Patent Citations and Licensing Value

by

Gregory P. Daines

B.A., Brigham Young University (1994)
M.Phil., University of Cambridge (1996)

Submitted to the Sloan School of Management
in partial fulfillment of the requirements for the degree of

Master of Business Administration

at the

Massachusetts Institute of Technology

June 2007

Signature of Author: ______________________
MIT Sloan Fellows Program in Innovation and Global Leadership
May 11, 2007

Certified by: ______________________
Fiona Murray
Class of 1922 Career Development Associate Professor
Technological Innovation & Entrepreneurship
Thesis Supervisor

Accepted by: ______________________
Stephen J. Sacca
Director, MIT Sloan Fellows Program in Innovation and Global Leadership

(this page intentionally left blank)

Patent Citations and Licensing Value

by

Gregory P. Daines

Submitted to the Sloan School of Management on May 11, 2007
in partial fulfillment of the requirements for the degree of
Master of Business Administration

ABSTRACT

Innovation has become the dominant economic idea of our time and has resulted in a proliferation of innovation-oriented rhetoric, and policies. There is a great need to understand and harness innovation, but it has proven to be as difficult to measure as it is to define. At present, the quest to understand innovation hinges on finding reliable ways to identify and measure it, the most promising of which is the analysis of patent information.

Patents have been increasingly used by economists to track inventiveness, the transmission of knowledge, and their economic impact. However, it is evident that the majority of patents have little or no economic potential, and so merely observing the number of patents provides little insight on innovation. It has become important, therefore, to develop reliable methods for measuring the true economic potential of patents. Of all the solutions proposed, the analysis of patent citations is the most promising.

This study examines the relationship between patent citations and the private economic value of patents, and makes both theoretical and empirical contributions. First, the previous literature is reviewed to further extend and clarify the theory of the economic meaning of patent citations. Second, a typology of patent value is proposed to contextualize the relevance of the theory under different appropriation regimes. Finally, this study tests the economic meaning of citations using a new dataset where the licensing value of a group of patents is observed directly. The findings confirm a consistent relationship between patent citations and two different measures of patent value.

Thesis Advisor: Fiona Murray, Class of 1922 Career Development Associate Professor

ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to the many individuals that contributed in one way or another to this research. First, I wish to gratefully acknowledge Memorial Sloan-Kettering Cancer Center and Dana-Farber Cancer Institute for providing the crucial licensing data that formed the basis for this study. In particular, special appreciation goes to Steve Bertha at Sloan-Kettering and Anthony del Campo at Dana-Farber for their vision to see the value of this research.

Special thanks also go to Kyle Jensen for providing the essential US patent data, and to Julian Kolev for his extensive tutoring and help in transforming and formatting the data and for his support and advice in specifying the models. Thanks also to Tony Briggs, Bronwyn Hall, Tom Stoker, Bob Sturim, and many other members of the Sloan Fellows class of 2007 for guidance, feedback, and enthusiastic support.

I especially wish to thank my thesis advisor, Fiona Murray, for her vital intellectual contributions, her very practical help in marshalling all the pieces of the puzzle, and for her tremendous moral support throughout the project.

Finally, I must acknowledge the untiring faith and support of my parents, family, and above all, my wife Leilani. Like me, this simply would not exist without you.

Table of Contents

List of Tables

List of Figures

Introduction: Innovation and Patent Citation Analysis

Innovation has emerged as the dominant economic idea of our time (Nicholas, 1999). It has disrupted the entrenched economic debate by changing the focus from *who* controls the economy (eg. the market or the state) to *what* drives the economy. The result is an astonishing new consensus reflected everywhere in the proliferation of innovation-oriented rhetoric and policy initiatives. Economists have both driven and responded to this growing awareness as evidenced by the burgeoning literature seeking answers to myriad questions about the economic role and characteristics of innovation (Warsh, 2006). Despite important advances, these efforts have been hindered by a scarcity of data and several important methodological limitations (Trajtenberg, 1990b). Innovation has proven to be as difficult to measure as it is to define (Jaffe, 1998; Nelson & Winter, 1977; von Hippel, 1987).

This predicament ultimately led researchers to use patents and publications as a potentially rich supply of information for a variety of analyses (Griliches, 1990). However, patents have also turned out to be recalcitrant research objects. For one thing, the extent to which patents accurately track inventive technological progress is not fully understood and is limited because of the very large proportion of patents that appear to have little or no economic relevance (Griliches, 1990; Lanjouw, Pakes, & Putnam, 1998; Trajtenberg, 1990b). Yet, while it is clear that economic attributes of interest are highly skewed, it has proven impossible as yet to accurately identify and measure them. The difficulty is that there have been no opportunities to systematically observe the actual economic value of patents (Schankerman, 1998; Scherer, 1993; Trajtenberg, 1990b).

Several novel solutions to this puzzle have been developed, and unlocking the critical economic information in patent data has been a growing thrust of the literature since at least the late 1980's (Pakes, Hall, & Griliches, 1988; Pakes & Simpson, 1989; Schankerman & Pakes, 1986). Of the solutions, the use of citations to trace the economic meaning of patents has emerged as the most promising (Hall, Jaffe, & Trajtenberg, 2001). As the patent data have become electronically encoded and the software to analyze them more sophisticated, citations have become accessible for econometric analysis. In particular, the creation of the NBER Patent Data File in the 1990s was a major breakthrough in enabling a new generation of citation studies (Hall et al., 2001).

While there are important institutional distinctions, patent citations serve a broadly similar purpose as citations in academic literature (Jaffe & Trajtenberg, 2002). By law, every patent is required to make reference to, or cite, previous inventions on which it is built and which are publicly known – referred to as "prior art" (Hall, Jaffe, & Trajtenberg, 2000). Citations are included by the authors of patents and are also inserted by patent examiners as part of the review of each application (Alcacer & Gittelman, 2004, 2006). As a result, patents often have more citations when they issue than when they are originally filed (Alcacer & Gittelman, 2006). The citations link patents together and are used primarily as a way of establishing and bounding what is and what is not novel about the subject matter of the new patent (Sampat, 2005).

The underlying economic intuition for studying the citations received by patents is that a relatively greater number of citations from other patents implies that the patent has had a greater technological impact (Griliches, 1990). The economic value of the patent is seen as a function of its application, with a greater number of citations implying a relatively larger number of applications. In this view, no distinction is made between public and private economic value. A number of studies have looked at citations in this generalized way (Henderson, Jaffe, & Trajtenberg, 1998; Trajtenberg, 1990b).

The patent system exists, among other things, to enable patent holders to capture a certain portion of the public value of inventions, to motivate investments in the inventive process, and to reward inventors, as well as to create incentives for disclosure. Therefore, it is of interest to determine the degree to which the number of citations a patent receives specifically tracks its economic value in much the same way that we would like to show that a highly cited scientific publication has contributed more substantially to our knowledge than a less cited publication. A brief review of the theory and literature on this topic is the subject of Chapter 1.

To the degree that a link can be established between citations and private economic value, a further step is needed to understand the specific types of private economic value that patents create. Patents are actually complex economic entities that generate value through a number of different mechanisms. Characterizing these modes is necessary for understanding the meaning of citation data, and is the subject of Chapter 2.

In the absence of direct measures of patent value, a number of proxies have been tested, but no systematic attempt has yet been made to relate them to the specific mechanisms by which patents create value. Chapter 3 uses a typology of the private economic value of patents developed in Chapter 2 to contextualize these methods. Chapter 4 describes the dataset compiled for this study and presents the design of the empirical analyses. Finally, Chapter 5 presents the analysis and findings of an original study in which the relationship between citations and value is tested on a dataset where one type of private economic value is observed directly, and discusses further implications of this research.

Chapter 1: Citations and Patent Value

Patents are a kind of economic good or asset, and as such, embody or enable the creation of value (Griliches, 1990). The traditional conception of the link between patents and value creation for the economy is quite old, although it has had many incarnations. A common version of the idea is that a government guaranteed monopoly establishes investments in technological development which would otherwise be sub-optimal in social terms (Jaffe, 1998). Private returns are seen as the natural outcome of this approach and indicative, to some degree, of the unmeasured social benefits. From another perspective, patents offer the promise of abnormal private rents for the purpose of enabling enhanced social benefits. The portion of the total economic value generated by the patent system that is captured by patent holders as private value is not known, and it represents an important economic question (Sampat & Ziedonis, 2004).

Citations have been used in a wide range of studies to estimate or proxy for the economic value of patents (see Figure 1 in Chapter 3 for a listing of relevant studies organized by their approach). In some cases, the distinction between social and private value has been explicit (Sampat & Ziedonis, 2004), but in many cases it has not (Guellec & Potterie, 2000). Many studies have looked at the value of patents without distinguishing whether the value that is being measured is the entire economic value, the public value enjoyed by all, or the private value captured by patent holders (Griliches, 1990; Jaffe, 1998). This has sometimes led to research findings which are difficult to interpret, or to uses of citation counts which have not been validated (Gay, Bas, Patel, & Touach, 2005). For instance, in studies demonstrating evidence of a relationship between citations and the likelihood of patent litigation, citations are used as the proxy for value (Harhoff & Reitzig, 2004; J. O. Lanjouw & M. A. Schankerman, 1997). However, it is possible to hypothesize that both above average forward citations and the presence of litigation are to some degree measurements of the same thing – competing interests in a particular technology. Also, citations to prior art are often inserted by attorneys and patent examiners and do not necessarily indicate the existence of a direct inventive contribution. Therefore, citations may merely indicate increasing patent activity in a specific field (Sampat & Ziedonis, 2004).

Given the widespread use of citation counts as a stand-in for patent value, there is a need to develop methods for validating the relationship between citations and economic value, and particularly private economic value. The use of citation counts as an indicator of value is valid only to the extent that citations actually track the phenomena of interest.

Theory of Citations and Patent Value

Validating the relationship between citations and patent value has been a major thrust of a number of studies (Chapter 3 provides an overview and discussion of this literature). There are two dimensions to this challenge. The first need is to develop a clear theoretical foundation for the relationship between citations and economic value, a topic insufficiently treated in the literature (Sampat & Ziedonis, 2004). The second need is for data and techniques to empirically validate the relationship.

The basic theory underlying the idea that citations contain information about value as described above was examined extensively by Trajtenberg (1990a). Harhoff et al. (1999) describe two distinct theoretical justifications. The first is the concept described by many (Griliches, 1990; Jaffe, Trajtenberg, & Fogarty, 2000; Lerner, 1994; Trajtenberg, 1990a) that a patent receiving relatively more citations has greater scope, or application in a broader range of areas. The disadvantage of this concept is that it conceptualizes value in a generalized way. This may not handicap the use of citations in broad studies of knowledge flows and externalities but fails to address the question of economic value so important to other research.

A more direct approach would be to show that citations are linked to measurable economic outcomes. The first study to do so was developed by Trajtenberg (1990a) in which he showed that citations are highly correlated with estimates of the social value of the patented inventions. Harhoff et al. (1999) demonstrated a relationship between citations and estimates of value from a survey of German patent holders. Hall (Hall, 1999) presented evidence of a relationship between citations and the market value of firms, the results of which have been replicated by others (H. Shane & Klock, 1997).

A second theory posits that citations reflect the impact a patented invention has on stimulating further inventions (Harhoff, Narin, Scherer, & Vopel, 1999). In a slightly different version, citations demonstrate entry into economically valuable new technological areas (Trajtenberg, 1990b). These concepts are not only more difficult to demonstrate empirically, but also lack specificity. It is not clear whether these effects are purely externalities, or if there is some private value associated with them.

There is sufficient evidence at this point to conclude that citations contain information about both the public and private economic value of patents. There is a need, and an opportunity therefore, to advance both the theoretical and empirical basis for citations analysis. Sampat and Ziedonis (2004) recently attempted to clarify the theoretical issues by proposing four hypotheses for the way in which citations represent private economic value.

The first hypothesis, "Citations represent the portion of social returns appropriated by the patent holder," attempts to address the problem noted above that the basic theory of the relationship between citations and value lacks clarity as to what citations actually represent in terms of private value. This hypothesis implies that a significant number of the citations will come either from the patent holder (eg. "self citations"), or from licensees (Sampat & Ziedonis, 2004). In their study of licensing data, they find licensee citations account for a relatively small proportion (16%) of the citations, concluding that this is not an important effect. However, what constitutes a significant proportion for the purposes of the hypothetical value effect is unresolved. It is possible that relatively few self-citations may indicate this effect, and this may even be expected given that a single organization is likely to produce a limited number of subsequent patents. It is important also to remember that very high citation counts suggests citing by patents across broad technological fields (Lerner, 1994). Therefore, it is possible that self-citations can be expected to represent less breadth and therefore be associated with relatively fewer citations.

Sampat and Ziedonis' (2004) second hypothesis, "Citations reflect entry into profitable areas of research," clarifies Trajtenberg's (1990b) idea by adding that the citations result from patenting are, therefore, economically-induced. That is, the economic success of the patent will incentivize further patenting with a direct impact on citations. Their third hypothesis, "Citations indicate technological opportunities or market interest in a technological area," captures the idea that citations represent profit potential as they arise out of expectations of profit in fields with significant commercial activity. Finally, hypothesis four states, "Citations result from public disclosure." This refers to the citations that arise from the publicity a patent receives through commercial success. Although their results provide limited support for the second, third, and fourth hypotheses (and particularly the third), additional research is needed to explore this further (Sampat & Ziedonis, 2004). In particular, difficulties in reliably matching patents with companies limits the extent to which these hypothesized effects may be measured precisely.

It is also possible that the citations a patent receives as a result of the drafting and examination processes may create a "publicity" effect in which the patented subject matter becomes exposed to related patentees. In this hypothesis, citations represent the value a patent creates as a result of its exposure to firms operating in the same technological field. The private component of value in this mode would be expected to accrue primarily via licensing.

All of these hypotheses are attempts to identify the mechanisms by which citations indicate economic value. However, they are limited to the extent that they are not specific as to the kind of value with which they are associated. In reality, patents are economically complex and create value in a variety of different ways. Theories about the relationship between citations and value need to be clarified to take into account not only the public/private division of value, but also distinctions regarding the types of value. In fact, different types of patent value relate to citations in different ways, and this is the subject of the next chapter. This detailed discussion provides a foundation for subsequently developing a more comprehensive theory of the relationship between citations and value.

Chapter 2: The Private Economic Value of Patents

From a private perspective, there are a variety of reasons for patenting and thus a range of different ways in which holders of patents or patent rights realize value (Cohen, Nelson, & Walsh, 2000). These modes or mechanisms translate into a complex picture for the larger economy and raise a number of questions regarding the way that patents interact with and promote innovation in the economy. It is useful in this context to characterize these different mechanisms as a way of providing a framework for understanding and evaluating the literature on patent value. Specifically, the following discussion offers some key insights on the use and meaning of citations, and provides context for further theoretical and empirical work.

Types of Patent Value

Perhaps the most important key to unlocking the full potential of patent data as an economic tool has been establishing a link with economic value. This is partly because patents are known to be highly heterogeneous economic assets. That is, the value distribution of patents in highly skewed (Schankerman, 1998) with only about 1 to 3 out of every 100 patents yielding significant private economic returns (Sapsalis & Potterie, 2007). This renders the use of simple patent counts useless as measures of innovation or as indicators of the value that patents embody (Trajtenberg, 1990b).

At least in theory, such a link could be established using standard econometric techniques. However, this effort has been hampered by a number of factors. The first problem is that there haven't been opportunities to directly observe the value of patents

(Trajtenberg, 1990b). The reason for this reveals the true complexity of patents as economic entities. Patents can be used in a number of different ways to create value (Gans & Stern, 2003). In general, these modes either cannot be measured, or cannot be observed.

The following discussion presents a stylized typology of the different modalities that patents have for producing economic value. The purpose of this section is not to be exhaustive and rigorous, but to provide a simple framework for thinking about the ways that patents can lead to economic value, where that value accrues, and how it might be measured or estimated. The overriding objective is to clarify the specific challenges associated with producing direct measures of patent value, and to clarify the interpretation of previous studies.

1. Practice Value

The first and most obvious method for realizing economic value from patents is producing and selling products which apply the patented subject matter. Value derives from the producer's ability to exclude competitors and therefore earn monopoly rents (Jaffe & Lerner, 2004). This is relevant whether the patent enables a new product or feature, whether it provides efficiencies in production and distribution, or any of a wide range of other potential benefits to the producer. In theory, the value of a patent in this scenario could be seen as the net present value of the enhanced profits directly attributable to the existence of the patent monopoly position.

In practice, this type of value is virtually impossible to measure directly. For one thing, it is difficult to disaggregate the additional value that patent protection confers over and above the value derived from all other inputs. In other words, it is nearly impossible to say with any degree of certainty exactly how much value was attributable to the patent versus other factors such as superior branding, timing, market power, other intellectual property, existing manufacturing efficiencies, etc. Clearly this challenge is compounded when attempting to aggregate data for econometric analyses. Another difficulty comes from the fact that any given company will likely apply patented subject matter in one or at most a small number of selected applications. For some patents there may be many potential applications, and therefore the private returns will not properly measure the patent's full economic potential.

Another dimesnion of this type of value is "blocking", which is the practice of patenting substitutes for a strategically vital product exclusively for the purpose of preventing the production of substitutes[1] (Cohen et al., 2000). However, the value from this kind of patenting activity derives from practicing the core patents, and therefore should be viewed in this context merely as a method for strengthening patent protection in a practicing regime.

2. License Value

A second mechanism for realizing value from patents is to license the rights to the patented subject matter. It is important to note that licensing patents is a relatively new phenomenon, and one that has grown dramatically, especially over the past two decades (AUTM, 2004). Licensing offers the promise of significantly reducing the friction associated with patenting and innovating, but fluid markets for patent rights have not yet emerged for several reasons (Arora, Fosfuri, & Gambardella, 2000).

Nevertheless, the use of licensing revenues as a measure of patent value offers certain benefits. For one, licensing allows the full potential of the patented subject matter to be

[1] This strategy is also frequently described as creating a "patent fence".

exploited for all applications. Second, the majority of licensing revenues come from royalties which are calculated as a percentage of the licensee's sales of the product/service in which the licensed rights were applied (Thursby, Jensen, & Thursby, 2001). Because of this, license royalties also offer the potential to indirectly observe the value conceptualized under Section 1 above. In this view, the licensing value of a patent is merely a function of the value to the licensee of practicing the patent rights. Finally, license transactions potentially provide an opportunity to observe market dynamics relating to price, volume, etc. In theory, the value of licensing could be calculated as the NPV of licensing revenues for all application areas. In practice, these values are not observed due to the fact that the terms of licenses are generally treated as confidential.

3. Litigation Value

The third stylized economic mechanism for realizing value from patents is actually a derivative of both practicing and licensing and arises from its unique legal status. An entity that does not practice or license a patent can sometimes receive value by leveraging the special legal status of patents. In particular, some patent holders (often referred to as "patent trolls") often use patents as an asset in a threatened or real patent infringement suit (McDonough, 2006). In this case, the patent holder does not intend to practice the patent, but seeks value through settlements, licenses, or even damages. The reason for distinguishing this from the standard licensing model above is that this has become a unique and increasingly visible business model (Orey, 2006). To the extent that patents are filed, litigated, and transacted for the primary purpose of leveraging them via litigation, or the threat of litigation, it is important to recognize this as a distinct mechanism for appropriating value from patents.

Measuring this kind of patent value would theoretically encompass the settlements, patent suit awards and damages, and revenues from resulting licenses. However, compiling these in a systematic way is virtually impossible. Patent settlements and resulting licenses are typically confidential and not tracked systematically. Patent suit awards and damages are also not reliably reported.

4. Defensive Value

"Defensive" patenting refers to the practice of filing patents primarily for the purpose of providing a basis for counter-infringement claims in patent infringement litigation. In a sense, it is the opposite of litigation value. Defensive patenting has often been compared to idea of "Mutually Assured Destruction" from counter-balancing nuclear threats (Orey, 2006). As patent infringement litigation has accelerated, many companies have sought to expand their patent portfolios so that, in response to an infringement claim, they could look for counter infringement claims that they could bring against the opposing party to reduce the potential value of the original action (Cohen et al., 2000). In this way, companies with large portfolios have a resource that can be used to reduce the motivation for bringing patent infringement claims against them.

The justification for including this in this typology of patent value is the fact that it has become an increasingly important reason for filing patents, especially by major corporations (Shapiro, 2001). Obviously, the measurement of this kind of value is exceedingly difficult in practice as the very kind of value that is generated is largely unknown even to the parties involved.

With this patent value typology, we proceed in the next chapter to relate the different measures and estimates of patent value that have been developed in previous research to the types of patent value that they measure or proxy for.

Chapter 3: Measuring Patent Value

In the absence of available direct measures of patent value, researchers have developed a wide range of indirect measures or proxies of patent value (see Figure 1 below). The majority of these indirect measures relate implicitly to the first type of patent value creation above (practicing value), though this link is rarely made explicit (Harhoff et al., 1999). Because any firm can potentially exploit any or all of the modes for creating private economic value, and because most indirect measures do not allow distinguishing between them, the majority of studies do not attempt to differentiate between the particular value creation mechanisms.

This is an important point as it relates to previous studies of patent value. A consistent feature of these studies has been that they have treated patent value in a generalized way, and in some cases the concept of value has been indistinguishable from "quality" or "importance" (Henderson et al., 1998; Lanjouw & Schankerman, 1999). Obviously, the difficulty in this is that economic value, and particularly, private economic value is very different from concepts such as scientific merit, scope, quality, and breadth, though it may be in part a function of all of these (Lerner, 1994).

Moreover, based on the previous typology of patent value modes, it is probable that the value of any given patent or patent grouping will be different in each mode. That is, a patent is likely to have a different value if practiced than if it is licensed, used in litigation, or in a defensive portfolio. The salient point is that there is a need to characterize, at least in a general way, the kind of value different indirect measures are capable of representing in order to be able to properly interpret the results.

Therefore, this section includes a review of previous patent value studies organized by three different dimensions. The first dimension uses the patent value typology above to examine what types of value each of the indirect measures connect to. The second dimension looks at the source of data that is being leveraged. Finally, the studies are organized into a framework that allows the different measures and methodologies to be compared.

Indirect Measures of Value

Following on the typology of patent value proposed above, this section organizes previous studies on patent value to characterize the various indirect measures of patent value that have been utilized in terms of how they relate to the patent value modes described. As mentioned above, most studies have utilized indirect measures of patent value that generally relate to the first and most important type of patent value: practicing the patent.

The first set of proxies emerged from the patent data itself. Theoretically they were derived from the idea that the very existence of patents is evidence of a deliberate investment decision and value judgment as indicated by the investment in prosecuting the patent from initial filing through issuance and then maintaining it potentially all the way to termination. Extensive expenditures must be made for filing fees, attorney fees, and maintenance fees to accomplish this. These indicate at the very least that the patent holder has an expected value which exceeds the cost of patent prosecution, and that multiple experts have assessed its technical and legal merit.

Patent Renewals

One of the first proxies for patent value to emerge in the literature was the data regarding patent renewals (Pakes et al., 1988; Pakes & Simpson, 1989). The intuition for this is that patent holders are more likely to pay the fees associated with renewing patents if they have produced, or are expected to produce, value (Schankerman & Pakes, 1986). There have been two distinct uses of patent renewals in the literature. The first were studies that designed and evaluated models based on the fundamental intuition as described above (Lanjouw, 1998; Schankerman & Pakes, 1986). Additional studies have utilized renewals data as an independent variable (Harhoff et al., 1999; Lanjouw & Schankerman, 1999). Findings have consistently evidenced a relationship between citations and renewals, though the results are somewhat noisy.

Family Size

The decision to file a patent in foreign jurisdictions has been hypothesized to be another indicator of existing or expected value, much like renewals data. These additional filings comprise a patent "family", the size of which roughly approximates its international scope and is a relatively good proxy for the total cost of patenting. This is discussed at more length below. Patent family data come from the patent data and have been used primarily as an independent variable and indicator of expected value (Guellec & Potterie, 2000; Harhoff, Frederic, & Katrin, 2003; Harhoff et al., 1999; Lanjouw & Schankerman, 1999; Sampat & Ziedonis, 2004; Sapsalis & Potterie, 2007).

Patent Grant

Some studies have suggested that a patent being granted is a strong indication of its value and have used this to test for relationship with other observable characteristics from the patent data such as, the number and nationality of inventors, fields, patenting procedures and others (Guellec & Potterie, 2000, 2002).

Litigation Data

Another source of data that have been leveraged come from databases drawn from the legal world including court documents, and databases of legal actions relating to patents. Some studies have examined the potential information value of the existence of patent infringement litigation as the dependent variable proxy for value (Harhoff & Reitzig, 2004; J. Lanjouw & M. A. Schankerman, 1997). Given the typology of patent value enumerated above, it is also worth noting here that there may be additional information in this type of litigation data given that it represents an embodiment of the litigation value mode described, and therefore could potentially be seen as a proxy for that type of value in particular. In practice, there is little promise for this kind of use as the majority of patent litigation arises within the course of practicing patents. However, over the past few years there has been a growing controversy regarding so-called "patent trolls" that acquire and hold patents presumably exclusively for the purpose of monetizing them via the litigation mode . This practice has been cast as predatory and has been used to justify proposed reforms of the patent system (Sag & Rohde, 2006). Despite this, it is very difficult to distinguish from the licensing mode which is also often motivated and enforced through legal action (McDonough, 2006).

Financial Data

A very different approach used in some studies has been to look for information about the value of patents in the value of the firms that hold them (Hall, 1999). The intuition is that some portion of the value of firms is contributed by the value of its patents. Firms with more valuable patent portfolios will be more valuable, and this differential in value should be detectable using econometric techniques (Hall, 1999; Hall, Jaffe, & Trajtenberg, 2005; Hall et al., 2000; Lerner, 1994; H. Shane & Klock, 1997; Trajtenberg, 1990a, 1990b). The data for this have been drawn primarily from stock valuations for publicly-traded companies.

Firm Creation

Using a related concept, Shane (2001) utilized the creation of new firms based on patented technologies as an indirect indicator of patent value. The data for this study were compiled from patents licensed by MIT's technology transfer office.

Defensive Patenting

No measures of the defensive value of patents have been developed or tested in the literature.

Direct Measurements of Patent Value

As noted, there have been virtually no opportunities to systematically measure patent value in a direct way for any of the value types defined above. This presents an important problem for the entire field of patent analysis because of the central role that value plays across a wide range of research thrusts. It is possible that some of the indirect measures that have been used are valid proxies for some kind of economic value. However, in order to be confident in utilizing them, they must first be validated using some kind of reliable direct measurement of value. One attempt to accomplish this (Harhoff et al., 1999) provided a tantalizing confirmation of forward citations using data from a survey of patent holders in which they were asked to estimate, in hind-sight, the minimum price they would have been willing to sell each patent. This can be seen as a kind of rough approximation of the licensing value, and benefits from its *ex post* perspective on value.

Licensing Data

Although it can be viewed as conceptually derivative of practicing value, licensing value has come to be a very important alternative route to commercializing and monetizing patent rights. Licensing value could theoretically be measured in two different ways. The first, and most direct, approach is to measure the revenues generated from licensing. The second approach would be to use data regarding the licensing status of the patent, or in other words, if the patent is licensed or not. This approach is actually an indirect measurement because it indicates only that the patent is expected to have value. This is a slightly stylized characterization of course, because as was noted above, not all types of patent value produce a measurable revenue stream.

Sampat & Ziedonis (2004) utilized a set of data drawn from two major licensing universities in which both direct (licensing revenues) and indirect (licensed status) data were observed (Sampat & Ziedonis, 2004). This study is also unique in that it analyzed patent value from multiple units of analysis including both the patent and the patent family level. This approach seems to offer the potential for further research if similar information from other contexts and/or more representative datasets can be compiled.

At a higher level of abstraction, this study offers data on whether a patent was licensed or not, which is actually an indirect measure of licensing value or expected practice value. While this may initially appear to be of much less interest than the more direct and continuous revenue data, it offers an intriguing opportunity. There are two reasons for this. The first is that gaining access to licensing revenue data is somewhat difficult due to the practice of keeping licensing transactions confidential. In addition, compiling this data is resource intensive. For both these reasons, it is likely that data on the licensing status of patents could be much more accessible.

The second reason is that interpreting revenue data presents certain theoretical and methodological difficulties (see Chapter 4). In particular, revenue data are truncated in time which makes comparing them like comparing revenues from different movies that have been out for different amounts of time. The problem is compounded in the case of licenses because the terms vary dramatically and because returns which are truly representative of the patent's ultimate economic potential cannot be expected in many cases for several years or more. License revenues often include reimbursements for past patenting expenses, and these are normally the only revenues that accrue in the early years of a license. Perhaps most problematic of all is the reality that in many cases licenses are effectively paid for through research funding that does not appear in the licensing revenues at all.

If licensing status can be validated as a meaningful correlate of patent value, then the effort to obtain a greatly expanded data set may be justified. In fact, there may also be a theoretical basis for pursuing this approach. First, given the problems mentioned above, it is possible that the existence of a license is a less noisy indicator of the intrinsic patent value than revenues. Transacting a license is actually a very costly exercise which not only entails assuming the full previous costs of patent prosecution, but also an often lengthy negotiation process. Therefore, the existence of a license can be seen as a strong indication that the licensee believes it will receive private economic value. In this way it is a similar measure to patenting and renewals. However, at least theoretically, it may be better calibrated to actual value since at this stage the patent has almost always issued (Dechenaux, Goldfarb, Shane, & Thursby, 2005), and the licensee is presumably in a position to have much greater visibility on its economic potential both by virtue of the more advanced stage (Dechenaux, Goldfarb, Shane, & Thursby, 2003) and their knowledge of its actual use and their position close to the field of application.

Sampat and Ziedonis (2004) tested both revenue and status data in their study using licensing data compiled from Columbia University and the University of California. The analysis demonstrated a strong relationship between forward citations and the probability a patent was licensed, but failed to find any meaningful relationship between revenue data and forward citations (Sampat & Ziedonis, 2004). This result may be a confirmation of the discussion above but it obviously does not validate the use of license status data as a viable alternative to revenue data.

Table 1: Direct & Indirect Measures of Patent Value

Value Mode	*Direct Measures*	*Indirect Measures*
1. Practice	Commercialization returns attributable to patent	Patent Renewals Family Size Patent Grant Litigation Firm Value Firm Creation
2. License	License revenues	Licensed (Y/N)
3. Litigation	Patent suit awards, damages, settlements	Litigation (Y/N)
4. Defensive	None	None

Units of Measure

One very important aspect of measuring patent value that has received little treatment in the literature is the unit of measure. Many patent value studies have emphasized the individual patent filing as the target of analysis primarily because the patent data that have been available from the US Patent and Trademark Office (USPTO) are in this form[2]. Each individual patent filing is entered as a record into the USPTO database and receives a unique identification number. Over the past few years, patent data have been improved by various service providers such as Delphion[3], and it is now possible to obtain a patent's related filings, or "family."

Individual patent filings are related in two ways. First, the process of issuing a patent, or patent prosecution, often results in filings being split into multiple derivative filings that

[2] Important exceptions to this are Harhoff et al., (2003), Sampat et al., (2004), Sapsalis and Potterie (2007), Lanjouw and Schankerman (1999).
[3] http://www.delphion.com/

cover more specific definitions of the subject matter. Second, patent filings in the US often precede filings of the same subject matter in foreign jurisdictions. The result of both of these processes is a patent "family"[4].

Patent families are very important in understanding value (Harhoff et al., 2003; Sapsalis & Potterie, 2007). This is because the value of patent rights is partly a function of the geographic scope of those rights. A patent family with greater geographic coverage will theoretically have greater value, all other things being equal. Another reason is that patents which have led to "continuations," or further related patents, often have more technological scope and/or greater temporal coverage.

The greater technological scope and temporal coverage is the result of splitting the subject matter of a single patent into more specific, and presumably more collectively comprehensive, filings. Greater technological scope has been associated in the literature with greater economic value (J. Lanjouw & M. A. Schankerman, 1997; Lerner, 1994; Merges & Nelson, 1990). This is actually the result of the common strategy of filing an initial provisional patent with somewhat generalized claims, and then filing a series of continuations that both split and refine the subject matter while extending the time until final patents issue (Lemley & Moore, 1905). Although some abuses of this approach have been curtailed through legal and policy innovations (Jaffe & Lerner, 2004), the basic approach remains in common practice.

In addition to family size representing the kind of intrinsic value described above, it also can be viewed from the perspective of expected value on the part of the patent holder, or

[4] In practice the exact definition of what constitutes a "family" varies between jurisdictions and database providers. For this study, family data were obtained from Delphion which includes only foreign filings.

"assignee"[5]. Obviously the strategy of making multiple continuations, or of filing in a number of foreign jurisdictions, significantly increases the investment (Sapsalis & Potterie, 2007)[6]. It is reasonable to expect that patent applicants undertaking these investments have a greater expectation of the value of the patents in the family as a group[7]. As Sapsalis and Potterie (2007) have also shown, the analysis of patent families can extend to weight specific jurisdictions or combinations thereof to indicate a specific pattern of value.

The trend for patents to become more "global" has been sharply on the rise in recent years, with growth in foreign patenting rising approximately 65% faster than the already rapid growth in domestic patenting (WIPO, 2006). It is important to note that patenting remains concentrated overall with the top five jurisdictions (US, Japan, Korea, China and the European Patent Office) accounting for 75% of worldwide patent applications (WIPO, 2006).

Another potential unit of analysis that emerges from examining patent licensing data is the level of the license agreement. Generally, a license agreement will include all of the family members for each of the licensed patents. Often this includes a single patent family, but in many cases combines several patent families. The importance of this is the way that the license agreement itself defines the unit of value in a way that is relevant to the transacting parties. Sampat and Ziedonis (2004) have conducted the only published study of which this author is aware utilizing this unit of measure. Many other potential approaches to grouping patents for value analyses can be conceptualized but are beyond the scope of this study.

---

[5] "Assignee" refers to the owner of the patent, rather than to the specific named inventors.
[6] This is often an amount that is many times the cost of individual US filings.
[7] It is also important to note that foreign filings must be made within a limited period of time (usually 18 months) of the original US filing, and this is normally prior to the realization of much value from the patents. Therefore, the investment can be viewed as a good indicator of expected, rather than realized, value.

Patent Field

In the context of the analysis contained in this study, it is important to mention the relevance of technological field to patent value. Many studies have demonstrated that different technological fields or sectors are associated with very different attributes in the patent data (Hall et al., 2001; Schankerman, 1998). Without attempting a comprehensive review of these effects, I wish to highlight two of particular relevance to this study.

First, when broadly defined technology fields are aggregated over the data, a consistent finding is that the category comprised primarily of biological, medical, pharmaceutical, and chemical patents exhibits some unique attributes (Mansfield, 1986; Schankerman, 1998). In particular, patents are thought to be relatively more enforceable and somewhat more important to businesses operating in the medical sector because of the highly scientific nature, but also because of the high investments associated with regulation and approvals (Cohen et al., 2000; Scherer, 1993). This contributes to the argument for "patent thickets" particularly in bio-medical sciences where patent rights are more important overall, and more patents may be required to bring any single project to fruition (Thumm, 2005).

Because of the importance of bio-medical patents to their related industries, they constitute a significant portion of licensing transactions, and especially of those licenses coming out of universities and other research organizations. As a result, some of the most active licensing organizations are medical schools and medical research hospitals (AUTM, 2004). Bio-medical/pharmaceuticals are the fields where patents are thought to be the most important to commercial success, where companies rely on them the most, and where they presumably serve the intended function the best (Harhoff et al., 2003; Schankerman, 1998).

Therefore, it is hypothesized that these patents offer the greatest potential for appropriability, particularly in a licensing regime. Figure 1 summarizes the previous literature on citations and patent value.

Figure 1: Patent Value Literature Matrix

RHS → *LHS* ↓	***Forward Citations***	***Backward Citations***	***Scope***	***Non-patent Citations***	***Claims***	***Family Size***	***Renewals***	***Upheld in opposition***	***Co-Assignees***
Grant						(Guellec & Potterie, 2002)			(Guellec & Potterie, 2002)
Opposition (Litigation)	(Harhoff & Reitzig, 2004) (J. O. Lanjouw & M. A. Schankerman, 1997)				(J. O. Lanjouw & M. A. Schankerman, 1997)				
Firm Value	(Hall, 1999; H. Shane & Klock, 1997) (Hall, 1999) (Hall et al., 2005) (Trajtenberg, 1990a) (Trajtenberg, 1990b) (Lerner, 1994)								
Firm Creation	(S. Shane, 2001)		(S. Shane, 2001)						
Value Estimates	(Harhoff et al., 2003) (Harhoff et al., 1999)	(Harhoff et al., 2003)				(Harhoff et al., 2003) (Harhoff et al., 1999)	(Harhoff et al., 1999)	(Harhoff et al., 2003)	
Forward Citations		(Sapsalis & Potterie, 2007) (Gay & Le Bas, 2005)	(Lerner, 1994)	(Sapsalis & Potterie, 2007)		(Sapsalis & Potterie, 2007)			(Sapsalis & Potterie, 2007)
Patent Revenue	(Sampat & Ziedonis, 2004)					(Sampat & Ziedonis, 2004)			
Patent License	(Sampat & Ziedonis, 2004)					(Sampat & Ziedonis, 2004)			
Quality	(Lanjouw & Schankerman, 1999)	(Lanjouw & Schankerman, 1999)			(Lanjouw & Schankerman, 1999)	(Lanjouw & Schankerman, 1999)			
Social Value	(Trajtenberg, 1990b)								

"LHS" refers to "Left Hand Side" or dependent variables. "RHS" refers to "Right Hand Side" or independent variables.

Chapter 4: Study Design and Data

This study examines the relationship between citations and private economic value by analyzing the patenting and licensing portfolios of two medical research hospitals specializing in cancer research. This data offer a unique opportunity to test the relationship using data that directly observe at least one type of private value. This chapter describes the data within the context of the theory and typology from the previous chapters and examines some of its attributes, characteristics, and idiosyncrasies.

Description of Dataset

This unique dataset was obtained from two public research institutions[8] containing information on their entire current stock of inventions, patents, and licenses. The data were extracted directly from each institution's patenting and licensing management database system in January, 2007. Extracting and merging the data was facilitated significantly by the fact that both institutions used the same commercial database management system and data were in exactly the same format.

Because the full dataset contained information on a large number of patent filings which were abandoned, never filed, or which have expired, the majority of records cannot be analyzed. The institutional data were then matched to data from the public patent database to add data on citation counts, claims, patent application and publication dates, and family size[9]. Out of a total of 5703 patent filing records from the combined institutional data, 686 patents were matched with the full data. Institutional data included whether each patent was licensed

[8] Data provided by Memorial Sloan-Kettering Cancer Center, and Dana-Farber Cancer Institute.
[9] Public patent data were obtained using Delphion (http://www.delphion.com).

and how much revenue had been received on related licenses, and the necessary information for grouping the patents according to license agreement. The resulting dataset form the basis of most of the analyses contained in this study.

Private Licensing Value

Within the context of citations research, this data has a very specific and limited application which can be characterized using the discussion from the preceding chapters. First, this data include only patents which were filed by public research institutions exclusively for the purpose of licensing[10]. Patenting and licensing by universities, research hospitals, non-profit research institutes, and government laboratories has one important driver of a significant expansion in patent licensing over the past 25 years since the passage of the Bayh-Dole Act in 1980 (Mowery, 2004). The dramatic expansion of patenting and licensing by public research institutions has been amply documented elsewhere, and for several reasons has become a major thrust of the literature in its own right (Henderson et al., 1998; Mowery, 2004). For the present study, it is important to note that because these institutions file patents solely for the purpose of licensing them, the resulting patents are expected to express only the "licensing" type of private economic value defined in Chapter 2. Previous research indicates that patents held by these institutions are different from those held by other organizations (Trajtenberg & Henderson, 1997), which both clarifies and limits the interpretation of the results from the analysis. Some implications of this are discussed in more detail with the results in Chapter 5.

---

[10] Since "licensing" can be subject to slightly varying definitions different technology licensing offices, this study defines license as any type of agreement which encumbers the patent rights for any length of time and which either produces or has the ability to produce revenues. This can include licenses, options, research and license agreements, and other instruments.

Unit of Analysis

This dataset includes both direct and indirect observations of patent value. The direct observation is the total revenue received by related license agreements. These agreements frequently contain entire patent families, and there is no basis either in the data or in theory for assigning particular shares of value to individual members of the patent family. Therefore, all analyses in which revenue is the dependent variable are conducted using the original US filing. The other observation of value is the licensing status of the patent (eg. "licensed" or "not licensed"). This is analyzed at the level of the individual patent.

Patent Field

The data for this study are drawn exclusively from medical research institutions operating in a somewhat narrowly-defined field. This further narrows the interpretation of results based on the expectation that patents from public research institutions may be qualitatively different from those filed by commercial entities (Griliches, 1990; Henderson et al., 1998; Jaffe & Trajtenberg, 1993; Trajtenberg & Henderson, 1997). Patents coming from these organizations are generally thought to be more "basic" or scientific and this is expected to increase the citations they receive, though the effect appears to be much less than expected (Henderson et al., 1998).

In addition, as noted above, many studies and substantial anecdotal information demonstrate that patents in the broad area of medical and pharmaceutical technologies exhibit distinct attributes (Schankerman, 1998). In particular, patents in this field are thought to be more important to production and commercial success (Hall et al., 2005), have been growing faster as a proportion of total patent filings since at least 1975, and receive relatively more citations overall than other categories except computers and communications (Hall et al., 2001). The common conclusion is that drugs and medical patents have greater appropriability and this is probably especially true for licensing.

Drugs and medical patents have special significance in patent licensing and for public research institutions in particular. This segment represents the most significant component in licensing and account for the majority of licensing revenues (AUTM, 2004). For example, 3 out of the top 5 public research institutions by licensing income in 2004 were medical research hospitals, not universities (AUTM, 2004). Of the top 25, only one (MIT[11]) did not have a medical school or was not a research hospital, as opposed to 64% of survey respondents overall (AUTM, 2004). The existence of strong medical research is considered to be the most important predictor of success in licensing for universities and other public research institutions (Bulut & Moschini, 2006). Therefore, by virtue of the narrow field of the participating institutions, the patent data examined in this study can be expected to contain disproportionately more patents with high licensing potential. Because of this, the present study effectively examines the "intensive margin" of patents located in the most active licensing segment.

Selection Bias

A particular issue regarding the data compiled for this study relates to a selection bias of the included patents. The technology licensing offices (TLO's) of public research institutions vary in their practices concerning filing patents on disclosed technologies. It is

---

[11] It can be argued that MIT is closely associated with a major medical school and medical research hospitals through its partnership with Harvard Medical School as well as the Broad and Whitehead Institutes.

common practice for many institutions to file less expensive provisional patent applications for most disclosed technologies and search for a licensee during the one year period before the application must be converted to the much more expensive status of a full application. Although, this practice is not universal, the proportion of disclosures receiving at least a provisional patent application has risen steadily from 26% in 1991 to 62% in 2004 (AUTM, 2004). This means that public research institutions are filing new applications[12] on the majority of technology disclosures they receive.

Based on the author's previous experience and anecdotal evidence (no systematic data on patent practice exist) a substantial proportion of the institutions only subsequently convert filings for which an interested licensee has been identified. Therefore, patents from these institutions for which citations data are available may be expected to have a disproportionately high licensing rate and be expected to display a less skewed value distribution than a broader sample of patents.

Table 2 shows the breakdown of the patents in the dataset. The results are actually remarkably similar to the results presented by Sampat and Ziedonis (2004)[13]. It is evident that a relatively high proportion of the patents have been licensed, though not as high as might be expected given the approach described above. One explanation for this may be that the data include many patents which are relatively new, and may not have had enough time to find a licensee. This relates to the very important issue of truncation which is discussed next.

[12] "New" filings are defined as the first filing of the patentable subject matter, and constituted 76% of total annual filings in 2004 (AUTM, 2004). The proportion of new filings by public research institutions being made initially in the US has gone down steadily from around 76% in 1994 to only 35% in 2004. Subsequent filings (not "new") include divisional filings and foreign applications with foreign applications constituting the largest share.

[13] Their data, drawn from the University of California and Columbia University showed the same proportion of licensed (41%) and unlicensed (59%) patents.

Table 2: Summary of Data Sample

Sample Data (n=686)	*Number*	*% of Total*
Licensed	278	41%
Unlicensed	408	59%
Revenues > 0	253	37%
Revenues = 0	433	63%
Citations > 0	449	65%
Citations = 0	237	35%

Truncation

An important characteristic of patent and licensing data is that ultimate outcomes are not observed, only the results to the time of observation. This is known as truncated or "right-censored" data (Hall et al., 2001). This has two important implications for citation analyses. First, the data report the licensing status of the patents only as of the date the data were extracted, not the ultimate licensing outcome. Therefore, patents listed as "not licensed" may ultimately license, and this cannot be observed in the data. As mentioned above, this is partly due to the fact that many of the patents in the dataset are relatively new. However, given that the TLO begins marketing a patent virtually as soon as it is filed, a medical or drug patent that issues has already been marketed to potential licensees for at least a few, and in many cases several years. Further research may be warranted to explore the licensing time lag after initial patent filing to determine the impact of this effect, if any.

In a similar way, revenues are also right-censored in time (Sampat & Ziedonis, 2004). That is, the true value of a license (or patent) is not observed via revenues, but only the fraction that has been earned to date. Since the actual distribution of revenues over time is not known, making adjustments for revenue truncation is problematic. In other words, license agreements may be "front-loaded" to realize the majority of the revenue early, or they may be designed to emphasize royalties producing significant returns over an extended period.

Two techniques have been used in this study to address revenue truncation[14]. In some tests, a dummy for revenue was used (revenue = 1, no revenue = 0). The purpose of this is to eliminate the impact that very few large observations have on the results. A second approach was to organize patents into quartiles by revenue, and assign each quartile a categorical value. The objective of this approach is to test for the general relationship of citations and revenues. Table 3 shows the upper bound of each quartile in dollars. Both of these techniques also partially address the skewness issue which is revisited below.

Table 3: Revenue Quartile Limits

Quartile	*Upper Bound*
Quartile 1	$33,779.59
Quartile 2	$171,882.30
Quartile 3	$724,999.99
Quartile 4	$620,758,764.83
Lower bound	*$36.00*

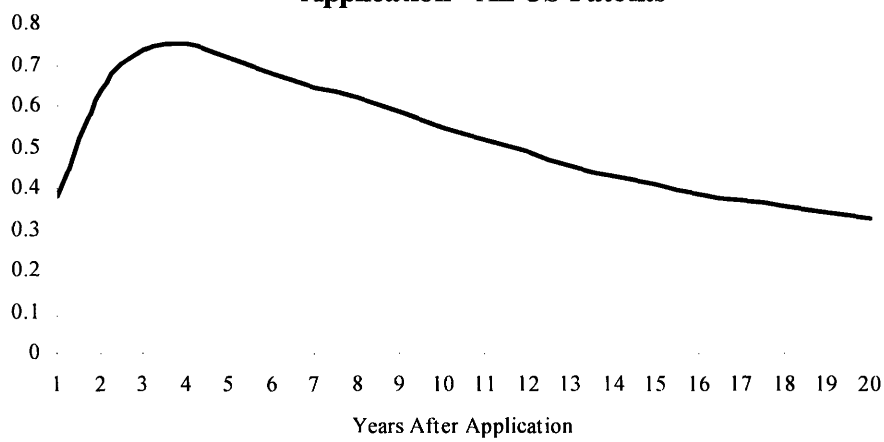
Truncation also impacts citations. Patents can continue to receive citations indefinitely, and therefore the citations that are observed represent some unknown fraction of the ultimate number. The impact is greater for younger patents since the fraction of lifetime citations earned in early years is very small (Hall et al., 2001). A number of methodological

[14] Because revenue is only observed at the level of the license agreement, patents were grouped by agreement for all calculations and regressions involving revenue.

solutions to the truncation problem as it relates to citations have been proposed (Hall et al., 2005; Hall et al., 2001). There are actually two effects that must be removed.

The first is the truncation (right-censoring) of the citation data that makes it inappropriate to compare the total citation counts for patents of different ages. The simplest method for resolving this is to deflate citations by the age of the patent, producing an "average citations per year" figure. The problem with this approach is that because citations are not received uniformly throughout the life of the patent, very young patents will be under-measured. Figure 2 displays the mean citations received annually after application for all patents.

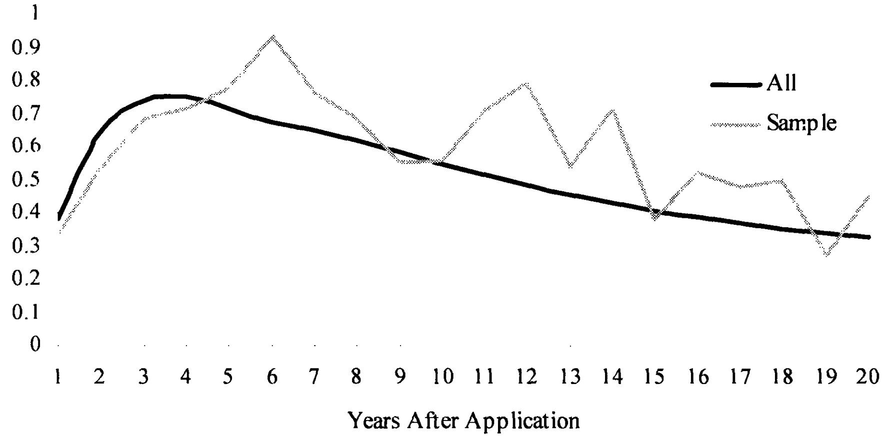
Figure 2: Mean Annual Citations Received After Application - All US Patents



Source: Data derived from USPTO, 2006. (approximately 4 mm patents)

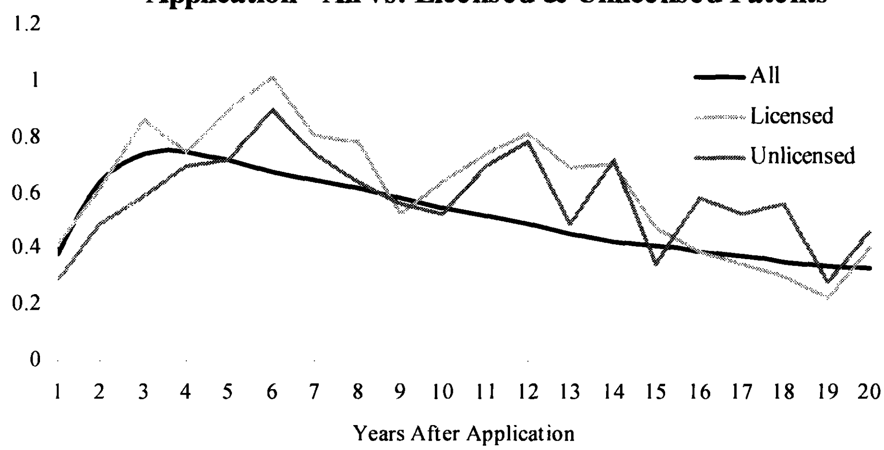
Comparing the distribution of all patents with those in the dataset (Figure 3) indicates a roughly similar distribution. Finally, separating the licensed and unlicensed patents in the sample (Figure 4) shows that the mean of licensed patents is consistently greater than unlicensed patents until they cross in the 15th year.

Figure 3: Mean Annual Citations Received After Application - All vs. Sample Patents



Source: Data derived from USPTO, 2006. (approximately 4 mm patents)

Figure 4: Mean Annual Citations Received After Application - All vs. Licensed & Unlicensed Patents



Source: Data derived from USPTO, 2006. (approximately 4 mm patents)

This raises the second truncation issue, which is the observation from previous studies that the propensity to cite and receive citations is not static but has changed over time (Hall et al., 2001). Selecting a technique to compensate for citation truncation relies on the development of a hypothesis regarding the meaning of citations data. For instance, one hypothesis is that changes in the propensity to cite over time merely reflect the changing nature of the patenting process, including: the advent of electronic patent databases that make identifying and citing prior art easier, the dramatic increase in patenting which increases the ratio of citing to cited patents, changes in patent rules, and changes in the resources of the patent office, just to name a few (Hall et al., 2001). An alternative hypothesis is that changes in citation frequencies are indicative of actual changes in the meaning and value of patents.

To accept the first hypothesis and adjust for this effect, citations received (forward citations) for each patent should be compared only to other patents of the same vintage. In this study, data were obtained[15] that provide annual citations for every patent, and this was used to calculate both the mean citations received by annual cohort, as well as annual means for citations received in each year following patent application for each cohort year[16]. This approach removes the effect of the changing propensity to cite over time by comparing patents only of the same age. It also eliminates the truncation problem[17]. Accepting the second hypothesis requires no such adjustment. Both approaches are implemented in this study.

However, there remains the difficulty relating to truncation of very young patents mentioned previously. Patents less than 1 year old have a likelihood of having received citations that is close to zero, and therefore it is not possible to infer anything from the observed zero values (Hall et al., 2005). In this way, the informational value of citation

[15] Many thanks to Kyle Jensen for assistance in accessing, matching, and compiling the annual citations data for all patents.
[16] Annual cohorts included all patents applied for in each given year. Thus, all patents with application dates during 1985 are included in the 1985 annual cohort.
[17] This technique does not remove the "field effect" on citations, which is the observation that the propensity to cite varies across fields identified by Hall et al., (2001). However, given the discussion above regarding the special nature of many patents in the context of licensing, this study hypothesizes that a greater propensity to receive citations by these patents is indicative of their greater appropriability (and therefore value) in a licensing regime, and so no attempt has been made in this study to adjust for this effect.

counts decreases for patents younger than 3 years, and drops precipitously for patents less than one year old. The number of patents in the present dataset that are likely to be significantly impacted for licensing and citations is relatively small. A total of 57 patents in the data (about 8%) had been published for less than one year, and 95 (about 14%) for less than two years as of the date of extraction.

Although it may be tempting to remove these young patents in an attempt to minimize the impact of truncation, it can potentially be more distorting than leaving them in. This is because very little is known about the relationship between licensing and citations. No previous study has examined the extent to which there may be simultaneous effects in the relationship between citations and licensing, and at present, data are not available to test for this effect. This relates to the hypothesis regarding the relationship between citations and private value in which a "publicity" effect results from the process of citing prior art. This may lead to increased awareness of the patent, and therefore have a positive impact on the probability of the patent being licensed (Sampat & Ziedonis, 2004).

This leads to the more general question of causation, or the degree to which citations create value or the reverse, and the extent to which there may be simultaneous effects between the two. Sampat and Ziedonis (2004) found that few citations were received by patents prior to licensing, providing possible support for their third and fourth hypotheses[18], and for the idea that citations may have a positive influence on licensing outcomes. Huang's (Huang, 2006) research also indicates that citations demonstrate a kind of simultaneity. However, answering the more fundamental question of causation is not within the scope of

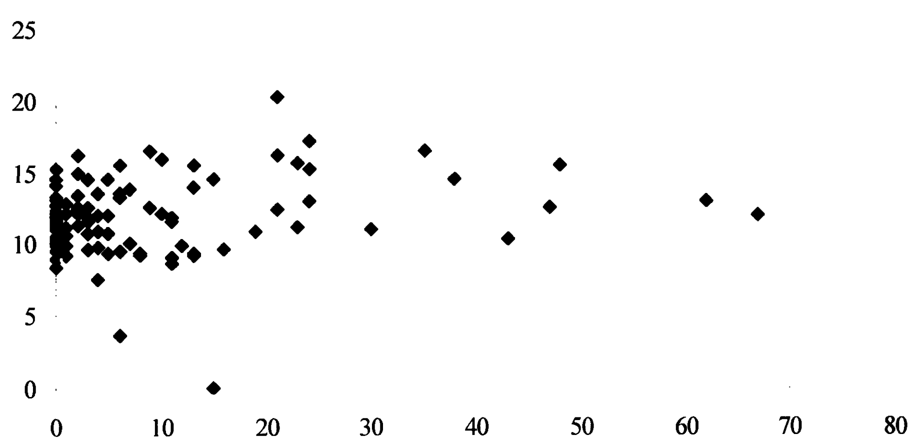
[18] "Hypothesis 3: Citations indicate technological opportunities or market interest in a technological area. Hypothesis 4: Citations result from public disclosure." (Sampat and Ziedonis, 2004)

this study, and therefore the data on relatively newer patents has not been removed from the analyses.

Skewness

Another challenge with the data is the highly skewed nature of both the citations and revenues data (Lanjouw et al., 1998). Overall, the top 5% of licenses account for 98% of the revenues. Therefore, a log transformation of revenues is used in this study as the dependent variable rather than gross revenues. The distribution of citations is also highly skewed, so log transformations of citations have also been used in some of the tests for comparative purposes with substantially the same results as the non-transformed variable. Figure 5 presents a scatter plot of log revenues versus citations.

Figure 5: Scatterplot of Log Revenues vs. Citations



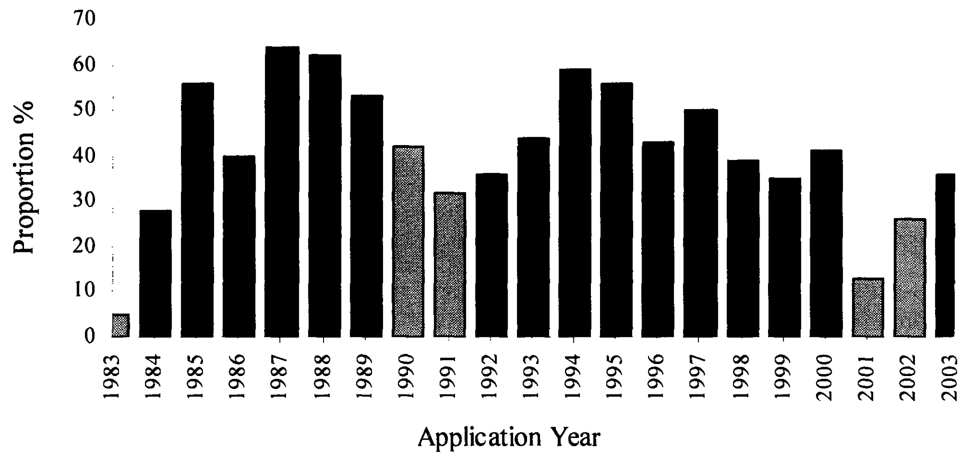
Licensing Outcomes

Overall, 42% of the patents in the dataset are licensed, and the data show no clear trend by the age of the patent (see Table 4 and Figure 6 below). However, there does seem to be a pattern, and it may not be a coincidence that it appears to match macroeconomic fluctuations. The years in which the NBER identifies economic recessions are shaded gray in the Figure 6 (NBER, 2007). This seems to indicate that licensing outcomes track the business cycle, and is a very interesting, though perhaps not surprising, result. This is an opportunity for further research.

Table 4: Distribution of Licensing Outcomes by Application Year

Application Year	*Proportion of Patents Unlicensed (%)*	*Proportion of Patents Licensed (%)*
1983	95	5
1984	72	28
1985	44	56
1986	60	40
1987	36	64
1988	38	62
1989	47	53
1990	58	42
1991	68	32
1992	64	36
1993	56	44
1994	41	59
1995	44	56
1996	57	43
1997	50	50
1998	61	39
1999	65	35
2000	59	41
2001	87	13
2002	74	26
2003	64	36
Total	58	42

Figure 6: Proportion of Patents Licensed by Application Year (1983-2003)

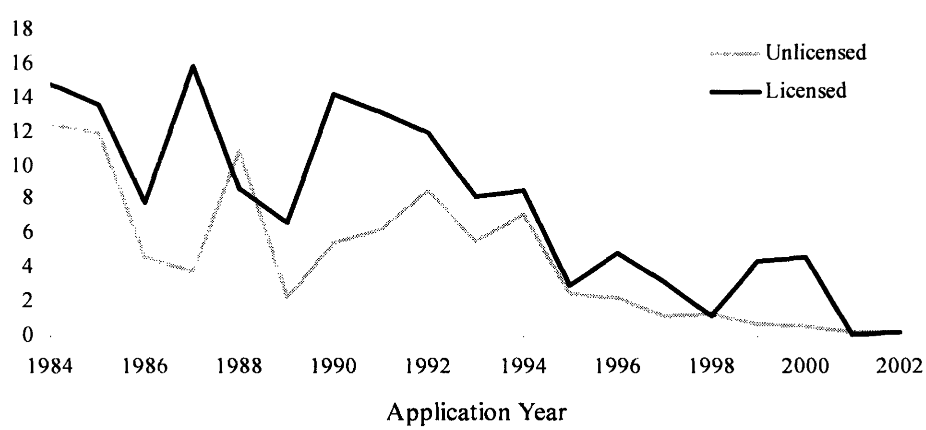


As an initial look at the question of whether citations are related to licensing outcomes, Table 5 and Figure 7 show the mean number of citations for both the licensed and unlicensed patents in the sample. There does appear to be a relationship with licensed patents displaying consistently higher levels of citations.

Table 5: Mean No. of Citations for Unlicensed and Licensed Patents

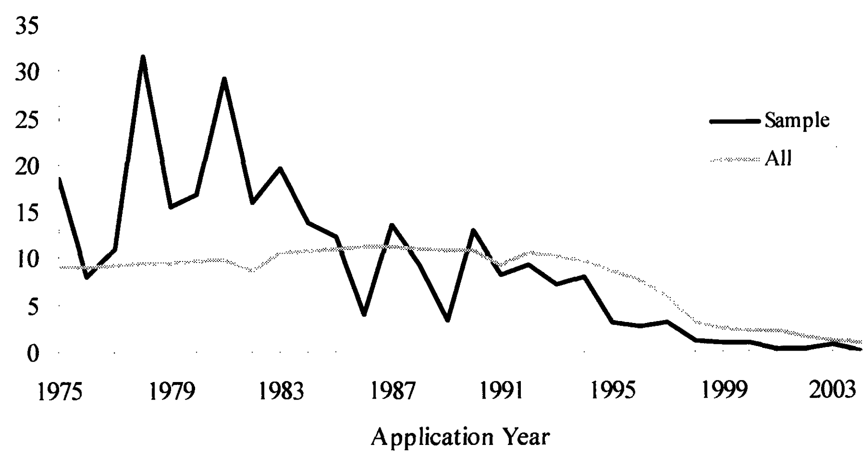
Application Year	*Mean No. of Citations Unlicensed Patents*	*Mean No. of Citations Licensed Patents*
1984	12.46	14.80
1985	12.00	13.56
1986	4.66	7.83
1987	3.80	15.89
1988	10.87	8.69
1989	2.25	6.67
1990	5.43	14.20
1991	6.26	13.14
1992	8.56	12.00
1993	5.60	8.22
1994	7.07	8.47
1995	2.53	2.94
1996	2.23	4.90
1997	1.23	3.18
1998	1.33	1.18
1999	0.75	4.36
2000	0.55	4.64
2001	0.21	0.17
2002	0.19	0.18
2003	1.22	0
Total	4.88	6.63

Figure 7: Mean Citations for Unlicensed and Licensed Patents by Application Year



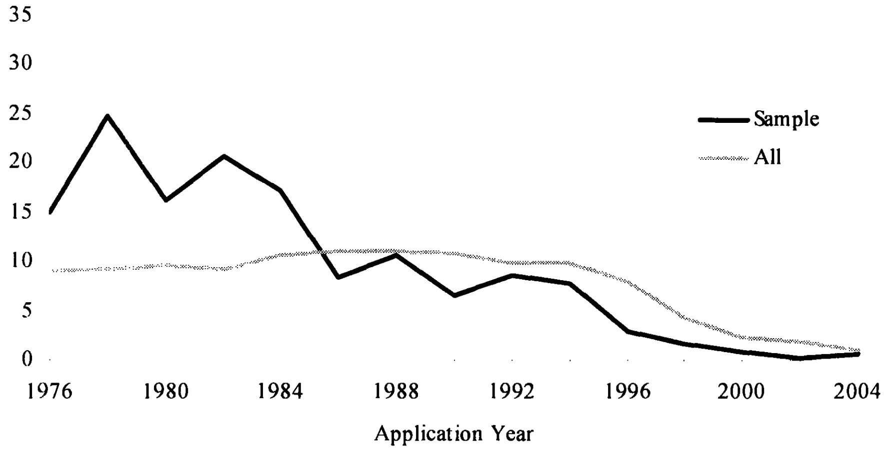
Figures 8 and 9 compare the citations received by the patents in the sample with all US patents for each year cohort. These results are interesting in light of the studies that have used citations to test whether patents from public research institutions have been trending downward in value or significance over time, and particularly since the Bayh-Dole Act triggered a significant increase in patenting for the institutions (Henderson et al., 1998; Jaffe & Lerner, 2004; Mowery, 2004; Sampat, Mowery, & Ziedonis, 2003; Sapsalis, Bruno van Pottelsberghe de la, & Ran, 2006). Citation counts for the sample group began a substantial decrease in the early 1980s, shortly after Bayh-Dole was enacted. This was the period when many institutions (including those providing the sample data used here) began active patenting and licensing activities (AUTM, 2004)[19], and during which was the largest increase in patenting by public research institutions (Henderson et al., 1998; Sampat et al., 2003).

Figure 8: Mean Citations Received by Application Year



[19] AUTM defines the start date of technology transfer for an institution as the date after which at least one half-time equivalent was devoted to patenting and licensing.

Figure 9: Mean Citations Received by Application Year - 2 Year Groupings



Figures 10 and 11 present the distribution of forward citations for the sample data over time. It is evident that, at least within this sample, the distribution curve peaks earlier and drops off more quickly for licensed patents than it does for unlicensed patents. This seems to confirm a previous study that showed that patents with value are those that are cited more quickly, or earlier than other patents (Gay et al., 2005).

Figure 10: Mean Annual Citations Received After Application - Sample Patents

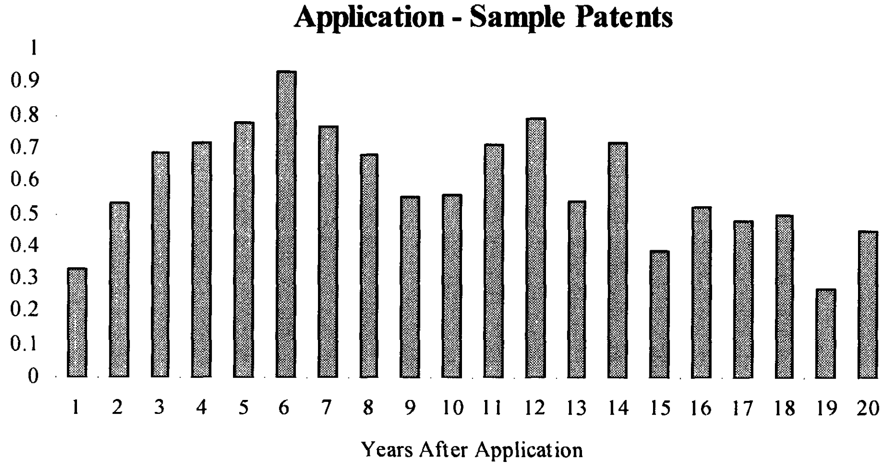
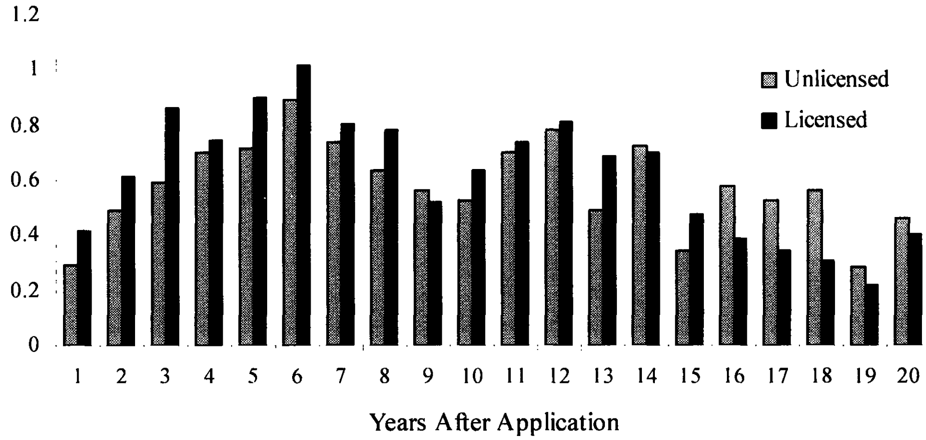


Figure 11: Mean Annual Citations Received After Application - Unlicensed vs. Licensed Patents



This may also demonstrate that licensing has a negative impact on citations received, but further research is required to explore this effect more fully. This kind of “chilling” effect could theoretically come from two different mechanisms. First, it could be that the investment subsequent to licensing by the licensee company reduces the incentive for competitors to enter the market and therefore file related patents thus reducing the pool of potentially citing patents. Second, once the license is public knowledge, subsequent filings may be drafted “around” the original patent so as to avoid citing it as prior art, and therefore, to avoid infringement and licensing. This may be one example of a simultaneous effect in the relationship between citations and licensing, and one that potentially applies to other types of value such as practicing, and especially litigation.

Chapter 5: Empirical Analyses and Findings

The purpose of this study is to examine the relationship between citations and the private value of patents. To accomplish this, a series of regression models have been estimated using data on a group of patents and their licensing outcomes obtained from two major public research institutions operating in the area of medical research[20]. This Chapter presents the empirical methods used to specify the models and the findings.

Variables

The data in this study provides two measures of economic value for the dependent variable. The first is the license status for each patent, or in other words, whether the patent has been licensed ("licensed"). The second measure is the total revenue earned on any licenses to date. These are somewhat different measures of value in both the theoretical and empirical dimensions.

Licensing Status (licensed)

Whether a patent is licensed may be the most basic measure of the licensing value of a patent. Because of the cost and effort required to negotiate and transact a license agreement, this can be seen as representing a substantial value expectation on the part of both the licensor and licensee (Sampat & Ziedonis, 2004). However, a number of licenses never

[20] In order to disguise information that is confidential in nature, there is no separation of the data by organization for analytical purposes in this study. The institutions were chosen specifically for their similarity on several levels and therefore the data were not expected to be substantively different.

produce significant revenues[21]. These are often licensed to startups that fail, but also include licenses to major companies that never reach the market as commercial products. Therefore, this measure must be understood as the *expected value*, and therefore an indirect measure of actual value. The value of licensed is binary (1=Yes, 0=No), and therefore probit models have been estimated for equations using this as the dependent variable.

Table 6: Licenses, Patents, & Outcomes

Total Licenses	110
Licenses Earning Revenues	86
Total Patents	686
Licensed Patents	278
Patents Earning Revenues	253
Unlicensed Patents	408

Revenues (rev_ln)

The second measure in the data allows for a more direct observation of the actual value of patents via the licensing revenues. Revenues are reported as total cumulative amounts received, and a log transformation is used in all models.

Citations

The primary independent variable is the total number of citations received. As described above, there are two versions used in the models. The first, "cites_defl_age," is simply the total count of forward citations divided by the age of the patent in years. This

[21] Reliable data are not available on the proportion of licenses that never produce income. However, data from the AUTM Annual Survey (2004) show that about 60% of active licenses to public research institutions were yielding income in 2004.

variable is designed to adjust for truncation. The second, "cites_defl_cohortcites," divides the total count of forward citations by the average number of forward citations for all patents within the age cohort. The age cohort is composed of all patents with application dates in the same calendar year. This variable is designed to adjust for both truncation and for the changing propensity to cite.

Controls

Four variables have been used as controls in the model specifications: backward citations, family size, claims, and publication lag. Table 8 lists the controls and provides descriptions.

Table 7: Listing of Control Variables

Control	*Description*
Backward Citations	The total count of citations made by the patent to other patents
Family Size	The total number of patents in the patent's family
Claims	The total number of claims in the patent
Publication Lag	The number of years between the application and publication dates

Descriptive Statistics

To gain an overall perspective on the dataset, Table 6 presents the descriptive statistics for the variables used in this study.

Table 8: Descriptive Statistics of Basic Patent Variables

	Age of Patents	*Forward Citations*	*Backward Citations*	*Family Size*	*Claims*	*Revenue*
Mean	11.91399	5.596210	4.916910	17.70991	15.02624	6353463.
Median	11.00000	2.000000	2.000000	9.000000	11.00000	0.000000
Maximum	31.00000	111.0000	158.0000	224.0000	101.0000	6.21E+08
Minimum	2.000000	0.000000	0.000000	1.000000	1.000000	0.000000
Std. Dev.	6.081193	9.972016	13.03136	30.09603	14.01352	57854764
Skewness	0.689218	3.950742	8.079388	4.702260	2.487522	10.49764
Kurtosis	2.950589	27.78000	83.14356	28.93657	11.52278	111.5728
Jarque-Bera	54.38061	19336.10	191053.7	21756.22	2783.699	349541.6
Probability	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
Sum	8173.000	3839.000	3373.000	12149.00	10308.00	4.36E+09
Sum Sq. Dev.	25331.93	68117.15	116324.3	620453.3	134519.5	2.29E+18
Observations	686	686	686	686	686	686

Empirical Results

Several models were specified to test both dependent variables against both versions of the citations variable, as well as additional models to include the controls. The first group of models (1-2) were specified using “licensed” as the dependent variable. All patents (686) were included in these equations because the level at which licensing status is observed is the patent. Both of the citation variables described above are tested separately. For the second group of equations (3-6), log revenues (“rev_ln”) was the dependent variable. Because revenues are only observed at the level of the license, only 422 “unique” patents are used in models 3 and 4. Finally, models 5 and 6 look only at “unique” patents with revenues greater than zero in order to examine the impact of citations on revenues. The results are summarized in Table 9 below and all estimation output is included in the Appendix.

Table 9: Estimation Results

Tests	(1-2)	(3-4)	(5-6)
Dependent Variable	License Status (licensed)	Revenues (rev_ln)	Revenues (rev_ln)
Sample Group	All patents	"Unique" Patents	"Unique" Patents with revenue
Sample Size	686	422	287
Method	Probit	OLS	OLS
Cite_defl_age	0.31	1.89	0.61
	(3.45)	(4.96)	(2.64)
Cite_defl_cohortcites	0.12	0.87	0.37
	(2.71)	(3.99)	(3.22)

Absolute value of z-statistics or t-statistics in parentheses.

Conclusions

The results confirm the existence of a consistent relationship between citation counts and both measures of licensing value. This is the first study to show evidence of a statistically significant relationship between citations and an objective, direct measure of private patent value such as licensing revenues[22]. In fact, the strongest results came from equations using licensing revenues as the dependent variable (see Appendix for detailed estimation output).

In addition, it is possible also to view licensing revenues as an indirect measure of the practicing value of the patents to the licensee, since licensing revenue comes primarily from royalties paid on the licensees earnings from practicing the patent (AUTM, 2004). In this way, these findings can also be seen as confirmatory of other studies that have used indirect measures of practicing value, such as firm valuation (Hall, 1999; Hall et al., 2005; H. Shane & Klock, 1997). In fact, it is likely that the relationship between the practicing value of patents and license royalties paid is closer than it is to the firm's stock price.

[22] This finding contrasts with Sampat and Ziedonis (2004) who found no statistically significant relationship in their study of licensing revenue data from Columbia University. In their study, the authors used Tobit regressions rather than the OLS method employed here, so the results are not directly comparable.

In general, within the context of the theoretical framework proposed above, these results can be interpreted as a confirmation that forward citations counts contain information about the private licensing value of patents. However, there are some important limitations to this conclusion mentioned above that bear repeating here. First, the patents in this dataset represent a particular field by virtue of their source, and therefore these results are not necessarily applicable to other fields and sectors. In fact, there is evidence (discussed above) that this field (medical) may be an important factor in these findings. However, more research is needed to examine the impact of this effect, perhaps by examining similar data with greater breadth across fields.

Second, these estimation results indicate that the impact of citation counts on the licensing value of patents is small (though consistently statistically significant). In fact, it would be surprising if it were otherwise. Chapter 2 of this study presented a typology of the ways in which patents produce private economic value that was intended to highlight the economic complexity of patents. Today patents are utilized, commercialized, leveraged, and monetized in a diversity of ways. In some cases the private value can be calculated, but in many cases it cannot. Therefore, these findings provide a confirmation of the idea suggested above that many other factors contribute to a patent's private value, and that only some types of that economic value can be measured.

Third, the patents analyzed here represent what may be a very unique type, patents filed by public research institutions exclusively for the purpose of licensing (Trajtenberg & Henderson, 1997). These patents are expected to be highly selected for their appropriability in a licensing regime and are likely to have had commercial interest even before they are published. They may also be more "basic" or scientific and, therefore, have broader applications. In theory this should produce higher citation counts on average. Therefore, we may expect these patents to tend toward the right tail of the private value distribution of all patents. The evidence presented here is not conclusive on this point, but may offer limited support for earlier studies indicating that the quality of patents filed by public research institutions has decreased as the volume of filings increased over the past two decades (Henderson et al., 1998).

One interesting finding was that the relationship between citation counts and licensing value is stronger in all tests for the simple age-adjusted citation measure than for the cohort-adjusted measure. One may conclude from this either that the changes in propensities to cite are indicative of real changes in patents over time, or that citations themselves have become increasingly noisy indicators of value. However these findings are not conclusive on this point and further theory development and empirical research is needed to address this question.

In general, the findings presented here can be interpreted as a validation of the relationship between patent citation counts and the private value of patents using two measures of licensing value. The findings here generally confirm results from previous studies and offer what may be the first evidence of a significant relationship between citations and a direct measurement of one type of patent value – licensing revenues. More work is needed in validating these results using larger and more heterogeneous datasets, in refining our theoretical understanding of the economic attributes and dynamics of patents, and in developing data and techniques for more precise measurement. Such work is vital in improving the economic "optics" that expose the dynamics of innovation to our view.

Appendix: Estimation Results

Table 10: Equation 1 Results

Dependent Variable: LICENSED
Method: ML - Binary Probit (Quadratic hill climbing)
Included observations: 686
Convergence achieved after 4 iterations
QML (Huber/White) standard errors & covariance

Variable	Coefficient	Std. Error	z-Statistic	Prob.
CITES_DEFL_AGE	0.310491	0.089937	3.452302	0.0006
FAMILY_SIZE	0.007698	0.002443	3.150690	0.0016
NUM_CLAIMS	0.005206	0.003537	1.471569	0.1411
PUB_LAG_YRS	0.090263	0.028239	3.196391	0.0014
NUM_DOM_REFS	-0.018688	0.006430	-2.906285	0.0037
C	-0.804134	0.130599	-6.157267	0.0000

Mean dependent var	0.405248	S.D. dependent var	0.491298
S.E. of regression	0.473817	Akaike info criterion	1.289707
Sum squared resid	152.6620	Schwarz criterion	1.329336
Log likelihood	-436.3695	Hannan-Quinn criter.	1.305040
Restr. log likelihood	-463.1064	Avg. log likelihood	-0.636107
LR statistic (5 df)	53.47382	McFadden R-squared	0.057734
Probability(LR stat)	2.69E-10		
Obs with Dep=0	408	Total obs	686
Obs with Dep=1	278		

Table 11: Equation 2 Results

Dependent Variable: LICENSED
Method: ML - Binary Probit (Quadratic hill climbing)
Included observations: 686
Convergence achieved after 4 iterations
QML (Huber/White) standard errors & covariance

Variable	Coefficient	Std. Error	z-Statistic	Prob.
CITES_DEFL_COHORTCITES	0.124366	0.045874	2.711007	0.0067
FAMILY_SIZE	0.007850	0.002394	3.278267	0.0010
NUM_CLAIMS	0.005386	0.003542	1.520556	0.1284
PUB_LAG_YRS	0.086163	0.028436	3.030021	0.0024
NUM_DOM_REFS	-0.019527	0.006566	-2.973891	0.0029
C	-0.757805	0.130244	-5.818357	0.0000

Mean dependent var	0.405248	S.D. dependent var	0.491298
S.E. of regression	0.475899	Akaike info criterion	1.298076
Sum squared resid	154.0064	Schwarz criterion	1.337705
Log likelihood	-439.2400	Hannan-Quinn criter.	1.313409
Restr. log likelihood	-463.1064	Avg. log likelihood	-0.640292
LR statistic (5 df)	47.73266	McFadden R-squared	0.051535
Probability(LR stat)	4.03E-09		

Obs with Dep=0	408	Total obs	686
Obs with Dep=1	278		

Table 12: Equation 3 Results

Dependent Variable: REV_LN
Method: Least Squares
Included observations: 422

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CITES_DEFL_AGE	1.894532	0.381398	4.967331	0.0000
NUM_CLAIMS	0.044665	0.016998	2.627701	0.0089
NUM_DOM_REFS	-0.037203	0.024443	-1.522047	0.1288
PUB_LAG_YRS	0.320539	0.147140	2.178465	0.0299
FAMILY_SIZE	0.013283	0.011784	1.127255	0.2603
C	0.130114	0.654128	0.198913	0.8424

R-squared	0.085639	Mean dependent var	2.674394
Adjusted R-squared	0.074649	S.D. dependent var	5.168094
S.E. of regression	4.971455	Akaike info criterion	6.059418
Sum squared resid	10281.59	Schwarz criterion	6.116930
Log likelihood	-1272.537	F-statistic	7.792535
Durbin-Watson stat	0.152356	Prob(F-statistic)	0.000001

Table 13: Equation 4 Results

Dependent Variable: REV_LN
Method: Least Squares
Included observations: 422

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CITES_DEFL_COHORTCITES	0.869843	0.217534	3.998646	0.0001
NUM_CLAIMS	0.044521	0.017168	2.593323	0.0098
NUM_DOM_REFS	-0.042731	0.024680	-1.731377	0.0841
PUB_LAG_YRS	0.308937	0.149163	2.071135	0.0390
FAMILY_SIZE	0.016761	0.011841	1.415483	0.1577
C	0.287268	0.664806	0.432109	0.6659

R-squared	0.067256	Mean dependent var	2.674394
Adjusted R-squared	0.056045	S.D. dependent var	5.168094
S.E. of regression	5.021183	Akaike info criterion	6.079324
Sum squared resid	10488.31	Schwarz criterion	6.136836
Log likelihood	-1276.737	F-statistic	5.999171
Durbin-Watson stat	0.126947	Prob(F-statistic)	0.000022

Table 14: Equation 5 Results

Dependent Variable: REV_LN
Method: Least Squares
Included observations: 253

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CITES_DEFL_AGE	0.612527	0.232130	2.638728	0.0089
NUM_CLAIMS	-0.013617	0.011835	-1.150570	0.2510
NUM_DOM_REFS	0.071643	0.027931	2.564969	0.0109
PUB_LAG_YRS	-0.193203	0.090184	-2.142314	0.0331
FAMILY_SIZE	-0.014813	0.004728	-3.133211	0.0019
C	13.17858	0.441537	29.84706	0.0000

R-squared	0.102815	Mean dependent var	12.46564
Adjusted R-squared	0.084653	S.D. dependent var	2.829399
S.E. of regression	2.706992	Akaike info criterion	4.852983
Sum squared resid	1809.968	Schwarz criterion	4.936779
Log likelihood	-607.9024	F-statistic	5.661087
Durbin-Watson stat	0.120951	Prob(F-statistic)	0.000058

Table 15: Equation 6 Results

Dependent Variable: REV_LN
Method: Least Squares
Included observations: 253

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CITES_DEFL_COHORTCITES	0.374150	0.116182	3.220383	0.0015
NUM_CLAIMS	-0.012995	0.011742	-1.106671	0.2695
NUM_DOM_REFS	0.066869	0.027760	2.408870	0.0167
PUB_LAG_YRS	-0.176195	0.089827	-1.961478	0.0509
FAMILY_SIZE	-0.015107	0.004686	-3.223949	0.0014
C	13.11964	0.433143	30.28939	0.0000

R-squared	0.114695	Mean dependent var	12.46564
Adjusted R-squared	0.096774	S.D. dependent var	2.829399
S.E. of regression	2.689010	Akaike info criterion	4.839653
Sum squared resid	1786.002	Schwarz criterion	4.923449
Log likelihood	-606.2161	F-statistic	6.399962
Durbin-Watson stat	0.124998	Prob(F-statistic)	0.000013

References

Alcacer, J., & Gittelman, M. (2004). *How do I Know what you Know? Patent Examiners and the Generation of Patent Citations*: SSRN.

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

Arora, A., Fosfuri, A., & Gambardella, A. (2000). *Markets for Technology and Their Implications for Corporate Strategy*: SSRN.

AUTM (Ed.). (2004). *AUTM Licensing Survey: FY 2004*. Northbrook, IL: Association of University Technology Managers.

Bulut, H., & Moschini, G. (2006). *U.S. Universities' Net Returns from Patenting and Licensing: A Quantile Regression Analysis*: Center for Agricultural and Rural Development (CARD) at Iowa State University.

Cohen, W. M., Nelson, R. R., & Walsh, J. P. (2000). *Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)*: SSRN.

Dechenaux, E., Goldfarb, B. D., Shane, S. A., & Thursby, M. C. (2003). *Appropriability and the Timing of Innovation: Evidence from MIT Inventions*: SSRN.

Dechenaux, E., Goldfarb, B. D., Shane, S. A., & Thursby, M. C. (2005). *Appropriability and Commercialization: Evidence from MIT Inventions*: SSRN.

Gans, J. S., & Stern, S. (2003). The product market and the market for "ideas": commercialization strategies for technology entrepreneurs. *Research Policy, 32*(2), 333-350.

Gay, C., Bas, C. L., Patel, P., & Touach, K. (2005). The determinants of patent citations: an empirical analysis of French and British patents in the US. *Economics of Innovation and New Technology, 14*(5), 339.

Gay, C., & Le Bas, C. (2005). Uses without too many abuses of patent citations or the simple economics of patent citations as a measure of value and flows of knowledge*. *Economics of Innovation & New Technology, 14*(5), 333-338.

Griliches, Z. (1990). Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature, 28*(4), 1661-1707.

Guellec, D., & Potterie, B. v. P. d. l. (2000). Applications, grants and the value of patent. *Economics Letters, 69*(1), 109-114.

Guellec, D., & Potterie, B. V. P. D. L. (2002). The Value of Patents and Patenting Strategies: Countries and Technology Areas Patterns. *Economics of Innovation & New Technology, 11*(2), 133.

Hall, B. H. (1999). *Innovation and Market Value*: SSRN.

Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market value and patent citations. *The Rand Journal of Economics, 36*(1), 16.

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2000). *Market value and patent citations : a first look*. Cambridge, MA: National Bureau of Economic Research.

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001). *The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools*: SSRN.

Harhoff, D., Frederic, M. S., & Katrin, V. (2003). Citations, family size, opposition and the value of patent rights. *Research Policy, 32*(8), 1343.

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

Harhoff, D., & Reitzig, M. (2004). Determinants of opposition against EPO patent grants - the case of biotechnology and pharmaceuticals. *International Journal of Industrial Organization, 22*(4), 443.

Henderson, R., Jaffe, A. B., & Trajtenberg, M. (1998). Universities as a Source of Commercial Technology: A Detailed Analysis of University Patenting, 1965–1988. *The Review of Economics and Statistics, 80*(1), 119-127.

Huang, K. G.-L. (2006). *Innovation in the Life Sciences: the Impact of Intellectual Property Rights on Scientific Knowledge Diffusuion, Accumulation and Utilization.* MIT, Cambridge.

Jaffe, A. B. (1998). Patents, patent citations, and the dynamics of technological change. *NBER Reporter*, 8.

Jaffe, A. B., & Lerner, J. (2004). *Innovation and its discontents : how our broken patent system is endangering innovation and progress, and what to do about it*. Princeton, N.J.: Princeton University Press.

Jaffe, A. B., & Trajtenberg, M. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics, 108*(3), 577.

Jaffe, A. B., & Trajtenberg, M. (2002). *Patents, citations, and innovations : a window on the knowledge economy*. Cambridge, Mass.: MIT Press.

Jaffe, A. B., Trajtenberg, M., & Fogarty, M. S. (2000). *The Meaning of Patent Citations: Report on the NBER/Case-Western Reserve Survey of Patentees*: SSRN.

Lanjouw, J., & Schankerman, M. A. (1997). *Stylized Facts of Patent Litigation: Value, Scope and Ownership*: SSRN.

Lanjouw, J. O. (1998). Patent protection in the shadow of infringement: Simulation estimations of patent value. *The Review of Economic Studies, 65*(225), 671.

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

Lanjouw, J. O., & Schankerman, M. A. (1997). *Stylized Facts of Patent Litigation: Value, Scope and Ownership*: SSRN.

Lanjouw, J. O., & Schankerman, M. A. (1999). *The Quality of Ideas: Measuring Innovation with Multiple Indicators*: SSRN.

Lemley, M. A., & Moore, K. A. (1905). *Ending Abuse of Patent Continuations*: SSRN.

Lerner, J. (1994). The importance of patent scope: An empirical analysis. *RAND Journal of Economics, 25*(2), 319-333.

Mansfield, E. (1986). Patents and Innovation: An Empirical Study. *Management Science, 32*(2), 173.

McDonough, J. F., III. (2006). *The Myth of the Patent Troll: An Alternative View of the Function of Patent Dealers in an Idea Economy*: SSRN.

Merges, R. P., & Nelson, R. R. (1990). On the Complex Economics of Patent Scope. *Columbia Law Review, 90*(4), 839-916.

Mowery, D. C. (2004). *Ivory tower and industrial innovation : university-industry technology transfer before and after the Bayh-Dole Act in the United States*. Stanford, Calif.: Stanford Business Books.

NBER. (2007). Business Cycle Expansions and Contractions. Retrieved May 8, 2007, from http://www.nber.org/cycles

Nelson, R. R., & Winter, S. G. (1977). In search of useful theory of innovation. *Research Policy, 6*(1), 36-76.

Nicholas, V. (1999). Survey: Innovation in industry: Industry gets religion. *The Economist, 350*(8107), S5.

Orey, M. (2006). The Patent Epidemic. *Business Week*(3966), 60.

Pakes, A., Hall, B. H., & Griliches, Z. (1988). *The Value of Patents as Indicators of Inventive Activity*: SSRN.

Pakes, A., & Simpson, M. (1989). Patent Renewal Data. *Brookings Papers on Economic Activity. Microeconomics, 1989.*

Sag, M. J., & Rohde, K. W. (2006). *Patent Reform and Differential Impact*: SSRN.

Sampat, B. N. (2005). *Determinants of Patent Quality: An Empirical Analysis*. Paper presented at the Fifth Annual Roundtable on Engineering Entrepreneurship Research (REER). from http://mgt.gatech.edu/news_room/news/2005/reer/.

Sampat, B. N., Mowery, D. C., & Ziedonis, A. A. (2003). Changes in university patent quality after the Bayh-Dole act: a re-examination. *International Journal of Industrial Organization, 21*(9), 1371-1390.

Sampat, B. N., & Ziedonis, A. A. (2004). Patent Citations and the Economic Value of Patents. In *Kluwer Handbook of quantitative science and technology research* (pp. 277-298): Dordrecht; Boston and London:

Kluwer Academic.

Sapsalis, E., Bruno van Pottelsberghe de la, P., & Ran, N. (2006). Academic versus industry patenting: An in-depth analysis of what determines patent value. *Research Policy, 35*(10), 1631.

Sapsalis, E., & Potterie, B. v. P. d. l. (2007). The Institutional Sources of Knowledge and the Value of Academic Patents. *Economics of Innovation and New Technology, 16*(2), 139.

Schankerman, M. (1998). How valuable is patent protection? Estimates by technology field. *The Rand Journal of Economics, 29*(1), 77.

Schankerman, M., & Pakes, A. (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. (1993). Pricing, Profits, and Technological Progress in the Pharmaceutical Industry. *Journal of Economic Perspectives, 7*(3), 97-115.

Shane, H., & Klock, M. (1997). The Relation Between Patent Citations and Tobin's Q in the Semiconductor Industry. *Review of Quantitative Finance and Accounting, 9*, 131-146.

Shane, S. (2001). Technological opportunities and new firm creation. *Management Science, 47*(2), 205.

Shapiro, C. (2001). *Navigating the Patent Thicket: Cross Licenses, Patent Pools, and Standard-Setting*: SSRN.

Thumm, N. (2005). Patents for genetic inventions: a tool to promote technological advance or a limitation for upstream inventions? *Technovation, 25*(12), 1410.

Thursby, J. G., Jensen, R., & Thursby, M. C. (2001). Objectives, Characteristics and Outcomes of University Licensing: A Survey of Major U.S. Universities. *Journal of Technology Transfer, 26*(1-2), 59.

Trajtenberg, M. (1990a). *Economic analysis of product innovation : the case of CT scanners*. Cambridge, Mass.: Harvard University Press.

Trajtenberg, M. (1990b). A Penny for Your Quotes: Patent Citations and the Value of. *The Rand Journal of Economics, 21*(1), 172.

Trajtenberg, M., & Henderson, R. (1997). University Versus Corporate Patents: A Window on the Basicness of Invention. *Economics of Innovation & New Technology, 5*(1), 19.

von Hippel, E. (1987). *The sources of innovation*. New York: Oxford University Press.

Warsh, D. (2006). *Knowledge and the wealth of nations : a story of economic discovery* (1st ed.). New York: W.W. Norton.

WIPO. (2006). *WIPO Patent Report 2006*. Geneva: WIPO.