns the latest innovation (patent) in sector j. The monopolist's production function takes the following simple form

$$y_{jt} = q_{jt} l_{jt} \tag{2}$$

where l_{jt} is the labor employed for production and q_{jt} is the variety-specific labor productivity. In what follows, new innovations improve labor productivity, which leads to aggregate growth in the economy. The linear production function implies that the marginal cost (M_{jt}) of producing 1 unit of y_{jt} is simply

$$M_{jt} = \frac{w_t}{q_{jt}}$$

where w_t is the market wage rate which is taken as given by the firm. Note that all monopolists hire from the same labor market, hence every monopolist faces the same wage rate w_t.

Labor productivity q_{jt} is improved through subsequent innovations in each product line j. Innovations belong to technology clusters. Let n index the order of an innovation in a technology cluster such that the very first patent that starts a new technology class has $n = 0$, the first follow-on innovation in the same technology cluster is indexed by $n = 1$, the second follow-on innovation by $n = 2$, and so on. Each innovation by a new entrant into j improves the previous incumbent's technology by a factor of $(1+\eta_n)$ which is only a function of the order n of the patent in the technology class and remains constant as long as the same firm is in charge of production. Consider a product line where productivity at time t is q_{jt} and a new innovation of step size η_n is received during $(t, t+\Delta t)$. Then the labor productivity evolves as:

$$q_{jt+\Delta t} = (1+\eta_n)\, q_{jt}. \tag{3}$$

When a new firm innovates and enters into j as the new market leader, the latest innovator and the previous incumbent compete in prices à la Bertrand.

2.1.1 Static Equilibrium: Production, Pricing and Profits

It is useful to solve the static production and pricing decisions before we describe the innovation technology. Consider the final good production in (1). Because the final good technology has a Cobb-Douglas form with respect to all varieties, the household will spend the same amount Y_t on each variety j. Hence the demand for each variety j can be expressed as

$$y_{jt} = \frac{Y_t}{p_{jt}} \tag{4}$$

where p_{jt} is the price charged by the monopolist j. Note that the Bertrand competition between the new monopolist and the previous incumbent, together with the unit elastic demand curve in (4) implies that the monopolist will follow limit pricing and charge a price that is equal to the marginal cost of the previous incumbent. If the productivity of the current monopolist in j is q_{jt} and the size of her innovation was η_n, then the marginal cost of the previous incumbent is simply $(1+\eta_n)\, w_t/q_{jt}$, which implies that the current monopolist's price is simply

$$p_{jt} = \frac{(1+\eta_n)\, w_t}{q_{jt}}.$$

Therefore we can express the equilibrium profit of the monopolist j as

$$\begin{aligned} \pi_t (q_{jt}) &= [p_{jt} - M_{jt}]\, y_{jt} \\ &= \pi_n Y_t \end{aligned} \tag{5}$$

where we define $\pi_n \equiv \frac{\eta_n}{1+\eta_n}$ as the normalized profit $(= \pi_t (q_{jt}) / Y_t)$. This is the first step in establishing the value of an innovation. Because a new innovation grants patent protection until another new innovation makes it obsolete through creative destruction, the value of an innovation (patent) will be the expected sum of future monopoly profits that will be generated by this innovation.

The following lemma summarizes the rest of the static equilibrium variables Y_t and w_t.

Lemma 1 *The aggregate output in this economy is equal to*

$$Y_t = Q_t$$

where Q_t is defined as a productivity index

$$Q_t \equiv \left[\int_0^1 (1+\eta_j)^{-1} \, dj \right]^{-1} \exp \int_0^1 \ln \frac{q_{jt}}{1+\eta_j} dj.$$

Moreover, the wage rate is equal to

$$w_t = Q_t \int_0^1 (1+\eta_j)^{-1} \, dj.$$

Proof. See Appendix F. ■

2.1.2 R&D and Productive Innovations

The economy has a measure of outside entrepreneurs who try to innovate and replace the existing incumbents endogenously. Outside entrepreneurs invest in R&D to produce a new innovation stochastically. When they are successful, they improve upon the prior highest quality as in (3) . Productive innovations come in clusters as in Akcigit and Kerr (2018). In particular, new entrants invest in two types of innovations:

1. *radical innovations*,

2. *follow-on innovations.*

When a new radical innovation occurs, it starts a new technology cluster with a step size $\eta_0 = \eta > 0$. Alternatively, when a new follow-on innovation occurs, it builds directly on existing technology and the marginal contribution of the new innovation depends on its rank in the sequence of follow-on innovations within the same technology cluster. Follow-on innovations run into diminishing returns within the cluster such that the n^{th} follow-up innovation has a step size of $\eta_n = \eta\alpha^n$ where $\alpha \in (0, 1)$. For mathematical convenience, we assume that after a certain number of follow-on innovations ($n > n^*$), the step size becomes a constant value $\eta_n = \eta\alpha^{n^*}$. In summary, the step size of the $n + 1^{st}$ patent in a given technology cluster can be summarized as follows:[17]

$$\eta_n = \begin{cases} \eta \text{ if radical innovation} \\ \eta\alpha^n \text{ if follow-on innovation and } n < n^* \\ \eta\alpha^{n^*} \text{ if follow-on innovation and } n \geq n^* \end{cases}.$$

Since innovations come in technology clusters and each new innovation utilizes the spillover from the previous patents from the same technology class, our model generates a natural interpretation for citations. When there is a major innovation in a technology class with a step size η, it produces spillovers for the subsequent innovations since the follow-on step size becomes $\eta\alpha$ which encourages new entry into the field. Innovations must cite previous innovations within the same technology cluster, acknowledging that the patents are technologically related. Therefore, new patents in a technology cluster will cite the previous patents that established and developed the cluster. The following illustrates the structure.

Example 1 *This example is provided to show the connection between our model and the data. In particular, we describe how technology clusters emerge and who cites whom in those clusters. The following chart provides an example of innovation patterns in a single product line:*

Step size:	η	$\eta\alpha$	$\eta\alpha^2$	$\eta\alpha^3$	η	$\eta\alpha$	η	$\eta\alpha$	$\eta\alpha^2$	η ...
Patent:	N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8	N_9	N_{10}
Tech Class:	*Tech Class 1*				*Tech Class 2*		*Tech Class 3*			*Tech Class 4*

An example of a sequence of innovations in a product line

The example starts with a radical innovation N_1 which has a step size η. Innovation N_2 follows on N_1 with a step size $\eta\alpha$. Since N_3 is the second follow-on innovation in cluster 1, it has a step size $\eta\alpha^2$ and so on. N_5, N_7 and N_{10} are radical innovations which start new technology clusters; therefore their step sizes are η. As a result, innovation step sizes follow

[17]Note that in principle, we can allow the step size η_j to be a function of the sector j. This would not have any major impact on the inverted-U relationship that our model predicts.

cycles. The citing-cited pairs can be summarized as follows:

Cited	***Citing***	***Cited***	***Citing***
N_1 :	N_2, N_3, N_4	N_6 :	*none*
N_2 :	N_3, N_4	N_7 :	N_8, N_9
N_3 :	N_4	N_8 :	N_9
N_4 :	*none*	N_9 :	*none*
N_5 :	N_6	N_{10} :	...

Consider N_2, for instance. Since it builds only on N_1, N_2 cites only N_1. However, there are two patents (N_3, N_4) in the cluster that are building on N_2. Hence, N_2 receives two citations from them.

Now we turn to the value of an innovation. Consider an innovation of step size $\eta_n = \eta\alpha^n$. Let the aggregate innovation arrival rate of the next follow-on innovation be denoted by $\bar{z}_{n+1}$ and the next radical innovation by $\bar{z}_0$. Then the steady-state value of the n^{th} innovation is summarized by the following continuous time Hamilton-Jacobi-Bellman (HJB) equation

$$V_{nt} = \frac{\eta_n}{1+\eta_n} Y_t \Delta t + (1 - r\Delta t) \left[\begin{array}{c} (\bar{z}_0 \Delta t + \bar{z}_{n+1} \Delta t) \times 0 \\ + (1 - \bar{z}_0 \Delta t - \bar{z}_{n+1} \Delta t) \, V_{nt+\Delta t} \end{array} \right].$$

This expression is intuitive. During a small Δt, the n^{th} innovation in a cluster delivers a profit of $\frac{\eta_n}{1+\eta_n} Y_t \Delta t$ to its owner (see equation (5)). The future period is discounted by $(1 - r\Delta t)$. After Δt, with probability $\bar{z}_{n+1}\Delta t$ there is a new follow-on entry, and with probability $\bar{z}_0 \Delta t$ there is a radical entry. In both cases, the incumbent exits the market because she is replaced by a new entrant and her firm value decreases to 0. With the remaining probability $(1 - \bar{z}_0 \Delta t - \bar{z}_{n+1} \Delta t)$, the incumbent survives the threat of entry and receives the continuation value $V_{nt+\Delta t}$ of being the incumbent. Subtracting $(V_{nt+\Delta t} - r\Delta t V_{nt})$ from both sides, dividing through Δt, and taking the limit $\Delta t \to 0$ leads to the following HJB equation:

$$rV_{nt} - \dot{V}_{nt} = \pi_n Y_t - (\bar{z}_{n+1} + \bar{z}_0) \, V_{nt}. \tag{6}$$

where $\pi_n \equiv \frac{\eta_n}{1+\eta_n}$.

In what follows, we will focus on a balanced growth path equilibrium where all aggregate variables and value functions (i.e., Y_t, w_t, and V_{nt}) grow at the constant rate g. Then the following lemma provides the exact form of the value function.

Lemma 2 *The normalized value of the* n^{th} *follow-on innovation at time* t *is equal to*

$$v_n \equiv \frac{V_{nt}}{Y_t} = \frac{\pi_n}{\rho + \bar{z}_{n+1} + \bar{z}_0} \tag{7}$$

where $\pi_n \equiv \frac{\eta_n}{1+\eta_n}$.

Proof. This result follows from substituting the household's Euler equation $r - g = \rho$ into (6). The Euler equation itself is derived in Appendix C equation (13). ■

This expression simply says that the value of an innovation depends mainly on four factors: First, a larger step size η_n implies larger mark-up and therefore higher value of innovation. Second, larger aggregate output Y_t (which can also be interpreted as the size of the aggregate economy), will lead each variety to receive a larger demand and hence generate higher per-period profit and innovation value. Third, a higher discount rate (in this case, the growth rate adjusted interest rate through the household problem in equation 13, $\rho = r - g$) decreases present discounted value. Finally, the rate of creative destruction of the next follow-on innovation $\bar{z}_{n+1}$ or radical innovation $\bar{z}_0$ lowers the value of the current innovation due to a decreased expected duration of monopoly power.

So far, we have determined the value of each innovation v_n, as a function of the next innovation's arrival rate $(\bar{z}_{n+1} + \bar{z}_0)$. In order to pin down the arrival rate of follow-on innovations and radical innovations, we now turn to the entry problem of outside entrepreneurs. Let z_n denote the innovation rate of an individual entrepreneur and $\bar{z}_n$ denote the aggregate innovation rate of all outside entrepreneurs who are trying to innovate in the same product line j. We assume that there are congestion externalities such that the individual cost of innovation $K(z_n)$ is increasing in the aggregate innovation rate such that

$$K(z_n) = z_n \zeta Q_t \bar{z}_n \text{ for } n \geq 0$$

in terms of the final good and $\zeta > 0$ is some constant. Then the free-entry condition for a new entrant can be summarized as

$$\max_{z_n} \left\{ z_n v_n Y_t - z_n \zeta Q_t \bar{z}_n \right\}. \tag{8}$$

The free-entry condition, together with Lemma 1, pins down the aggregate entry rate as

$$\bar{z}_n = \frac{v_n}{\zeta}. \tag{9}$$

As expected, the entry rate is increasing in the value of a new innovation and decreasing in the cost parameter ζ.

Now, combining this last expression (9) with (7) gives us the recursive solution of patent value

$$v_n = \frac{\pi_n}{\rho + (v_0 + v_{n+1})/\zeta}.$$

Finally, the limit value of patents with $n > n^*$ is

$$\bar{v} = \frac{\zeta}{2}\left[\sqrt{\left(\rho + \frac{v_0}{\zeta}\right)^2 + \frac{4}{\zeta}\pi_{n^*}} - \left(\rho + \frac{v_0}{\zeta}\right)\right].$$

This model generates a positive relationship between patent value and average forward citations (subsequent entry). The following assumption is a sufficient condition to prove this result formally.

Assumption 1 *The model's parameters satisfy the following inequality,*

$$\frac{\zeta\rho^2}{\eta} \geq \frac{\alpha}{1-\alpha}.$$

This is a fairly weak assumption that asks for a sufficient level of decreasing returns in innovation, i.e., sufficiently small α. Again, this assumption is only a *sufficient condition* to prove Proposition 1.[18] Here are the main results emerging from this model:

Proposition 1 *Under Assumption 1, v_n decreases in n, and therefore the average number of forward citations received by an η_n patent during any time interval $[t_1, t_2]$ decreases in n.*

Proof. See Appendix F. ■

Assumption 1 ensures that the patent values are decreasing in n which implies that being an early inventor in a technology cluster (smaller n) is more profitable. Hence, entrants invest more aggressively (higher $\bar{z}_n$) in younger clusters which in turn generate a higher entry rate. The following theorem follows trivially from Proposition 1.

Theorem 1 *In the case of productive patents, patent value and forward citations are **positively correlated**.*

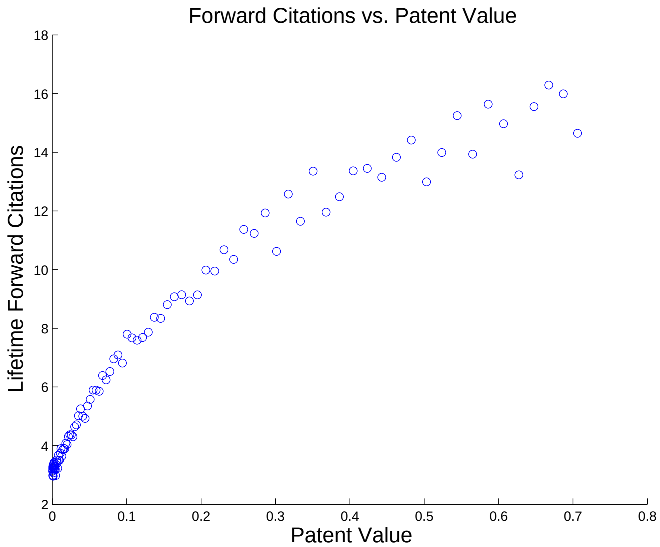
The intuition behind this result is straightforward: when a path-breaking innovation occurs, it creates a new technology cluster which generates spillovers for subsequent innovations. These spillovers generate a large number of entrants which all then cite the prior art

[18] We verified numerically that Proposition 1 holds for much broader set of parameters.

in the cluster. Since the path-breaking major innovation also has the largest mark-up (and value, accordingly), the positive correlation follows.

Figure 2 illustrates this positive correlation. We simulate the above model for 50,000 patents for 100 years, plotting patent value on the x-axis and forward citations on the y-axis.

FIGURE 2: FORWARD CITATIONS AND PATENT VALUE



NOTE: PARAMETERS USED IN THIS FIGURE: $\zeta = 1, \rho = 0.02, \eta = 1, \alpha = 0.9, \gamma = 0.05, n^* = 100$. EACH CIRCLE REPRESENTS THE AVERAGE NUMBER OF CITATIONS OF THE PATENTS WITHIN THE CORRESPONDING PATENT VALUE PERCENTILE.

2.2 The Case of Strategic Innovations

In the previous model, incumbents were passive in terms of protecting their monopoly position. In this section, we relax this assumption and introduce the possibility of incumbents producing *strategic innovations* in order to secure their position. The idea is that if an incumbent has a high value productive innovation, then she can potentially invest in a strategic innovation in order to make it harder for the next outside entrepreneur to leapfrog and steal the high monopoly rents. A strategic patent is one that decreases the likelihood that a prior productive patent will be improved upon, thereby increasing the value of the prior innovation and decreasing the expected number of citations it receives due to lack of entry. Hence, we should expect a negative relationship between patent value and citations when we add strategic patents, as illustrated in Figure 3.

Formally, upon each productive innovation, an incumbent has the opportunity to produce a single strategic (defensive) innovation. The technology for strategic innovation is such that by paying a fixed cost $\psi > 0$, a new entrant who just invented a productive patent can also obtain a strategic patent. To simplify the analysis, assume that ψ is high enough that it is profitable to invest in strategic innovations only for radical inventions (i.e., inventions with step size η). When a firm engages in strategic innovation, it raises the cost of innovation for the subsequent innovator by a multiplier $m > 1$ which is an iid random variable (realized upon innovation) from a distribution $\mathcal{F}(m)$ such that the cost to the next outsider is

$$K\left(z_{nm}\right) = \begin{cases} m z_{0m} \zeta Q_t \bar{z}_{0m} \text{ for radical inventors} \\ m z_{1m} \zeta Q_t \bar{z}_{1m} \text{ for follow-on inventors} \end{cases} .$$

Note that we index the entry rate of a radical inventor z_0 and a follow-on inventor z_1 by m since their cost is a function of the height of the fence m of the current incumbent.

Let us denote the value of an $m-$type strategic patent by v_m^d. Since a strategic patent is produced only by radical inventors, the profit collected in every instant is π_0. Therefore the HJB equation (after drawing m) is simply $\rho v_m^d = \pi_0 - \left(\bar{z}_{0m} + \bar{z}_{1m}\right) v_m^d$. This value function is expressed as

$$v_m^d = \frac{\pi_0}{\rho + \bar{z}_{0m} + \bar{z}_{1m}}. \tag{10}$$

Now consider the free-entry condition of an outsider who tries to enter after a strategic patent of size m. For $n \in \{0, 1\}$ the entry problem is simply

$$\max_{z_{nm}} \left\{ z_{nm} v_n - m z_{nm} \zeta \bar{z}_{nm} \right\}$$

where $v_0 \equiv \mathbb{E}_m v_m^d - \psi$ is the value of becoming a new incumbent through radical innovation. Note that this is an expected value over all possible values of m since the radical entrant will pay the fixed cost ψ and produce an additional innovation upon entry. Similarly, v_1 denotes the value to an incumbent that enters into the technology cluster with the *first* follow-on innovation. For $n \geq 2$, follow-on entry follows the same process as in (8) therefore we do not repeat it here.

As a result of the free-entry condition we get

$$\bar{z}_{nm} = \frac{v_n}{\zeta m}. \tag{11}$$

An important result here is that as the cost of innovation increases with m, the entry rate (and the potential forward citation rate) decreases.

Next, combining this entry rate with (10) we get the value of a strategic patent of type

m:

$$v_m^d = \frac{\pi_0}{\rho + \frac{v_0+v_1}{\zeta m}}. \tag{12}$$

Now we have the new results.

Proposition 2 *The value from strategic patents increases and the subsequent entry rate (forward citations) decreases in m.*

Proof. The first part of this proposition follows directly from equation (12) and the second part follows from equation (11). ■

Now we can state the main result of this subsection.

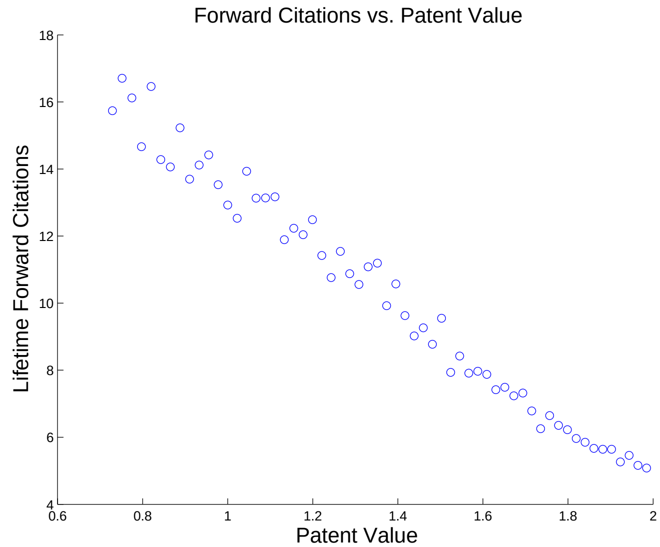
Theorem 2 *When strategic patents are allowed, patent value and forward citations are **negatively correlated**.*

The underlying reason for this negative relationship stems from the fact that the most successful strategic patents are the ones that increase the cost of entry the most (high m), which will reduce forward citations due to lower entry. The lower entry rate also allows the current incumbent to enjoy monopoly power longer, raising the value of being the incumbent. Hence we get a negative relationship between patent value and citations, as illustrated in Figure 3. The combination of radical, productive innovation and strategic innovation is very valuable, but because the strategic innovation alters the entry rate of new entrepreneurs through the endogenous citation dynamic, forward citations are dramatically reduced. Put another way, since forward citations enumerate all previous innovations since the most recent radical innovation, the reduction in citations is not due to less valuable technology, but to a higher cost of entry for new entrepreneurs.

Figure 4 illustrates the overall relationship between patent value and citations. The pattern is a very clear inverted-U, and echoes what we observe empirically in the data.

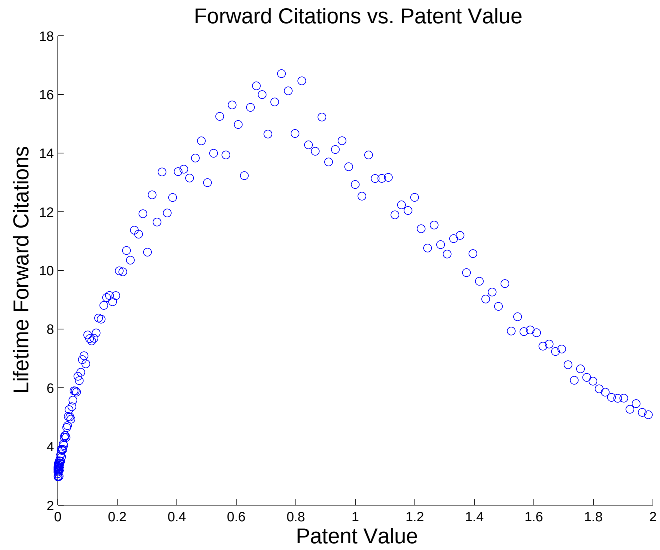
Our model suggests that incumbents with high-value patents will rely on strategic patenting to protect their existing market shares. Therefore we should expect the patents on the decreasing side of the inverted-U to come with greater frequency from large corporations with big market shares and those in emerging industries with higher profits. Moreover, the model implies that the strategic patenting is done to protect existing patents of the firm, so that we should also expect to see more divisional and continuation patents (as opposed to first-time patents) on the downward sloping side of the inverted-U curve. Section 3 tests these predictions.

Figure 3: Forward Citations and Patent Value



Note. Parameters used in this figure: $\zeta = 1, \rho = 0.02, \eta = 1, \gamma = 0.05, m \sim U[1, 4]$. Each circle represents the average number of citations of the patents within the corresponding patent value percentile.

Figure 4: Forward Citations and Patent Value



Note. Parameters used in this figure: $\zeta = 1, \rho = 0.02, \eta = 1, \gamma = 0.05, m \sim U[1, 4]$. Each circle represents the average number of citations of the patents within the corresponding patent value percentile.

2.3 Extensions

2.3.1 Social Value and Citations

To this point we have concentrated on the relationship between private patent value and forward citations, which we test empirically in Section 3. The literature has also largely focused on private patent value, but for policy decisions, it is valuable to consider the social value of patents as well. We extend the model to consider social value in Appendix D and show that there is a positive, monotonic relationship between social value and forward citations, a fact that is borne out by the most prominent empirical study that has examined this relationship (Trajtenberg (1990)). The intuition of this result is as follows. Forward citations are indications of subsequent inventive activity. While this could be bad news for an incumbent firm (due to shorter expected monopoly duration), such subsequent inventive efforts are socially valuable because they mean a faster pace of technological progress, no matter who produces the innovations. Therefore a larger number of citations is associated with higher social value. An important implication that emerges from this analysis is that the inverted-U is an indication of strategic use of the patent system, one that causes technological progress to diverge from its socially optimal level. Firms generate additional profit by using patents for strategic purposes, but this slows down subsequent technological progress - as indicated by the negative relationship between private value and citations. It thus may be desirable to direct policies toward sectors where this negative relationship (strategic behavior) is more pronounced.

2.3.2 NPEs and Opportunistic Use of Patents

Patents may be used strategically in ways other than to protect valuable innovations, as the model in Section 2.2 has explored. For example, some NPEs are accused of using the threat of patent infringement suits even where a finding of infringement is unlikely in an attempt to extract rents – we call this *opportunistic* use of patents. Although this scenario in unlikely to be of substantial impact in our data set, as we mentioned in Section 2, for completeness we extend our model to consider the opportunistic use of patents (Appendix E) for interested readers.

We show that the opportunistic use of patents by NPEs will also generate a negative relationship between private value and citations which buttresses our main hypothesis: when patents are used for productive purposes, the relationship between private value and forward citations is positive, whereas when patents are produced (Section 2.2) or used (Appendix E) for strategic purposes, the relationship between private value and citations become negative, delivering an overall inverted-U relationship.

The intuition for this result is as follows. When an opportunistic patent appears in a field that is used to sue some existing firms, the fear of future litigation discourages subsequent entry which then lowers the forward citations given to the opportunistic patent. The stronger the opportunistic patent's power to extract revenues, the lower the incentives for subsequent entry. In other words, the higher the monetary value of the opportunistic patent, the fewer citations.

While opportunistic patenting could lead to the same inverted-U prediction, strategic (defensive) patenting of Section 2.2 is the much more likely explanation in this data set. Almost all of the patents are originally granted to practicing companies or individuals and later sold to one fo the NPEs. In Section 3 we provide strong evidence that the use of continuation and divisional applications, which are associated with strategic patenting (as recently documented by Hegde, Mowery and Graham (2009)), are used more prominently exactly for the most valuable patents. This would be predicted by the strategic patenting story, but not the opportunistic one.

The revenue in this data set is also generated from licensing deals, not litigation. The licensing deals are non-exclusive and each patent is typically licensed to a large number of counterparties. The licenses also generally cover a portfolio of patents and not simply an individual patent. All of these facts suggest that the opportunistic hold-up of patentees is unlikely to explain the inverted-U in the data.

3 Empirical Analysis

We have seen how productive and strategic patents can combine to produce an inverted-U relationship between citations and patent value. We now expound upon and then expand upon the empirical results first presented in the introduction to assess the predictions of the model.

3.1 Inverted-U Results

Figure 1 displays the empirical relationship between forward citations and patent value for the data set described in Section 1. The figure plots the mean number of citations for each of 200 quantiles of patent value, with the top 5% trimmed. The figure shows an increasing relationship between patent value and citations for lower value patents and then a decreasing relationship for higher value patents: the inverted-U that our model predicts.

In Table 3 we report results of regressions of citations on a linear or quadratic function of patent value. Each column is a separate regression, with no controls, but with robust standard errors. The two pairs of columns in the table vary the share of the overall data set that is included, excluding the top 5% and 1% in the first and second pair of columns, respectively. Such exclusions help to ensure that extremely high value patents do not drive our statistical inferences. The coefficients show that there is indeed an overall positive relationship between forward citations and patent value. The even columns show that adding a quadratic term improves the fit, and the impression of an inverted parabola from Figure 1 is borne out by the statistically significant negative coefficients on the quadratic terms.

Table 3: Forward Citations and Patent Value

	(1)	(2)	(3)	(4)
Share of Most Valuable Patents Excluded	5%	5%	1%	1%
Revenue ($100,000s)	7.002**	17.04**	1.788**	4.804**
	(0.275)	(0.630)	(0.201)	(0.309)
Revenue Squared		-5.592**		-0.149**
		(0.293)		(0.0107)
R-squared	0.04	0.06	0.04	0.06

Robust standard error in parentheses
** p<0.01, * p<0.05

Note. Separate regressions are reported in each column with robust standard errors in parentheses. The dependent variable is forward citations while under control of one of the NPEs. The data are normalized so that the mean annual revenue is $30,000 (2010 dollars).

Figure 1 and Table 3 omit potentially relevant covariates, and one may be concerned that variation of both patent value and forward citations by these covariates drives the observed relationship. In Figure 5 the residual from a regression of forward citations on a set of indicator variables for technology category, application year, and NPE-acquisition year is presented. The same inverted-U relationship in Figure 5 is apparent that was previously seen in Figure 1. This suggests that there is still evidence for productive and strategic patenting, even within technology categories and years. Next, we include even more controls in our regression specifications.

Table 4 reports the results of regressions of forward citations on a quadratic of patent value and the following covariates: dummies for individual inventor, whether the patent was a parent patent, fraction of backward citations within the last three years, self-citations, claims, originality, and generality (Harhoff, Scherer and Vopel, 2003; Lanjouw and Schankerman, 2004; Moser, Ohmstedt and Rhode, 2018). The regressions also include NPE-acquisition

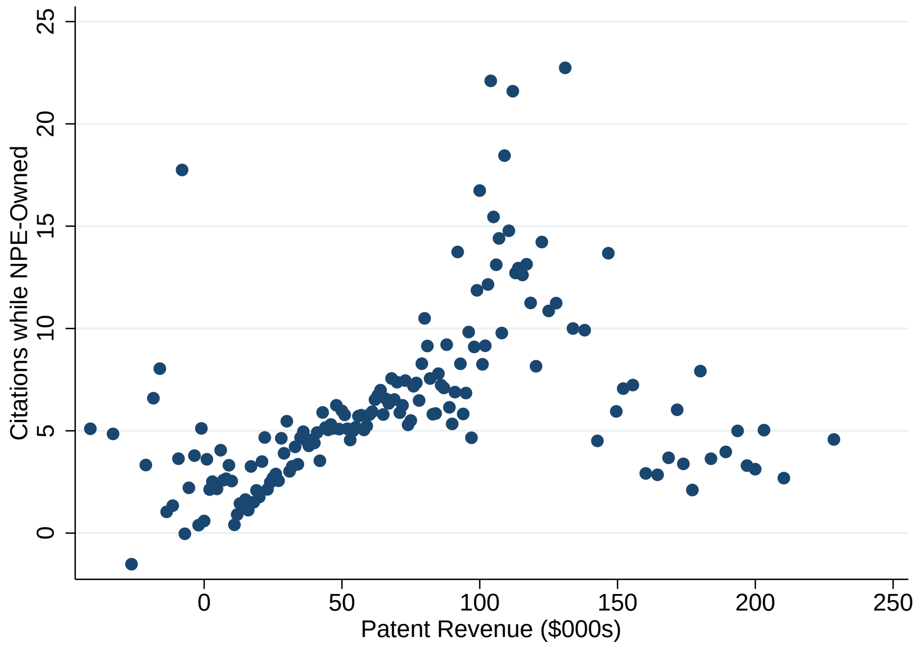


Figure 5: Forward Citation Residuals and Patent Value

Note. The data are normalized so that the mean annual revenue is $30,000 ($2010). Residuals are from a regression on a set of indicator variables for technology category, application year, and NPE-acquisition year.

year, application year, and subcategory fixed effects. The point estimate on the coefficients for the linear and quadratic value terms vary somewhat by which covariates are included, but in general consistently indicate an inverted-U shaped relationship between citations and value. Taken together, our empirical tests suggest a robust, inverted-U relation between patent value and citations.

While we document what appears to be an inverted-U relationship both graphically and through regressions with a quadratic functional form, we want to rule out the possibility that the significance of the square term is coming from observations outside of the downward sloping or strategic portion of the curve. To that end, we examine the relation between citations and value on the strategic portion of the curve (i.e., 75th to 95th percentile of the value distribution given that we exclude the top 5% of observations). As reported in Table 5, there is a significantly negative relationship between revenue and citations on the strategic portion of the curve. The more revenue the patent receives, the fewer citations it will receive. We also introduce an indicator variable for a patent being in this strategic range of the value distribution and test the model by interacting our indicator with patent value. As reported in columns 3 and 4 of Table 5, the interaction term is significant and negative,

TABLE 4: FORWARD CITATIONS AND PATENT VALUE

	(1)	(2)	(3)	(4)	(5)
Share of Most Valuable Patents Excluded	5%	5%	5%	5%	5%
Revenue ($100,000s)	13.40**	13.28**	11.11**	9.491**	6.856**
	(0.972)	(0.969)	(0.975)	(1.028)	(1.000)
Revenue Squared	-4.398**	-4.366**	-4.095**	-3.404**	-2.428**
	(0.400)	(0.399)	(0.397)	(0.423)	(0.414)
Individual Inventor	3.111**	2.962**	3.297**	2.803**	2.199**
	(0.584)	(0.586)	(0.589)	(0.584)	(0.559)
Parent Patent		3.597**	3.754**	3.590**	2.994**
		(0.577)	(0.594)	(0.582)	(0.568)
Fraction of Backward Citations in Last 3 Years			1.036**	1.040**	1.165**
			(0.366)	(0.368)	(0.364)
Self-Citations			1.350**	1.337**	1.224**
			(0.0566)	(0.0575)	(0.0601)
Claims			0.117**	0.0937**	0.0632**
			(0.0157)	(0.0159)	(0.0154)
Originality					0.771*
					(0.328)
Generality					12.33**
					(0.372)
NPE-acquisition Year Fixed Effect	Yes	Yes	Yes	Yes	Yes
Application Year Fixed Effect	Yes	Yes	Yes	Yes	Yes
Subcategory Fixed Effects	No	No	No	Yes	Yes
R-squared	0.089	0.093	0.147	0.164	0.205

Robust standard errors in parentheses
** p<0.01, * p<0.05

NOTE. SEPARATE REGRESSIONS REPORTED IN EACH COLUMN, WITH ROBUST STANDARD ERRORS IN PARENTHESES. THE DEPENDENT VARIABLE IS FORWARD CITATIONS WHILE UNDER CONTROL OF ONE OF THE NPES. THE INDEPENDENT VARIABLE IS REVENUE OF EACH PATENT. THE DATA ARE NORMALIZED SO THAT THE MEAN ANNUAL REVENUE IS $30,000 (2010 DOLLARS).

confirming that the non-linearity is larger. These findings hold when both controlling for a host of covariates and when we do not.

3.2 Robustness of Inverted-U

Establishing reliable empirical evidence of an inverted-U shape between patent value and citations requires careful calculation of the inputs. In our baseline specification presented in this paper, we do not do any truncation procedure because of its potential to add noise and the similarity between revenue and citation profiles in our data (see Figure 10). However, we want to try other methods to assure this is not a spurious correlation.

In the Appendix, we consider multiple permutations of the data and the regression spec-

Table 5: Forward Citations and Patent Value Along Strategic Range

	(1)	(2)	(3)	(4)
	Strategic Range	Strategic Range	Main Sample	Main Sample
Range of Valuable Patents Included	75th–95th	75th–95th	0–95th	0–95th
Revenue ($100,000s)	-2.141**	-2.863**	8.198**	2.030**
	(0.648)	(0.793)	(0.427)	(0.763)
Strategic Range			14.20**	9.583**
			(1.009)	(1.074)
Revenue x Strategic Range			-10.34**	-5.464**
			(0.776)	(1.020)
Individual Inventor		4.840**		2.730**
		(1.248)		(0.583)
Parent Patent		4.420*		3.551**
		(1.762)		(0.580)
Fraction of Backward Citations in Last 3 Years		2.039		1.021**
		(1.244)		(0.367)
Self-Citations		1.361**		1.340**
		(0.0638)		(0.0576)
Claims		0.118**		0.0937**
		(0.0444)		(0.0160)
NPE-acquisition Year Fixed Effect	No	Yes	No	Yes
Application Year Fixed Effect	No	Yes	No	Yes
Subcategory Fixed Effects	No	Yes	No	Yes
R-squared	0.00	0.19	0.06	0.17

Robust standard errors in parentheses
** p<0.01, * p<0.05

Note. Separate regressions are reported in each column with robust standard errors in parentheses. The dependent variable is forward citations while under control of one of the NPEs. "Strategic range" is defined as the 75th–95th percentile of the revenue distribution while the patent is held by one of the NPEs. The interaction regression includes 0–95th percentile of the revenue distribution. The data are normalized so that the mean annual revenue is $30,000 (2010 dollars).

ification. To address the issue that forward citations are non-negative, we consider a log transformation and a Poisson functional form. Appendix Figure 12 shows that in a log-log setup we would see the same inverted-U relationship. Appendix Table 8 reports regression evidence when using the log transformation for forward citations. We find a significant negative coefficient for the quadratic term for revenue, and evidence that suggests the patents in the strategic range of the patent value distribution are driving the inverted-U shape. Using a Poisson regression of citations on revenue, we also see similar results and we report those findings in Appendix Table 9.

We address remaining concerns about the robustness of our findings through additional checks on our measures for revenue and citations. Given the well-established skewness in the patent data, we verify that the inverted-U is not just the result of greater noise in the tail of the data. Appendix Table 10 shows significant negative coefficients on the quadratic term for revenue when we exclude a different portion of the valuable patents (only the top 1%). Second, given that citation counts are inherently truncated, we follow the method proposed by Hall, Jaffe and Trajtenberg (2005) to minimize measurement error. Specifically, we estimate the shape of the citation-lag distribution in each major technology subcategory using all U.S. utility patents with an application year between 1987 and 2014, which covers the scope of our data set, and use these coefficient estimates to generate lifetime citations. Our baseline specification for patent value as discussed earlier is a simple aggregation of revenue across licensing agreements while the patent is held by one of the set of NPEs. For our lifetime adjustments we try both a uniform revenue adjustment (i.e. multiply patents by the inverse of the proportion of years held by NPE), and a non-parametric model that estimates lifetime revenue in similar fashion to Hall, Jaffe and Trajtenberg (2005) for citations. Appendix Figures 13 and 14 reveal the same inverted-U relationship for both revenue adjustment procedures plotted against the lifetime citation measure. We evaluate these lifetime adjustments in Appendix Table 11 and report coefficient estimates that are statistically significant and consistent with the inverted-U hypothesis. Overall, across a range of data permutations, the inverted-U shape remains meaningful.

3.3 Strategic Patenting Results

In this section, we look at both the strategic and technological components of the relevant patents. Figure 6 reports the forward citation-patent value relationship as it relates to the prevalence of parent patents in the right tail of the revenue distribution. Parent patents are defined as patents that subsequently have follow-on patent applications via continuation or division. Using a lot of continuation or divisional applications is likely an indicator of sophisticated and strategic behavior by the applicant. For example, when a foundational technology is being applied to specific commercial applications, the priority application will become subdivided and amended. It is worth noting that this action is taken by the applicant developing the technology and not the NPEs, which suggests it is not subject to measurement error associated with NPEs opportunism. As is shown in Figure 6, the parent patents are more prevalent in the high value/low citation region of the graph. This evidence supports the prediction from the model that strategic innovative efforts generate the downslope.

We provide additional empirical support for strategic patenting driving the inverted-U shape between patent value and citations in Table 6. Using parent patents as our proxy for strategic patents, we find the interaction between the quadratic term for patent value and the strategic proxy is negative and statistically significant. This finding holds across

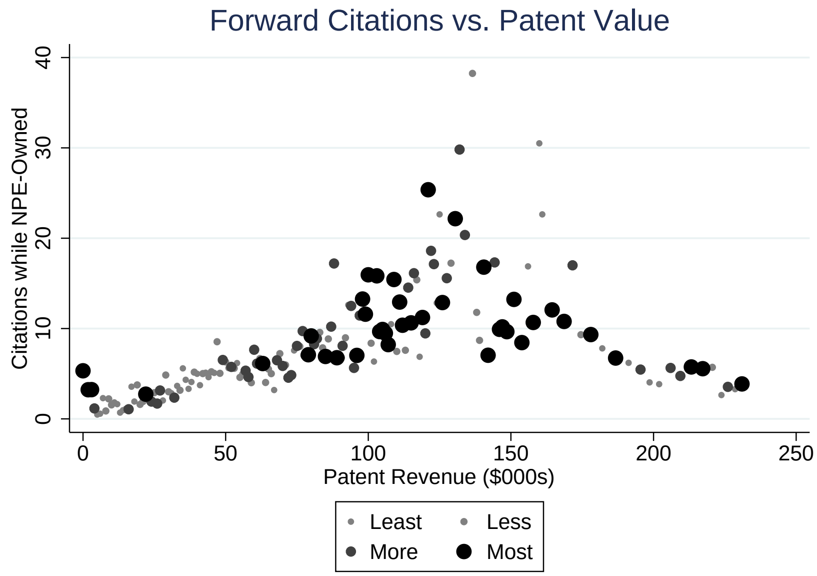


FIGURE 6: FORWARD CITATIONS AND PATENT VALUE BY SHARE OF PARENT PATENTS

NOTE. THE DATA ARE NORMALIZED SO THAT THE MEAN ANNUAL REVENUE IS $30,000 (2010 DOLLARS). PARENTS ARE TAKEN TO BE PATENTS THAT HAVE CHILDREN WITHIN PATENTS GATHERED BY ONE OF THE SET OF NPES.

a variety of specifications and suggests that the inverted-U shaped relationship between citations and value is more pronounced for strategic patents. Another patent characteristic that may influence the likelihood of strategic patenting is the rate of innovative growth in the field. Consistent with this intuition, areas of rapid innovation are likely to generate greater expected profits, and thus greater incentives for the patent applicant to engage in strategic patenting to protect its valuable, productive patents. Our measure for the growth of innovation is the share of backward citations within the prior three years. Given that this measure is based on the environment at the time of the application, it is also free from measurement error associated with NPE opportunism. Using this second, alternative proxy for strategic patenting, we find additional evidence to support the conjecture that inverted-U relationship stems from strategic patenting. We report these results in Appendix Table 12.

3.4 Inverted-U Across Technologies

We have noted above that the inverted-U relationship is not driven by differences across technology categories. We now investigate whether the relationship holds across technologies. In Table 7, we report results from six regressions of forward citations on patent value and patent value squared, one for each technology class. We find that the same overall relation-

TABLE 6: FORWARD CITATIONS AND PATENT VALUE WITH STRATEGIC PROXY

	(1)	(2)	(3)	(4)
Share of Most Valuable Patents Excluded	5%	5%	5%	5%
Parent Patent	1.185	1.204	1.005	0.897
	(1.286)	(1.286)	(1.327)	(1.308)
Revenue	12.59**	12.47**	11.85**	9.907**
	(0.860)	(0.859)	(0.873)	(0.914)
Parent x Revenue	7.830*	7.506*	7.454*	7.336*
	(3.587)	(3.592)	(3.670)	(3.638)
Revenue Squared	-4.019**	-3.975**	-3.732**	-2.884**
	(0.365)	(0.365)	(0.373)	(0.390)
Parent x Revenue Squared	-3.824**	-3.631*	-3.575*	-3.549*
	(1.475)	(1.479)	(1.510)	(1.495)
Individual Inventor		2.918**	2.670**	2.192**
		(0.588)	(0.589)	(0.584)
Fraction of Backward Citations in Last 3 Years			0.272	0.339
			(0.376)	(0.377)
Claims			0.119**	0.0925**
			(0.0158)	(0.0160)
NPE-acquisition Year Fixed Effect	Yes	Yes	Yes	Yes
Application Year Fixed Effect	No	Yes	Yes	Yes
Subcategory Fixed Effects	No	No	No	Yes
R-squared	0.093	0.094	0.098	0.117

Robust standard errors in parentheses

** $p<0.01$, * $p<0.05$

NOTE. SEPARATE REGRESSIONS ARE REPORTED IN EACH COLUMN WITH ROBUST STANDARD ERRORS IN PARENTHESES. THE DEPENDENT VARIABLE IS FORWARD CITATIONS WHILE UNDER CONTROL OF ONE OF THE SET OF NPES. THE INDEPENDENT VARIABLE IS THE REVENUE OF EACH PATENT WHILE HELD BY THE NPE. THE DATA ARE NORMALIZED SO THAT THE MEAN ANNUAL REVENUE IS $30,000 (2010 DOLLARS).

ship in each category: the now-familiar inverted-U. The coefficients vary across technologies, which may result from differences in the use of strategic patenting as well as overall citation practices and patent values. Figures 7 and 8 show that the inverted-U relationship holds for software and computer architecture patents. While the pattern is noisier for each technology category individually, due to smaller number of observations, it is unmistakable both in the figures and the regression results.

3.5 Discussion

We have now verified several predictions of the model and seen that the inverted-U relationship between citations and value appears to be robust to various specifications. While the data under study here have a number of features that make it ideally suited for learning

Table 7: Forward Citations and Patent Value by Major Subcategory

Technology Subcategory	(1) Comm.	(2) Hardware	(3) Peripherals	(4) Info Storage	(5) Semiconductors	(6) Other
Revenue	11.33**	15.81**	5.731	19.26**	7.282**	12.49**
	(1.426)	(2.393)	(3.146)	(3.814)	(1.404)	(2.534)
Revenue Squared	-4.114**	-3.161**	-2.682	-7.881**	-2.741**	-3.640**
	(0.654)	(1.177)	(1.497)	(1.573)	(0.580)	(1.162)
Application Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
NPE-acquisition Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.100	0.148	0.107	0.129	0.082	0.071

Robust standard errors in parentheses
** p<0.01, * p<0.05

Note. Separate regressions reported in each column, with robust standard errors in parentheses. The dependent variable is forward citations while under control of one of the set of NPEs. The data are normalized so that the mean annual revenue is $30,000 (2010 dollars). The patent subcategories reflect those initially introduced in Hall, Jaffe and Trajtenberg (2001).

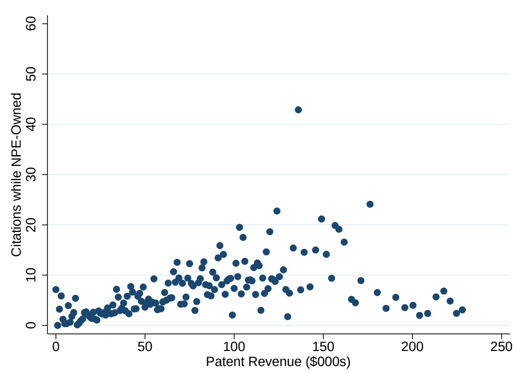


Figure 7: Forward Citations and Patent Value, Communication

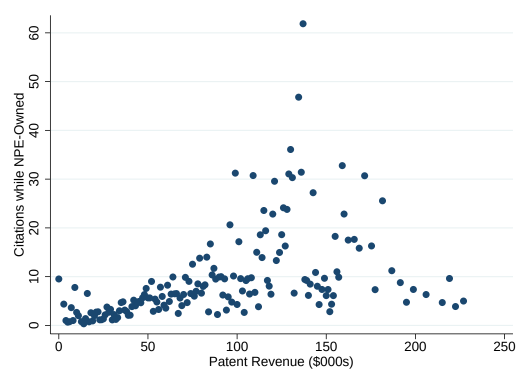


Figure 8: Forward Citations and Patent Value, Hardware & Software Patents

Note. The data are normalized so that the mean annual revenue is $30,000 (2010 dollars).

about this relationship, like all data sets, it has limitations. One concern may be that of representativeness - to what extent are the patents studied representative of the universe? Clearly, NPEs are in the business of selecting patents that they believe will be most valuable and thus selection may be a serious concern. That being said, of the tens of thousands of patents under study, less than 1% were specifically targeted for acquisition. Many were acquired as part of large portfolios and thus are closer to a random draw.

To further investigate the role selection might play in these data, we compare features of our data with the universe of patents in the same technology classes. Figure 9 displays the relationship between aggregate forward citations and age for our data and a comparable subset of the universe of patents, matched on age and technology category. While there is a slight divergence in the last five years of patent life, it is clear that the lifetime profile of citations is very similar between the patents held by our NPEs and the universe.

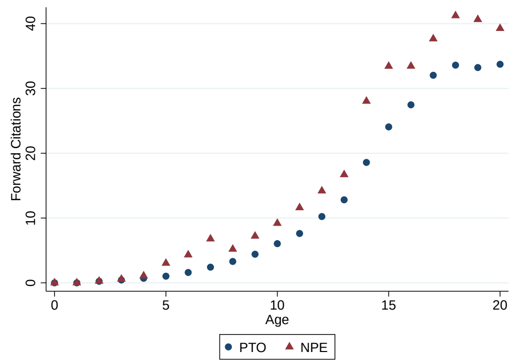


Figure 9: Citation Profiles of NPE Vs. USPTO

Note. To match the NPE data set, we use USPTO data through 2015 to calculate citations.

Still, the pool from which these patents were randomly drawn is particularly focused on technology patents and to that extent the results in the paper may not hold more broadly. That being said, we find the basic inverted-U relationship holds across technology categories within our data set (see e.g. Figures 7 and 8). Further, given that much of the value of innovation is concentrated in technology, even if it cannot be generalized to all patents, it is an important area of interest in itself.

A further concern may be about whether the model we put forth uniquely predicts the patterns we observe in the data. The basic inverted-U shape could no doubt be generated by a host of models of the innovative process. We attempt to address this concern in the previous section by testing further predictions of the model. We have seen that breaking up the data by parent status and level of innovative activity in a field bolster the view that the inverted-U is due to a combination of innovative and strategic patents.

Finally, concerns may remain that the revenues generated are specific to our set of NPEs and the way they use patents, and may not generalize to patents held by practicing entities. As discussed in Section 1 there are several reasons to believe the results presented here may be broader. The data under study come exclusively from licensing revenue, not litigation, and the licensing agreements are usually made with numerous licensees. Additionally, almost all of the patents were initially granted to practicing entities, either individuals or firms, and subsequently sold to one of the set of NPEs. These facts make it more likely that the revenues studied reflect real patent value.

4 Conclusion

Given that innovation and the value of innovation are difficult to measure in a consistent, systematic way, this has limited economists' ability to make reliable inferences about what policies are the best for productivity gains and ultimately economic growth. Given that patent records contain a wealth of information about inventions, patent counts and citation-weighted patent counts are popular proxies for the value of innovation. This paper brings the necessary data and theoretical model to external validate these existing proxies for valuing innovation. Using a new data set with tens of thousands of observations and licensing revenue from many customers, we show that the relationship between forward citations and patent value is an inverted-U. While we find evidence that the traditionally posited positive correlation between citations and value holds at the lower end of the value distribution, the full pattern is more complex. Once patent value exceeds a certain threshold, the citation-value relationship becomes negative. Taken together, this forms an inverted-U relationship between forward citations and patent value.

We explain this pattern in the data with a new theory of how citations accrue that allows for two types of innovative effort: productive and strategic. Productive innovations are more familiar. Innovators that make major, early contributions to a field, earn substantial profits and their patents are cited frequently by those who come after and make incremental, and less valuable improvements. This leads to a positive relationship between forward citations and patent value. In addition to this familiar type of patenting, we add a new type: strategic. Strategic patents have the property of reducing the likelihood that a firm's patents are improved upon by a competitor. This has the simultaneous effects of increasing the original patent value and also making it less likely to be cited. The incentive to invest in strategic patenting increases with patent value, which leads to a negative relationship between citations and value for strategic patents. When we allow for both types of innovation, we expect productive patenting to dominate up to a point, after which strategic patenting becomes more prevalent, which is why we observe the inverted-U relationship.

For studies focusing on relatively low value patents, the assumption that citation-weighting is a good proxy for patent value is a good first approximation. But analyses that focus on higher value patents, or the full range of patent value, or where there is no good indication of the likely value distribution, may not be able to use this simple proxy and obtain reliable results. By testing the reliability of the citation proxies for the value of innovation, we hope to broaden and deepen our understanding of the innovation process, with an eye ultimately towards informing policy decisions to better foster it. To that end, the model introduced here creates the potential to rigorously analyze specific innovation-related policy proposals.

If our understanding of the innovative process is correct, it will be able to guide decisions on questions such as broadening patent rights and increasing R&D subsidies. There is also an opportunity to learn a great deal more about the innovation process, by combining the data introduced here with further information about assignees, such as industry structure and concentration, corporate structure and history, and more.

References

Abrams, David S., and Bhaven Sampat. 2018. "Drug Patent Citations and Value." Working Paper.

Abrams, David S., Ufuk Akcigit, Gokhan Oz, and Jeremy Pearce. 2018. "The Patent Troll: Benign Middleman or Stick-Up Artist?" Working Paper.

Aghion, Philippe, and Xavier Jaravel. 2015. "Knowledge Spillovers, Innovation and Growth." *The Economic Journal*, 125: 533–573.

Aghion, Philippe, Ufuk Akcigit, and Peter Howitt. 2014. "What Do We Learn from Schumpeterian Growth Theory?" In *Handbook of Economic Growth*. Vol. 2, 515–563. Elsevier.

Akcigit, Ufuk, and William R Kerr. 2018. "Growth through heterogeneous innovations." *Journal of Political Economy*, 126(4): 1374–1443.

Allison, John R., Mark A. Lemley, Kimberly A. Moore, and R. Derek Trunkey. 2009. "Valuable Patents." *Georgetown Law Journal*, 92: 435–479.

Arora, Ashish, Marco Ceccagnoli, and Wesley M Cohen. 2008. "R&D and the Patent Premium." *International Journal of Industrial Organization*, 26(5): 1153–1179.

Bernstein, Shai. 2015. "Does Going Public Affect Innovation?" *The Journal of Finance*, 70(4): 1365–1403.

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

Bessen, James, and Michael James Meurer. 2008. *Patent Failure: How Judges, Bureaucrats, and Lawyers Put Innovators at Risk.* Princeton University Press.

Bessen, James, and Michael J. Meurer. 2014. "The Direct Costs from NPE Disputes." *Cornell Law Review*, , (99): 387–424.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen. 2013. "Identifying Technology Spillovers and Product Market Rivalry." *Econometrica*, 81(4): 1347–1393.

Boldrin, Michele, and David K. Levine. 2013. "The Case Against Patents." *The Journal of Economic Perspectives*, 27(1): 3–22.

Bresnahan, Timothy F., and M. Trajtenberg. 1995. "General purpose technologies Engines of growth?" *Journal of Econometrics*, 65(1): 83 – 108.

Budish, Eric, Benjamin Roin, and Heidi Williams. 2015. "Do Firms Underinvest in Long-Term Research? Evidence from Cancer Clinical Trials." *American Economic Review*, , (105): 2044–2085.

Caballero, Ricardo J., and Adam B. Jaffe. 1993. "How High are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth." *NBER Macroeconomics Annual 1993, Volume 8*, 15–86. MIT Press.

Cockburn, Iain M, and Megan J MacGarvie. 2011. "Entry and Patenting in the Software Industry." *Management Science*, 57(5): 915–933.

Cohen, Lauren, Umit Gurun, and Scott D. Kominers. 2018. "Patent Trolls: Evidence from Targeted Firms." *Management Science*. forthcoming.

Cohen, Wesley M., and Daniel A. Levinthal. 1989. "Innovation and Learning: The Two Faces of R & D." *The Economic Journal*, 99(397): 569–596.

Farrell, Joseph, and Carl Shapiro. 2008. "How Strong Are Weak Patents?" *American Economic Review*, 98(4): 1347–1369.

Feng, Joshua, and Xavier Jaravel. 2018. "Crafting Intellectual Property Rights: Implications for Patent Assertion Entities, Litigation and Innovation." unpublished manuscript.

FTC. 2016. "Patent Assertion Entity Activity: An FTC Study." Federal Trade Commission.

Galasso, Alberto, and Mark Schankerman. 2010. "Patent Thickets, Courts, and the Market for Innovation." *RAND Journal of Economics*, 41(3): 472–503.

Galasso, Alberto, and Mark Schankerman. 2015. "Patents and Cumulative Innovation: Causal Evidence from the Courts." *Quarterly Journal of Economics*, 130(1): 317–369.

Gallini, Nancy T. 1992. "Patent Policy and Costly Imitation." *The RAND Journal of Economics*, 23(1): 52–63.

Geroski, P.A., J. Van Reenen, and C.F. Walters. 1997. "How persistently do firms innovate?" *Research Policy*, 26(1): 33–48.

Green, Jerry R., and Suzanne Scotchmer. 1995. "On the Division of Profit in Sequential Innovation." *The RAND Journal of Economics*, 26(1): 20–33.

Griliches, Zvi. 1981. "Market Value, R&D, and Patents." *Economics Letters*, 7(2): 183–187.

Hall, Bronwyn H, Adam B Jaffe, and Manuel Trajtenberg. 2001. "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools." National Bureau of Economic Research Working Paper 8498.

Hall, Bronwyn H, Adam Jaffe, and Manuel Trajtenberg. 2005. "Market Value and Patent Citations." *RAND Journal of Economics*, 36(1): 16–38.

Hall, Bronwyn H, and Dietmar Harhoff. 2012. "Recent Research on the Economics of Patents." *Annual Review of Economics*, 4(1): 541–565.

Hall, Bronwyn H, and Rosemarie Ham Ziedonis. 2001. "The Patent Paradox Revisited: An Empirical Study of Patenting in the US Semiconductor Industry, 1979–1995." *RAND Journal of Economics*, 32(1): 101–128.

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

Harhoff, Dietmar, Frederic M Scherer, and Katrin Vopel. 2003. "Citations, family size, opposition and the value of patent rights." *Research Policy*, 32(8): 1343 – 1363.

Hegde, Deepak, David C Mowery, and Stuart J H Graham. 2009. "Pioneering Inventors or Thicket Builders: Which US Firms Use Continuations in Patenting?" *Management Science*, 55(7): 1214–1226.

Jaffe, Adam B, and Joshua Lerner. 2004. *Innovation and Its Discontents: How Our Broken Patent System is Endangering Innovation and Progress, and What To Do About It.* Princeton University Press.

Jaravel, Xavier. 2018. "The Unequal Gains from Product Innovations: Evidence from the U.S. Retail Sector." *Quarterly Journal of Economics*, forthcoming.

Jaravel, Xavier, Neviana Petkova, and Alex Bell. 2018. "Team-Specific Capital and Innovation." *American Economic Review*, 108(4-5): 1034–1073.

Kerr, William R. 2010. "Breakthrough inventions and migrating clusters of innovation." *Journal of Urban Economics*, 67(1): 46 – 60. Special Issue: Cities and Entrepreneurship.

Kesan, Jay P., and David L. Schwartz. 2014. "Analyzing the Role of Non-Practicing Entities in the Patent System." *Cornell Law Review*, , (99): 425–456.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman. 2017. "Technological Innovation, Resource Allocation, and Growth*." *The Quarterly Journal of Economics*, 132(2): 665–712.

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

Lerner, Josh. 1995. "Patenting in the Shadow of Competitors." *The Journal of Law & Economics*, 38(2): 463–495.

Lerner, Josh, and Amit Seru. 2017. "The Use and Misuse of Patent Data: Issues for Corporate Finance and Beyond." National Bureau of Economic Research Working Paper 24053.

Lerner, Joshua, and Jean Tirole. 2002. "Some Simple Economics of Open Source." *Journal of Industrial Economics*, 50(2): 197–234.

Lerner, Joshua, and Jean Tirole. 2005. "The Scope of Open Source Licensing." *Journal of Law, Economics, and Organization*, 21(1): 20–56.

Levin, Richard C, Alvin K Klevorick, Richard R Nelson, Sidney G Winter, Richard Gilbert, and Zvi Griliches. 1987. "Appropriating the Returns from Industrial Research and Development." *Brookings Papers on Economic Activity*, 18(3): 783–831.

Moser, Petra, Joerg Ohmstedt, and Paul W. Rhode. 2018. "Patent CitationsAn Analysis of Quality Differences and Citing Practices in Hybrid Corn." *Management Science*, 64(4): 1926–1940.

Nicholas, Tom. 2013. "Are Patents Creative or Destructive?" Harvard Business School Working Paper.

Noel, Michael, and Mark Schankerman. 2013. "Strategic Patenting and Software Innovation." *Journal of Industrial Economics*, 61(3): 481–520.

O'Donoghue, Ted, Suzanne Scotchmer, and Jacques-Franois Thisse. 1998. "Patent Breadth, Patent Life, and the Pace of Technological Progress." *Journal of Economics & Management Strategy*, 7(1): 1–32.

Pakes, Ariel. 1986. "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks." *Econometrica*, 54(4): 755–784.

Sampat, Bhaven, and Heidi Williams. 2018. "How Do Patents Affect Follow-On Innovation? Evidence from the Human Genome." *American Economic Review*.

Schankerman, Mark, and Ariel Pakes. 1986. "Estimates of the Value of Patent Rights in European Countries During the Post-1950 Period." *Economic Journal*, 96(384): 1052–1076.

Scherer, Frederic M. 1967. "Market Structure and the Employment of Scientists and Engineers." *American Economic Review*, 57(3): 524–531.

Serrano, Carlos J. 2010. "The Dynamics of the Transfer and Renewal of Patents." *RAND Journal of Economics*, 41(4): 686–708.

Shapiro, Carl. 2012. "Patent Assertion Entities: Effective Monetizers, Tax on Innovation, or Both?" *Transcript from United States Department of Justice, Antitrust Division Workshop on Patent Assertion Entities.*

Trajtenberg, Manuel. 1990. "A Penny for Your Quotes: Patent Citations and the Value of Innovations." *RAND Journal of Economics*, 21(1): 172–187.

Von Graevenitz, Georg, Stefan Wagner, and Dietmar Harhoff. 2013. "Incidence and Growth of Patent Thickets-The Impact of Technological Opportunities and Complexity." *Journal of Industrial Economics*, 61(3): 521–563.

Williams, Heidi L. 2013. "Intellectual Property Rights and Innovation: Evidence from the Human Genome." *Journal of Political Economy*, 121(1): 1–27.

Williams, Heidi L. 2017. "How Do Patents Affect Research Investments?" *Annual Review of Economics*, 9(1): 441–469.

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

Online Appendix

A Variable Normalization

Since the major focus of this paper is better understanding the relationship between patent value and citations, it is important to clearly define how these values are calculated. Doing so requires some understanding of the business model of the NPEs from which the data were acquired. The NPEs acquire patents either by purchasing them from patent assignees or entering into revenue-sharing agreements with them. The patent portfolios generate revenue through licensing agreements which may be on an entire portfolio or a subset thereof.

Revenue is allocated on a patent-year-licensee level based on the prominence the patent played in negotiations with the licensee. Those patents that were most heavily focused upon in licensing negotiations are placed in category 1, which is allocated the largest revenue share. All patents within category 1 are given equal revenue for a particular licensee. In an analogous way, for each licensing deal, patents are also assigned to categories 2, 3, and 4. The categories denote declining relevance to the particular licensing deal and also declining revenue share. Each patent in the same category for a deal receives the same revenue allocation. Given that the firms licensing the patents operate in a variety of industries, a patent that is assigned a category 3 or 4 for one licensing agreement can be assigned a category 1 or 2 for licensing agreement.

While there is certain to be imprecision in revenue assignment, this allocation scheme is disciplined by competing interests on two sides. Patent owners who are due a share of future revenues seek to maximize the revenue allocated, while the incentive of shareholders in the NPEs is for larger revenue allocation to patents in which they have a stake and less to others, since total revenue is fixed.

In our baseline specification for revenue and citations, we use simple measures. For revenue, we aggregate revenue across many licensing agreements to get patent-year revenue and then we use the consumer price index (CPI) to deflate the annual revenues per patent. Finally, we sum these revenues for each patent. We normalize all revenue amounts so that mean annual revenue is $30,000 (2010 dollars) in order to fulfill the terms of our confidentiality agreement. For forward citations, we use forward citations while the NPEs hold the patent. This allows for a comparison group for our revenue measure.

While these baseline measures are intuitive and transparent, more complicated procedures have been proposed in the literature. Given that citations and revenue while held by the

NPEs are inherently truncated and for more recent patents, the problem is more severe, in the appendix we consider alternatives. Figure 10 presents the incremental patent citation-age profile as well as the revenue-age profile in aggregate. To account for the changing distribution over time, we introduce lifetime adjustments. Another important missing factor from the simple calculation is technology subcategory. Figure 11 shows substantial variation in patenting rates by technology class; therefore, we consider such adjustments. In the appendix, we consider two alternative adjustments for lifetime revenue. We estimate both a uniform revenue adjustment (i.e. multiply patents by the inverse of the proportion of years held by the NPEs), and a non-parametric model that estimates lifetime revenue in similar fashion to (Hall, Jaffe and Trajtenberg, 2005) for citations. Specifically, we calculate the coefficients from the shape of the citation (or revenue)-lag distribution in each major technology subcategory using all U.S. utility patents (or all NPE patents) with an application year between 1987 and 2014. Then, we use these coefficient estimates to generate lifetime citations and revenue. While this procedure will understate the number of lifetime citations for any patent that has zero in our data set, the mean number of lifetime cites should still be correct. If anything, this would lead us to find an excess of high value, zero citation patents, which is something we do not observe.

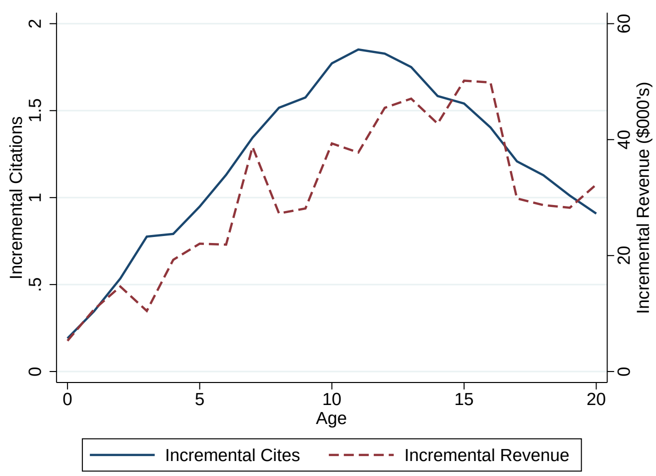


Figure 10: Incremental Forward Citations and Revenue By Patent Age
Notes: The data are normalized so that the mean annual revenue is $30,000 (2010 dollars).

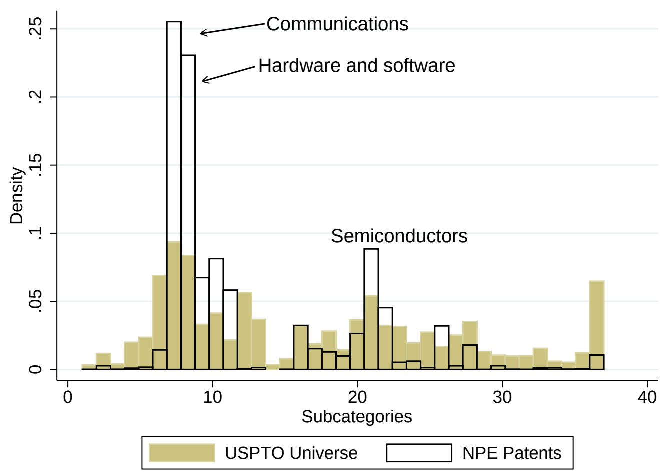


FIGURE 11: PATENT DISTRIBUTION ACROSS TECHNOLOGY SUBCATEGORIES, NPE VS. USPTO

B Robustness Checks

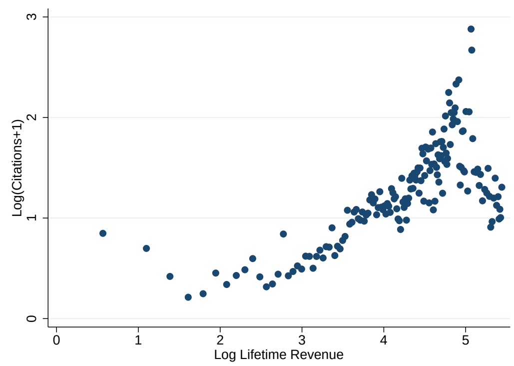


FIGURE 12: FORWARD CITATIONS AND PATENT VALUE (LOG TRANFORMATION)

NOTE. THE DATA ARE NORMALIZED SO THAT THE MEAN ANNUAL REVENUE IS $30,000 (2010 DOLLARS).

TABLE 8: FORWARD CITATIONS AND PATENT VALUE (LOG OF CITATIONS)

	(1)	(2)	(3)	(4)
	Main Sample	Main Sample	Strategic Range	Main Sample
Range of Valuable Patents Included	0–95th	0–95th	75th–95th	0–95th
Revenue ($100,000s)	2.056**	0.899**	-0.186**	0.408**
	(0.0422)	(0.0580)	(0.0438)	(0.0450)
Revenue Squared	-0.702**	-0.323**		
	(0.0222)	(0.0263)		
Strategic Range				0.783**
				(0.0626)
Revenue x Strategic Range				-0.619**
				(0.0606)
Individual Inventor		0.309**	0.456**	0.305**
		(0.0360)	(0.0713)	(0.0359)
Parent Patent		0.285**	0.260**	0.282**
		(0.0253)	(0.0555)	(0.0252)
Fraction of Backward Citations in Last 3 Years		0.121**	0.140*	0.120**
		(0.0251)	(0.0660)	(0.0250)
Self-Citations		0.764**	0.826**	0.770**
		(0.0123)	(0.0192)	(0.0124)
Claims		0.0102**	0.00728**	0.0102**
		(0.000791)	(0.00167)	(0.000794)
NPE-acquisition Year Fixed Effect	No	Yes	Yes	Yes
Application Year Fixed Effect	No	Yes	Yes	Yes
Subcategory Fixed Effects	No	Yes	Yes	Yes
R-squared	0.144	0.305	0.322	0.306

Robust standard errors in parentheses
** p<0.01, * p<0.05

NOTES: SEPARATE REGRESSIONS ARE REPORTED IN EACH COLUMN WITH ROBUST STANDARD ERRORS IN PARENTHESES. THE DEPENDENT VARIABLE IS LOG (1 + FORWARD CITATIONS WHILE UNDER CONTROL OF ONE OF THE SET OF NPEs). THE DATA ARE NORMALIZED SO THAT THE MEAN ANNUAL REVENUE IS $30,000 (2010 DOLLARS).

TABLE 9: FORWARD CITATIONS AND PATENT VALUE (POISSON REGRESSION)

	(1)	(2)	(3)	(4)
	Main Sample	Main Sample	Strategic Range	Main Sample
Range of Valuable Patents Included	0–95th	0–95th	75th–95th	0–95th
Revenue ($100,000s)	3.583**	1.707**	-0.217**	0.573**
	(0.127)	(0.194)	(0.0778)	(0.155)
Revenue Squared	-1.283**	-0.620**		
	(0.0546)	(0.0795)		
Strategic Range				1.025**
				(0.137)
Revenue x Strategic Range				-0.768**
				(0.175)
Individual Inventor		0.341**	0.357**	0.344**
		(0.0684)	(0.0910)	(0.0681)
Parent Patent		0.413**	0.299**	0.407**
		(0.0604)	(0.102)	(0.0600)
Fraction of Backward Citations in Last 3 Years		0.0351	0.0449	0.0382
		(0.0602)	(0.0953)	(0.0605)
Self-Citations		0.0492**	0.0467**	0.0501**
		(0.00396)	(0.00409)	(0.00402)
Claims		0.0129**	0.0106**	0.0132**
		(0.00151)	(0.00283)	(0.00151)
NPE-acquisition Year Fixed Effect	No	Yes	Yes	Yes
Application Year Fixed Effect	No	Yes	Yes	Yes
Subcategory Fixed Effects	No	Yes	Yes	Yes

Robust standard errors in parentheses
** p<0.01, * p<0.05

NOTE. SEPARATE REGRESSIONS ARE REPORTED IN EACH COLUMN WITH ROBUST STANDARD ERRORS IN PARENTHESES. THE DEPENDENT VARIABLE IS FORWARD CITATIONS WHILE UNDER CONTROL OF NPE. THE DATA ARE NORMALIZED SO THAT THE MEAN ANNUAL REVENUE IS $30,000 (2010 DOLLARS).

Table 10: Forward Citations and Patent Value (1% Cutoff)

	(1)	(2)	(3)	(4)	(5)
Share of Most Valuable Patents Excluded	1%	1%	1%	1%	1%
Revenue ($100,000s)	3.725**	3.682**	3.320**	3.182**	2.767**
	(0.339)	(0.340)	(0.342)	(0.349)	(0.340)
Revenue Square	-0.113**	-0.111**	-0.0988**	-0.0946**	-0.0819**
	(0.0112)	(0.0112)	(0.0111)	(0.0112)	(0.0109)
Individual Inventor	4.131**	4.008**	4.295**	3.798**	3.115**
	(0.846)	(0.849)	(0.858)	(0.859)	(0.836)
Parent Patent		3.331**	3.440**	3.267**	2.673**
		(0.587)	(0.602)	(0.590)	(0.576)
Fraction of Backward Citations in Last 3 Year			1.016*	1.008*	1.174**
			(0.397)	(0.396)	(0.394)
Self-Citations			1.348**	1.322**	1.193**
			(0.0564)	(0.0574)	(0.0601)
Claims			0.134**	0.105**	0.0710**
			(0.0157)	(0.0159)	(0.0153)
Originality					1.034**
					(0.356)
Generality					13.46**
					(0.409)
NPE-acquisition Year Fixed Effect	Yes	Yes	Yes	Yes	Yes
Application Year Fixed Effect	Yes	Yes	Yes	Yes	Yes
Subcategory Fixed Effects	No	No	No	Yes	Yes
R-squared	0.104	0.107	0.150	0.166	0.205

Robust standard errors in parentheses

** $p<0.01$, * $p<0.05$

Note. Separate regressions are reported in each column with robust standard errors in parentheses. The dependent variable is forward citations while under control of NPE. The data are normalized so that the mean annual revenue is $30,000 (2010 dollars).

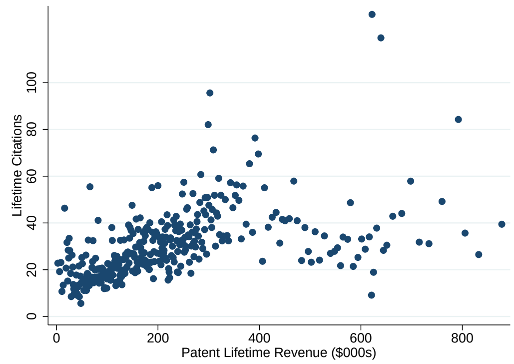


Figure 13: Lifetime Forward Citations and Patent Value)

Note. The data are normalized so that the mean annual revenue is $30,000 (2010 dollars). Lifetime forward citations are calculated using Hall, Jaffe and Trajtenberg (2005) coefficients estimated from the shape of the citation-lag distribution in each major technology subcategory using all U.S. utility patents with an application year between 1987 and 2014, which covers the scope of our data set. Lifetime revenue is adjusted in a uniform fashion such that if the NPEs hold the patent for X years garnering revenue R, lifetime revenue is $R * \frac{20}{X}$

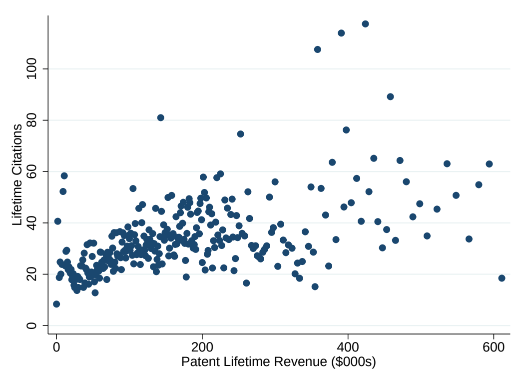


Figure 14: Lifetime Forward Citations and Patent Value

Note. The data are normalized so that the mean annual revenue is $30,000 (2010 dollars). Lifetime forward citations are calculated using Hall, Jaffe and Trajtenberg (2005) coefficients estimated from the shape of the citation-lag distribution in each major technology subcategory using all U.S. utility patents with an application year between 1987 and 2014, which covers the scope of our data set. Lifetime revenue is also adjusted in a similar manner.

Table 11: Lifetime Forward Citations and Patent Value

	(1)	(2)	(3)	(4)
	Main Sample	Main Sample	Strategic Range	Main Sample
Range of Valuable Patents Included	0–95th	0–95th	75th–95th	0–95th
Revenue ($100,000)	7.626**	3.272**	-1.994**	6.145**
	(0.425)	(0.473)	(0.587)	(0.651)
Revenue Squared	-5.027**	-2.529**		
	(0.687)	(0.780)		
Strategic Range				32.49**
				(2.977)
Revenue x Strategic Range				-9.372**
				(0.859)
Individual Inventor		14.23**	16.90**	13.96**
		(2.248)	(4.358)	(2.232)
Parent Patent		13.45**	16.04**	13.33**
		(1.849)	(4.805)	(1.841)
Fraction of Backward Citations in Last 3 Years		6.797**	7.789*	6.243**
		(1.298)	(3.282)	(1.286)
Self-Citations		2.454**	2.495**	2.555**
		(0.0928)	(0.0869)	(0.0929)
Claims		0.411**	0.420**	0.382**
		(0.0472)	(0.127)	(0.0471)
NPE-acquisition Year Fixed Effect	No	Yes	Yes	Yes
Application Year Fixed Effect	No	Yes	Yes	Yes
Subcategory Fixed Effects	No	Yes	Yes	Yes
R-squared	0.03	0.19	0.27	0.20

Robust standard errors in parentheses
** p<0.01, * p<0.05

Note. Separate regressions are reported in each column with robust standard errors in parentheses. Lifetime forward citations are calculated using Hall, Jaffe and Trajtenberg (2005) coefficients estimated from the shape of the citation-lag distribution in each major technology subcategory using all U.S. utility patents with an application year between 1987 and 2014, which covers the scope of our data set. Lifetime revenue is adjusted in a uniform fashion such that if the NPEs hold the patent for X years garnering revenue R, lifetime revenue is $R * \frac{20}{X}$. "Strategic range" is defined as the 75th–95th percentile of the revenue distribution given the lifetime revenue calculated by uniformly inflating the revenue to match the time the patent held by the NPEs. The data are normalized so that the mean annual revenue is $30,000 (2010 dollars).

Table 12: Forward Citations and Patent Value with Alternative Strategic Proxy

	(1)	(2)	(3)	(4)
Share of Most Valuable Patents Excluded	5%	5%	5%	5%
Strategic Proxy	-4.335**	-4.312**	-4.623**	-4.288**
	(0.555)	(0.554)	(0.562)	(0.557)
Revenue	8.939**	8.723**	7.691**	5.842**
	(1.170)	(1.172)	(1.187)	(1.245)
Strategic Proxy x Revenue	14.16**	14.29**	15.06**	14.60**
	(1.942)	(1.940)	(1.950)	(1.932)
Revenue Squared	-2.529**	-2.427**	-1.998**	-1.121*
	(0.507)	(0.508)	(0.514)	(0.540)
Strategic Proxy x Revenue Squared		-5.989**	-6.347**	-6.377**
	(0.947)	(0.947)	(0.947)	(0.942)
Individual Inventor		3.108**	2.861**	2.374**
		(0.582)	(0.584)	(0.580)
Claims			0.128**	0.101**
			(0.0159)	(0.0160)
NPE-acquisition Year Fixed Effect	Yes	Yes	Yes	Yes
Application Year Fixed Effect	No	Yes	Yes	Yes
Subcategory Fixed Effects	No	No	No	Yes
R-squared	0.090	0.091	0.096	0.115

Robust standard errors in parentheses

** p<0.01, * p<0.05

Note. Separate regressions arereported in each column with robust standard errors in parentheses. The dependent variable is forward citations while under control of NPE. The independent variable is the revenue of each patent during the time it is held by the NPEs. The data are normalized so that the mean annual revenue is $30,000 (2010 dollars).

C Closing the model: Household Problem

In this section, we close our model by solving the household's maximization problem. The representative household consists of a fixed measure of 1 production workers, each of which supplies one unit of labor inelastically. The household holds a balanced portfolio of assets of all the firms in the economy $\mathcal{A}_t$, earns $r_t\mathcal{A}_t$ from it, collects the labor income w_t and chooses consumption C_t to maximize the following lifetime utility

$$U = \int_0^\infty e^{-\rho t} \ln C_t dt$$

subject to the following budget constraint

$$w_t + r_t\mathcal{A}_t = C_t + \dot{\mathcal{A}}_t.$$

Note that the household discounts the future at the rate $\rho > 0$. The household's intertemporal maximization yields the standard Euler equation

$$g_t = r_t - \rho. \tag{13}$$

Finally, the resource constraint of the economy is expressed as

$$Y_t = C_t + R_t. \tag{14}$$

This expression says that the final good produced in the economy (Y_t) is used for household consumption (C_t) and R&D expenses (R_t).

D Social Value and Patent Citations

We now analyze the relationship between social value of a patent and its forward citation counts. We show that our model predicts a positive relationship between the social value of patents and forward citations, as documented empirically by Trajtenberg (1990).

To this end, we express the social value function for each innovation, which requires us to define the household's welfare. Since the welfare analysis is lengthy and technically involved, we shorten it by making the following assumptions without affecting the main trade-offs in the model.

1. Since the main surprising outcome was due to strategic patents which led to a negative relationship between private value and forward citations (Section 2.2), we focus only

on the relationship between the social value of strategic patents and their forward citations.[19]

2. Recall that the impact of the height of the incumbent's fence, m, in Section 2.2 was to reduce the subsequent entry rate ($\partial \bar{z}_{0m}/\partial m < 0$ and $\partial \bar{z}_{nm}/\partial m < 0$), which was increasing in the expected duration of the monopoly power of the incumbent. We replicate the same dynamics of the free entry condition by assuming that a measure 1 of entrants pays the same fixed cost ζQ_t and enters with a *radical* innovation at the rate $\bar{z}_{0m}$ or a follow-on innovation at the rate $\bar{z}_{1m}$, which are both decreasing functions of the height of the fence of the incumbent (m) as in the model of Section 2.2. Finally, we assume $\alpha \to 1$ such that all innovations generate the same return with η innovation size.

These assumptions keep the welfare analysis brief and notationally clean while preserving the backbone of the model. Now we are ready to go into the technical details.

Let us start with the household welfare, which is simply the discounted sum of future utilities

$$\text{Welfare} = \int_0^\infty e^{-\rho t} \ln C_t dt. \tag{15}$$

The resource constraint of the economy (14) describes how total output is allocated:

$$Y_t = C_t + 2\zeta Q_t + \bar{z}_0 Q_t \psi$$

where C_t is household consumption, $2\zeta Q_t$ is the total R&D cost by radical and follow-on innovators, and $\bar{z}_0 Q \psi$ is the flow cost of strategic innovation by incumbents. We can therefore express the household consumption as

$$C_t = (1 - 2\zeta - \bar{z}_0 \psi)\, Y_t$$

where the term in parenthesis is simply a constant. Using this expression for C_t, we can rewrite welfare in (15) as

$$\text{Welfare} = \frac{\ln(1 - 2\zeta - \psi \bar{z}_0)}{\rho} + \int_0^\infty e^{-\rho t} \ln Y_t dt.$$

This expression indicates that household welfare is determined by two factors. The first factor is the propensity to consume, $(1 - 2\zeta - \psi \bar{z}_0)$, which governs the fraction of output

[19] The positive relationship between social value and patent citations is trivially satisfied under productive innovations.

that is consumed at every instant. The second -and for our purpose more important- factor is the total output component of welfare (Y_t) , which grows through innovations in our model. We can now express welfare as a sum of product line–specific social values W_{jt} :

$$\text{Welfare} = \frac{\ln(1-2\zeta-\psi\bar{z}_0)}{\rho} + \int_0^1 W_{jt}dj$$

where

$$W_{jt} \equiv \int_0^\infty e^{-\rho t}\ln y_{jt}dt.$$

Recall that the output of variety j in equilibrium is $y_{jt} = \frac{Y_t q_{jt}}{(1+\eta_n)w_t}$, the aggregate output is $Y_t = Q_t$ and the wage rate is $w_t = Q_t \int_0^1 (1-\eta_j)^{-1}$, as shown in Lemma 1. Since we restrict our attention to $\alpha \to 1$, each product line will have the same markup, $(\eta_n = \eta)$, and thus $w_t = \frac{Q_t}{1-\eta}$. We may then write equilibrium production in j as $y_{jt} = \frac{(1-\eta)}{(1+\eta)}q_{jt}$.

We are now ready to express the social value function $W_{jt} = W(q_{jt}, m_{jt})$ as a function of the quality level of the product line, (q_{jt}) , and the strength of the defensive patent, (m_{jt}) , which will determine subsequent entry. Henceforth, we suppress the product line index j and the time index t on q and m when it causes no confusion. Then the social value function of each product line as a function of its quality level and the strength of the current strategic patent $(W(q,m))$ is equal to:

$$W(q,m) = \Delta t \ln\frac{(1-\eta)}{(1+\eta)}q + (1-\rho\Delta t)\left[\begin{array}{c}(\bar{z}_{0m}+\bar{z}_{1m})\,\Delta t\mathbb{E}_m W(q(1+\eta),m) \\ +(1-\bar{z}_{0m}\Delta t-\bar{z}_{1m}\Delta t)\,W(q,m)\end{array}\right]$$

The intuition is as follows. During any small time interval Δt, the household generates a flow utility of $\Delta t \ln \frac{(1-\eta)}{(1+\eta)}q$ for its consumption which is a function of the quality level q. After a small time interval Δt three continuation events can happen: First, there could be a new radical innovation with probability $\bar{z}_{0m}\Delta t$, second there could be a follow-on innovation with probability $\bar{z}_{1m}\Delta t$ (in both cases, productivity improves by a factor of $1+\eta$), finally there could be no new innovation with the remaining probability. Note that all new successful entrants will also produce a subsequent strategic innovation and draw m (hence the expected term, $\mathbb{E}_m W$). Clearly the forward-looking social value function takes all these contingencies into account. The following lemma provides the exact form of this social value function.

Lemma 3 *The form of the social value function is*

$$W(q,m) = A\ln q + B_m + C$$

where

$$A \equiv 1/\rho, \qquad C \equiv \frac{1}{\rho}\ln\frac{(1-\eta)}{(1+\eta)}, \quad \textit{and} \quad B_m \equiv \frac{\bar{z}_{0m}+\bar{z}_{1m}}{\rho+\bar{z}_{0m}+\bar{z}_{1m}}\left[A\ln(1+\eta)+\mathbb{E}_m B_m\right]$$

Proof. See Appendix F. ■

Since the entry rate is decreasing in the size of the strategic innovation m, the following proposition follows from Lemma 3.

Proposition 3 *The social value of an innovation is decreasing in m.*

Proof. This follows directly from the fact that B_m and therefore $W(q,m)$ is decreasing in m. ■

The intuition behind this proposition is straightforward. When the strength of the strategic patent is bigger (large m), it leads to less entry as described in Proposition 2. Clearly this is detrimental for social welfare since there is less technological progress and consumption growth due to the blocked entry. Since forward citations are a result of subsequent entry, the following corollary follows.

Theorem 3 *In the case of strategic patents, social patent value and forward citations are positively correlated.*

Note that our model predicts a monotonic positive relationship between the social value of a strategic patent and its forward citations, whereas the relationship was negative between the private value and its forward citations in Section 2.2. This shows that our model is in line with Trajtenberg (1990)'s finding. In addition, our model also predicts a new result which has been vocally raised during recent economic debates. Our model generates a stark negative relationship between private value and citations in the case of strategic patenting. This is because social value and private incentives are at odds when it comes to competition generated through creative destruction. An incumbent raises her patent's private patent value by producing a strategic innovation while this behavior is detrimental to overall welfare.

E NPEs and the Opportunistic Use of Patents

In this section, we extend our model to consider an alternative scenario where instead of incumbents engaging in strategic patenting to defend existing technologies, an NPE uses an existing patent (which is not necessarily used for production) to extract rents from other firms until the whole cluster becomes obsolete through a radical innovation. We show that this alternative strategic use of a patent by the NPE generates a negative relationship between private patent value and citations (which would then generate an overall inverted-U if productive patents are also included). Thus our main conclusion, that strategic patents lead to a negative relationship between private patent value and citations, is strengthened by this model.

In this new setup, we assume that in each technology cluster, there is a single opportunistic patent that is owned by an NPE. This opportunistic patent is used to extract a fraction $m \in (0,1)$ of the instantenous profits from the producing incumbent in the same technology cluster due to the threat of a patent infringement. Claim m is determined once a new technology cluster is initiated by a radical innovation and is drawn from a distribution $\mathcal{M}([0,1])$. Therefore the incumbent has to pay a licensing fee $P_t^{fee}(m)$ to the NPE at every instant until the technology cluster becomes obsolete through radical innovations, which arrive at the rate $\bar{z}_0$.

The licensing fee paid to the NPE for its opportunistic patent of strength m is determined simply

$$P_t^{fee}(m) = m\pi_n Y_t \tag{16}$$

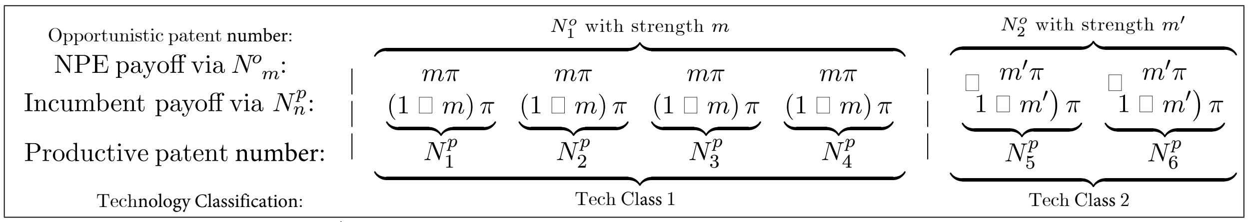
where $m \in (0,1)$ is the strength of the NPE's opportunistic patent and n is the rank of the innovation in the technology cluster. As in the previous extension, we simplify our analysis by considering the case where $\alpha \to 1$ so that all the profits in the economy are equal, i.e,

$$\begin{aligned} \pi_n &= \frac{\eta}{1+\eta} \\ &\equiv \pi \end{aligned}$$

and all follow-on entrants choose the same entry rate

$$\bar{z}_{nm} = \bar{z}_{1m}. \tag{17}$$

Example 2 *The following example illustrates these dynamics. Each technology group is associated with an opportunistic patent that is unkown ex-ante. Technology group 1 is initiated by a radical innovation N_1^p. Upon entry, the radical entrant learns about the existence of an opportunistic patent (which is a realization of $m \sim \mathcal{M}([0,1])$) and the entrant learns that it is infringing an opportunistic patent N_1^o that is owned by an NPE. Note that this setup is general enough to incorporate the possibility that there might be no opportunistic patent in this group ($m = 0$). Once m is realized, the incumbent pays a fraction m of its per-period return to the NPE as a licensing fee and keeps $1 - m$ to herself. Next, through free-entry,* *a new productive follow-on entrant replaces the incumbent with a new patent N_2^p. This time the NPE collects the fee from N_2^p (and not anymore from N_1^p). Later in the game, there are two more successful follow on entries labeled as N_3^p and N_4^p. The NPE keeps collecting these fees until the technology group is made obsolete by a radical entrant labeled N_5^p. In the new technology grouping (technology classification 2), an opportunistic patent N_2^o is drawn with strength $m' \sim \mathcal{M}([0,1])$, and so on.*



An example of a sequence of innovations in a product line

Now we go back to the solving the model. Let us denote the value of the infringing productive patent to the incumbent by $V_t^p(m)$ and the value of the infringed opportunistic patent to the NPE by $V_t^o(m)$. Then the value of an incumbent productive patent is

$$V_t^p(m) = (1-m)\pi Y_t \Delta t + (1 - r\Delta t)\left[\begin{array}{c} (\bar{z}_0 + \bar{z}_{1m})\Delta t \times 0 \\ + (1 - \bar{z}_0 \Delta t - \bar{z}_{1m}\Delta t) V_{t+\Delta t}^p(m) \end{array}\right].$$

Since our focus is on the balanced growth path equilibrium where aggregate variables grow at a constant rate and the entry rates depend only on the payoff relevent variable m, simple algebra leads to the following patent value:

$$V_t^p(m) = \frac{(1-m)\pi Y_t}{\bar{z}_{1m} + \bar{z}_0 + \rho}.$$

Finally, from the free entry condition for follow-on inventors we have

$$\max_{z_{1m}} \left\{ z_{1m} V_t^p(m) - z_{1m} \zeta Q_t \bar{z}_{1m} \right\},$$

and for radical inventors

$$\max_{z_0} \left\{ z_0 \mathbb{E} V_t^p(m) - z_0 \zeta Q_t \bar{z}_0 \right\}.$$

Note that since the radical innovator is starting a new technology grouping, he does not know the existence of an opportunistic patent and therefore forms an expectation over it. The equilibrium entry rates are simply

$$\bar{z}_{1m} = \frac{(1-m)\pi}{(\bar{z}_{1m} + \bar{z}_0 + \rho)\zeta} \tag{18}$$

and

$$\bar{z}_0 = \mathbb{E}\left[\frac{(1-m)\pi}{(\bar{z}_{1m} + \bar{z}_0 + \rho)\zeta}\right].$$

where the latter is independent of m. Now we solve (18) for the equilibrium follow-on entry rate

$$\bar{z}_{1m} = \frac{-(\bar{z}_0 + \rho) + \sqrt{(\bar{z}_0 + \rho)^2 + 4(1-m)\pi/\zeta}}{2}$$

which clearly decreases in the strength of the opportunistic patent m

$$\frac{\partial \bar{z}_{1m}}{\partial m} < 0.$$

Now we have the first part of our main result.

Proposition 4 *The entry rate (forward citations) is decreasing in m.*

This is due to the fact that when an NPE extracts a larger fraction of value in a technology grouping, new entrants account for this and invest less in follow-on innovations.

We may now solve for the value of the opportunistic patent $V_t^o(m)$:

$$\rho V_t^o(m) = P_t^{fee}(m) + \bar{z}_0 \left[0 - V_t^o(m)\right].$$

This expression simply states that the NPE collects the license fee P_t^{fee} until the technology grouping becomes obsolete by a new radical innovation, which arrives at the rate $\bar{z}_0$. Using (16) the value of an opportunistic patent is expressed as

$$V_t^o(m) = \frac{m\pi Y_t}{\bar{z}_0 + \rho}.$$

Clearly the value of the opportunistic patent increases in its ability to extract rents m

$$\frac{\partial V_t^o(m)}{\partial m} > 0$$

This leads to our main result.

Theorem 4 *In the case of opportunistic use of patents, the private value of the patent $V_t^o(m)$ are negatively correlated with forward citations.*

The intuition for this result is that when a patent is used for an opportunistic goal by an NPE, the value of that patent increases in the ability of the patent to extract rents. For the same reason, new entrants get discouraged and technological progress in that cluster slows and the reduction in subsequent entry results in fewer future forward citations. The end result is that the relationship between patent value and citations is negative when patents are used for opportunistic reasons.

The general message from our theoretical analysis follows. The relationship between social value and forward citations is positive since citations imply technological progress by subsequent researchers, which increases the social value of the current innovations. However, our theory highlights an important distinction when it comes to private value. When patents are used for productive purposes, the relationship between private value and citations is positive. However, when a patent is used for strategic reasons, say as a strategic patent by an incumbent or an existing patent being used opportunistically for litigation purposes by an NPE, the usual positive relationship between patent value and citations breaks, becomes negative, and results in an overall inverted-U relationship.

F Proofs and Derivations

Proof of Lemma 1. The monopolist's equilibrium quantity is

$$y_{jt} = \frac{Y_t q_{jt}}{(1+\eta_j)\, w_t}.$$

Substituting this into the final good production function (1) we get

$$w_t = Q_t \int_0^1 (1+\eta_j)^{-1}\, dj$$

where $Q_t \equiv \left[\int_0^1 (1+\eta_j)^{-1}\, dj\right]^{-1} \exp \int_0^1 \ln \frac{q_{jt}}{1+\eta_j} dj$. Moreover, the labor used for production in each line is

$$l_{jt} = \frac{y_{jt}}{q_{jt}} = \frac{Y_t}{(1+\eta_j)\, w_t}.$$

Using this expression in the labor market clearing condition

$$1 = \int_0^1 l_{jt} dj$$

delivers $Y_t = Q_t$. ■

Proof of Proposition 1.

We need to show that $v_n > v_{n+1}$. This condition can be expressed as

$$v_n = \frac{\pi_n}{\rho + (v_0 + v_{n+1})/\zeta} > \frac{\pi_{n+1}}{\rho + (v_0 + v_{n+2})/\zeta} = v_{n+1}.$$

Since $\frac{\pi_{n+1}}{\rho+v_0/\zeta} > v_{n+1}$, we can write

$$v_n = \frac{\pi_n}{\rho + (v_0 + v_{n+1})/\zeta} > \frac{\pi_n}{\rho + \left(v_0 + \frac{\pi_{n+1}}{\rho+v_0/\zeta}\right)/\zeta}.$$

Moreover we have

$$\frac{\pi_{n+1}}{\rho + v_0/\zeta} > v_{n+1} = \frac{\pi_{n+1}}{\rho + (v_0 + v_{n+2})/\zeta}.$$

Our result will be proven if we show

$$\frac{\pi_n}{\rho + \left(v_0 + \frac{\pi_{n+1}}{\rho+v_0/\zeta}\right)/\zeta} > \frac{\pi_{n+1}}{\rho + v_0/\zeta} \tag{19}$$

since we will then have

$$v_n = \frac{\pi_n}{\rho + (v_0 + v_{n+1})/\zeta} > \frac{\pi_n}{\rho + \left(v_0 + \frac{\pi_{n+1}}{\rho+v_0/\zeta}\right)/\zeta} > \frac{\pi_{n+1}}{\rho + v_0/\zeta} > \frac{\pi_{n+1}}{\rho + (v_0 + v_{n+2})/\zeta} = v_{n+1}.$$

The inequality in (19) is true if

$$\rho + \frac{v_0}{\zeta} > \frac{\alpha(1+\eta_n)}{1+\alpha\eta_n}\left[\rho + \frac{v_0}{\zeta} + \frac{\pi_{n+1}}{\zeta\rho + v_0}\right].$$

A sufficient condition is

$$\rho \geq \frac{\alpha(1+\eta_n)}{1+\alpha\eta_n}\left[\rho + \frac{\pi_{n+1}}{\zeta\rho}\right].$$

where we eliminated v_0. Likewise, a sufficient condition would be

$$\rho\left[\frac{(1-\alpha)}{1+\alpha\eta_n}\right] \geq \frac{\alpha+\alpha\eta_n}{1+\alpha\eta_n}\frac{\pi}{\zeta\rho}$$

since the right-hand side (RHS) is increasing in π_n. Finally, since RHS is also increasing in η_n, we get the desired condition

$$\frac{\zeta\rho^2}{\eta} \geq \frac{\alpha}{1-\alpha}.$$

■

Proof of Lemma 3. Substitute in the conjecture and cancel the common terms on both sides to get

$$0 = \Delta t \ln q + \Delta t \ln \frac{(1-\eta)}{(1+\eta)} + \left\{ \begin{array}{c} (\bar{z}_{0m}\Delta t + \bar{z}_{1m}\Delta t) \left[A \ln (1+\eta) + \mathbb{E}B_m - B_m\right] \\ -\rho \Delta t \left[A \ln q + A \ln + B_m + C\right] \end{array} \right\}$$

where we ignored the second order terms since they vanish as $\Delta t \to 0$. Dividing both sides by Δt we get

$$\rho A \ln q + \rho B_m + \rho C = \ln q + \ln \frac{(1-\eta)}{(1+\eta)} + (\bar{z}_{0m} + \bar{z}_{1m}) \left[A \ln (1+\eta) + \mathbb{E}B_m - B_m\right]$$

Equating the coefficients we get

$$\rho A = 1, \quad \rho C = \ln \frac{(1-\eta)}{(1+\eta)}, \quad \textit{and} \quad (\rho + \bar{z}_{0m} + \bar{z}_{1m}) B_m = (\bar{z}_{0m} + \bar{z}_{1m}) \left[A \ln (1+\eta) + \mathbb{E}B_m\right]$$

This completes the proof. ■