Abstract

Intuition suggests that more valuable patents are cited more, and this is a standard assumption in the empirical literature. However, using a proprietary dataset with patent-specific revenues we find that the relation of cites to value is an inverted U, with fewer cites at the high end of value than 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 defensive patents. We find evidence of greater use of defensive patents where it would be most expected: among corporations, in fields of rapid development, in more recent patents and where divisional and continuations 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.

[STILL PRELIMINARY, COMMENTS WELCOME]


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 1965; Griliches 1979) 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 last 30 years, two primary metrics have been used to proxy for the value of innovation, patent counts and citation-weighted patent counts. The intuition is simple: fields with greater innovative activity will have more value to protect and will do so by applying for more patents. Weighting patent counts by forward citations\(^1\) is a natural augmentation to simple patent counts, given the well-known fact that patents vary tremendously in value\(^2\). The use of this measure, however, is based on the assumption that a larger number of citations corresponds to higher value.

![Forward Citations vs. Patent Value](image)

Figure 1: LIFETIME FORWARD CITATIONS VS. PATENT VALUE

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

---

\(^1\)Forward citations is the number of citations received by a particular patent by subsequent patents.

\(^2\)Fewer than 10 percent of patents are worth the money spent to secure them (Allison, Lemley, Moore, Trunkey 2009), but the most valuable ones are thought to be worth hundreds of millions of dollars (Hall, Jaffe, and Trajtenberg 2005).
Until now there has been no good way to test this assumption. Several problems have held back this inquiry: the reluctance of companies to share proprietary patent data, the lack of generality (and sufficient observations) from any single patent portfolio, 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 patents, we are able to examine the patent value - citation relationship empirically in a level of detail that was previously impossible.

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. While there is still an overall positive correlation between citations and value, it comes primarily from lower value patents, and the full pattern is more complex. This striking finding calls for a deeper understanding of the process of innovation, patenting, and citations, which we explore empirically and theoretically in this paper.

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 defensive. Productive innovative effort leads to the traditional increasing relationship between patent value and citations; defensive innovative effort, however, leads to a negative relationship. Defensive 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.

In order to further test our theory of defensive patents, we examine the citation-value relationship using four characteristics that should be related to defensive patenting. Defensive 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 defensive patenting is more prominent in these categories.

This is certainly not the first paper that has attempted to examine the relationship between patent value and citations, but it is the first not severely constrained by data availability, for the reasons mentioned above. Trajtenberg (1990) is perhaps the leading prior work on the subject, and was the first paper to rigorously examine the citation-value relationship. This paper focuses on a relatively small number of patents in the computed tomography field, with

---

3 We discuss details of the data sets in Section 2. Confidentiality agreements limit our ability to disclose actual revenue numbers or number of observations.

4 Further details on the normalization and other aspects of the production of this figure are discussed in Section 4.
values imputed from a structural model of the CT device market. He finds an approximately linear relationship between citations and patent value. Harhoff et al. (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 found that the number of citations received is correlated with the patent’s value, but that the relationship is quite noisy. For example, patents in the $1 million value category elicited 8 forward citations on average but with a two standard error range from 0.7 to 90 citations. Several excellent studies examine the patent value distribution using the patent renewal decision to infer value (Pakes 1986; Schankerman and Pakes 1986; Bessen 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 an individual patent and thus is not informative about the citation-value relationship.

In the legal literature, defensive 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 defensive 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 defensive 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 linear 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 efforts. The introduction of defensive patents adds to prior models that generally have a single type of patent. The defensive 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 defensive disruption is important for corporate finance scholars interested in valuing firms’ intellectual property assets. Both our empirical and theoretical findings show that surprisingly not every patent leads to creative destruction and economic growth rather some of them are defensive 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 B contains theoretical proofs and derivations and Appendix B
contains additional data descriptions.

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 char-
acteristics in some detail. The data in this paper was provided by large non-practicing entities
(NPEs), with a focus on the technology sectors. NPEs are firms whose revenue primarily de-
rives not from producing products based on patented technology, but from licensing patents.
These companies employ a range of different business models ranging from aggressive litiga-
tors 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.

Before presenting summary statistics, it is important to note several distinctive characteris-
tics 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 B) we have also normalized the
data such that mean annual revenue is $10,000. 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 rela-
tionship or other correlates of value. Appendix B also discussed 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, we restrict the data accordingly.

<table>
<thead>
<tr>
<th>Table 1: SUMMARY STATISTICS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Patent Value ($000s)</td>
</tr>
<tr>
<td>Lifetime Forward Citations</td>
</tr>
<tr>
<td>Backward Citations</td>
</tr>
<tr>
<td>Fraction of Backward Cites in Past 3 Years</td>
</tr>
<tr>
<td>Fraction of Backward Cites in Past 5 Years</td>
</tr>
<tr>
<td>Original Indicator</td>
</tr>
<tr>
<td>Application Year</td>
</tr>
<tr>
<td>Individual Inventor Indicator</td>
</tr>
</tbody>
</table>

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.

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 $251,720 (all figures are 2010 dollars). The interesting thing to note is that the standard deviation of $1.8 million is more than 7 times the mean and more than 27 times the median value of $66,000. 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 with 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.

Our other main variable of interest, lifetime forward citations, also has a skewed distribution with a mean of 29, standard deviation of 53 and median of 1. This degree of skewness in the distribution of forward citations, the very wide range of forward citations, and the concentration of patents with 1 or fewer citations replicate familiar patterns such as those reported in Trajtenberg (1990); Harhoff et al. (1999); and Hall, Jaffe, and Trajtenberg (2005). In the sample of 456 CT scanner patents analyzed by Trajtenberg (1990), the mean number of citations is 2, the standard deviation is 5, and the median is 1. In a sample of 192 U.S. patents
analyzed by Harhoff et al. (1999), the mean is 16, the standard deviation is 21, and median is 9. Hall, Jaffe, and Trajtenberg (2005) analyzing the patents held by 12,118 public firms report a mean of 8, a standard deviation of 7, and a median of 6.

Backward citations are also skewed, with a mean of 23, median of 8 and standard deviation of 60. About 46% of backward citations are for patents issued within the prior 3 years and 64% 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 (68%) patents are original applications and the remainder are divisionals or continuations. Under the law in the U.S., 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. 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 2000, meaning the median patent had about 8 years of protection left by the end of our revenue data. Individual inventors account for 14% of the patents. This figure is similar to that reported in Bessen (2008).

In Table 2 one may see that value and forward citations vary substantially by technology class. The most valuable patents are found in the Circuits category with a mean value of $367,130, but only an average of 7.1 citations. Computer Architecture patents also have a high average value at $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>
<thead>
<tr>
<th>Technology</th>
<th>Patent Value</th>
<th>Lifetime Forward Citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circuits</td>
<td>$367,130</td>
<td>7.1</td>
</tr>
<tr>
<td>Computer Architecture</td>
<td>$283,773</td>
<td>6.0</td>
</tr>
<tr>
<td>Internet &amp; Software</td>
<td>$273,093</td>
<td>12.6</td>
</tr>
<tr>
<td>Wireless Communications</td>
<td>$174,605</td>
<td>35.4</td>
</tr>
<tr>
<td>Network Communications</td>
<td>$146,974</td>
<td>9.4</td>
</tr>
<tr>
<td>Semiconductor Devices</td>
<td>$115,824</td>
<td>7.8</td>
</tr>
<tr>
<td>Peripheral Devices</td>
<td>$99,801</td>
<td>8.1</td>
</tr>
<tr>
<td>Electro-Mechanical</td>
<td>$62,018</td>
<td>7.4</td>
</tr>
<tr>
<td>MEMS &amp; Nano</td>
<td>$58,860</td>
<td>11.1</td>
</tr>
<tr>
<td>Optical Networking</td>
<td>$56,425</td>
<td>16.5</td>
</tr>
</tbody>
</table>

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

In our subsequent theoretical and empirical analyses, where we attempt to provide a theo-

---

5These classifications are not part of the Patent and Trademark Office classification system but rather are used by the portfolio owners.
tical foundation for the inverted U-shape in the data, we focus on a few variables characterized by productive and defensive innovations. Intuitively, this distinction suggests that non-original and less productive patent applications with a higher concentration of backward citations in recent years are more likely to be strategic or defensive patents. Around 16% of the patents in our sample are non-original\textsuperscript{6} and only 20% of the backward citations are in the recent past.

3 Theory of Patent Valuations and Citations

In the previous section, we provided a striking 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.

While building our theoretical 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 help us rationalize the non-monotonic relationship between patent value and subsequent entry, identified by forward citations.\textsuperscript{7} Our model emphasizes the decision to innovate productively or defensively.

Our model features two distinct types of innovation efforts – productive and defensive. The intuition for productive innovation follows the traditional economic view that patents are offered as a contract between society and the inventor. In return for a limited period of exclusivity, the inventor agrees to make his invention public rather than keeping it secret. This institutional arrangement promotes the diffusion of ideas (spillover) and economic growth. New big ideas generate a higher profit for the original inventor and also generate more spillovers for subsequent innovations. Hence a positive relationship occurs between patent value and subsequent entry (forward citations). However, this is likely not the full story. Therefore, we also introduce the notion of the defensive innovation, a type of destructive creation. This idea seeks to capture the fact that when firms and individuals are endowed with a complex

\textsuperscript{6}Within the intellectual property legal framework, an “original” patent is an application that establishes its own filing date and does not have an effective filing date based upon another previously filed application. If an “original” application is then used to establish an effective filing date of a later filed application, it becomes known as a parent application and the later filings are either divisions or continuations. There can be many strategic advantages to non-original patents if the first-to-file is important or if one desires to prolong the original patents disclosure.

\textsuperscript{7}Relatedly, models presented by 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.
legal instrument, 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 instrument in the first place. Hence a valuable defensive innovation is the 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 defensive 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 defensive 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 are positive for productive innovation efforts and negative for defensive 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 A.1 for the interested readers.

The economy features a large number of outside entrepreneurs who invest in productive innovations. These productive innovations enable the entrepreneurs to innovate, to replace existing incumbents, and to obtain 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, certain markets become hot and attract all the 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, it is also consistent with empirical evidence (Jaffe and Lerner 2004). In terms of the model, what follows from this logic is an endogenous-citation dynamic.

The positive link between the citations and patent value comes from the fact that more novel innovations will have larger mark-ups due to their originality, denoted by the step size of a new innovation. In the model, this then translates into larger patent values. At the same time, more novel innovations will generate larger spillovers for the subsequent innovations, which will encourage new innovations by outside entrepreneurs. With more entrepreneurs entering

---

8Household’s saving decision pins down the equilibrium interest rate in this model and provided for completeness.
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 patent citations and patent
value. Given the intuition and logic underlying this first model of productive innovation, we
now turn to the details.

Basic Environment  Consider the following continuous time economy that admits a represen-
tative household. The household consumes a unique consumption basket $C_t$ that consists
of large set of varieties indexed by $j \in [0, 1]$ as follows:

$$ C_t = \exp \int_0^1 \ln c_{jt} dj, $$

(1)

In this expression, $c_{jt}$ is the quantity of variety $j$ at time $t$. We normalize the price of the
final good $C_t$ to be 1 in every period without loss of generality. The consumption basket 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

$$ c_{jt} = q_{jt} l_{jt} $$

(2)

where $l_{jt}$ is the labor employed for production and $q_{jt}$ is the variety-specific labor productivity.
In what follows, new innovations will improve labor productivity, which leads to an aggregate
growth in this economy. The linear production function implies that the marginal cost of
producing 1 unit of $c_{jt}$ 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 in the economy, 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 $\eta_n$ is received during $(t, t + \Delta t)$. Then the labor productivity evolves as:

$$q_{jt + \Delta t} = (1 + \eta_n) q_{jt}.$$  \hfill (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 consumption basket in (1). Because the consumption basket has a Cobb-Douglas form with respect to all varieties, the household will spend the same amount $C_t$ on each variety $j$. Hence the demand for each variety $j$ can be expressed as

$$c_{jt} = \frac{C_t}{p_{jt}}$$

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 $\eta_n$, 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

$$\pi_t (q_{jt}) = [p_{jt} - M_{jt}] c_{jt}$$

$$= \pi_n C_t$$

where we define $\pi_n \equiv \frac{\eta_n}{1 + \eta_n}$ as the normalized profit ($= \pi_t (q_{jt}) / C_t$). This is the first step in establishing the value of an innovation. Because a new innovation grants a 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 $C_t$ and $w_t$.

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

$$C_t = Q_t$$
where $Q_t$ is defined as a productivity index

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

Moreover, the wage rate is equal to

$$w_t = Q_t \int_0^1 (1 + \eta) \, dj.$$  

### 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. Outside entrepreneurs invest in R&D to produce a new innovation stochastically. When they are successful, they improve the latest quality as in (3). However 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 re-starts a new technology cluster with a step size $\eta_0 = \eta > 0$. Alternatively, if a new follow-on innovation occurs, it directly builds on the existing technology and the marginal contribution of this new innovation depends on how exploited the technologies are within the same technology cluster. In other words, 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:

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

Since innovations come in technology clusters and that each new innovation utilizes the spillover from the previous patents from the same technology class, our model generates a natural interpretation of citations. When there is a major innovation in a technology class with a step size $\eta$, it produces spillovers for the subsequent innovations since the follow-on step size becomes

---

Note that in principle, we can allow the step size $\eta_j$ to be a function of the sector $j$. This would not have any major impact on the inverted-U relationship that our model predicts.
which encourages new entry into the field. Innovations must cite previous innovations within the same technology cluster, acknowledging that the patents are technologically related. Therefore, patents from the same technology cluster will cite the initial major patent that opened the field. The following example will elaborate this structure further.

**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 who in those clusters. The following chart illustrates an example of some innovation patterns in a single product line:

<table>
<thead>
<tr>
<th></th>
<th>( P_1 )</th>
<th>( P_2 )</th>
<th>( P_3 )</th>
<th>( P_4 )</th>
<th>( P_5 )</th>
<th>( P_6 )</th>
<th>( P_7 )</th>
<th>( P_8 )</th>
<th>( P_9 )</th>
<th>( P_{10} )</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tech Cluster 1</td>
<td>( \eta ) ( \eta \alpha ) ( \eta \alpha^2 \eta \alpha^3 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tech Cluster 2</td>
<td></td>
<td>( \eta ) ( \eta \alpha )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tech Cluster 3</td>
<td></td>
<td></td>
<td>( \eta ) ( \eta \alpha ) ( \eta \alpha^2 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tech Cluster 4</td>
<td></td>
<td></td>
<td></td>
<td>( \eta )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

An example of a sequence of innovations in a product line

Example starts with a radical innovation \( P_1 \) which has a step size \( \eta \). Then innovation \( P_2 \) follows on \( P_1 \) with a step size \( \eta \alpha \). Since \( P_3 \) is the second follow-on innovation in cluster 1, it has a step size \( \eta \alpha^2 \) and so on. Note that \( P_5, P_7, \) and \( P_{10} \) turn out to be a radical innovations which start new technology clusters; therefore their step sizes are \( \eta \). As a result, innovation step sizes follow cycles. Finally, the citing-cited pairs can be summarized as follows:

<table>
<thead>
<tr>
<th>Cited</th>
<th>Citing</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_1 )</td>
<td>( P_2, P_3, P_4 )</td>
</tr>
<tr>
<td>( P_2 )</td>
<td>( P_3, P_4 )</td>
</tr>
<tr>
<td>( P_3 )</td>
<td>( P_4 )</td>
</tr>
<tr>
<td>( P_4 )</td>
<td>none</td>
</tr>
<tr>
<td>( P_5 )</td>
<td>none</td>
</tr>
<tr>
<td>( P_6 )</td>
<td>( P_7, P_8, P_9 )</td>
</tr>
<tr>
<td>( P_7 )</td>
<td>( P_8, P_9 )</td>
</tr>
<tr>
<td>( P_8 )</td>
<td>( P_9 )</td>
</tr>
<tr>
<td>( P_9 )</td>
<td>none</td>
</tr>
<tr>
<td>( P_{10} )</td>
<td>...</td>
</tr>
</tbody>
</table>

Consider \( P_2 \), for instance. Since it builds only on \( P_1 \), \( P_2 \) cites only \( P_1 \). However, there are two patents \((P_3, P_4)\) in the cluster that are building on \( P_2 \). Hence, \( P_2 \) receives two citations from them.

Now we can 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 \( \tilde{z}_{n+1} \) and the next radical innovation by \( \tilde{z}_0 \). Then the steady-state value of the \( n^{th} \) innovation is summarized by the following continuous time Hamilton-Jacobi-Bellman (HJB) equation

\[
V_{nt} = \frac{\eta_n}{1 + \eta_n} C_t \Delta t + (1 - r \Delta t) \left[ (\tilde{z}_0 \Delta t + \tilde{z}_{n+1} \Delta t) \times 0 + (1 - \tilde{z}_0 \Delta t - \tilde{z}_{n+1} \Delta t) V_{nt+\Delta t} \right].
\]
This expression is intuitive. During a small $\Delta t$, $n^{th}$ innovation in a cluster delivers a profit of $\frac{\eta_n}{1+\eta_n}C_t\Delta t$ to its owner. The future period is discounted by $(1-r\Delta t)$. After $\Delta t$, 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 $(1-\bar{z}_{n+1}\Delta t-\bar{z}_0\Delta t)$, the incumbent survives the threat of entry and receives the continuation value $V_{t+\Delta t}$ of being the incumbent. Subtracting $(V_{nt+\Delta t} - r\Delta t V_{nt})$ from both sides, dividing through $\Delta t$, and taking the limit $\Delta t \to 0$ leads to the following HJB equation:

$$rV_n - \dot{V}_n = \pi_n C_t - (\bar{z}_{n+1} + \bar{z}_0) V_n.$$  \hspace{1cm} (5)

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

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

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

**Proof.** This result follows from using the household’s Euler equation $r - g = \rho$ in (5). The Euler equation itself is derived in Appendix A equation 9. This expression simply says that the value of an innovation depends mainly on four factors: First, a larger step size $\eta_n$ implies larger mark-up and therefore higher innovation value. Second, if the aggregate consumption $C_t$ is larger, each variety will receive a larger demand and hence generate higher per-period profit and innovation value. Third, present discounted value of future profits depends on growth rate adjusted interest rate $r - g$, which boils down to the discount rate $\rho$ through the household problem in equation 9. 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 shorter expected duration of monopoly power.

So far, we 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 innovation rate of an individual entrepreneur and $\bar{z}_n$ denote the aggregate innovation rate by the outside entrepreneurs who are trying to innovate in the same product line $j$. We assume that there are some 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 ζ > 0 is some constant. Then the free-entry for a new entrant can be summarized as
\[ \max_{z_n} \{ z_n v_n C_t - z_n \zeta Q_t \bar{z}_n \}. \]
Free-entry condition pins down the aggregate entry rate as
\[ \bar{z}_n = \frac{v_n}{\zeta}. \quad (7) \]
As expected, entry rate is increasing in the value of a new innovation and decreasing in the cost parameter ζ.

Now combining this last expression (7) with (6) 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
\[ \bar{v} = \frac{\zeta}{2} \left[ \sqrt{\left( \rho + \frac{v_0}{\zeta} \right)^2 + \frac{4}{\zeta^2} \pi_n^*} - \left( \rho + \frac{v_0}{\zeta} \right) \right]. \]

Here are the main results emerging from this first model:

**Proposition 1** The average number of forward citations received by an \( \eta_n \) patent during any time interval \([t_1, t_2]\) decreases in \( n \).

**Corollary 1** Hence, in the case of productive patent, patent value and forward citations are positively correlated.

The intuition behind this result is very straightforward: when a new path-breaking innovation occurs, it creates a new technology cluster which then generates spillovers for the 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. On the x-axis, we list the patent valuations and on the y-axis, we have the number of citations that was received by each patent.

### 3.2 The Case of Defensive 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 doing defensive innovations by the incumbents to secure their position. The idea is that if an incumbent has a
Figure 2: MEAN LIFETIME FORWARD CITATIONS VS. PRODUCTIVE PATENT VALUE

high value productive innovation, then she can potentially invest in some defensive innovation in order to make it harder for the next outside entrepreneur to leapfrog and steal the high monopoly rents. If a defensive patent is very successful from the patenting firm’s point of view, the probability of being invented on would be very small which would increase the value of a defensive innovation and decrease the expected number of citations received due to lack of entry. Hence, we should expect a negative relationship between patent valuations and citations in the case of defensive patents.

Formally, upon each productive innovations, an incumbent has also the opportunity to do a single defensive innovation. The technology for defensive innovation is such that by paying a fixed cost $\psi > 0$, and new entrant who just invented a productive patent can also obtain a defensive patent. To simplify the analysis, assume that $\psi$ is high enough such that it is profitable to invest in defensive innovations only for the radical inventors (i.e., inventors with step size $\eta$). When a firm does defensive 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) such that the cost of the next outsider is

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

Consider the value $v_m^d$ of an $m$–type defensive patent. Since a defensive patent is done only
by radical inventors, the profit collected every instant is $\pi_0$. Therefore the HJB equation is simply $\rho v^d_m = \pi_0 - (\bar{z}_{0m} + \bar{z}_{1m}) v^d_m$. This value function is expressed as

$$v^d_m = \frac{\pi_0}{\rho + \bar{z}_{0m} + \bar{z}_{1m}}. \quad (8)$$

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

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

which implies

$$\bar{z}_{nm} = \frac{v_{nm}}{\zeta m}.$$  

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

Next, combining this entry rate with (8) we get the value of a defensive patent of type $m$:

$$v^d_m = \frac{\pi_0}{\rho + \frac{v_{nm}}{\zeta m} + \bar{z}_{1m}}.$$ 

Now we have the new results.

**Proposition 2** The value of defensive patents increases in $m$. 

Figure 3: MEAN LIFETIME FORWARD CITATIONS VS. DEFENSIVE PATENT VALUE
Figure 4: MEAN LIFETIME FORWARD CITATIONS VS. PATENT VALUE

**Proposition 3** The entry rate (forward citations) decreases in $m$.

**Corollary 2** Hence, in the case of defensive patents, patent value and forward citations are negatively correlated.

Clearly, the underlying reason for this negative relationship stems from the fact that more successful defensive patents are the ones that increase the cost of entry the most (high $m$). When this is the case, the subsequent number of forward citations will decrease due to lower entry. In the meantime, lower entry rate means that the current incumbent can enjoy the monopoly power longer, which raises the value of the defensive patent. Hence we get the negative relationship between defensive patent valuation and citations, as illustrated in Figure 3.

Figure 4 illustrates the overall relationship between patent value and citations. The pattern is a very clear inverted-U, and repeats what we observe empirically in the data. The basic intuition is that the most valuable innovations, which are the radical ones that leapfrog the competition, are owned by entrepreneurs, who are willing to also invest a fixed cost to defensively innovate in order to protect their productive innovation. The combination of the radical, productive innovation and the defensive innovation is very valuable, but because the defensive innovative alters the entry rate of new entrepreneurs through our 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 a less valuable technology, but rather it is due to a more costly entry rate for new entrepreneurs.

Our model suggests that incumbents with high-value patents will rely on defensive 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. Moreover, the model implies that the defensive 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. The next section will test these predictions.

4 Empirical Results

We have seen how productive and defensive 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.

Figure 1 displays the empirical relationship between forward citations for the data set described in Section 2. In order to make the visualization easier, we split the data into patent value percentiles and plot the mean number of citations for each value percentile. The figure shows an increasing relationship between patent value and citations for values under $150,000 and then a decreasing relationship for values above this threshold: the inverted-U that our model of two types of innovation predicts. In Table 3 we report results of regressions testing this relationship. Each column is a separate regression of forward citations on a function of patent value, with no controls. The three 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, with no controls. The three 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 scores of recent papers. 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.

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 1 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

10The figure excludes the 5% of highest value patents since the long right tail of the value distribution obscures the important relationship in the bulk of the data.
Table 3: FORWARD CITATIONS vs PATENT VALUE

<table>
<thead>
<tr>
<th>Share of most valuable patents excluded</th>
<th>10%</th>
<th>5%</th>
<th>1%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patent Value ($100,000s)</td>
<td>9.047**</td>
<td>22.497**</td>
<td>7.104**</td>
</tr>
<tr>
<td></td>
<td>(0.256)</td>
<td>(0.654)</td>
<td>(0.232)</td>
</tr>
<tr>
<td>Patent Value Squared</td>
<td>-6.036**</td>
<td>-2.193**</td>
<td>-0.139*</td>
</tr>
<tr>
<td></td>
<td>(0.288)</td>
<td>(0.195)</td>
<td>(0.070)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.04</td>
<td>0.05</td>
<td>0.04</td>
</tr>
</tbody>
</table>

** Significant at the 1% level; * Significant at the 5% level
Note: Separate regressions reported in each column, with 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.

in Figure 5 is apparent that was previously seen in Figure 1, although here the relationship is a bit noisier. The takeaway is that there is still compelling evidence for productive and defensive patenting, even within technology categories, and accounting for inventor type and original status, although these variables also impact the observed relationship to some extent.

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 patent value and patent value squared that include the covariates listed above plus one addition: dummies for inventor type (individual or company), whether the patent was original (not a continuation or divisional), technology category, and whether the patent was applied for prior to 2000. 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, even when including controls. Individual inventor status has a substantially negative impact on number of forward citations. Earlier patents appear to have more lifetime citations - this should be a cohort effect, as patent age is accounted for in the normalization methodology (see Appendix B). Original patents tend to receive fewer citations than continuations or divisionals.

Table 4: FORWARD CITATIONS AND PATENT CHARACTERISTICS

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patent Value ($100,000s)</td>
<td>7.569**</td>
<td>9.272**</td>
<td>8.669**</td>
<td>8.444**</td>
</tr>
<tr>
<td></td>
<td>(0.622)</td>
<td>(0.637)</td>
<td>(0.631)</td>
<td>(0.615)</td>
</tr>
<tr>
<td>Patent Value Squared</td>
<td>-0.906**</td>
<td>-1.254**</td>
<td>-1.213**</td>
<td>-1.130**</td>
</tr>
<tr>
<td></td>
<td>(0.205)</td>
<td>(0.206)</td>
<td>(0.206)</td>
<td>(0.201)</td>
</tr>
<tr>
<td></td>
<td>(0.388)</td>
<td>(0.385)</td>
<td>(0.406)</td>
<td>(0.399)</td>
</tr>
<tr>
<td></td>
<td>(0.332)</td>
<td>(0.330)</td>
<td>(0.332)</td>
<td>(0.332)</td>
</tr>
<tr>
<td>Indicator Original Patent</td>
<td>-7.583**</td>
<td>-5.384**</td>
<td>(0.682)</td>
<td>(0.659)</td>
</tr>
<tr>
<td>Tech Category (Computer Architecture)</td>
<td>3.632**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.565)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tech Category (Electro-Mechanical)</td>
<td>4.03**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.642)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tech Category (Internet &amp; Software)</td>
<td>19.87**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.872)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tech Category (MEMS &amp; Nano)</td>
<td>3.798**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.314)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tech Category (Networking &amp; Communications)</td>
<td>9.808**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.734)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tech Category (Optical Networking)</td>
<td>2.1**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.472)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tech Category (Peripheral Devices)</td>
<td>2.508**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.413)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tech Category (Semiconductors)</td>
<td>3.387**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.431)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tech Category (Wireless Communications)</td>
<td>7.22**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.524)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| $R^2$ | 0.12 | 0.12 | 0.13 | 0.16 |

** Significant at the 1% level; * Significant at the 5% level

Note: Separate regressions reported in each column, 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 $10,000 (2010$).

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.\textsuperscript{[11]}

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 defensive patenting. This decision will vary by observable characteristics of the innovator and we now test several implications of the model empirically.

Since there is a fixed cost to defensive patenting, and firms tend to have greater resources than individuals on average, we would expect individuals to engage in less defensive patenting than companies. We use the type of original assignee to test this hypothesis, and expect that corporations should be more represented in the "defensive" 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. 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.

Another patent characteristic that may influence likelihood of defensive patenting is the rate of innovative growth in the field. Areas of rapid innovation are likely to also generate greater

\textsuperscript{[11]}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).
current and expected profits, and thus greater incentive to engage in defensive patenting. Our measure for the growth of innovation is the share of backward citations within the prior several years. We would expect that patents with more recent backward citations are in fields of rapid growth, generating greater profits and thus incentives for defensive patenting.

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 defensive 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 defensive 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.

Finally, we examine whether there has been recent growth in the use of defensive patenting. There has been much written about the increasing use of patents for defensive 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 defensive 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 variation in the use of defensive patenting as well as overall citation practices and patent values. 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.
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 defensive. 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: defensive. Defensive 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 defensive patenting increases with patent value, which leads
Table 5: FORWARD CITATIONS AND PATENT VALUE BY TECHNOLOGY CLASS

<table>
<thead>
<tr>
<th>Technology Class</th>
<th>Patent Value ($100,000s)</th>
<th>Patent Value Squared</th>
<th>R^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circuits</td>
<td>6.233</td>
<td>(6.89)**</td>
<td>0.05</td>
</tr>
<tr>
<td>Computer Architecture</td>
<td>14.497</td>
<td>(11.28)**</td>
<td>0.09</td>
</tr>
<tr>
<td>Electro-Mechanical</td>
<td>10.917</td>
<td>(6.60)**</td>
<td>0.04</td>
</tr>
<tr>
<td>Internet &amp; Software</td>
<td>23.542</td>
<td>(10.95)**</td>
<td>0.05</td>
</tr>
<tr>
<td>MEMS &amp; Nano</td>
<td>17.051</td>
<td>(4.75)**</td>
<td>0.06</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Networking Communication</th>
<th>Patent Value ($100,000s)</th>
<th>Patent Value Squared</th>
<th>R^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communication</td>
<td>19.107</td>
<td>(8.64)**</td>
<td>0.08</td>
</tr>
<tr>
<td>Networking</td>
<td>13.496</td>
<td>(11.43)**</td>
<td>0.07</td>
</tr>
<tr>
<td>Optical</td>
<td>9.847</td>
<td>(14.64)**</td>
<td>0.02</td>
</tr>
<tr>
<td>Peripheral Devices</td>
<td>9.329</td>
<td>(9.60)**</td>
<td>0.06</td>
</tr>
<tr>
<td>Semiconductors</td>
<td>18.007</td>
<td>(12.04)**</td>
<td>0.07</td>
</tr>
<tr>
<td>Wireless Communications</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communications</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

** Significant at the 1% level; * Significant at the 5% level

Note: Separate regressions reported in each column, t-statistics in parentheses. Dependent variable is lifetime forward citations. Data is normalized so that the mean annual revenue is $10,000 (2010$).

To a negative relationship between citations and value for defensive patents. When we allow for both types of innovation, we expect productive patenting to dominate up to a point, after which defensive 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.

While we believe the data analyzed for this paper is the finest that has been used to understand patent value, it has limitations like all data sets. 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 should 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. 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. Further, given that much of the value of innovation is concentrated in technology this is perhaps the most important subset of the patent universe to study.
Figure 10: FORWARD CITATIONS VS. REVENUE, COMPUTER ARCHITECTURE PATENTS
Notes: Data is normalized so that the mean annual revenue is $10,000.

Figure 11: FORWARD CITATIONS VS. REVENUE, SOFTWARE PATENTS
Notes: Data is normalized so that the mean annual revenue is $10,000.
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 defensive patents.

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 or 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


Appendix

A Proofs and Derivations

A.1 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 $\rho > 0$. Household’s intertemporal maximization delivers the standard Euler equation

$$g_t = r_t - \rho.$$  \hfill (9)

B 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 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 A1 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 12: INCREMENTAL FORWARD CITATIONS AND REVENUE BY PATENT AGE

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