STYLIZED FACTS OF PATENT LITIGATION: VALUE, SCOPE AND OWNERSHIP

Department of Economics, Yale University

and

Mark Schankerman
Department of Economics, London School of Economics
and Political Science

Contents:

Discussion Paper
No.EI/20
January 1998

The Toyota Centre
Suntory and Toyota International Centres for
Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel.: 020 – 7955 6674

We thank Dietmar Harhoff, Robert Merges, Suzanne Scotchmer and Brian Wright for very constructive comments on an earlier draft. We gratefully acknowledge the financial support of the Nuffield Foundation and the Suntory and Toyota International Centres for Economics and Related Disciplines at the London School of Economics.

Abstract

This paper investigates the characteristics of litigated patents by combining for the first time information about patent case filings from the U.S. district courts and detailed data from the U.S. Patent and Trademark Office. We construct a series of indicators for the factors which the theoretical literature suggests contribute to litigation: frequency of disputes, the size and asymmetry of stakes, the structure of information, and costs. Compared to a random sample of U.S. patents from the same cohorts and technology areas, we find that more valuable patents and those with domestic owners are considerably more likely to be involved in litigation. Patents owned by individuals are at least as likely to be the subject of a case as corporate patents and litigation is particularly frequent in new technology areas. We interpret the results with reference to theoretical models of litigation and settlement and discuss what they suggest about the effect of patent litigation on the incentives to invest in R&D.

Keywords: Patents, litigation, R&D.

Litigation is an expensive way to resolve disputes. There is serious concern that civil litigation imposes significant costs on business, and this has led to institutional reforms to mitigate these costs such as the introduction of damage award ceilings, mediation requirements (Farber and White, 1991), and rules to shift legal fees which facilitate settlement (for example, Rule 68 of the Federal Rules of Civil Procedure; see Kaplow, 1993, and Spier, 1994). And there have been market responses too, such as the emergence of patent litigation insurance and the development of investment vehicles to finance intellectual property rights litigation (*Business Insurance*, 1995).

Economic analysis has been used to understand the factors that impede parties from settling their disputes out of court. These factors include the relative costs of litigation compared to settlement, the size of stakes, the probability that the plaintiff will prevail at trial, and the structure of information. The theoretical predictions have been tested indirectly by examining whether the relationship between different endogenous outcomes of the legal process, such as the trial rate and the win proportion, are consistent with those predicted by theory (Priest and Klein, 1984; Wittman, 1988; Waldfogel, 1996). Direct tests have been hampered by the difficulty of finding suitable proxies for the key determinants of litigation identified by the theory. Some successful examples include the studies of medical malpractice reforms by Hughes and Snyder (1989), and Farber and White (1991) who study the settlement-litigation decision using various proxies for potential damages (age of patient, severity of injury), litigation costs (contingency fee limits), and probability of the plaintiff prevailing (the quality of medical care). Siegelman and Waldfogel (1996) examine the probability of trials across different classes of civil suits, using proxy variables for the importance of reputation (repeat play) and other factors. Other empirical studies include Fournier and Zuehlke (1989) on general civil litigation, and Eisenberg and Henderson (1992) on product liability.

In this paper we analyze patent cases, distinguishing between infringement suits brought by patentees and patent invalidity suits brought by competitors.[2] We bring together, for the first time, litigation data and the detailed information about inventions and their owners found in patent documents. The richness of this data set allows us to construct new proxies for variables thought to contribute to the likelihood of litigation. Most importantly, for each patent we will have several indicators for (i) the number of *potential disputes* - the number of claims in the patent, the diversity of technology classes into which the patent falls, and the technological similarity of future patents that cite the original one; (ii) the

---

[2] Specifically case filings. Throughout the paper, the term 'litigation' is used interchangeably with case filing, and does not imply that the case went to trial.

size of the stakes in patent litigation - the number of future patent citations the patent receives, and the extent of self-citation; and (iii) the *cost of litigation* and the *structure of information* - which we will infer from whether the patent ownership is individual or corporate, and domestic or foreign.

From an R&D perspective, the central question is whether patent litigation seriously dilutes or distorts the incentives provided by the patent system. Studies of patent renewal data and other sources have demonstrated that property rights from patents provide a significant R&D incentive, and there is evidence that these incentives vary across technology fields (see Lanjouw, Pakes and Putnam, 1996, for a review). But in addition to providing an incentive to invest in R&D, Waterson (1990) also emphasizes that patent protection changes firms' product location decisions as they try to avoid costly litigation, and that this can have negative welfare effects. This underlines the broader point that the costs of enforcing patent rights can significantly alter the R&D incentives provided by the patent system. We need to understand more about this interaction in designing patent and legal policy. How common is litigation? Are valuable patents more vulnerable to infringement litigation? Is the value associated with broader patents at least partly offset by a greater likelihood of litigation? Does litigation distort the structure of R&D incentives by technology area, nationality or other characteristics of the patent owner (such as individual or corporation)? Such differences are potentially important because firms will adapt their strategies to the prevailing incentives. For example, there is evidence that small firms tend to rely more heavily on trade secrets rather than patents for protection (Baldwin, 1996; Lerner, 1995a), and that the strength of patent rights affects the way firms structure their contracts and relationships with other firms for technology transfer (Anand and Khanna, 1996). Finally, is litigation more pervasive in new technology areas? If inventors in new technology areas get ensnared in frequent litigation it may retard the development of new areas, particularly if small, capital constrained firms are a leading source of inventions.

The paper is organised as follows. Section 1 summarises the theoretical hypotheses about the determinants of litigation, derived from existing literature, and relates them to the specific context of patent litigation. The data set is described in Section 2. In Section 3 we develop the key stylized facts of patent litigation and discuss them in the context of the hypotheses identified in Section 1. In this section we do not distinguish between patent infringement and patent invalidity (hereafter, challenge) suits because there are no important qualitative differences in the results. In Section 4 we present separate econometric estimates of the determinants of infringement and challenge litigation. This allows us to quantify the importance of the various factors, identified in Section 3, and to identify a striking difference between the two types of cases. We conclude by summarising the main findings and their implications for the R&D incentives generated by the patent system.

1. Determinants of Litigation Activity

Existing theoretical models identify four key determinants of the likelihood of observing a filed case. We summarise them here, drawing heavily on the excellent survey of theoretical models of litigation by Cooter and Rubinfeld (1989), and discuss them in the context of patent litigation.

A. The likelihood of a potentially litigious situation, or 'event', occurring and being detected by the plaintiff. In this context, an event is any action which could be considered an infringement of patent rights.

The likelihood of a potentially litigious 'event' increases when patenting is relied upon extensively to appropriate returns to inventions. If all inventions are patented, rival inventors will more often make competing claims. Since survey evidence shows that the reliance on patents differs sharply across technology areas (Levin, Klevorick, Nelson and Winter, 1987; Cohen, Nelson and Walsh, 1996), we would expect significant technology group differences in litigation rates. Second, the likelihood of an event increases when there are many firms innovating or producing in the same technology area. Thus the probability of litigation should be a function of the "crowdedness" of the area in which the invention applies.[3] Third, patents that are broader in scope, in the sense that they embody a larger number of claims, will be more exposed to potential infringement and thus litigation. Fourth, patents that are more general, in the sense of being applicable in many fields, may be more likely to be infringed simply because there are more potential infringers. In any case, we would expect detection to be more difficult for foreign firms, for individual patentees as compared to corporate owners, and for firms which have less experience in the technology area of the specific invention. On this account, the likelihood of

---

[3] Legal rules and procedures also affect the likelihood of an "event" by changing the incentives for parties to avoid conflict. Survey evidence indicates that small firms, which are likely to face higher legal costs, make less use of the patent system for appropriating returns to inventions (Levin *et. al.*, 1987; Lerner, 1995a). Lerner (1995b) suggests that the cost of trials dissuades small biotechnology firms from pursuing R&D in areas where they might come into conflict with larger firms. Anand and Khanna (1996) show that the strength of patent rights affects the way firms structure their contracts and relationships with other firms for technology transfer. A number of theoretical papers have explored how features of the legal system influence the level of care taken to avoid harm (for example, Shavell, 1996.)

litigation should be lower for such patentees.

Given a dispute, the parties may settle out of court or the plaintiff may file a case.[4] This choice is determined by the following sets of factors:

B. Asymmetric information or a divergence in the parties' expectations regarding the likelihood that the plaintiff will prevail at trial.

If the patentee and the infringer have sufficiently divergent expectations about the probability that the patentee will win at trial, they will be unable to reach a settlement agreement (Priest and Klein, 1984; Siegelman and Waldfogel, 1996). This can arise from uncertainty about whether infringement has occurred or about how courts will treat the property rights. The former will be greater in technology areas where inventions are difficult to define precisely in patent documents. Uncertainty about how courts will interpret patent rights will be greater in new technology fields, such as biotechnology, and in areas where patent protection is new, such as computer software. This arises both because there is little precedent on which to judge whether certain actions constitute infringement, and because patent claims cannot be as well defined in relation to prior art.[5] There will also be more uncertainty when legal procedure itself is changing, as in the period after the introduction of the specialised patent court in 1982 in the United States.

Asymmetric information models of litigation show that the likelihood of trial (relative to settlement) depends delicately on the particular information available to the parties and the sequence of offers. Bebchuk (1984) and Png (1983) show that, when there is asymmetric information, one or both of the parties may make unacceptable settlement offers with positive probability as a signalling device, and this probability (and thus the likelihood of trial) rises with the degree of asymmetry. Spier and Spulber (1993) show that when the plaintiff has private information about case quality and the defendant about damages, the likelihood of trial falls in the accuracy of information. Since litigants in patent suits are primarily firms (see Siegelman and Waldfogel, 1996), we expect greater asymmetry of information when the patent owner is an individual.

---

[4] In the theoretical literature, the decision whether to settle or proceed to trial is usually modelled from a point after case filing. In the discussion here, we assume that the factors which affect settlement probabilities after a case has been filed, identified in that literature, have the same qualitative effect on the likelihood that a dispute is settled before a case is filed.

[5] For detailed discussion of the evolution of legal rules regarding computer software patents, and the uncertainty it created, see Samuelson, Davis, Kapon and Reichman (1994).

C. The size of and differences in the expectations of the parties about the damages to be awarded at trial (the stakes), and whether there are returns to filing a case not associated with damages.

With imperfect information, the probability of trial increases with the size of the stakes - unless trial costs increase in tandem, as with contingency fees proportional to damages (Waldfogel, 1995). In a patent dispute, the stakes should be proportional to the private value of the patent rights, that is on the *incremental* returns associated with the patent monopoly relative to the next best means of protecting the innovation. Thus, more valuable patents should be litigated more frequently.

There may also be private economic benefits from winning a trial, beyond any direct damages that might be more efficiently collected in a settlement. In such cases, there can be motive for litigation even when there is perfect information. The leading example is the beneficial reputation effect that may come from filing (and winning) a suit. Aggressively pursuing infringers now may dissuade potential infringers in the future (for a review of strategic models of reputation, see Tirole, 1989). Similarly, if a firm mounts an active defense when accused of an infringement it may avoid future suits. A reputation for toughness should be more valuable to inventors working in a crowded technology area. Another example of dynamic returns to litigation arises when inventions occur in "cumulative technology" areas - where subsequent developments build closely on earlier inventions. In such cases, it may be critical for a firm to enforce patent protection on its early patent in order to extract rents from its own or others' subsequent inventions (Scotchmer, 1991; Merges and Nelson, 1990).

D. The cost of trial relative to the cost of settlement.

Legal costs are probably higher for foreign firms. Even if they engage domestic legal representation, they will incur higher costs in communications and in translating business documents into a form which will be understood by a U.S. court. Costs of litigation are also likely to be higher for smaller firms and individuals because of higher financing costs and their greater reliance on external legal counsel (Lanjouw and Lerner, 1996). However, their costs of settlement are also likely to be higher. Corporate disputants may have repeated commercial interactions or other relationships -- such as cross-license or patent pooling arrangements -- which encourage cooperation and reduce the likelihood of litigation. These considerations suggest that foreign firms and individuals would be less likely to litigate than their domestic counterparts, but there is no unambiguous prediction about individual versus corporate patentees.

The generic factors discussed in this section should influence the likelihood of both patent infringement and challenge suits. But there is one important difference between these two types of litigation: successful patent challenges generate positive externalities while infringement suits do not. If the plaintiff in a challenge suit is active in related R&D, it may be well-placed to appropriate these gains if the disputed patent is declared invalid or is restricted in scope. But other R&D performing firms may also innovate in any technology space opened up, and all firms would be able to use the original innovation freely. By contrast, the gains from a successful patent infringement suit accrue primarily, if not exclusively, to the patentee. Besides making infringements suits more frequent than challenges, this externality has the testable implication that the determinants of litigation discussed in this section would have a stronger impact on the likelihood of infringement suits than on patent challenges (see Section 4).

2. Description of the Data

The data source used to identify litigated patents is the Patent History CD-ROM produced by Derwent.[6] This database is constructed from information collected by the U.S. Patent and Trademark Office (PTO). The data include 5,452 patent cases during the period 1975-1991, involving 3,887 U.S. patents. Although the U.S. Federal courts are required to report to the PTO whenever a case is filed which involves a U.S. patent, they often fail to do so. This means that the PTO data is a subset of all patent cases. We calculated the number of missing cases by linking the information from the PTO to comprehensive data on court activity available from the Inter-university Consortium for Political and Social Research (Federal Judicial Center, 1991). Only 22 percent of patent disputes recorded by the Federal courts (code 830) were reported to the PTO during the years 1977-79. This had increased to 85 percent of all cases by the years 1985-87, with the most substantial improvement in reporting occurring between 1983 and 1984. The incomplete reporting by the courts to the PTO seems to be mainly a clerical issue (reporting rates differ by districts). We checked for selection bias by comparing various features of federal cases reported to the PTO to those of unreported cases, and found no discernable difference between the two groups in a range of variables in the federal database.[7] Thus, although the data used

[6] We thank Derwent and Micro-Patent for making available to us these data and those discussed below, and Orli Arav for providing assistance in obtaining the data.

[7] The variables included the method of disposition of the case, the type of judgement, the percentage of large (greater than ten million dollars) monetary judgements received, and the party favoured in

here include only about half of all cases filed during the period of analysis, there is no reason to expect selection bias in the empirical results.

In order to understand what is different about patents, and patentees, that get involved in legal actions, we need a population sample as a control group. We generated a 'matched' set of patents from the population of all U.S. patents (both litigated and unlitigated). For each litigated patent, a patent was chosen at random from the set of all U.S. patents with the same application month and a common 4-digit International Patent Classification (IPC) sub-class assignment (see below). The selection was done with replacement and, as a result, 12 of the 3,887 matched patents are duplicates. By constructing the population sample in this way, the comparisons we will present between litigated patents and matched patents control both for technology and cohort effects.

The next step was to obtain information on a range of characteristics for each litigated and matched patent. These include the number of claims, the area of technology of the patented innovation, the patterns of backward and forward citations, the type of ownership and whether there is a third party interest in the patent. We briefly describe each of the variables used in the subsequent analysis.

A patent is comprised of a set of claims which delineate what is protected by the patent. The principal claims define the essential novel features of the invention in their broadest form and the subordinate claims are more restricted and may describe detailed features of the innovation claimed. The patentee has an incentive to claim as much as possible in the application, but the patent examiner may require that the claims be narrowed before granting. The number of claims for each patent was obtained from the Patent Scan CD-ROM (front page) database produced by Micro-Patent.

Each patent is assigned by the patent examiner to 9-digit categories of the IPC system. Our data contain assignments at the more aggregated 4-digit sub-class level (614 sub-classes). The IPC is a technology-based classification system and patents may be assigned to more than one sub-class. For example, pharmaceutical inventions are often classified both in A61K and C07K. In the empirical analysis, we use the set of all 4-digit IPC sub-classes to which each patent was assigned. An inventor must cite all related prior U.S. patents in the patent application. A patent examiner who is an expert in the field is responsible for insuring that all appropriate patents have been cited. Like the claims, the citations in the patent document help to define the rights of the patentee. For each patent in the litigated and matched data we obtained the number of prior patents cited in the application (backward citations)

---

cases were a final judgement was rendered.

and their IPC sub-class assignments. There are 55,205 backward citations in the data. We obtained the same information on all subsequent patents which had cited a given patent in their own applications, as of 1994 (forward citations). There are 69,248 forward citations in the data.[8] For recent patents there is substantial truncation in the number of forward citations as citation lags can stretch over decades (Jaffe and Trajtenberg, 1996). We discuss this more fully below. For older patents there is considerable missing information on the IPC sub-class assignments of backward citations, as comprehensive IPC data are only available from about 1970, but the *number* of backward citations is complete for all patents.

We construct a variable for the nationality and the type of ownership (corporate or individual) of each patent as follows. First, we classify the inventor as domestic, Japanese or other foreign, based on the address of the inventor. The nationality of the assignee of the patent is defined in the same way. Not all inventors assign their patent rights to others: 66 percent of litigated and 73 percent of matched patents have an assignee. The nationality of patent ownership is defined as that of the assignee, if there is one, and otherwise as the nationality of the inventor. With few exceptions, assignees are firms, whereas inventors are individuals. Thus we identify a patent as corporate-owned if there is an assignee.

For the litigated sample, we can obtain a rough indicator of whether there is a third party interest in the patent. When neither litigant in a case is identified as the patentee, it means that a third party, invariably a firm, is engaged in questioning or defending the patent rights, even though it does not own the patent. Third parties may be parent companies, exclusive licensees or joint venture partners. There are third party litigants for 26 percent of corporate-owned patents and 53 percent of patents owned by individuals. This measure understates the extent of third party involvement because even in the 47 percent of cases where an individual owner *is* identified as one of the litigants, there are, frequently, firms acting as co-litigants (in roughly half the cases). This implies that individuals are actually litigating in only about 25 percent of cases which involve individual ownership. It will be important to bear this in mind when interpreting the effect of ownership on litigation.

Finally, we distinguish patent infringement from patent challenge suits. We have no direct information on whether a filed case is an infringement suit or some other form of litigation, but we check whether the patent owner is the plaintiff or defendant in the case. We are able identify one of the litigants as the patent owner for about two-thirds of the patents in the data set. Among these, the patent owner was the plaintiff 84 percent of the time. We interpret these as infringement suits. The other cases,

---

[8] We thank Jonathan Putnam for generating the matched data set and for obtaining the citations information. The information was made available by Mogee Research and Analysis of Reston, Virginia, and was accessed through their PCAD database.

where the patent owner was the defendant, are almost surely suits for patent invalidation (in whole or part) brought by competitors.

3. The Stylized Facts

3.1 The Prevalence of Litigation

We first look at how frequently patents are litigated. Panel A of Table 1 provides estimates of the number of cases which have been or will be filed per thousand patents applied for during the period 1980-1984, broken down by technology field and ownership.[9] For example, the aggregate figure of 10.7 means that for every thousand granted patents applied for during the years 1980-84, they will eventually become the subject of 10.7 filed cases. It is an estimate because of the need to adjust both for the under-reporting of cases (discussed in Section 2) and for truncation. Based on the comprehensive ICPSR data, the number of cases reported to the patent office by 1991 is assumed to be three-quarters of the true number of cases involving these patents by that date. Because the litigation data end in 1991, there is truncation at seven to eleven years after application. Some of the patents from cohorts 1980-84 which appear as unlitigated in 1991, will become the subject of a dispute late in life. To estimate the size of the truncation effect we examine the first few cohorts in the data (1975-77), where one would expect truncation to be minimal, and find that 45 percent of all case filings occur within seven years of the application date of the patent, and 75 percent within eleven years. We use the midpoint of this range, and the estimated 25 percent under-reporting, and gross up the litigation rates accordingly.[10]

Turning to the breakdown by ownership, the case filings involving foreign-owned patents (including Japanese) tend to occur later than for domestically-owned patents, so the truncation effect is larger. Following the procedure above, we adjust by a truncation rate of 48 percent for foreign-owned patents and 41 percent for domestically-owned patents. To compute total patents, in the denominator, ownership is imputed based on the country of first application (the priority country). Inventions are often

---

[9] We thank Sam Kortum for providing data on total patenting activity.

[10] Because some patents are litigated more than once, the number of cases per thousand patents is larger than the number of individual patents which are litigated. The latter statistics, adjusted for truncation in *first* filings, are 6.3 (total), 9.1 (domestic), and 2.0 (foreign). That is, for every thousand patents, 6.3 will be litigated and they will generate 10.7 cases.

patented in multiple countries but, with few exceptions, the priority country is the country of the inventor (Putnam, 1996). Thus a patent is considered to have a domestic owner if the priority country is the U.S. and it is counted as foreign if the priority country is elsewhere. Most exceptions involve cases where a foreign patentee chooses the U.S. as the priority country.[11] Thus the share of total patents with domestic owners is slightly over-estimated and, as a result, the trial rate for domestic patentees is under-estimated, and vice versa for foreign owners.

Despite this bias, we find that domestically-owned patents are involved much more frequently in litigation. The aggregate litigation rate is nearly five times as large for domestic patentees, and this finding holds (with some variation in the magnitude) in all five technology areas. There are several explanations. First, foreign patentees may find it more difficult to detect infringements occurring in the U.S. This is suggested by the fact, noted above, that there is a longer lag between the patent application date and the first case when the owner is foreign. Second, it may be more costly for a foreign firm to pursue a case in the U.S. courts. Third, foreign patents may be less valuable. We will try to discriminate between these hypotheses below.

Note first that, although domestically-owned patents are more likely to be litigated, they are no more likely than foreign-owned patents to be litigated *more than once*. Japanese patentees own 1.6 percent of patents litigated once and 1.4 percent of patents litigated more than once; other foreign patentees own 9.6 percent of patents in both categories. The distribution of the number cases per patent is not significantly different across ownership categories ($\chi^2/2=0.084$, p-value=0.96). And finally, among litigated patents the mean number of cases per patent is 1.39 (0.02) for domestic owners, 1.30 (0.10) for Japanese owners and 1.53 (0.10) for non-Japanese foreign owners.[12] These means are not statistically different from each other at the 0.01 level.

This evidence strongly suggests that the nationality of the patent owner influences the *probability* of litigation, but it does not affect the *frequency* that a given patent is litigated. This finding is consistent with there being a fixed cost to litigating a given patent, which is higher for foreign firms. It is not consistent with the idea that foreign firms are at a disadvantage in detecting infringements in the U.S. market since, if that were true, we would also expect the number of cases per litigated patent to be lower for foreign firms. An alternative to the fixed-cost interpretation is that there are two types of foreign

---

[11] Putnam (1996) calculates that 98 % of U.S. inventors have their residence as priority country. Figures for other leading OECD countries include: France, 84 %; Germany, 88 %; the U.K., 83 %; and Japan, 87 %.

[12] Throughout the paper, figures in parentheses are standard errors.

patent owners, those with low costs of detecting infringement and litigation (for example, owners with a domestic presence by virtue of a local subsidiary) and those with high costs. If most foreign patent owners are of the first type, the probability of litigation will be low. But the low-cost type will be highly represented among the foreign patent owners who do get involved in litigation, so the frequency of cases, conditional on a first filing, may be high.

There are also sharp differences in litigation rates across technology fields, holding ownership constant. The most notable are the very low litigation rate for chemical patents and the high rate for drugs and health. There is a cased filed for every fifty drug patents. In part this reflects the fact that patents are relied upon more heavily to protect inventions in pharmaceuticals than in other industries, as confirmed by extensive survey evidence (Levin, *et. al.*, 1987; Cohen, Nelson and Walsh, 1996). The low litigation rate for chemical patents is consistent with the hypothesis that there is less litigation when it is clear whether an action constitutes an infringement, leaving little room for discrepancies in expectations between parties. Interviews with patent lawyers indicate that it is relatively straightforward to describe chemical inventions in patent claims. Finally, some of the technology group differences in litigation rates may be associated with variation in the value of patent protection. (See Lanjouw, 1993, and Cohen, Nelson and Walsh, 1996, for evidence of large differences in the value of protection across technologies.)

A comparison of the litigation rates in the table with those for biotechnology presented by Lerner (1995a) strongly supports the hypothesis that uncertainty makes it more likely that disputes end up in court. Based on a sample of 530 new biotechnology firms, Lerner estimates that as much as 6 percent of U.S. corporate biotechnology patents are litigated.[13] This figure is much higher than the litigation rates in Table 1 and, even allowing for different estimation procedures, it clearly indicates that there is a great deal of litigation in this field relative to other technology areas. This is not surprising since biotechnology invention is very recent, and there is little precedent to guide disputants in assessing how the courts will judge their case.

Finally, there are sharp differences between litigation rates for individual and corporate owners (not shown in the table). Litigation rates for domestic and non-Japanese foreign individuals are 16

[13] The figure is only a rough estimate, obtained by dividing the number of patent suits in which these firms were involved in Massachusetts during January 1990-June 1994 by the number of patents they were awarded during that period. The litigation rate should be expressed relative to the existing *stock* of patents, so there is likely to be some downward (upward) bias to the extent that the number of patent awards in that period was abnormally high (low). Since biotechnology is a young field, Lerner's estimate is likely to be too low.

percent *higher* than corporate owners in those countries. Even more striking, Japanese individuals are 3.3 times more likely than Japanese corporations to become engaged in litigation involving their U.S. patents. This evidence strongly suggests that, while corporate owners may have lower litigation costs, they must have even greater advantages in reaching acceptable settlement agreements. But this comparison does not control for other characteristics of the patent, and we will see in Section 4 that the conditional effect of ownership is less pronounced.

Panel B in Table provides information on the type of litigation occurring in each technology group and ownership category. We have already defined infringement (challenge) suits as those where the patent owner is the plaintiff (defendant). The unclassified category refers to cases where the patent owner is neither the plaintiff nor defendant. Under U.S. patent law, an exclusive licensee (but not a non-exclusive licensee) can sue for infringement on behalf of the patentee, but the defendant in a challenge suit must be the patent owner. Thus we conclude that the unclassified category is primarily be infringement suits brought by an exclusive licensee or the parent company of the patent owner. In each technology group, infringements account for the bulk of litigation; about 60 percent if unclassified cases are excluded, and 90 percent if they are treated as infringement suits. This is consistent with the fact that the plaintiff is better able to internalize the gains from a successful infringement suit than from a patent challenge. Turning to the breakdown by ownership category, the unclassified category is considerably larger for foreign-owned patents than for domestic patentees (48.8 versus 29.2 percent). This is not surprising since one would expect foreign patentees to be more likely to license their U.S. patents (relative to domestic owners), and this category includes infringement suits brought be exclusive licensees.

3.2 Patent Citations

Future citations received by a patent are one indication that it has contributed to the development of subsequent inventions. For this reason, they have been used as one measure of the social value of an invention. Trajtenberg (1990a and 1990b) first presented compelling evidence that citations were related to the gains in profit and consumer surplus due to innovation occurring over time in CAT scanner technology. Patent citations have also been used successfully to trace R&D spillovers and to relate them to growth and other performance measures (Jaffe, Henderson and Trajtenberg, 1993; Caballero and Jaffe, 1993). We use citations to investigate whether litigation is systematically related to the (private) value of patents.

The link between citations and litigation in these data is striking. Litigated patents are far more heavily cited than a randomly chosen patent; the mean number of citations among the litigated patents is 11.9 (0.2), but only 5.8 (0.1) for matched patents. To some extent this is driven by the fact that domestically-owned patents tend to be more heavily cited in the data, combined with the weight of domestic patents in the litigated sample. In the matched sample, domestically-owned patents are cited 6.3 (0.16) times on average, compared to 4.8 (0.19) citations for non-Japanese foreign-owned patents and 5.8 (0.32) for Japanese-owned patents.[14] However, litigated patents are more heavily cited even when we control for ownership (Table 2). We can easily reject the null hypothesis that litigated and matched patents have the same distribution of the number of citations, both for domestically-owned patents ($\chi^2/6 = 190.8$, p-value < 0.01) and foreign-owned patents ($\chi^2/6 = 492.9$, p-value < 0.01). Considering Japanese-owned patents separately does not have any appreciable effect on the results. Table 2 clearly shows that litigation is much more likely to be a feature of maintaining property rights for more valuable (heavily cited) inventions.

These data, together with findings from previous research, can help distinguish between the alternative explanations for the lower litigation rate among foreign-owned patents. A number of studies of the private value of patent protection in different countries have found that foreign-owned patents tend to be more valuable than those owned by domestic patentees (Pakes and Simpson, 1989; Lanjouw, 1992; Putnam, 1996. Schankerman, 1997, provides mixed evidence).[15] Thus, it does not appear to be the case that foreign-owned patents are litigated less because they are less valuable. Second, since foreign-owned patents are less heavily cited in the population (matched sample), we can infer that they receive fewer U.S. citations per unit of patent value. While domestic patents are more heavily cited than patents owned

[14] This difference may simply be due to the truncation in the data as foreign-owned patents are cited more slowly than domestically-owned patents. The mean number of forward citations for the subset of data which includes only patent cohorts before 1978 is 6.8 (0.18) for domestically-owned patents, 5.4 (0.25) for non-Japanese foreign patents, and 6.9 (0.47) for Japanese owners. A related finding is presented by Jaffe and Trajtenberg (1996), who show that foreigners cite domestically-owned U.S. patents more slowly. They interpret this result as an indication of lags in the geographical diffusion of information, which would also explain our finding.

[15] This finding is due, at least in part, to self-selection in the patent application process -- foreign patentees are less likely to take out patent protection in other countries for low-valued inventions, given the substantial cost of applying for and maintaining patent protection. The finding could reflect differences across nationalities in the underlying distribution of the value of their inventions, and hence differences in the value of the associated patent rights (i.e., U.S. inventions may simply be less valuable on average). But studies of patent systems *in different countries* tend to find foreign-owned patents are higher valued, suggesting that it is more to do with selection.

by foreign patentees for the matched sample, this is *not* the case among litigated patents. Those with a foreign owner are just as heavily cited as those with a U.S. owner (11.6 and 12.0 citations on average, respectively). Using the same breakdown of citations categories shown in Table 2, the test for equality of the distributions in columns (2) and (4) yields a $\chi^2/6 = 2.45$ with a p-value of 0.88. The fact that citation distributions are the same in the litigated sample implies that litigated patents owned by foreigners are *more* valuable than their domestic counterparts. This conclusion is consistent with the hypothesis that foreign patentees face higher costs of litigation than do domestic patentees, and thus only litigate particularly valuable patents. But the finding is not consistent with the hypothesis that foreigners have more difficulty in detecting infringements, since there is no particular reason to expect the latter to have a selective (value related) effect.

As noted earlier, one explanation for the finding that more heavily cited patents are more likely to be litigated is that citations are an indicator of value and litigation is positively related to the value of the stakes. Another, mutually consistent, hypothesis is that when many competitors are innovating in the same area, they are likely both to generate citations and to make competing claims which lead to litigation. This would also tend to make reputation more important. We investigate this hypothesis by noting that where many inventors are operating in the same area, their patents will tend to be more similar in technological characteristics to the patents they cite and to those patents that cite them. To measure the similarity between a patent and one of its citing patents, we calculate the percentage of 4-digit IPC assignments of the citing patent which overlap with those of the patent itself. The degree of similarity between a patent and one of its backward citations is measured analogously. Similarity is defined for pairs of patents and measures the technological closeness of a patent to its children and its parents. Greater overlap implies greater similarity. Our similarity index is the mean degree of similarity taken over a group of citations.[16] We treat self-citations separately since, when a firm cites its own past

---

[16] Our similarity index differs from the "generality" of a patent as defined by Henderson, Jaffe and Trajtenberg (1995). Generality refers to the diversity of IPCs into which citing (or cited) patents fall without reference to the IPC sub-class assignments of the patent itself. It measures how different citing patents are *from each other*. Generality is defined as $G = 1-\Sigma x_i^2$ where x_i is the percentage of the total number of a patent's 4-digit IPC sub-class assignments of its citing patents which are in sub-class i ($0 \leq G \leq 1$). The summation is over all sub-classes to which at least one citing (cited) patent is assigned. We computed the generality index for our data and find two results of interest. First, as one would expect, self-citations are less general than citations by others. Second, the backward self-citations of litigated patents are much less general that those of matched patents: G=0.15 (0.014) for litigated patents, G= 0.21 (0.013) for matched. In other words, inventions which are litigated seem to draw on a less diverse set of past patents of the firm. There are no significant differences between matched and litigated patents in the generality of their forward or backward citations by others.

patents, they are more likely to be technologically similar because they arise from the same research programme.

Table 3 presents the results broken down by domestic and foreign owners. For backward citations, we cannot reject the hypothesis that the similarity index is the same for litigated and matched patents. This holds both for self-citations and citations to others' patents, and for domestic and foreign owners. Somewhat surprisingly, backward self-citations are not significantly more similar than citations to others. The striking result is for forward citations, where there are large and significant differences in the similarity index between litigated and matched patents. The similarity between litigated patents and those which cite them is much higher than the similarity between matched patents and their subsequent citations. This holds both for forward self-cites and citations by others, and for domestic and foreign patentees. This finding supports the idea that crowdedness in a research area contributes to litigation by increasing the number of events and the importance of reputation. Here we also observe the expected pattern that future self-citations are more similar to the patent than citations by others. Finally, the similarity index is not significantly different by ownership type - either domestic/foreign or individual/corporate (not shown).

It is also striking that there is significantly greater similarity between patents and their future self-citations among the litigated patents than there is among the matched patents - 0.67 versus 0.58 for domestic, 0.72 versus 0.63 for foreign owners. Thus patentees are more likely to litigate when they have subsequent inventions in the same technology areas. One interesting explanation for this finding is that greater similarity of future self-citations is an indication of cumulative or sequential invention. In such cases, the ability to appropriate returns from later inventions may depend on having effective proprietary control over the earlier invention. This can arise because stronger control of earlier inventions changes the bargaining position for subsequent licensing agreements (Scotchmer, 1991 and 1996; Scotchmer and Green, 1995). This gives a firm trying to control a technology as it develops a strong incentive to prosecute infringers of the early patents. Not only does this increase the stakes of the dispute, but, perhaps more important in explaining litigation, it is also likely to cause an *asymmetry* in the stakes. The positive externalities across patents created in the protection of cumulative innovations of the same patentee will only be enjoyed by the patent owner.

We have shown that the number of forward citations and the likelihood of litigation are positively related, but what is the direction of causality? The explanations proposed earlier assume that the causality runs from citations to litigation. But when deciding which previous inventions to cite, the patent applicant (or patent examiner who screens the application) is faced with a large body of patents to search.

When a patent is disputed, it may receive more attention and, as a result, more citations than it would have otherwise. To investigate this hypothesis, we compute the mean number of citations received by litigated patents at various lags after the application date. The first column in Table 4 presents the mean number of citations at various lags for patents which have not yet had a first case filing by the time of the citation. For example, the value 0.96 at lag six indicates that, for patents which will eventually be litigated but which have not been involved in litigation by six years after application, the mean number of citations in that year is 0.96. The second column gives the mean number of citations T years after application for patents which had a first case filing T-1 years after application. The following columns give the same figure when the first case filing date is progressively further from the year of citation. The final row of the table presents the mean of the preceding rows and gives clear evidence that there is a "publicity effect" associated with a patent dispute going to court. The mean number of citations for given lags is significantly higher in the two years following a case filing. The effect of the case filing wears off over time and by the fourth year after filing no longer increases citations.[17] However, while there *is* an observable publicity effect, it is much too small to explain the higher number of citations for litigated patents. The point estimates in the last row of Table 4 imply that the publicity effect raises the number of citations received by a litigated patent by an average of just 0.50, a half a citation.[18] The actual difference in the mean number of citations between litigated and matched patents is 6.1, so the publicity effect can account for less than 10 percent of the observed difference in litigation rates.

3.3 Patent Claims

The first part of Table 5 shows that, in terms of the typical number of claims embodied in a single patent, technologies break into two distinct groups: drugs and health, chemical, and electronic inventions have more claims per patent, while patents protecting mechanical and other types of inventions have fewer claims. Equality across technology groups in the distribution of claims is easily rejected

---

[17] Columns (2)-(5) are increasingly heavily weighted towards patents for which a first case is filed early after application. One might suspect that the higher number of citations shown after a case is filed is simply the result of patents which are litigated early being the more valuable patents. But the fact that the effect tapers off suggests that it represents news associated with bringing a case.

[18] The publicity effect is somewhat greater for more highly cited patents. Considering all patents, the publicity effect accounts for 4% of mean citation. Restricting attention to increasingly highly cited patents, the percentage of mean citation attributable to the publicity effect of litigationg grows but is never more than 10%.

($\chi^2/20 = 4.56$, p-value < 0.01, based on the claims categories 1-5, 6-10, 11-15, 16-20, 21-50 and > 50).[19] The mean number of claims per patent also differs significantly across ownership types: the mean for domestic owners is 13.2 (0.2), it is 10.7 (0.2) for non-Japanese foreign owners, and 9.0 (0.3) for Japanese owners. This conclusion holds within each technology field as well.

The mean number of claims per patent has increased over time (Table 5). It was stable prior to 1983, averaging 12.2 (0.15) over the period 1975-1982. But in 1983 the mean number of claims began to drift upward and averaged 13.5 (0.24) for the cohorts after 1982. This evidence is consistent with the hypothesis that the number of claims is related to the private value of a patent, because in 1982 the cost of patent protection in the U.S. increased sharply. In that year, the PTO raised application and issuance fees and introduced the requirement that patentees pay renewal fees to keep their patents in force.[20] This increase in the cost of protection should have led inventors to cease applying for patent protection on less valuable inventions, and the observed rise in the number of claims per patent is consistent with self-selection.[21] There is also some supportive, cross-sectional evidence that the number of claims is related to the value of the invention: the number of claims is positively correlated with the number of forward and backward citations in all technology areas in these data (Lanjouw and Schankerman, 1997).

In addition to being an indicator of patent value, the number of claims is associated with the technology or product "space" being protected by the patent. All else equal, a patentee making many claims runs a larger risk of conflict with competitors. For both reasons one would expect to such patents end up in court, and this is what we observe. Litigated patents have far more claims than matched patents, both when they are domestically-owned and foreign-owned (recall that this comparison controls for IPC classification). The differences are large and statistically significant, and are also evident in the median and mode (Table 6). Controlling for ownership is necessary because, as with citations, foreign-

---

[19] The test done deleting outliers (claims >55) gave the same results.

[20] Expressed in 1995 dollars, filing and issuance fees were $1260 in 1982 and $2000 in 1995. For a patent applied for in 1982, renewal fees for full term patent protection would be about $5600 in 1995 dollars.

[21] An alternative explanation is that the increase in the fixed cost of patenting to lead inventors to try to package their new bits of information (claims) into fewer patents, raising the number of claims per patent. However, we cannot empirically distinguish between this repackaging hypothesis and the self-selection described in the text. This increase is all the more striking because, during this period, the new specialized appeals court liberally interpreted the doctrine of "means and functions claims," which allows patentees to widen the coverage of individual claims. We thank Robert Merges for pointing this out.

owned patents tend to have significantly fewer claims in general (row 1). Not only is the mean number of claims larger for litigated patents, but the number of forward citations per claim is also higher (controlling for ownership). Thus litigated patents have both more claims, and more valuable claims. Moreover, we cannot reject the hypothesis that, *among litigated patents*, domestic and foreign-owned patents have the same number of citations per claim. As in our earlier analysis of citations per patent, this highlights the self-selection at work among foreign-owned, litigated patents.

In Table 6 we also examine the number of prior patents cited in the patent documents of litigated and matched patents (backward citations). Controlling for ownership, we find that a litigated patent is likely to cite *fewer* prior patents per claim than a randomly selected patent. One interpretation is that a small number of backwards citations indicates that the invention is in a relatively new technology area. In this situation there is little information available to either the patentee or his competitors about how the courts will view the claims made in the patent, increasing the likelihood of divergent expectations and, as a result, litigation. An alternative hypothesis is that a small number of backward citations just reflects a failure by the patentee to cite relevant patents. This would increase the probability of the new patent infringing on earlier ones, and thus litigation. But under this hypothesis we would expect to observe challenges to the new patent, not infringement suits by the patentee. To distinguish between these hypotheses, we group patents into those where the owner is the plaintiff (infringement suits) and those where the owner is the defendant (patent challenges). Patents where the owner is neither plaintiff nor defendant are dropped for this exercise. Under the "new area" hypothesis, the number of backward citations per claim should not be significantly different for infringements and challenges, but under the "failure to cite" hypothesis, backward citations per claim should be larger for infringement suits. We use citations per claim because of the large variance across patents in the number of claims. This comparison was made for domestic, non-Japanese foreign and Japanese patents and there was no significant difference for any of the three groups.[22] This evidence does not support the "failure to cite" hypothesis but it is consistent with the idea that fewer backward citations per claim is an indication of patenting in relatively new areas, which in turn is associated with more frequent litigation.[23]

---

[22] The number of backward citations per claim for infringements and challenges are: 1.53 (.09) vs 1.15 (0.14) for domestic owners, 0.76 (0.07) versus 0.95 (0.16) for non-Japanese foreigners, and 0.88 (0.14) versus 0.73 (0.31) for Japanese owners.

[23] This finding is not surprising since, under U.S. patent law, "gross negligence" in citing prior art is sufficient grounds for rendering the entire patent unenforceable (not just those claims under dispute). See Merges (1997), Ch. 7.

3.4 International Patent Classification Assignments

We use the IPC sub-class assignments to construct Lerner's (1994) measure of the breadth of a patent: a simple count of the number of 4-digit IPC sub-classes to which a patent is assigned by the patent examiner. We also develop a second, more refined measure that takes into account how specific a patent is to a particular technology area. All else equal, we would expect an invention which has uses in many areas -- a *non*-specific patent -- to be faced with more potential infringers and thus be litigated more frequently.

Table 7 summarises the results for Lerner's measure of patent breadth. We easily reject the null hypothesis that the distribution of the number of IPC sub-classes to which litigated and matched patents are assigned is the same. The tests for individual technology areas show that differences in the breadth of litigated and matched are driven by mechanical patents. What is very surprising is not that there are differences, but that they go the "wrong" way. It is the *narrower* patents which tend to be litigated more often. This sharply contrasts with Lerner (1994) who studies a sample of biotechnology patents and finds a large, statistically significant positive relationship between the probability a patent is litigated and the number of its 4-digit IPC sub-class assignments. Table 7 shows that there is variation across technologies. For example, for the chemicals technology group we also find that litigated patents are (insignificantly) broader. But the more comprehensive data for patents from all technology areas clearly point to the opposite conclusion: broader patents, as measured by the number of distinct 4-digit sub-class assignments, are *less* likely to be litigated. This finding is surprising since one would expect broader patents to have a higher probability of infringement, other things equal. But it may also be more difficult for the patentee to detect infringement in diverse technology areas.

Our second measure of patent breadth is defined as one minus the maximum percentage of patent's 4-digit IPC sub-class assignments which fall into one of twenty-two 2-digit technology groups. Unlike a simple count of 4-digit IPC assignments, this measure recognizes that 4-digit sub-classes may be closely related (a patent with three sub-class assignments all within chemicals is less broad by this measure than a patent with two sub-class assignments where one assignment is in chemicals and the other is in electronics). Using only patents with at least two 4-digit IPC sub-class assignments (if there is only one, breadth is zero), the breadth measure for litigated patents is 0.308 (.010) and for matched patents it is 0.315 (.009). The point estimates confirm again that broader patents are *less* likely to be litigated, but here the difference is not statistically significant. This conclusion also holds for each of the five technology groups (not shown).

4. Econometric Analysis

In this section we present probit regressions that relate the probability of infringement and challenge suits to the following regressors: the number of claims, forward citations per claim, backward citations per claim, the number of 4-digit IPC's to measure patent breadth, similarity indices for backward and forward citations to measure crowdedness, the percentage of backward and forward citation which is *self*-citation to capture cumulative technology, and a set of ownership dummy variables to allow for nationality and individual/corporate differences. Technology group dummies are included as controls since the 50:50 breakdown of litigated and matched patents by IPC group does not hold for the separate samples of infringement and challenge suits. This analysis allows us both to confirm the stylized facts from Section 3 in a multivariate setting, and to quantify the impact of those factors on infringement and challenge suits.

Table 8 summarises the results. The first two columns present the parameter estimates and the implied marginal effect of each covariate on the litigation probability for infringement suits; the other columns provide estimates for challenge suits. Where there are quadratic terms, the reported marginal effect includes the full effect of a change in the variable. For the similarity of forward citation, the marginal effect indicated for the interaction is the full effect of a change in similarity for corporate owners. The marginal effects for ownership are calculated as follows: for ownership type i it is the difference between the probability of litigation given i and a weighted average of the probabilities of litigation given ownership type j, where the weights are the probability that a patent is type j given that it is not type i. Because the characteristics of patents (and their owners) involved in infringement and challenge suits differ, we evaluate the marginal effects at the *population means* for all covariates (see Appendix I for details). Thus the reported marginal effects correspond to a randomly drawn patent and can be compared between the infringement and challenge regressions. In each regression, the sample includes the litigated patents of the indicated type plus all of the matched patents (6151 and 4227 observations, respectively).

We first analyze infringement suits, which are the predominant type of patent litigation. The results strongly confirm that the probability of litigation rises with the number of claims and forward citations per claim, and the effects are substantial. A ten percent rise in the number of claims (1.0 claim at the mean) implies a 1.4 percentage point increase in the probability of litigation. Because the number of claims is very skew, a one standard deviation increase in the number of claims increases the probability of litigation by 12.0 percentage points. One additional forward citation per claim raises the probability of an infringement suit by 8.1 percentage points (standard error of 0.57). A one standard deviation increase in forward citations per claim raises the probability of litigation by 13.1 percentage points. These findings clearly show the importance of the value of a patent in determining infringement disputes.

The point estimates suggest that the likelihood of an infringement suit declines with the number of backward citations per claim, however, the effect is not significant. The point estimate is consistent with the hypothesis, proposed in Section 3, that a small number of backward citations indicates patenting in a relatively new area and that the associated uncertainties lead to more frequent patent disputes. However the effect is insignificant.

There is no evidence that Lerner's index of patent breadth (NO4IPC) increases the probability of infringement litigation. On the contrary, as we found in Section 3, the point estimate indicates that broader patents, on this measure, are *less* likely to be litigated. But the effect is marginally significant and quantitatively small: a one standard deviation increase in breadth lowers the litigation probability by only 1.1 percentage points (standard error of 0.6). Our second measure of breadth, that takes into account the relatedness of IPC classifications, had no detectable effect on the probability of litigation.

The similarity of forward citations by others (SIMFWD) significantly raises the probability of infringement litigation, and the effect is substantial. Moreover, the effect of the similarity index is almost twice as large for corporate owners than for individual owners: a one standard deviation increase in the similarity index raises the likelihood of infringement litigation by 3.6 percentage points for individual owners and 6.2 percent for corporate owners. Recall that greater similarity, conditional on the number of citations, is a measure of the "crowdedness" of a technology area. This causes litigation because it increases both the likelihood of disputes, as well as the importance of reputation in dealing with disputes. Our finding that similarity increases the probability of litigation more sharply for corporate owners, who are likely to be more sensitive to reputational concerns, suggests that this latter explanation may be an important factor in patent litigation.[24]

In Section 3 we suggested that forward *self*-citations for a patent (given its total number of forward citations) may be an indicator that the patent owner is engaged in subsequent inventions that build

---

[24] The similarity of backward citations by others does not affect litigation (not reported in the table). The conclusions do not change when we also include similarity measures for forward and backward *self*-citations in the analysis. The inclusion of these variables reduces the sample size by more than two-thirds; the point estimate on similarity of forward self-citations is also positive and significant at the 0.10 level, but backward self-citations have no effect.

on this earlier patent and that, as a result, he would have a greater incentive to protect his property rights in this area. This hypothesis is supported by the positive and significant coefficient on the variable FWDSELF, the percentage of citation which is self citation. The point estimate implies that a one standard deviation increase in this variable raises the probability of an infringement suit by 4.0 percentage points. This evidence supports the idea that there are complementarities among technologically-linked inventions in a firm's R&D portfolio, and that this raises the willingness to protect the property rights of inventions in this chain. But at the same time, we find that greater *backward* self-citation (BWDSELF) significantly reduces the likelihood of litigation: a one standard deviation rise in this variable lowers the litigation probability by 4.4 percentage points. Other things equal, greater backward self-citation in a patent indicates that an invention builds more extensively on one's own past research and is thus more likely to be a "derivative" invention. It is not surprising that we find that such patents are less likely to be involved in infringement litigation.

Finally, there is some evidence of ownership differences in infringement litigation, but the effects are smaller than was suggested by the simple comparisons in Section 3. We can easily reject that there are no nationality differences ($\chi^2/4=375$, p-value$<.01$). Foreign individuals and corporations are less likely to engage in infringement suits than their domestic counterparts.[25] Comparing the marginal effect of DINDOWN, for example, with that of FINDOWN and JINDOWN we see that, at the population (matched sample) means, non-Japanese foreign and Japanese individual owners are about 30 percentage points, respectively, less likely to become engaged in litigation than domestic individual owners. A similar pattern is seen among corporate owners. Second, we cannot reject the null hypothesis that the litigation probability is the same for corporate and individual owners of a given nationality ($\chi^2/3=7.34$, p-value $= 0.06$). But if we look only at domestic patentees, it appears that corporate owners are more likely become engaged in litigation than individual owners ($\chi^2/1=4.34$, p-value$=0.04$). This reverses the finding in Section 3, and is primarily due to the inclusion of the similarity index, which has a stronger effect for corporate owners. Most, if not all, of the corporate ownership effect appears to be linked to the crowdedness of the technology area, which we interpret as evidence for the importance of reputation in patent litigation.

---

[25] As discussed in Section 3, foreign-owned patents are litigated more slowly, that is, first filings come later after the patent application date. Thus, truncation, with respect to the probability of litigation, is greater for foreign-owned patents. We checked whether this was driving the nationality of ownership results presented in Table 8 by re-estimating the equations using data for various numbers of the earlier cohorts. There were no dramatic changes in the estimates, although, as one would expect, some of the coefficient estimates lost significance as the sample size was reduced.

The empirical results for patent challenges are very similar to those for infringement suits so we do not discuss them extensively. The new and striking finding is that the *marginal effects* of the covariates in the patent challenge regression are nearly proportional to those in the infringement regression: the factor of proportionality varies between 18 and 50 percent. We cannot reject the null hypothesis of a constant factor of proportionality across the sets of marginal effects ($\chi^2/12=13.77$, p-value=0.33) and, under this constraint, the obtain a point estimate of 0.31 (see Appendix I). This finding is *not* an artifact of the data generation process--i.e., that patents tend to have three infringement suits and one challenge. In fact, there are very few patents in the data set that are involved in both infringement and challenge suits. A simple explanation for this finding is based on the observation that a successful patent challenge generates positive externalities for other firms and thus free rider problems. Suppose, for argument's sake, that the probability of litigation is proportional to the expected benefits that it generates (as a reduced form of some underlying strategic litigation game.) Then, for any given win probability, an increase in the value of patent (or one of its proxy variables) will increase the expected value of a suit and thus the probability of litigation. But if a firm bringing a challenge suit can only appropriate α of the benefits (relative to an infringement suit), then the marginal effect of the increase in patent value would only be α percent as large for a challenge suit. Under this interpretation, our point estimate of α suggests that a plaintiff can appropriate about a third of the benefits from a successful patent challenge (relative to an infringement suit). Another explanation of the finding could be that the probability of winning a challenge suit is lower, so that the expected value of litigation for any level of patent value is reduced. The development of models of infringement and challenge suits will be useful in identifying the explanation for this clear empirical difference in the effect of patent value on litigation.

5. Concluding Remarks

We conclude by summarizing the main findings as they relate to the hypotheses discussed in Section 1. The first hypothesis was that litigation is associated with the number of 'events'. This was confirmed in these data by the prevalence of litigation among patents protecting drug innovations, an area in which the propensity to patent is high. We also find that patents in 'crowded areas', as indicated by a high level of similarity between the patent and its forward citation, are more likely to be litigated. But surprisingly, patent breadth does not increase the likelihood of litigation.

The second hypothesis was that uncertainty and asymmetric information might impede settlement and therefore contribute to disputes ending up in court. Given the difficulties in measuring information, this is not easy to measure directly. However, the quite dramatic difference between the litigation rates in all technology areas found in these data, and the much higher litigation rate found by Lerner (1995a) for the new area of biotechnology, strongly suggests that the uncertainty associated with emerging technologies encourages litigation. More speculatively, the fact that individual patentees are as likely to get engaged in litigation as corporate owners may be explained in part by the fact that they have greater uncertainty about how they will fare if they go to court. Further, they are typically in dispute with a firm. One might suppose that the information readily available to individuals, particularly about the results of similar cases, is different than that available to firms, contributing to a divergence in expectations.[26]

The third hypothesis was that disputes with higher stakes, or with a divergence in the stakes for the two parties, are more likely to be litigated. There is considerable evidence that this an important factor. Both the number of claims, and more valuable claims, as measured by citations by subsequent inventors, are associated with a greater probability of litigation and a somewhat higher frequency of litigation. The results regarding the similarity of forward citations are suggestive that the desire to build up a reputation contributes to making litigation attractive. A reputation for aggressively pursuing infringers is important when there are many other agents active in the same area. It is also important when the patent owner expects future disputes. It is precisely in the *interaction* between a high degree of similarity in forward citations and corporate ownership that we see the highest probability of litigation. Finally, the importance of the size of stakes, and, in particular, an asymmetry in the size of stakes, is supported by the finding that patents on innovations which appear to be the basis of a series of cumulative innovations, as indicated by subsequent *self*-citations in the same area, are more likely to become the subject of litigation.

The final hypothesis is that patent owners with higher litigation costs, in particular, higher litigation costs relative to the costs of arranging a settlement agreement, are less likely to become engaged in litigation. The evidence we can muster here is tentative because we do not have a direct measure of costs. However, the results based on type of ownership are suggestive. First, foreign patent owners have

---

[26] Experimental studies suggest that individuals' expectations are biased-there exists a propensity to view one's own case too favourably. The resultant divergent expectations about outcomes may impede settlement. It is quite likely that self-serving expectations are more common when the disputant is the inventor rather than a firm to which the patent has been assigned.

a lower probability of engaging in litigation. While there are a number of possible explanations for this, such as type of technology area, lower value, and difficulties in detection, we are able to reject these alternatives. The fact that nationality does not have a large effect on the number of cases filed once a patent has become the subject of litigation, suggests that there may be a substantial fixed cost to litigation. Corporate patent owners appear less likely to become engaged in litigation than individual patent owners. We interpret this as meaning that, while corporate owners have lower costs of litigation, which would make them more inclined to go to court, their advantages in settling disputes are even greater. This may be because they have repeated interactions or other relationships with the opposite party.

Our findings that litigated patents are wider in scope, in the sense of having more claims, and that they tend to be in 'crowded' technology areas, provide some support for the perspective in recent theoretical analyses of the patent system. In particular, Waterson (1990) argues that the patent system changes firms' incentives in their choice of product location (R&D) decisions because locating close to others' protected space runs the risk of infringement and costly litigation. The evidence suggests that these are empirically relevant concerns.

We have seen that foreign patentees are less likely to engage in litigation, and the results of the analysis suggest that this is because they face higher costs of accessing the U.S. courts to enforce their rights and therefore are more selective about which patents they defend. To the extent that this is true, it means that foreign owners of U.S. patents are likely to tolerate more infringement and consequently receive a lower return to investments in R&D from the country's patent system than do domestic inventors. Since this is quite likely to be true in other countries, for similar reasons, it is not necessarily the case that this domestic advantage in enforcement results in a large bias when global patent rights are considered (although the size of the U.S. market would tend to make this advantage run in favour of U.S. inventors).

Probably more important from the point of view of incentives is the finding that individual-owned patents are at least as likely to be litigated as are corporate-owned patents. This means that, not only are individual owners more likely to incur higher costs for the legal services that they employ in any given case, but, controlling for the characteristics of the patent, they are just as likely to require those services. Together with the fact that the uncertainty associated with new technology fields is likely to result in more litigation, it appears that enforcement issues will weigh most heavily on small firms and individuals who are active in emerging technologies.

References

Anand, Bharat and Tarun Khanna, 1996, "Intellectual Property Rights and Contract Structure," Mimeo, Yale School of Management.

Baldwin, James, 1996, "The Use of Intellectual Property Rights by Canadian Manufacturing Firms: Findings from the Innovation Survey," Mimeo, Micro-Economic Analysis Division. Statistics Canada.

Bebchuk, Lucian, 1984, "Litigation and Settlement under Imperfect Information," *RAND Journal of Economics*. 15, 404-415.

Business Insurance. June 15, 1995. p. 29.

Caballero, Ricardo and Adam Jaffe, 1993, "How High are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth," in Olivier Blanchard and Stanley Fischer, eds., *NBER Macroeconomics Annual 1993*. (Cambridge: MIT Press).

Cohen, Wesley, Richard Nelson and John Walsh, 1996, "Appropriability Conditions and Why Firms Patent and Why They Do Not in the American Manufacturing Sector," Paper presented at the Conference on New Science and Technology Indicators for the Knowledge-Based Economy (OECD, June 1996).

Cooter, Robert and Daniel Rubinfeld, 1989, "Economic Analysis of Legal Disputes and their Resolution," *Journal of Economic Literature*. 27, 1067-97.

Eisenberg, Theodore and James Henderson, 1992, "Inside the Quiet Revolution in Products Liability,"

Farber, Henry and Michelle White, 1991, "Medical Malpractice: An Empirical Investigation of the Litigation Process," *RAND Journal of Economics*. 22, (Summer), 199-217.

Federal Judicial Center, *Federal Court Cases: Integrated Data Base, 1970-89*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research. Tapes updated to 1991.

Fournier, G. and T. Zuehlke, 1989, "Litigation and Settlement: An Empirical Approach," *The Review of Economics and Statistics*. 71, 189-95.

Green, Jerry and Suzanne Scotchmer, 1995, "On the Division of Profit in Sequential Innovation," *RAND Journal of Economics*. 26, (Spring), 20-33.

Harhoff, Dietmar, Francis Narin, F.M. Scherer and Katrin Vopel, 1997, "A Million for your Cites: Patented Invention Value and Citation Frequency," Mimeo, Harvard University.

Hughes, James and Edward Snyder, 1989, "Policy Analysis of Medical Malpractice Reforms: What Can We Learn from Claims Data?" *Journal of Business and Economic Statistics*. 7, (October).

Jaffe, Adam B., Rebecca Henderson and Manuel Trajtenberg, 1993, "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics*. 108 (1993), 577-598.

Jaffe, Adam and Manuel Trajtenberg, 1996, "Flows of Knowledge from Universities and Federal Labs: Modelling the Flow of Patent Citations over Time and Across Institutional and Geographic Boundaries," NBER Working Paper No. 5712.

Kaplow, Louis, 1993, "Shifting Plaintiffs' Fees versus Increasing Damage Awards," *RAND Journal of Economics*. 24, (Winter).

Lanjouw, Jean O., 1992, *The Private Value of Patent Rights: A Dynamic Programming and Game Theoretic Analysis of West German Patent Renewal Data, 1953-1988.* Ph.D thesis. London School of Economics.

Lanjouw, Jean O., 1993, "Patent Protection: Of What Value and for How Long?" NBER Working Paper no. 4475.

Lanjouw, Jean O. and Josh Lerner, 1996, "Preliminary Injunctive Relief: Theory and Evidence from Patent Litigation," NBER Working paper no. 5689.

Lanjouw, Jean O., Ariel Pakes and Jonathan Putnam, 1996, "How to Count Patents and Value Intellectual Property: Uses of Patent Renewal and Applications Data," NBER Working Paper no. 5741.

Lanjouw, Jean O. and Mark Schankerman, 1997, "Measuring Innovation: Patent Claims, Citations and Families," Mimeo, Yale University (forthcoming).

Lerner, Josh, 1995a, "The Importance of Trade Secrecy: Evidence from Civil Litigation," Working Paper, Harvard University.

Lerner, Josh, 1995b, "Patenting in the Shadow of Competitors," *Journal of Law and Economics*. 38, 463-96.

Lerner, Josh, 1994, "The Importance of Patent Scope: An Empirical Analysis," *The RAND Journal of Economics*. 25, 319-333.

Levin, Richard, Alvin Klevorick, Richard Nelson and Sidney Winter, 1987, "Appropriating the Returns from Industrial Research and Development," *Brookings Papers on Economic Activity*. 783-820.

Merges, Robert, 1997, *Patent Law and Policy: Cases and Materials*, 2nd Edition (Charlottesville, PA: MICHIE Law Publishers).

Merges, Robert and Richard Nelson, 1990, "On the Complex Economics of Patent Scope," *Columbia Law Review*. 90, 839-916.

Pakes, Ariel and Margaret Simpson, 1989, "Patent Renewal Data," *Brookings Papers: Microeconomics*. 331-410.

Png, I.P.L., 1983, "Strategic Behavior in Suit, Settlement and Trial," *Bell Journal of Economics*. 14, 539-550.

Priest, George and Benjamin Klein, 1984, "The Selection of Disputes for Litigation," *Journal of Legal Studies* . 8, 1-56.

Putnam, Jonathan, 1996, *The Value of International Patent Protection*. Ph.D. Thesis. Yale University.

Samuelson, Pamela, Randall Davis, Mitchell Kapon and J Reichman, 1994, "A Manifesto Concerning the Legal Protection of Computer Programs," *Columbia Law Review*. 94, (December), 2308-2431.

Schankerman, Mark, "How Valuable is Patent Protection: Estimates by Technology Field," forthcoming in *RAND Journal of Economics*.

Scotchmer, Suzanne, 1996, "Protecting Early Innovators: Should Second-Generation Products be Patentable," *RAND Journal of Economics*. 27 (Summer), 322-331.

Scotchmer, Suzanne, 1991, "Standing on the Shoulders of Giants: Cumulative Research and the Patent Law," *Journal of Economic Perspectives*. 5 (Winter), 29-41.

Shavell, Steven, 1996, "The Level of Litigation: Private versus Social Optimality," Discussion Paper no. 184, John M. Olin Center for Law, Economics and Business, Harvard Law School.

Siegelman, Peter and Joel Waldfogel, 1996, "Toward a Taxonomy of Disputes: New Evidence Through the Prism of the Priest/Klein Model," Mimeo, Yale University.

Spier, Kathryn, 1994, "Pretrial Bargaining and the Design of Fee-Shifting Rules," *RAND Journal of Economics*. 25 (Summer).

Spier, Kathyrn and Daniel Spulber, 1993, "Pretrial Bargaining under Asymmetric Information: The Mechanism Design Approach," Mimeo, Northwestern University.

Tirole, Jean, 1989, *The Theory of Industrial Organization*. (Cambridge, MA: MIT Press).

Trajtenberg, Manuel, 1990a, "A Penny for Your Quotes: Patent Citations and the Value of Innovation," *RAND Journal of Economics*. 21.

Trajtenberg, Manuel, 1990b, *Economic Analysis of Product Innovation*. (Cambridge, MA: Harvard University Press).

Waldfogel, Joel, 1995, "The Selection Hypothesis and the Relationship between Trial and Plaintiff Victory," *Journal of Political Economy*. 229-60.

Waldfogel, Joel, 1996, "Reconciling Asymmetric Information and Divergent Expectations Theories of Litigation," Mimeo, Yale University.

Waterson, Michael, 1990 "The Economics of Product Patents," *American Economic Review*. 80.

Wittman, Donald, 1988, "Dispute Resolution, Bargaining, and the Selection of Cases for Trial: A Study of the Generation of Biased and Unbiased Data," *Journal of Legal Studies*. 17, (June), 313-352.

Table I
Litigation Rates and Composition

Panel A: Filed Cases per 1,000 Patents, 1980-1984

By Technology Group and Ownership[a,b]

	Total	Domestic	Foreign
Drugs and Health	20.1	26.6	6.5
Chemical	5.4	6.1	1.4
Electronic	9.6	12.7	3.3
Mechanical	11.8	20.1	3.4
Other	15.2	23.4	9.9
Total	10.7	16.4	3.5

Panel B: Composition of Litigation by Technology Group and Ownership

	Infringements	Challenges	Unclassified
By Technology Group			
Drugs and Health	63.2	10.5	26.2
Chemical	57.0	11.5	31.5
Electronic	58.1	10.5	31.4
Mechanical	59.4	9.8	30.8
Other	57.9	8.1	34.0
By Ownership			
Domestic	60.7	10.1	29.2
Foreign	43.8	7.4	48.8
Total	58.7	9.8	31.5

Notes:
[a] The IPC categories included in each of these groups are: Drugs and Health: A61 and A01N; Chemical: A62, B31, C01-C20, D-; Electronic: G01-G21, H-; Mechanical: B21-B68 not incl. B31, C21-C30, E01-F40; Other: A- not incl. A61 or A01N, B01-B20, F41-F42, G21.
[b] Panel A is based on the sample of all filed cases. Panel B deletes cases where the patent owner is listed as both the plaintiff and defendant, which are almost surely misreported.

Table 2
The Distribution of Forward Citations (percent)

Number of Citations	Domestic		Non-Japanese Foreign		Japanese	
	Matched	Litigated	Matched	Litigated	Matched	Litigated
	(1)	(2)	(3)	(4)	(5)	(6)
0	9.1%	3.3%	17.9%	2.7%	9.1%	1.6%
1-5	60.4	36.0	70.7	37.0	62.9	32.8
6-15	31.1	39.8	24.0	39.8	29.5	37.7
16-50	8.2	22.2	5.6	19.0	7.6	27.8
>50	0.3	2.1	0.0	2.1	0.0	1.6
Mean Citations	6.3 (0.16)[a]	12.0 (0.24)	4.8 (0.19)	11.5 (0.78)	5.8 (0.32)	12.0 (1.55)
No. of Observations	2,573	3,440	899	373	396	61

Notes:

[a] Standard errors are in parentheses.

Table 3
Similarity Index for Backward and Forward Citations

	Backward Citations		*Forward Citations*	
	Matched	Litigated	Matched	Litigated
Domestic Owners				
Self-citations	0.57	0.62	0.58	0.67
	(.019)[b]	(.014)	(.017)	(.012)
Citations by others[a]	0.55	0.58	0.51	0.59
	(.010)	(.008)	(.008)	(.006)
Foreign Owners				
Self-citations	0.64	0.62	0.63	0.72
	(.026)	(.041)	(.024)	(.032)
Citations by others	0.57	0.60	0.54	0.62
	(.014)	(.025)	(.011)	(.017)

Note:
[a] Citations by others includes all backward or forward citations where the owner of the citing or cited patent could not be directly matched to the owner of the patent. Because of entry errors in the original data and format differences, some self-citations are probably included.
[b] Standard errors are in parentheses.

Table 4
Publicity Effect of Case Filings on Mean Annual Forward Citations

Citation Lag	Not Yet Litigated (1)	Litigated: 1-yr Back (2)	Litigated: 2-yrs Back (3)	Litigated: 3-yrs Back (4)	Litigated: 4-yrs Back (5)
6	0.96 (.05)[a]	1.22 (.11)	1.04 (.07)	0.95 (.08)	0.84 (.08)
7	0.96 (.06)	1.22 (.12)	1.17 (.09)	1.21 (.12)	0.96 (.09)
8	0.90 (.06)	1.15 (.17)	1.14 (.11)	0.90 (.08)	1.08 (.10)
9	0.87 (.07)	1.42 (.16)	1.08 (.13)	1.10 (.10)	0.73 (.06)
10	0.93 (.08)	1.37 (.13)	1.25 (.14)	0.84 (.09)	1.00 (.11)
11	1.22 (.11)	1.03 (.21)	1.24 (.19)	1.10 (.17)	0.89 (.10)
12	0.93 (.11)	1.46 (.15)	0.77 (.11)	1.38 (.17)	0.97 (.14)
13	1.25 (.18)	0.94 (.17)	0.88 (.14)	0.85 (.13)	0.89 (.13)
14	0.72 (.17)	1.20 (.16)	1.09 (.19)	0.84 (.11)	1.03 (.16)
15	0.67 (.26)	1.10 (.45)	1.36 (.22)	1.11 (.20)	0.79 (.14)
Mean[b]	0.94 (.04)	1.18 (.07)	1.10 (.05)	1.03 (.05)	0.95 (.04)

Notes:
[a] Standard errors are in parentheses.
[b] These the overall mean numbers of annual citations taken over citation lags 6 through 15.

Table 5
Mean Claims per Patent

	Drugs & Health[a]	Chemicals	Electronic	Mechanical	Other
Mean Claims	13.6 (0.5)[b]	13.5 (0.5)	14.2 (0.3)	11.9 (0.2)	11.2 (0.2)

	1975	1980	1983	1985	1990
	11.6 (0.4)	12.5 (0.5)	12.6 (0.4)	13.7 (0.7)	17.2 (1.6)

Notes:
[a] See notes to Table 1 for technology group definitions.
[b] Standard errors are in parentheses.

Table 6
Claims per Patent and Citations per Claim

	Domestic		Non-Japanese Foreign		Japanese	
	Matched	Litigated	Matched	Litigated	Matched	Litigated
Claims						
Mean	11.2 (0.18)	14.7 (0.22)	9.7 (0.23)	13.2 (0.60)	8.8 (0.31)	10.2 (1.41)
Median	9	11	8	10	7	8
Mode	3	8	4	7	4	4
Forward Cites per Claim[a]	1.04 (0.08)	1.57 (0.07)	0.84 (0.09)	1.88 (0.29)	1.20 (0.19)	1.61 (0.32)
Backward Cites per Claim	1.21 (0.04)	1.01 (0.03)	0.98 (0.04)	0.90 (0.06)	1.10 (0.07)	0.79 (0.08)

Note:
[a] The forward citation statistic is calculated using cohorts 1975-77 only.
[b] Standard errors are in parentheses.

Table 7
Chi-Square Tests of Differences in the Breadth of Matched and Litigated Patents

Pooled Data	χ^2/df	df	p-value	Mean Number Matched	Mean Number Litigated
Number of 4-digit IPC sub-classes	6.68	3	<0.001	1.243	1.197
				(.009)[b]	(.008)
By Technology Group					
Drugs and Health[c]	1.15	2	0.315	1.081	1.052
				(.018)	(.019)
Chemicals	2.41	3	0.065	1.327	1.365
				(.032)	(.040)
Electronics	1.31	3	0.270	1.207	1.182
				(.016)	(.017)
Mechanical	4.48	3	0.004	1.259	1.191
				(.014)	(.012)
Other	1.30	3	0.270	1.276	1.229
				(.019)	(.016)

Notes:
[a] The 2-digit IPC sub-class aggregates the 4-digit sub-classes into 22 technology groups.
[b] Standard errors are in parentheses.
[c] In Drugs and Health the maximum number of 4-digit IPC assignments in the sample is three, so the test has two degrees of freedom.

Table 8
Probit Estimation for the Patent Infringements and Challenges[a,b]

	Infringements		Challenges	
	Parameters[a]	Marginal Effects[b]	Parameters[c]	Marginal Effects
	(1)	(2)	(3)	(4)
Log Claims	0.405**	0.136**	0.433**	0.058**
	(.030)	(.010)	(.050)	(.007)
FWD Cites/Claim	0.256**	0.081**	0.230**	0.029**
	(.019)	(.006)	(.027)	(.003)
FWD Cites/CLM^2	-0.0083**		-0.005**	
	(.0011)		(.001)	
BWD Cites/Claim	-0.033	-0.010	-0.043	-0.005
	(.023)	(.007)	(.040)	(.005)
BWD Cites/CLM^2	0.0012		0.003	
	(.0011)		(.002)	
NO4IPC	-0.062*	-0.021*	-0.052	-0.007
	(.036)	(.012)	(.057)	(.008)
SIMFWD	0.278**	0.093**	0.129	0.017
	(.087)	(.029)	(.144)	(.019)
SIMFWD*CORP	0.203**	0.162**	0.223	0.047**
	(.103)	(.020)	(.170)	(.013)
FWDSELF	0.674**	0.226**	0.585**	0.078**
	(.106)	(.036)	(.165)	(.022)
BWDSELF	-1.015**	-0.341**	-0.936**	-0.125**
	(.173)	(.058)	(.307)	(.041)
DINDOWN	-1.419**	0.103**	-2.320**	0.022
	(.126)	(.024)	(.205)	(.019)
FINDOWN	-2.321**	-0.214**	-2.911**	-0.059**
	(.158)	(.024)	(.250)	(.014)
JINDOWN	-2.305**	-0.195**	NE	NE
	(.349)	(.071)		
DCORPOWN	-1.562**	0.195**	-2.466**	0.060**
	(.123)	(.016)	(.200)	(.013)
FCORPOWN	-2.321**	-0.181**	-2.963**	-0.050**
	(.129)	(.016)	(.209)	(.011)
JCORPOWN	-2.893**	-0.287**	-3.510**	-0.084**
	(.156)	(.015)	(.266)	(.008)

Table 8 Continued

	Infringements		Challenges	
No. Observations	6151		4227	
Log-likelihood	-3506.6		-1222.5	
Pseudo-R^2	0.141		0.115	
χ^2 Tests (df)				
No Nationality	374.7 (4)	(p-value < 0.001)	66.47 (3)	(p-value < 0.001)
No Corp/Individual	7.3 (3)	(p-value $= 0.062$)	1.7 (2)	(p-value $= 0.42$)

Notes:

[a] Statistical significance at the 0.01 (0.05) level is denoted by ** (*).

[b] Marginal effects are calculated at the population means of the variables. For dummy variables the partial effect is calculated as the increase in the probability of a case filing with a change in the dummy variable from zero to one. See text for details.

[c] NE denotes not estimable.

Appendix I
Marginal Effects Calculations and the Chi-Square Test of Proportionality

The Definition and Estimated Variance-Covariance Matrix of the Marginal Effects

Let $m(\hat{\beta})$ be vector of the marginal effects of changes in the independent variables on the probability of litigation, implied by the coefficient estimates, $\hat{\beta}$, and evaluated at the population (matched sample) variable means, $\bar{X}$. For a representative continuous variable i, the marginal effect is:

$$m_i(\hat{\beta}) = \Phi(\bar{X}'\hat{\beta})\hat{\beta}_i. \tag{A.1}$$

In the case of a variable with linear and quadratic effects, say β_i and β_{i2},

$$m_i(\hat{\beta}) = \Phi(\bar{X}'\hat{\beta})(\hat{\beta}_i + 2\hat{\beta}_{i2}). \tag{A.2}$$

In the case of a variable with an ownership interaction term, again β_{i2}, the marginal effect when the ownership dummy variable is one is,

$$m_i(\hat{\beta}) = \Phi(\bar{X}'\hat{\beta})(\hat{\beta}_i + \hat{\beta}_{i2}). \tag{A.3}$$

There are six discrete variables indicating the six possible ownership types. For a representative discrete variable i, the marginal effect is a change in the dummy variable from zero to one. The probability of litigation when i is zero is taken to be a weighted average of the probability of litigation for each of the other ownership categories. Thus:

$$m_i(\hat{\beta}) = \Phi(\bar{X}_.'\hat{\beta},\ i) - \textstyle\sum_{j\neq i} Pr(j|i=0)\Phi(\bar{X}_.'\hat{\beta},\ j), \tag{A.4}$$

where $\bar{X}_.$ is the vector of independent variable means but with ownership dummy variable mean values replaced by zeros except for those appropriate to the ownership category indicated by the second argument, which are one. $Pr(j|i=0)$ is the sample proportion of patents falling in ownership category j, given that they are not in category i. Taking a Taylor's expansion around $m(\beta)$, the estimated variance-covariance matrix of $m(\hat{\beta})$ is:

$$V(m(\hat{\beta})) = [\partial m(\hat{\beta})/\partial\hat{\beta}]'\ V(\hat{\beta})\ [\partial m(\hat{\beta})/\partial\hat{\beta}], \tag{A.5}$$

where $V(\hat{\beta})$ is the estimated variance-covariance matrix of the parameters.

The Chi-Square Test of the Proportionality of the Marginal Effects

The chi-square statistic was calculated by minimizing with respect to α the following chi-square statistic:

$$\chi^2(\alpha) = [\alpha\ m(\hat{\beta}_I) - m(\hat{\beta}_{II})]'\ [\alpha^2 V(m(\hat{\beta}_I)) + V(m(\hat{\beta}_{II}))]\ [\alpha\ m(\hat{\beta}_I) - m(\hat{\beta}_{II})], \tag{A.6}$$

where $\hat{\beta}_I$ and $\hat{\beta}_{II}$ are vectors of parameter estimates from the infringement and challenge probit regressions, respectively. This two-step procedure makes the test of proportionality conservative because the β parameters are not allowed to adjust to improve the fit under the constraint.