Penn Institute for Economic Research
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104-6297
pier@econ.upenn.edu
http://economics.sas.upenn.edu/pier

PIER Working Paper 13-065

"Patent Value and Citations: Creative Destruction or Strategic Disruption?

by

David S. Abrams, Ufuk Akcigit, and Jillian Popadak

http://ssrn.com/abstract=2351809

Patent Value and Citations: Creative Destruction or Strategic Disruption?*

David S. Abrams
University of Pennsylvania

Ufuk Akcigit
University of Pennsylvania & NBER

Jillian Popadak
University of Pennsylvania

November 5, 2013

Abstract

Prior work suggests that more valuable patents are cited more and this view has become standard in the empirical innovation literature. Using an NPE-derived dataset with patent-specific revenues we find that the relationship of citations to value in fact forms an inverted-U, with fewer citations at the high end of value than in the middle. Since the value of patents is concentrated in those at the high end, this is a challenge to both the empirical literature and the intuition behind it. We attempt to explain this relationship with a simple model of innovation, allowing for both productive and strategic patents. We find evidence of greater use of strategic patents where it would be most expected: among corporations, in fields of rapid development, in more recent patents and where divisional and continuation applications are employed. These findings have important implications for our basic understanding of growth, innovation, and intellectual property policy.

JEL Codes: O3, L2, K1.

Keywords: productive innovation, defensive innovation, patents, creative destruction, citations, patent value, competition, intellectual property, entrepreneurship, strategic patenting, defensive patenting, patent thickets, fencing patents.

*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), Works in Progress Intellectual Property Conference (Newark), and particularly our discussants Josh Lerner and Sharon Belenzon. We also thank Daron Acemoglu, Nick Bloom, Bill Kerr, and Heidi Williams for very helpful discussions and suggestions. Salomé Baslandze provided excellent research assistance.

1 Introduction

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. While the importance of such a metric has long been recognized (Scherer (1967), Griliches (1981)) so too have the inadequacies of the proxies for value that are in widespread use (Schankerman and Pakes (1986), Hall and Harhoff (2012)).

Over the past 30 years, two primary metrics have been used to proxy for the value of innovation, patent counts and citation-weighted patent counts.[1] The intuition is thus: ceteris paribus, fields with greater innovative activity will have more value to protect and will do so by applying for more patents.[2] Weighting patent counts by forward citations[3] is a natural augmentation to simple patent counts, given the well-known fact that patents vary tremendously in value.[4] The use of this measure, however, is based on the assumption that a larger number of citations corresponds to higher value.

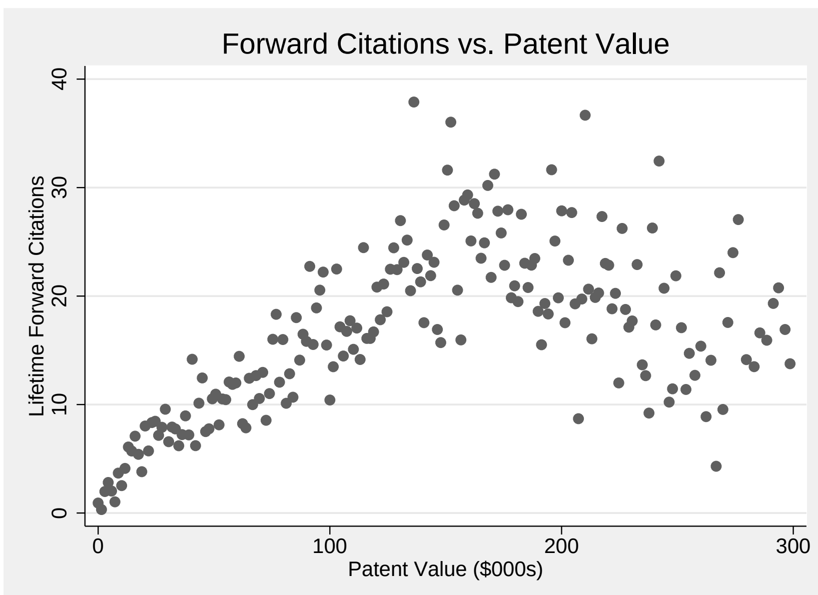


Figure 1: LIFETIME FORWARD CITATIONS VS. PATENT VALUE
Note: Data is normalized so that the mean annual revenue is $10,000.

Until now this assumption has been very challenging to test. Several problems have held back this inquiry: the reluctance of companies to share proprietary patent data, the lack of generality

[1] Throughout the paper we generally use the term value to mean private value, but we do explore social value as well in Appendix C and discuss it in Section 3.3.1.

[2] 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, Klevorick, Nelson, Winter, Gilbert, and Griliches (1987).

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

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

(and sufficient observations) from any single patent portfolio, and the fact that almost no companies allocate revenues to specific patents. The proliferation of non-practicing entities (NPEs) in the past ten years has led to the availability of patent-specific revenue data covering a range of technologies. By using a proprietary data set of tens of thousands of NPE-owned patents, we are able to examine the patent value - citation relationship empirically in a level of detail that was previously impossible.[5] While the patents analyzed are held by NPEs, almost all were originally granted to individual inventors or those employed by practicing firms.

Using this data we find strong evidence that the standard approach to valuing innovation is imperfect. Indeed, the relationship between lifetime 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.[6] There is still an overall positive correlation between citations and value, which has been found in prior work, but it comes primarily from lower value patents, and the full pattern is more complex.

This suggests that patents may not have a purely productive role in technological progress, an assumption that most prior models have made. The non-monotonic relationship could be evidence of the use of patents for strategic purposes, an idea that has been much discussed in recent years. Recent papers on the subject include Farrell and Shapiro (2008) and Noel and Schankerman (2013) on the theory side and Hall and Ziedonis (2001), Ziedonis (2004), Hegde, Mowery, and Graham (2009), Galasso and Schankerman (2010), Cockburn and MacGarvie (2011), and Von Graevenitz, Wagner, and Harhoff (2013) on the empirical side.[7] Our goal is to shed light on this issue systematically both through the lens of the theory and data.

We introduce a theoretical model that suggests that the inverted-U shape is the result of 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 producing fencing patents, which seek to expand the area of protection available to previously granted patents. In an economy that exhibits both of these types of innovative effort, the link between patent value and citations will be the inverted-U that we observe empirically.[8]

In order to further test our theory of strategic patents, we examine the citation-value relationship using four characteristics that should be related to strategic patenting. Strategic patents should be more prevalent among larger entities, for divisional and continuation patents, for newer patents, and in technology classes with rapid growth. Each of these predictions is borne out in the data and we find evidence that strategic patenting is more prominent in these categories.

In this paper we add to the literature that has attempted to examine the relationship between

[5] We discuss details of the data sets in Section 2 as well as implications of NPEs as its source. Confidentiality agreements limit our ability to disclose actual revenue numbers or number of observations.

[6] Further details on the normalization and other aspects of the production of this figure are discussed in Section 4.

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

[8] In Section 3.3.2, we discuss an additional model focused on strategic choices that may be more common among NPEs. That model as predicts a non-monotonic relationship between forward citations and value and may be found in Appendix D.

patent value and citations. Trajtenberg (1990) is perhaps the leading prior work on the subject, and was the first paper to rigorously examine the citation-value relationship. The paper focuses on a relatively small number of patents in the computed tomography field, with values imputed from a structural model of the CT device market and finds an approximately linear relationship between citations and patent value. Harhoff, Narin, Scherer, and Vopel (1999) obtain categorical measures of value on 772 patents from a survey of German patents with 1977 priority dates, all of which were renewed to full term. They find a positive correlation between forward citations and patent value, but that the relationship is noisy. Several excellent 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 make use of the contingent claim valuation method pioneered by Pakes and Schankerman. This approach is useful for understanding the distribution of patent value, but conveys little information for individual patents, in particular high-value patents, which is where we find the greatest deviation from the assumed monotonic citation-value relationship. In a more recent work, Moser, Ohmstedt, and Rhode (2012) link the forward citations of patents for hyrid corn with yields and find a positive correlation between the two. In a parallel study, Kogan, Papanikolaou, Seru, and Stoffman (2012) show that stock market response to news about patents is a good predictor of future forward cites.

In the legal literature, strategic patenting has received a great deal of attention in recent years as allowable subject matter has widened to include software and business methods patents. As the number of patents granted has increased, technological progress has led to devices that implicate thousands of separate patents. 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 (2012)). 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.

Building upon these previous findings, we contribute to several lines of literature. Our primary finding is that the citation-value relationship has an inverted-U shape, rather than the monotonic relationship that has previously been assumed. We contribute to the innovation literature by showing how the inverted-U relationship can arise naturally in an economy with two types of innovative effort. The introduction of strategic patents adds to prior models that generally have a single type of patent. The strategic patent explanation for the observed relationship is borne out by an examination of correlates of high value-low citation patents. In particular, our empirical evidence that corporate assignees are more likely to engage in strategic disruption is important for corporate finance scholars interested in valuing firms' intellectual property assets and in understanding how firm performance is affected by defensive product market strategies. Both our empirical and theoretical findings show that surprisingly not every patent leads to creative destruction and economic growth rather some of them are strategic disruptions.

The rest of the paper proceeds as follows. In Section 2 we provide substantial detail about incentives to patent and cite, the business models of NPEs and further description of the data. Section 3 introduces our model which we believe captures some of the key elements of innovation and the patenting and citing processes. In Section 4 we present the main empirical results and a discussion of them. Section 5 concludes 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 B-E contain additional theoretical proofs and derivations.

2 Background and Data Description

Since the major impediment to greater understanding of patent value has been the lack of available data on individual patent revenues, it is worth discussing the data source and characteristics in some detail. The data in this paper was provided by large non-practicing entities (NPEs), with a focus on the technology sector. NPEs are firms whose revenue primarily derives not from producing products based on patented technology, but from licensing patents. The incentives associated with patents within the technology sector encourages licensing (Lerner and Tirole (2002, 2005)). NPEs employ a range of different business models ranging from aggressive litigators to passive licensors, 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 is also likely to be substantially more useful for researchers, as they tend to determine patent-specific revenues. This is something that almost no 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.

While the data analyzed here is unique, there may be concern about how representative it is of the entire patent universe. We leave to the reader to assess the findings but point out several aspects of the data that suggest they may be fairly general. The 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. Additionally, the NPEs have no evidence that even those patents they target tend to be particularly valuable. Their strategy is to accumulate patents in a broad range of technology areas that they believe may be important, without targeting specific patents. We further compare characteristics of the data with the universe of patents in Section 4.2.

An additional concern about the data source may be that NPEs use patents differently than practicing entities and as such the revenue data we use may not reflect the value of patents to practicing entities. In particular, there may be concern that 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 3.3.2 and Appendix D. 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 multiple times and usually as part of portfolios. Thus it is far more likely that the licensing fees paid to the NPEs reflect real value for the licensees. Additional detail about the revenue data may be found in Appendix A.

Before presenting summary statistics, it is important to note several 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 \$10,000.[9] 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 forward 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 2008 - 2012 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 not active at all during this period, it is excluded.

We define patent value as the sum of the normalized annual revenues realized by a patent during the 20 years from application to expiration. The mean lifetime patent value is \$204,200 (all figures are 2010 dollars). Note that the standard deviation of \$1.9 million is more than 9 times the mean and more than 35 times the median value of \$52,190. 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 4.0, 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.

[9] 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.

Table 1: SUMMARY STATISTICS

	Mean	Standard Deviation	Median
Patent Value ($000s)	204.2	1904.7	52.19
Lifetime Forward Citations	29.1	52.5	11.5
Backward Citations	23.1	59.9	8.0
Fraction of Backward Cites in Past 3 Years	0.20	0.30	0.00
Fraction of Backward Cites in Past 5 Years	0.28	0.37	0.00
Original Indicator	0.84	0.36	1.00
Application Year	1999	4.7	1999
Individual Inventor Indicator	0.14	0.35	1.00

Note: Data is normalized so that the mean annual revenue is $10,000 (2010$). Original patent applications are those which are not divisionals or continuations.

Our other main variable of interest, lifetime forward citations, also has a skewed distribution with a mean of 29.1, standard deviation of 52.5 and median of 11.5. 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, Narin, Scherer, and Vopel (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 23.1, median of 8.0 and standard deviation of 59.9. About 20% of backward citations are for patents issued within the prior 3 years and 28 of cited patents are 5 years old or less. We use both of these measures as indicators of how active or ”hot” a field is.

Most (84%) 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 1999, meaning the median patent had about 7 years of protection left by the end of our revenue data. Individual 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.[10] The most valuable patents are found in the Circuits category with a mean value of $367,130 but with only an average of 7.1 citations. Computer Architecture patents also have a high average value at

[10]These classifications are not part of the Patent and Trademark Office classification system but rather are used by the portfolio owners.

$283,773 but the lowest average number of forward citations with 6. At the low end are MEMS and Nanotechnology patents which average $58,860 and 11.1 citations and Optical Networking patents with 16.5 citations and an average value of $56,425.

TABLE 2: VALUE AND FORWARD CITATIONS BY TECHNOLOGY

Technology	**Patent Value**	**Lifetime Forward Citations**
Circuits	$367,130	7.1
Computer Architecture	$283,773	6.0
Internet & Software	$273,093	12.6
Wireless Communications	$174,605	35.4
Network Communications	$146,974	9.4
Semiconductor Devices	$115,824	7.8
Peripheral Devices	$99,801	8.1
Electro-Mechanical	$62,018	7.4
MEMS & Nano	$58,860	11.1
Optical Networking	$56,425	16.5

Note: Data is normalized so that the mean annual revenue is $10,000 (2010$).

3 Theory of Patent Valuations and Citations

In the introduction we provided a new empirical finding which is at odds with the received wisdom about the link between patent value and citations. How can we reconcile the two and account for the inverted-U? 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.

In our model, we rely on the *Schumpeterian theory of creative destruction* (see the recent survey by Aghion, Akcigit, and Howitt (2013) 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.[11] 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

[11] 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.

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 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 3.1, 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 3.2 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.

3.1 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 B for the interested readers.[12]

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 in-

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

novations will have larger mark-ups due to their originality, denoted by the step size of a 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.

3.1.1 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 E. ■

3.1.2 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 (2010). 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:[13]

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

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

[13]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.

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 ID:	N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8	N_9	N_{10}	
Tech Cluster ID:	*Tech Cluster 1*				*Tech Cluster 2*		*Tech Cluster 3*			*Tech Cluster 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 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 B 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.[14] 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 E. ■

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.

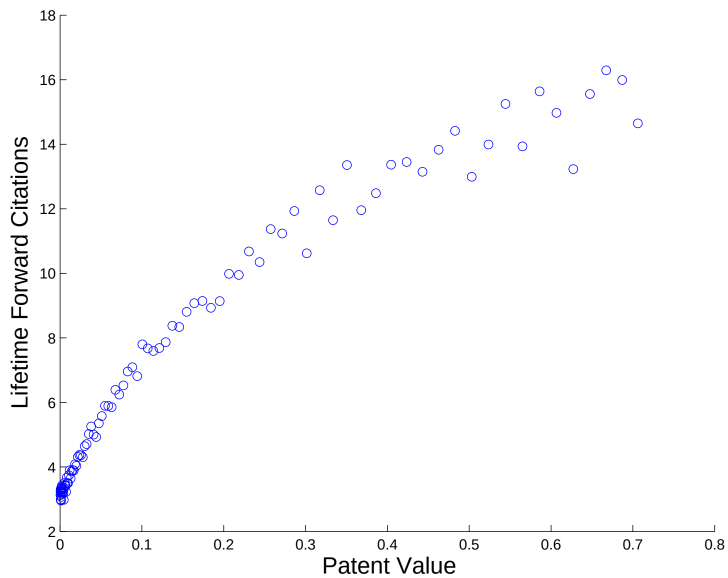
Theorem 1 *In the case of productive patents, patent value and forward citations are* ***positively correlated****.*

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

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 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 VS. 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.

3.2 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

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

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 $\rho v_m^d = \pi_0 - (\bar{z}_{0m} + \bar{z}_{1m}) v_m^d$. 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

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

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.*

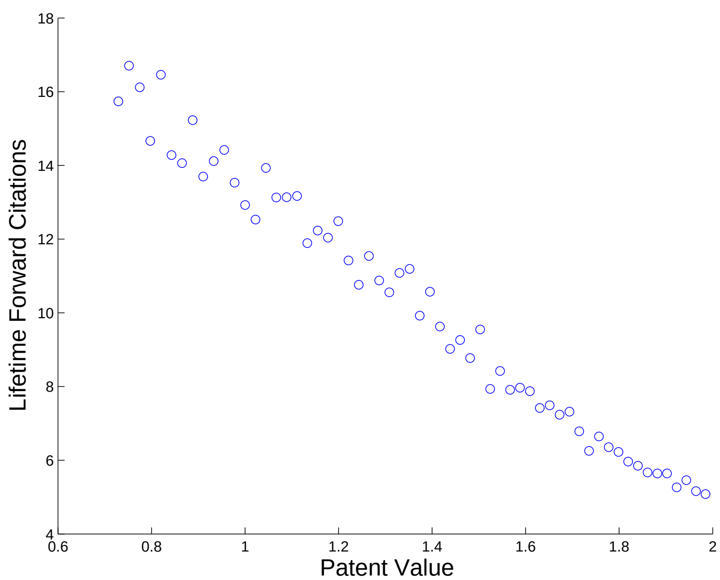
Proof. The first part of this proposition follows directly from equation (12) and the second part follows from equation (11). ■

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.

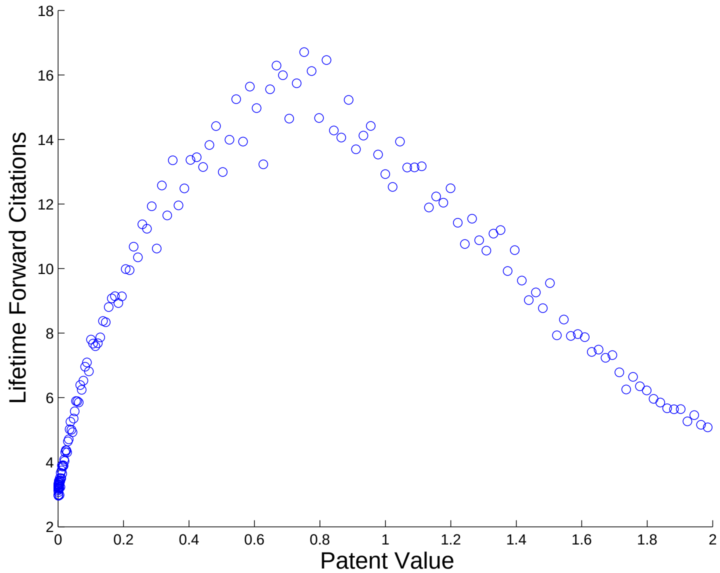
FIGURE 3: FORWARD CITATIONS VS. PATENT VALUE



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 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.

Figure 4: LIFETIME FORWARD CITATIONS VS. PATENT VALUE



Notes: 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.

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 4 tests these predictions.

3.3 Extensions

3.3.1 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 4. 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 C 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.

3.3.2 NPEs and Opportunistic Use of Patents

Patents may be used strategically in ways other than to protect valuable innovations, as the model in Section 3.2 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 D) 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 3.2) or used (Appendix D) 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 3.2 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 the NPEs. In Section 4 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.

4 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 that test various predictions of the model.

4.1 Main Results

Figure 1 displays the empirical relationship between forward citations and patent value for the data set described in Section 2. The figure plots the mean number of citations for each of 200 quantiles of patent value, with the top 5

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 standard errors corrected for heteroskedasticity. The three pairs of columns in the table vary the share of the overall dataset that is included, winsorizing the top 10, 5, and 1% in the first, second and third pair of columns, respectively. The coefficients show that there is indeed an overall positive relationship between forward citations and patent value, one that has been assumed in much of the recent literature. But 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 coefficients on the quadratic terms.

Table 3: FORWARD CITATIONS vs PATENT VALUE

	Share of most valuable patents excluded					
	10%		5%		1%	
Patent Value ($100,000s)	9.047**	22.497**	7.104**	14.402**	6.961**	8.016**
	(0.256)	(0.654)	(0.232)	(0.566)	(0.246)	(0.432)
Patent Value Squared		-6.036**		-2.193**		-0.139*
		(0.288)		(0.195)		(0.070)
R^2	0.04	0.05	0.04	0.05	0.09	0.09

** Significant at the 1% level; * Significant at the 5% level
Note: Separate regressions reported in each column, with robust standard errors in parentheses. Dependent variable is lifetime forward citations. Data is normalized so that the mean annual revenue is $10,000 (2010$). Regression excludes indicated top percent of patents by value.

Figure 1 and Table 3 omitted covariates, and one may be concerned that variation of both patent value and forward citations by these covariates drives the observed relationship. Thus the dependent variable in Figure 5 is the residual from a regression of forward citations on a set of dummies for technology category, whether the inventor was an individual, and whether the patent was original (i.e., not a divisional or continuation). The same inverted-U relationship in Figure 5 is apparent that was previously seen in Figure 1, although slightly noisier. The takeaway is that there is still compelling evidence for productive and strategic patenting, even within technology categories, and accounting for inventor type and original status.

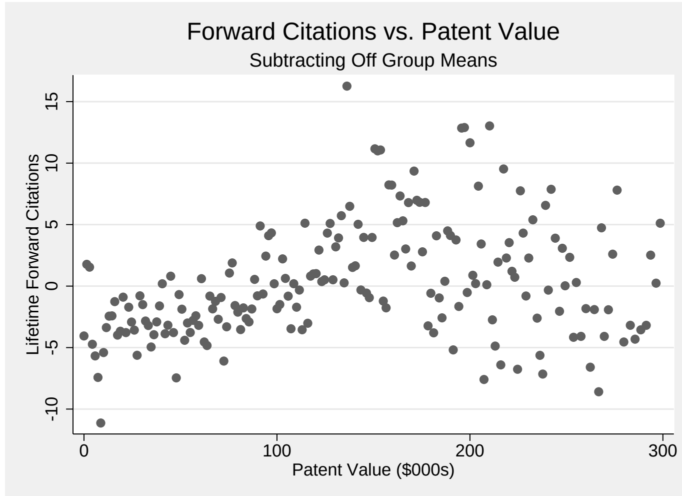


FIGURE 5: FORWARD CITATION RESIDUALS VS. PATENT VALUE

NOTES: DATA IS NORMALIZED SO THAT THE MEAN ANNUAL REVENUE IS $10,000 ($2010). RESIDUALS ARE FROM A REGRESSION ON DUMMIES FOR TECHNOLOGY CATEGORY, INVENTOR TYPE, AND ORIGINAL PATENT STATUS.

Table 4 reports the results of regressions of forward citations on a quadratic of patent value and the following covariates: dummies for inventor type (individual or corporate), whether the patent was original (not a continuation or divisional), fraction of backward citations within the last three years, and technology category. Each regression also includes application-year fixed effects. The coefficients on 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. Individual inventor status has a substantially negative impact on number of forward citations. Original patents tend to receive fewer citations than continuations or divisionals and patents in "hot" fields - those with more backward citations in the past 3 years - are cited more.

There is a good amount of variation in the number of forward citations by technology category, but all receive significantly more than circuits, which is the excluded category. Internet & software patents have the largest coefficient, followed by networking & communications patents. Variation across technology class may be driven by differences in norms, patent examiner practices, or a number of other causes and thus it is important to note the relationship between citations and value still holds even after controlling for technology class.[15]

To this point, the evidence has largely bolstered the central empirical finding of the paper, of an inverted-U relationship between citations and value. But this sort of finding might be generated by a number of models of the innovation process. One of the important features of the model we propose is the decision of an innovator to engage in strategic patenting. This decision will vary by observable characteristics of the innovator and we now test several implications of the model empirically.

[15] In additional regressions, not shown here, we find that the number of claims and dependent claims are statistically insignificant, which echoes recent findings by Moser, Ohmstedt, and Rhode (2012).

Table 4: FORWARD CITATIONS AND PATENT CHARACTERISTICS

	(1)	(2)	(3)	(4)
Patent Value ($100,000s)	9.561**	8.877**	7.739**	7.475**
	(0.648)	(0.641)	(0.630)	(0.614)
Patent Value Squared	-1.292**	-1.238**	-1.040**	-0.948**
	(0.208)	(0.208)	(0.206)	(0.201)
Individual Inventor	-18.399**	-17.052**	-11.795**	-12.130**
	(0.386)	(0.407)	(0.682)	(0.663)
Indicator Original Patent		-9.436**	-9.938**	-8.096**
		(0.717)	(0.721)	(0.696)
Fraction of Backward Cites in Past 3 Years			12.357**	11.924**
			(1.060)	(1.025)
Tech Category (Computer Architecture)				3.336**
				(0.567)
Tech Category (Electro-Mechanical)				4.001**
				(0.632)
Tech Category (Internet & Software)				20.085**
				(0.861)
Tech Category (MEMS & Nano)				2.996**
				(1.309)
Tech Category (Networking & Communications)				9.684**
				(0.734)
Tech Category (Optical Networking)				1.896**
				(0.476)
Tech Category (Peripheral Devices)				2.403**
				(0.414)
Tech Category (Semiconductors)				3.453**
				(0.430)
Tech Category (Wireless Communications)				7.391**
				(0.523)
Application Year Fixed Effects	Yes	Yes	Yes	Yes
R^2	0.12	0.13	0.14	0.17

** Significant at the 1% level; * Significant at the 5% level
Note: Separate regressions reported in each column, robust standard errors in parentheses. Dependent variable is lifetime forward citations; circuits is the excluded technology category. Data is normalized so that the mean annual revenue is $10,000 (2010$).

Since there is a fixed cost to strategic patenting, those entities that hold more valuable patents should be more likely to engage in strategic patenting. Thus we would expect individuals to engage in less strategic patenting than companies. We test this hypothesis using the type of original assignee to test this hypothesis, and expect that corporations should be more represented in the "strategic" part of the citations-value relationship, namely where the curve is downward sloping. Figure 6 shows the relationship between the citations and value, where the size and darkness of the data points is split into quartiles according to the corporate share of assignees. The biggest, darkest points are those with the highest share of corporate assignees. As the theory predicts, we find that the highest value patents, with fewer citations tend to be composed of a larger proportion of corporate assignees than the less-cited and less valuable patents.

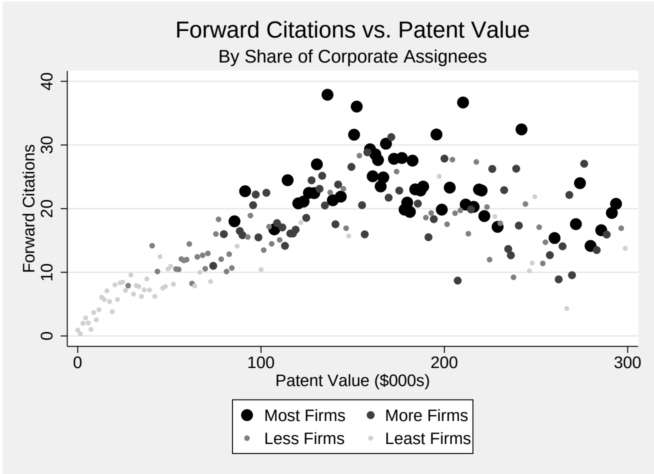


Figure 6: FORWARD CITATIONS VS. REVENUE BY SHARE OF CORPORATE ASSIGNEES

Notes: Data is normalized so that the mean annual revenue is $10,000.

Another patent characteristic that may influence likelihood of strategic patenting is the rate of innovative growth in the field. Consistent with this intuition, areas of rapid innovation are likely to also generate greater current and expected profits, and thus greater incentive to engage in strategic patenting to protect valuable, productive patents. Our measure for the growth of innovation is the share of backward citations within the prior three years.

This is in fact the relationship we find in Figure 7, which has an analogous structure to Figure 6. Here the size and darkness of a data point is determined by its quartile in the distribution of the share of backward citations in the previous three years. As predicted by the model, we find those patents with the greatest share of recent backward citations to be toward the right end of the figure, with high revenues, but not particularly high citations. This is consistent with greater use of strategic patenting in fields of rapid innovation.

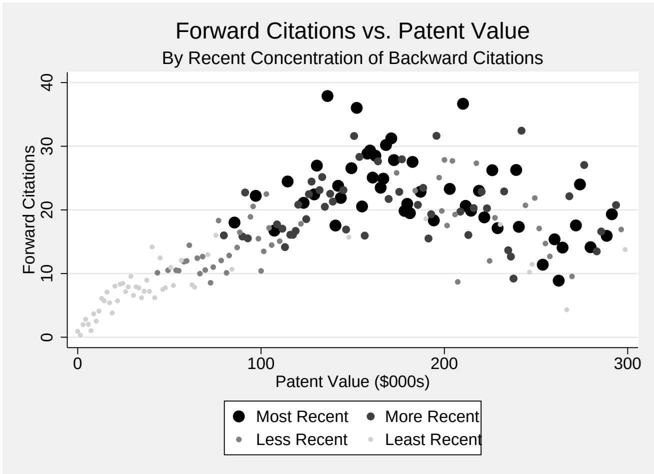


Figure 7: FORWARD CITATIONS VS. REVENUE BY RECENT CONCENTRATION OF BACKWARD CITATIONS

Notes: Data is normalized so that the mean annual revenue is $10,000.

Figures 8 and 9 report the forward citation-patent value relationship in two additional ways, that provide further support to the theory of innovative and strategic patenting. Continuation and divisional patents are frequently employed strategically by sophisticated patentees in order to extend the duration of patent prosecution. These uses are less likely to be employed for truly innovative patents, because a truly innovative patent's value will be less dependent on market conditions and thus extending patent prosecution has less value. In Figure 8 we observe that divisional and continuation patents are more prevalent in the high value/low citation region of the graph.

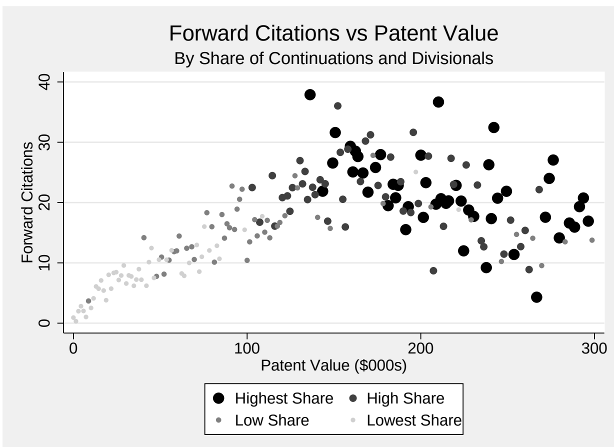


Figure 8: FORWARD CITATIONS VS. REVENUE BY SHARE OF CONTINUATIONS AND DIVISIONALS

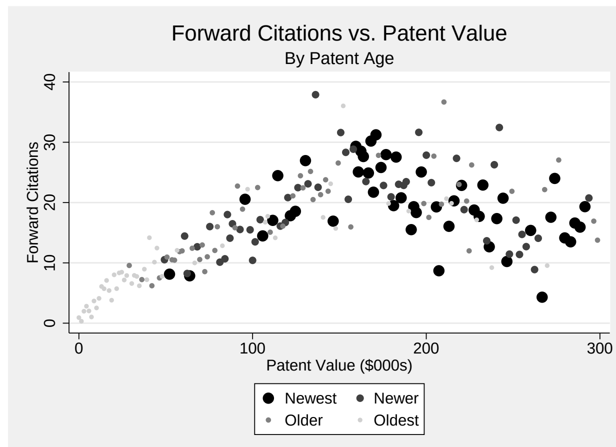


Figure 9: FORWARD CITATIONS VS. REVENUE BY PATENT AGE

Notes: Data is normalized so that the mean annual revenue is $10,000.

Finally, we examine whether there has been recent growth in the use of strategic patenting. There has been much written about the increasing use of patents for strategic purposes in the past several years. We divide patents by their application date in order to investigate these claims. In fact, we find support for this view in Figure 9, which shows that the share of patents newer than the median is higher where revenues are higher and citations relatively lower. This does not provide an explanation for this trend, but does provide the first direct evidence that strategic patenting is more prevalent among newer patents.

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 5 we report results from 10 regressions of forward citations on patent value and patent value squared, one for each technology class. We find that the same overall relationship 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.

Table 5: FORWARD CITATIONS AND PATENT VALUE BY TECHNOLOGY CLASS

	Circuits	Computer Architecture	Electro-Mechanical	Internet & Software	MEMS & Nano
Patent Value ($100,000s)	6.233	14.497	10.917	23.542	17.051
	(0.905)**	(1.285)**	(1.655)**	(2.151)**	(3.592)**
Patent Value Squared	-0.777	-2.212	-2.341	-3.184	-4.325
	(0.244)**	(0.353)**	(0.597)**	(0.726)**	(1.140)**
R^2	0.05	0.09	0.04	0.05	0.06
	Networking Communication	**Optical Networking**	**Peripheral Devices**	**Semiconductors**	**Wireless Communications**
Patent Value ($100,000s)	19.107	13.496	9.847	9.329	18.007
	(2.211)**	(1.181)**	(0.673)**	(0.972)**	(1.496)**
Patent Value Squared	-2.328	-2.114	-2.355	-1.020	-3.292
	(0.802)**	(0.462)**	(0.212)**	(0.339)**	(0.557)**
R^2	0.08	0.07	0.02	0.06	0.07

** Significant at the 1% level; * Significant at the 5% level
Note: Separate regressions reported in each column, robust standard errors in parentheses. Dependent variable is lifetime forward citations. Data is normalized so that the mean annual revenue is $10,000 (2010$).

Figures 10 and 11 shows 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.

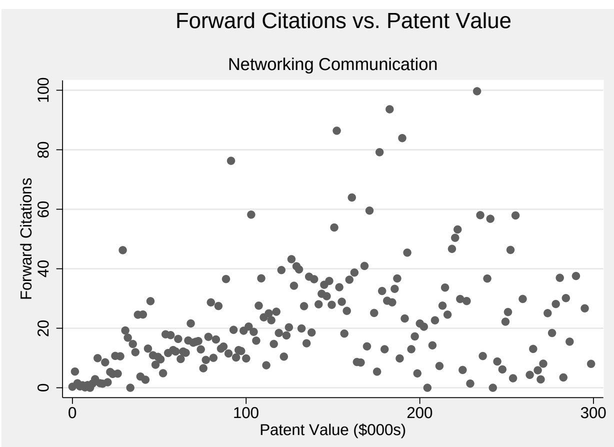


Figure 10: FORWARD CITATIONS VS. REVENUE, NETWORKING COMMUNICATION

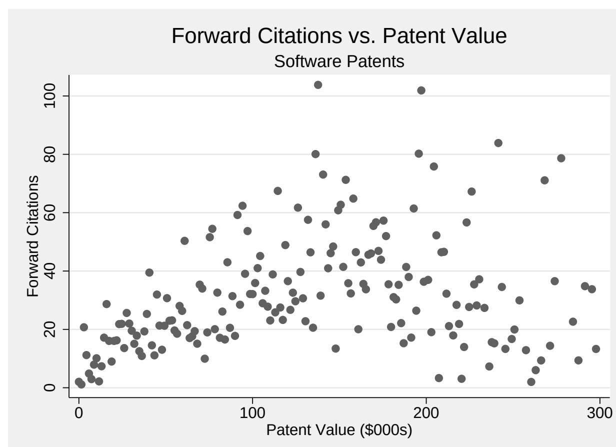


Figure 11: FORWARD CITATIONS VS. REVENUE, SOFTWARE PATENTS

Notes: Data is normalized so that the mean annual revenue is $10,000.

4.2 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 has a number of features that make it ideally suited for learning 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, fewer 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 this data, we compare features of our data with the universe of patents in the same technology classes. Figure 12 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 5 years of patent life, it is clear that the lifetime profile of citations is very similar between the NPE data and the universe. We additionally find that the median number of forward citations (unadjusted for patent unlike elsewhere in the paper) is 8.75 for our data and 8.0 for the universe within the same technology category and of similar age.

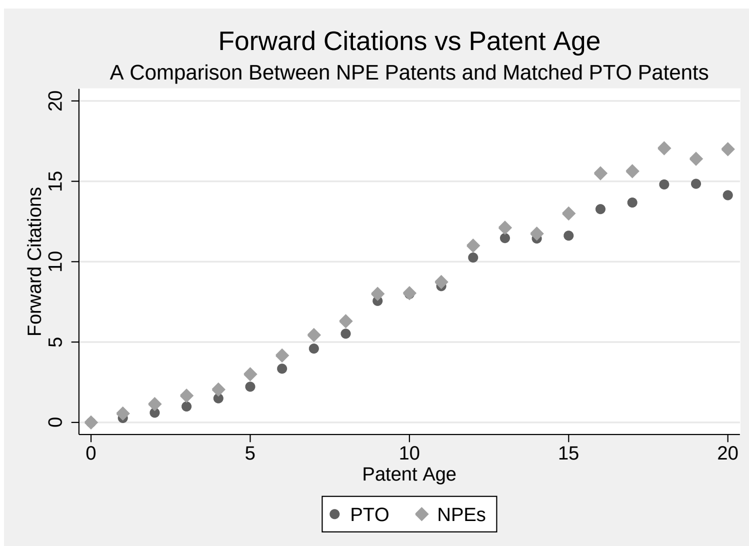


Figure 12: LIFETIME CITATION PROFILES OF NPE VS. USPTO PATENTS

We perform several other robustness checks that are available upon request. We calculate a correlation between the forward citations received within broad technology classes for our data and all patents in the same class and we find most correlations over 0.9. We perform several tests to investigate whether the inverted-U may be driven by patent age. We produce the basic citation-value figure and the regression analog for only patents under 10 years old and separately for only older patents over 10 years old. We find the inverted-U in both specifications. We run regressions of citations on a quadratic function of patent value separately for 10 years of patent grant dates and find a robust inverted-U for each. These results help solidify the robustness and generality of the main result.

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 10 and 11). 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 individual inventor status, original versus continuation patents, age of patent and level of 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 NPEs and the way they use patents, and may not generalize to patents held by practicing entities. As discussed in Section 2 there are several reasons to believe the results presented here may be more broad. The data under study comes 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 the NPE's. These facts make it more likely that the revenues studied reflect real patent value.

5 Conclusion

Using a new dataset with an unprecedented number of observations and patent-specific revenue we have found a surprising result for the relationship between forward citations and patent value. This finding should impact a large array of literature that has relied on citation-weighted patent counts to proxy for innovation. While we find evidence that the long-believed positive correlation between citations and value is correct, the story is substantially more complicated than has previously been assumed. For lower value patents, there appears to be something like a linear relationship between value and citations, but that relationship does not hold once patent value exceeds a certain threshold, at which point 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 two types of innovation: 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. In forthcoming papers we plan to give greater guidance on how to make use of these new findings to better proxy for value using forward citations and other patent characteristics.

The real potential for this work is yet to come. 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. The goal of this line of work is to broaden and deepen our understanding of the innovation process, with an eye ultimately towards informing policy decisions to better foster it.

References

Aghion, P., U. Akcigit, and P. Howitt (2013): "What Do We Learn From Schumpeterian Growth Theory?," National Bureau of Economic Research WP 18824.

Akcigit, U., and W. R. Kerr (2010): "Growth through Heterogeneous Innovations," National Bureau of Economic Research WP 16443.

Allison, J. R., M. A. Lemley, K. A. Moore, and R. D. Trunkey (2009): "Valuable Patents," *Georgetown Law Journal*, 92, 435–479.

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

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

Bessen, J., and M. J. Meurer (2008): *Patent Failure: How Judges, Bureaucrats, and Lawyers Put Innovators at Risk*. Princeton University Press.

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

Boldrin, M., and D. Levine (2012): "The Case Against Patents," FRB of St. Louis Working Paper No. 2012-035A.

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

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

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

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

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

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

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

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

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

Jaffe, A. B., and J. 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.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman (2012): "Technological Innovation, Resource Allocation, and Growth," National Bureau of Economic Research WP 17769.

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

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

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

Moser, P., J. Ohmstedt, and P. Rhode (2012): "Patent Citations and the Size of Innovation – Evidence from Hybrid Corn," Stanford University Working Paper.

Nicholas, T. (2013): "Are Patents Creative or Destructive?," Harvard Business School Working Paper.

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

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

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

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

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

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

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

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

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 was 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.

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 order to compute patent value, we aggregate revenues to the patent-year level and then compute the mean revenue profile over the life of a patent separately for each of the 10 primary technology categories. We estimate patent value for each patent by inflating the observed cumulative revenue by the ratio of the mean lifetime revenue to the mean cumulative revenue for patents of the same age and technology category. We then normalize all revenue amounts so that mean annual revenue is $10,000 in order to maintain the confidentiality of the revenue data.

Lifetime citations are computed in a similar manner. We obtain data on forward citations, defined as the total number of times a patent has subsequently been cited. By definition, newer patents will have less time to acquire citations than old ones and this must be accounted for. We define "lifetime citations" as the total number of citations we expect a patent to have by its expiration. We compute this by first producing the forward citation- patent age profile for each of our ten technology categories. Figure 13 presents the incremental patent citation-age profile as well as the revenue-age profile in aggregate. There is substantial variation by technology class; therefore, we create separate revenue and citation profiles for each technology class. We calculate lifetime citations by inflating the total citations already received by the ratio of the total mean citations of the same technology class divided by the mean for the average patent of the same age and technology class as the one in question. While this procedure will understate the number of lifetime citations for any patent that has zero in our dataset, 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.

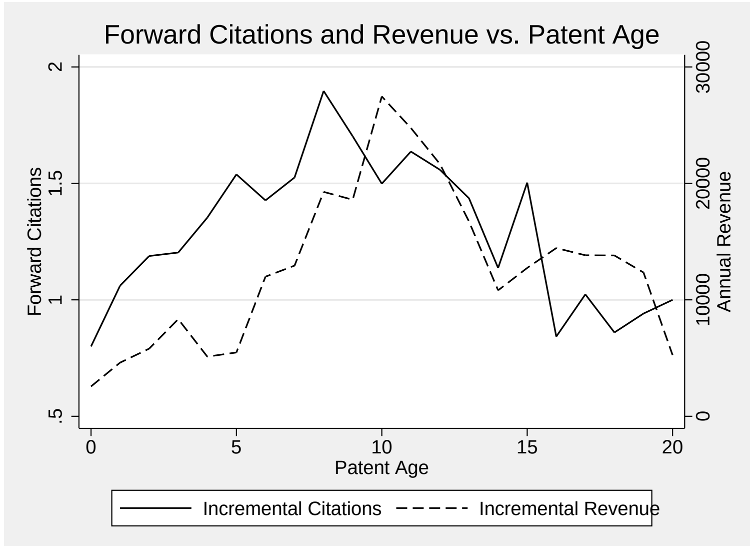


Figure 13: Incremental forward citations and revenue by patent age

Notes: Data is normalized so that the mean annual revenue is $10,000.

B 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).

C 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 3.2), we focus only on the relationship between the social value of strategic patents and their forward citations.[16]

2. Recall that the impact of the height of the incumbent's fence, m, in Section 3.2 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 3.2. 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

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

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

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

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

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 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} :

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

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

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

Proof. See Appendix E. ∎

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 3.2. 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.

D 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 cluster is associated with an opportunistic patent that is unkown ex-ante. Technology cluster 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 cluster ($m = 0$). 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. NPE keeps collecting these fees until the technology cluster is made obsolote by a radical entrant labeled N_5^p. In the new technology cluster (tech cluster 2), an opportunistic patent N_2^o is drawn with strength $m' \sim \mathcal{M}([0,1])$, and so on.*

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

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

$$
\begin{array}{r|cccc|cc}
\text{Opportunistic patent ID:} & \multicolumn{4}{c|}{\overbrace{\hspace{8cm}}^{N_1^o \text{ with strength } m}} & \multicolumn{2}{c}{\overbrace{\hspace{4cm}}^{N_2^o \text{ with strength } m'}} \\
\text{NPE payoff via } N_m^o: & m\pi & m\pi & m\pi & m\pi & m'\pi & m'\pi \\
\text{Incumbent payoff via } N_n^p: & \underbrace{(1-m)\,\pi} & \underbrace{(1-m)\,\pi} & \underbrace{(1-m)\,\pi} & \underbrace{(1-m)\,\pi} & \underbrace{(1-m')\,\pi} & \underbrace{(1-m')\,\pi} \\
\text{Productive patent ID:} & N_1^p & N_2^p & N_3^p & N_4^p & N_5^p & N_6^p \\
\text{Tech Cluster ID:} & \multicolumn{4}{c|}{\underbrace{\hspace{8cm}}_{\text{Tech Cluster 1}}} & \multicolumn{2}{c}{\underbrace{\hspace{4cm}}_{\text{Tech Cluster 2}}}
\end{array}
$$

An example of a sequence of innovations in a product line

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 cluster, 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 cluster, 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 cluster 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.

E 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

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

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

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

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

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

Equating the coefficients we get

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

This completes the proof. ■