LAW AND ECONOMICS SEMINAR **Professor Polinsky**
Fall Term 2007

Thursday, November 1, 2007
4:15 - 5:45 p.m.
Stanford Law School
Room 272

"The Private Cost of Patent Litigation"

by

Michael Meurer

(Boston University Law School)

Note: It is expected that you will have reviewed the speaker's paper before the Seminar. Actually, there are two papers attached, neither especially long. Please look at one or the other. The following guidance from the author may be helpful: "The papers are addressed to different audiences. The book chapter is an easy read for all students regardless of their background in economics. 'The Private Cost of Patent Litigation' is written for economists, but students should not have trouble getting the gist -- some may have trouble with the details. So I recommend that students start with the book chapter and read as much of the other article as time permits. For faculty, I recommend reading the Private Cost of Patent Litigation."

The Private Costs of Patent Litigation

Working Paper
Original version: March 2006
Current version: September 2007

By James Bessen and Michael J. Meurer*

Abstract: This paper estimates the total cost of patent litigation to alleged infringers. We use a large sample of stock market event studies around the date of lawsuit filings for US public firms from 1984-99. Even though most lawsuits settle, we find that the total costs of litigation are large in aggregate—over $16 billion per year—and large compared to legal fees and to estimates of patent value. Moreover, infringement risk rose sharply during the late 1990s to over 14% of R&D spending. These estimates support the view that infringement risk should be a major concern of policy.

Keywords: patent, litigation, litigation cost, property rights

JEL Classifications: O31, O34, K41

*Research on Innovation and Boston University School of Law, and Boston University School of Law, respectively. Thanks to comments from Megan MacGarvie, Jesse Giummo and participants at the IIOC and NBER seminars. Thanks also to research assistance from Debbie Koker and Dan Wolf.

Contact: jbessen@bu.edu

1 Introduction

Like any regulatory mechanism, the patent system has benefits and costs, both private and social. Yet little empirical evidence exists about the magnitude of some of these costs, leaving policy analysts to sometimes rely on guesstimates. For example, recent policy consideration of a patent opposition proceeding in the US (Levin and Levin 2002, Hall et al. 2004) has been based on rough estimates of the costs of patent litigation and the social costs of inappropriately-granted patents.

In contrast, there is a significant literature estimating benefits of the patent system, especially private benefits in the form of estimates of patent value[1] or of the patent premium (Arora et al. 2005). However, without comparable estimates of private and social costs, it is difficult to conduct either analysis of specific policy changes or a normative analysis of the patent system in comparison to other means of encouraging innovation. For example, Schankerman (1998) suggests that the ratio of aggregate patent value to R&D constitutes an upper bound measure of the subsidy that patents provide to R&D. He asserts that this ratio can be used to compare patents to other forms of appropriating returns on invention. But surely this is only an estimate of a *gross* subsidy against which private costs of patents need to be netted out.

This paper takes a step toward quantifying costs by estimating the private costs of patent litigation to alleged infringers. To the extent that costly patent litigation is mainly the result of inadvertent infringement—and we argue elsewhere that it is (Bessen and Meurer 2005, 2006, 2008)—then the costs of defending against inadvertent infringement represent a disincentive to investing in innovation. The risk of unavoidable infringement acts like a "tax" on innovation. Litigation can also be costly to plaintiffs, however, the expected costs of patent enforcement are presumably incorporated into patentee's valuation of patents and these have been estimated.

Litigation costs are not, of course, the only private cost of patents. There are costs of obtaining and maintaining patents; inadvertent infringement also leads to costly disputes that may be settled without lawsuits. However, during the 1990s we found evidence of a dramatic increase in the hazard of patent litigation for publicly traded firms (Bessen and Meurer 2005), so this seemed to be a potentially important cost that is also measurable.

Often, when policy analysts discuss litigation costs, they only consider the out-of-pocket legal costs of litigation. Moreover, since most lawsuits settle without going to trial, it might seem that the average legal costs of litigation are not substantial.

1This literature began with Pakes and Schankerman (1984). See Bessen (2006a) for a survey of this literature.

But legal costs are not the only costs of litigation that affect innovating firms. Business costs of litigation can be much larger and can take many forms. Business can be disrupted as managers and researchers spend their time producing documents, testifying in depositions, strategizing with lawyers, and appearing in court. Litigation strains the relationship between the two parties and may jeopardize cooperative development of the patented technology or cooperation on some other front. Firms in a weak financial position might see their credit costs soar because of possible bankruptcy risk created by patent litigation. Preliminary injunctions can shut down production and sales while the litigation pends. But even without a preliminary injunction, customers may stop buying a product. Frequently, products require customers to make complementary investments; they may not be willing to make these investments if a lawsuit poses some risk that the product will be withdrawn from the market. Furthermore, patent owners can threaten customers and suppliers with patent lawsuits because patent infringement extends to every party who makes, uses, or sells a patented technology without permission, and sometimes to those who participate indirectly in the infringement. Some of these costs can be quite substantial even when the lawsuit is settled. In addition, in cases of inadvertent infringement, the payments necessary to achieve a settlement from unavoidable disputes can constitute a substantial risk.

Even simple delay can impose large business costs. Consider, for example, litigation against Cyrix, a startup firm that introduced Intel-compatible microprocessors. Intel, the dominant microprocessor maker, sued Cyrix and the suit lasted a year and a half. During that time Cyrix had difficulty selling microprocessors to computer manufacturers, who were almost all also customers of Intel and who were reluctant to break ranks to go with a product that might be found to infringe. In the meantime, Intel responded by accelerating its development of chips that would compete against Cyrix's offerings. In the end, Cyrix won the lawsuit, but lost the war, having lost much of its competitive advantage. In effect, Cyrix lost the window of opportunity to establish itself in the marketplace. Litigation exacted a heavy toll indeed.

Although we gather some original information about legal costs of litigation, our main emphasis is on estimating the total business cost to alleged infringers in order to estimate firm risk of infringement. We do this using a large number of event studies on lawsuits involving publicly traded firms from 1984 through 1999.

The event study methodology has been used before to study litigation, beginning with Cutler and Summers (1988) in the context of litigation over a merger. Several papers have

performed event studies of patent litigation, both the event of the initial filing and the terminating event (settlement, judgment or verdict), including Bhagat, et al. (1994), Lerner (1995), Bhagat et al. (1998), Lunney (2004) and Haslem (2005).

These studies of initial filings, however, do not provide the best estimates from which to calculate the aggregate risk of infringement to the firms that perform R&D. They use small, selective samples and their estimates of wealth loss are not especially precise. Our contribution is to work with a much larger set of disputes: our sample covers most patent lawsuits filed against US public firms over a 16 year period, a sample responsible for the lion's share of R&D spending. This gives our results greater precision and also makes them highly representative of R&D-performing firms, permitting us to calculate a variety of cost and risk measures to inform policy. We find, in fact, that the estimates of wealth loss reported in some earlier studies appear to be overstated.

A key assumption of this literature is that the change in firm value that occurs around a lawsuit filing reflects investors' estimates of the direct and indirect effects of the lawsuit on the profits of the firm on average and do not systematically reflect any unrelated information. We show evidence below that the revelation of unrelated information does not overstate our estimates for defendants in infringement suits and that, therefore, we may associate the loss in wealth with the effective total cost of litigation for defendants.

We find that firms lose about half a percentage point of their stock market value upon being sued for patent infringement. This corresponds to a mean cost of $28.7 million in 1992 dollars (median of $2.9 million), much larger than mean legal fees of about half a million. In aggregate, infringement risk rose sharply during the late 1990s, exceeding $16 billion for US public firms. This amounts to 19% of these firms' R&D spending, a ratio that exceeds some estimates of the value of patents granted relative to R&D.

The next section describes the data and methods used for estimating cumulative abnormal returns. Section 3 reports average returns and some analysis of factors that affect returns. Section 4 calculates litigation cost, Section 5 calculates some broader measures of infringement risk and Section 6 concludes.

2 Data and Methods

2.1 Data Sources

Our research matches records from three data sources: lawsuit filings from Derwent's Litalert database, firm financial data from Compustat, and CRSP data on securities prices. In addition, we searched the electronic archives of the Wall Street Journal to locate any articles announcing lawsuit filings and also any announcements of other events that might confound our analysis.

Using these sources, we constructed two main samples. The first, small sample included just those lawsuits where we could identify one or more parties on both sides of the dispute as public firms. The second, large sample included all cases where the alleged infringer (defendant in an infringement suit or plaintiff in a declaratory action) was a publicly traded firm, but the patentee litigant need not be public.

Our primary source of information on lawsuit filings is Derwent's Litalert database, a database that has been used by several previous researchers (Lanjouw and Schankerman, 2004, Ziedonis, 2004, Bessen and Meurer 2005). Federal courts are required to report all lawsuits filed that involve patents to the U.S. Patent and Trademark Office (USPTO) and Derwent's data is based on these filings. Beginning with the Derwent data from 1984 through 2000, we removed duplicate records involving the same lawsuit as identified by Derwent's cross-reference fields. We also removed lawsuits filed on the same day, with the same docket number and involving the same primary patent. Sometimes firms respond to lawsuits by filing counter-suits of their own, perhaps involving other patents. Since our main focus is on initial disputes rather than on lawsuit filings *per se,* we also removed filings made within 90 days of a given suit that involved the same parties.

The Derwent data does not distinguish whether the suit filed is an infringement suit or a declaratory judgment suit. A firm threatened with an infringement suit can file a declaratory action which aims for a judgment that the patent is uninfringed or invalid. To classify each suit, we first identified whether the patent assignee of the main patent at issue matched one of the parties to the suit. If the assignee matched a plaintiff, the suit was classified as an infringement suit; if the assignee matched a defendant, the suit was classified as a declaratory action. We were able to match the assignee for 83% of the suits, and of these, only 17% were declaratory actions.[2]

2 These numbers are quite similar to findings by Moore (2000) and Lanjouw and Schankerman (2004).

If the assignee did not match a party to the suit, then it was classified as an infringement suit because there are relatively few declaratory actions. This classification then allowed us to identify whether the subject firm was a "patentee litigant" (that is, plaintiff in an infringement suit or defendant in a declaratory action) or an "alleged infringer" (the reverse).

To explore characteristics of firms involved in these lawsuits, we matched the listed plaintiffs and defendants to the Compustat database of U.S. firms from 1984-99 that report financials (excluding American Depository Receipts of foreign firms traded on US exchanges). These data were based on merged historical data tapes from Compustat and involved an extensive process of tracking firms through various types of re-organization and eliminating duplicate records for firms (e.g., consolidated subsidiaries listed separately from their parent companies).[3]

The lawsuit data were matched to the Compustat data by comparing the litigant names with all domestic firm names in Compustat and also a list of subsidiary names used in Bessen and Hunt (2007).[4] To check the validity and coverage of this match, we randomly selected a number of parties to suits and then checked them manually using various databases including PACER, LexisNexis, the Directory of Corporate Affiliations and the LexisNexis M&A databases. Although we were not able to definitively identify all parties, the rate of false positives was not more than 3% (no more than 5 of 165 parties were found to have been falsely matched) and the rate of false negatives was no more than 7% (no more than 34 of 502 public companies were not matched). The Compustat firms were then also matched to the CRSP file of daily security prices.

We identified 2,648 suits with sufficient data on alleged infringers, some of these having multiple alleged infringers, for a total of 2,887 events in our large sample. We also selected all those lawsuits where we could identify at least one party on each side as a publicly listed firm. This left us with asample of 750 plaintiffs and 747 defendants in lawsuits where public firms were parties on both sides.

Summary statistics of our samples are shown in Table 1 and further details from a closely related sample are reported in Bessen and Meurer (2005). Parties to patent lawsuits tend to be larger than average firms with large R&D budgets. Moreover, our large sample captures the

3 This work was conducted by Bob Hunt and Annette Fratantaro at the Federal Reserve Bank of Philadelphia for an earlier project and we thank them for graciously sharing it with us.

4 A software program identified and scored likely name matches, taking into account spelling errors, abbreviations, and common alternatives for legal forms of organization. These were then manually reviewed and accepted or rejected. Note that this match is based on the actual parties to litigation, not the original assignee of the patent at issue.

bulk of patent litigation against R&D performers. In 1999, US public firms in Compustat spent $150 billion on R&D, while total industrial R&D spending reported by the National Science Foundation was $160 billion.[5] Aside from issues of under-reporting, our large sample constitutes a comprehensive sample with which we can obtain a lower bound measure of the aggregate risk of infringement to R&D performers.

Finally, each lawsuit in the small sample was checked against the Wall Street Journal archive to identify those suits that were announced in the Journal within one month of the filing data and to identify possible confounding news about either party to the suit within one week of the filing date.

In Section 4 we discuss a supplemental dataset of lawsuits that reports legal fees.

2.2 Estimating Cumulative Abnormal Returns

To estimate the impact of a lawsuit filing on the value of a firm, we use event study methodology (see Mackinlay 1997 for a review). In particular, we use the dummy variable method described by Michael Salinger (1992).[6] This assumes that stock returns follow a market model,

$$(1) \qquad r_t = \alpha + \beta r_t^m + \epsilon_t$$

where r_t is the return on a particular stock at time t, r_t^m is the compounded return on a market portfolio, and ϵ_t is a stochastic error. If an event, such as a lawsuit filing, occurs on day T, then there may be an "abnormal return" to the particular stock on that day. This can be captured using a dummy variable,

$$(2) \qquad r_t = \alpha + \beta r_t^m + \delta I_t + \epsilon_t$$

where I_t equals 1 if $t=T$ and 0 otherwise. Equation (2) can be estimated using OLS for a single event. In practice, this equation is estimated over the event period and also over a sufficiently long pre-event window. In this paper we use a 200 trading-day pre-event window.[7] The

5There are important differences in the scope of what was included in these two measures, nevertheless, they suggest that public firms account for the lion's share of R&D spending.

6Salinger shows that this model is mathematically equivalent to the OLS market model described in Brown and Warner (1985) and widely used.

7We also ran regressions with a 180 day pre-event window that ended 30 days before the lawsuit filing. Cumulative abnormal returns were very close to those with a 200 day window that last up to the day before the event window.

coefficient estimate of δ obtained by this procedure is then an estimate of the abnormal return on this particular stock. For different stocks, the precision of the estimates of δ will vary depending on how well equation (2) fits the data. The estimated coefficient variance from the regression provides a measure of the precision of the estimate of the abnormal return.

We want to obtain a representative estimate of the abnormal returns from lawsuit filings for multiple stocks, under the assumption that these represent independent events and that they share the same underlying "true" mean. Previous papers estimating abnormal returns from patent lawsuits have simply reported unweighted means for the group of firms. Although the unweighted mean is an unbiased estimator, it is not efficient. Since we are concerned with obtaining the best estimate to use in policy calculations (and not just testing the sign of the mean), we use a weighted mean to estimate the "average abnormal return," where the weight for each observation is proportional to the inverse of the variance of the estimate of δ for that firm.[8]

When we test our means against the null hypothesis that the true mean is zero, we report both the significance of *t*-tests using the weighted mean and also the significance of the *Z* statistic (see Dodd and Warner 1983), a widely used parametric test of significance that incorporates the variation in precision across events.[9] In any case, the significance test results are closely similar as are those of some non-parametric tests.

As Salinger (1992) notes, this procedure assumes that the returns for each event are independent of each other. However, when there are multiple defendants in a suit, returns may be systematically related. For example, one defendant may be a supplier to another or two defendants may be unequal rivals. Since the large sample has 188 lawsuits with multiple defendants, in these cases we estimate the returns for the defendants to each suit jointly, estimating common abnormal returns for this group of defendants.

Finally, (2) describes the abnormal return for a single day. It is straightforward to design dummy variables to estimate a "cumulative abnormal return" (CAR) over an event window consisting of multiple consecutive days. In the following, for instance, if the suit is filed on date $t=T$, then we may use a window from day T-1 to T+24.

8In any case, we find that for our entire sample, the weighted mean is quite close to the unweighted mean and also to the median. However, there are significant differences in the averages for sub-samples.

9The *Z* statistic is a joint test of the individual firm *t*-tests. We use a robust version described in Kramer (2001).

2.3 The Event

This paper also differs from previous research in the nature of the events we study. Previous studies have used the *announcement* of the lawsuit in a newspaper or wire service as the event. We use the *filing* of the lawsuit, instead. This may seem to be a minor difference, but it is significant for two reasons.

First, most patent lawsuits are not announced in newspapers or wire service reports at all. Various factors may influence whether a lawsuit is announced or not. Firms may choose to issue a press release or not. The SEC requires reporting of major lawsuits in quarterly and annual filings, but lawsuits will be reported separately only if they materially affect the profits of the firm. And news sources may not report all lawsuits even if the firms issue press releases.

We took a random sample of patent lawsuits against US public firms and searched LexisNexis for news stories that mention the lawsuits within one month of the filing date, before and after. We found that only 19% of the lawsuits were mentioned in the Dow Jones Newswire, one of the most comprehensive reporting services; only 7% were mentioned in the Wall Street Journal, which was used by several of the previous studies. Since one of our objectives is to tally the combined risk of lawsuits for public firms, clearly we cannot obtain comprehensive estimates relying only on announced lawsuits.

Moreover, announced lawsuits are a select group that may be qualitatively different from other lawsuits. That is, samples of announced lawsuits may suffer from sample selection bias. To test this, we performed a series of Probit regressions in our small sample on whether a lawsuit was reported in the Wall Street Journal (see Appendix). Among other things, we find that the probability of a Wall Street Journal announcement is strongly correlated with the defendant firm's stock market beta. This means that a Wall Street Journal announcement is more likely if the news is more likely to have a larger effect on the defendant's stock price or, perhaps, word of the lawsuit is already affecting the defendant's stock price. This, in turn, suggests that estimates made on a sample of announced lawsuits may have abnormal returns with a larger absolute magnitude than those from a more representative sample.

Below we compare estimates of abnormal returns from samples of lawsuits announced in the Wall Street Journal with estimates from our comprehensive sample. We find that our estimates from the announced sample are quite similar to those reported in the previous literature. However, these estimates are substantially larger in absolute magnitude than those for our comprehensive sample, suggesting considerable sample selection bias.

On the other hand, our estimates may be understated for another reason: investors may not receive news of the lawsuit within an event window around a filing date. With an announcement in a newspaper or major newswire, we can be reasonably sure that investors hear the news of the lawsuit within a day or two of the announcement. But we cannot be sure that investors hear the news about a legal filing in a district courthouse. Indeed, depending on how long it takes to serve papers, the defendant may not be aware of the lawsuit for a day or so after the filing date. In other words, news of an unannounced patent lawsuit filing may leak out more slowly and investors may not learn of a lawsuit within a specified event window.

We see evidence of this slower diffusion of information in the lawsuits that were announced in the Wall Street Journal. Figure 1 displays the frequency of these news stories relative to the actual court filing date. Event studies based on public announcements typically use an event window of two or three days (often one day before the announcement). Although many lawsuits are announced within two days of filing, such a small event window around a *filing date* would clearly miss a very large share of lawsuit announcements. Moreover, it seems likely (given the role of stock market beta in the likelihood of a Wall Street Journal article) that the lawsuits that are announced within a few days of the filing may be qualitatively different from those for which the news leaks out more slowly and are either announced later or not announced at all. Indeed, we find evidence within our data that stocks with beta above 1 react to the filing faster than lower beta stocks.[10] In order to have representative and comprehensive estimates, we use a longer event window than would be appropriate in an announcement event study. Specifically, we use a 25 day window (from T-1 to T+24), which, based on the data in Figure 1, should capture 96% of the announced events and, we hope, a large share of the unannounced filings. We show some CARs from shorter windows in the Appendix.

There are two possible concerns with such a longer window. First, the longer window introduces more "noise" into the estimation reducing precision and possibly attenuating the estimates. Because we have a much larger sample size than earlier studies, this is not such a significant concern and our estimates are reasonably precise, although they may be slightly attenuated. Second, research on long horizon event studies—that is, studies with *multi-year* event windows—find certain biases that arise for a variety of reasons (Barber and Lyon 1997, Kothari

10At day 2, the higher beta stocks for defendant firms have a CAR that is significantly lower than the CAR for lower beta stocks (at the 5% level) and the lower beta CAR is not significantly different from zero. At day 24, the CARs for these two groups are not significantly different and both are significantly different from zero. One explanation for the faster speed of diffusion for high beta stocks is that the opportunities for investors to make returns from the information about the lawsuit filing are relatively greater for these stocks.

and Warner 1997).[11] However, it seems highly unlikely that these concerns can exert a substantially greater influence in a 25 day window than they exert in a three day window.

In summary, restricting the study to events announced in news services likely introduces substantial sample selection bias. Our estimates, based on a larger window following the filing of the lawsuit, are smaller, although they might be biased toward zero.

3 Empirical Findings

3.1 Estimates of cumulative abnormal returns

Since previous studies have used samples where parties on both sides of a lawsuit were public firms and the suits were reported newspapers or wire services, we begin by exploring a sub-sample. Table 2 shows estimates of cumulative abnormal returns for just those suits from our small, matched sample that were reported in the Wall Street Journal. In this table, we exclude suits that had a potentially confounding news story in the Wall Street Journal within a month of the suit filing date. Two previous studies have reported on event study estimates on announcements of patent lawsuits filings. Bhagat et al. (1998) examine lawsuits filed between 1981 and 1983 (51 plaintiffs and 33 defendants) and Lerner (1995) obtains estimates for 26 biotech lawsuits from 1980 to 1992.[12] To maintain consistency with the previous literature, in this Table (but not in the next) we report simple unweighted means of cumulative abnormal returns.[13] The mean and median values are reported for two different event windows, one around the Wall Street Journal publication date, the other a longer window around the actual suit filing date reported in court records (these dates occasionally differed significantly).

Consistent with most of the previous literature on litigation, we find that patentee litigants do not show a positive response to a lawsuit filing. Bhagat et al. (1998) report a CAR of -0.31%, and we find a similar value. For defendants (alleged infringers), we find a substantial loss in market value of around 2%. Bhagat et al. report a loss of 1.50%. For the combined loss of wealth, we find a mean of 2.5-2.6%, although smaller median values. Bhagat et al. (1994) report a mean loss of 3.13% and Lerner (1995) reports a mean loss of 2.0%. All three results are

11These reasons include: 1.) with a long window, the composition of the market index may change with the addition of new entrants or from rebalancing, 2.) compounding of returns leads to a highly skewed distribution, 3.) not all firms survive to the end of a long event window, and, 4.) the market model or its variance may change or may be sensitive to specification errors over long windows. We find that our measured returns are not highly skewed and there are few cases of firms failing to survive the event window.

12Bhagat et al. (1998) includes the data from the Bhagat et al. 1994 paper, so we do not list that separately. Lerner searched the Wall Street Journal as well as news wire services for announcements. The other studies just used articles in the Wall Street Journal.

13For this reason, this table does not report standard errors or significance tests.

broadly similar and quite substantial. Lerner reports a mean absolute loss of shareholder wealth of $67.9 million, a median loss of $20.0 million. In general, there does not appear to be a major difference between the results reported in the event window around the Wall Street Journal publication date and the longer window around the filing date.

As noted above, estimates for this sub-sample may be unrepresentative of most patent litigation, however, because most lawsuits are not reported in the Wall Street Journal. Table 3 reports cumulative abnormal returns for all lawsuits in the matched sample (top) as well as those for the large sample (bottom). The base result for the matched sample uses a 25 day event window (T-1 to T+24) and excludes possibly confounding events. The table also reports CARs for suits that were positively identified as infringement suits (that is, the plaintiff was the patent assignee), and for a sample that included lawsuits with possibly confounding news events. The reported means and standard errors use weights based on the variance of the dummy variable coefficient in the event regression. Several results stand out.

First, the estimated percentage losses for alleged infringers are substantially less than those for lawsuits reported in the Wall Street Journal in Table 2. We cannot tell, however, whether the percentage loss estimates in the Journal are larger because of a selection effect or because of the greater information conveyed by publication in the Journal. Even though some learning takes place, we suspect that in most lawsuits, investors remain relatively uninformed compared to those cases where an announcement is published in the Wall Street Journal. The SEC requires reporting of major lawsuits in quarterly and annual filings, but lawsuits will be reported separately only if they materially affect the profits of the firm. For a handful of suits, we checked published sources and typically found no mention of the suit. For this reason, estimates for the non-Journal sample should be interpreted as lower bound estimates of defendant firms' loss of wealth—significant numbers of investors likely became informed about the suit either after our event window or, if there were pre-filing interactions, before.

Second, patentee litigants/plaintiffs appear to suffer some losses as well. These losses are smaller than those for alleged infringers/defendants, but they are statistically significant. This is consistent with previous research and it indicates that lawsuits do not represent simple transfers of wealth on average. Instead, there is dissipation of wealth to consumers, to rivals or to deadweight loss.

Finally, the magnitudes of returns for definite infringement suits are generally larger than for those of all suits and they show a higher level of statistical significance. This may be because among those cases where we could not match the patent to one of the parties, some plaintiffs are

mistakenly classified as defendants and vice versa. Or it could be because declaratory actions may be more likely when the stakes at issue are smaller.

The bottom of Table 3 reports results for our large sample. The CARs for alleged infringers are similar to those obtained from the smaller sample—a loss of 0.5% to 0.6%—but here they have statistical significance at the 1% level, except for those lawsuits involving multiple defendants.

When multiple defendants are involved the returns are negligible, suggesting that something is fundamentally different about these estimates. There are several possible explanations for this. It may be that suits naming multiple defendants are more frivolous, so that investors do not expect serious losses. Alternatively, some defendants may have been contractually indemnified, diluting the estimates. A higher percentage of defendants in lawsuits with multiple defendants come from retail and wholesale industries, suggesting that these suits more frequently involve downstream resellers who have less at stake. Costs may be shared among multiple defendants, reducing the individual firm costs. Also, antitrust authorities may be more likely to scrutinize settlements when multiple defendants are involved.

The estimates in the lower portion of the table do not control for possibly confounding events. However, we find that excluding observations with possibly confounding events does not seem to substantially alter the mean estimated CARs in the top portion of the table (the matched parties sample). To check this further, we repeated the estimates for the large sample of all alleged infringers, but we terminated the pre-event window 30 days prior to the filing of the lawsuit. This made little difference in our estimates, suggesting that confounding events may add noise, but do not bias our estimates.[14]

Figure 2 shows histograms for the cumulative abnormal returns for all lawsuits from the matched sample. The curve for alleged infringers/defendants clearly falls to the left of the curve for patentee litigants/plaintiffs, but both curves are quite diffuse. The distributions are significantly leptokurtic (kurtosis of 7.2 and 9.7 for plaintiffs and defendants, respectively), meaning that they have long tails. This suggests that outliers may be influential. To make sure that our results are not driven by outliers, we also conducted non-parametric tests (the binomial probability test and the Wilcoxon signed rank test) on the large sample and several sub-samples. All of these tests rejected the null hypothesis of a CAR of zero at either the 5% or the 1% level of statistical significance. In addition, the close correspondence between the means and the medians suggests that our mean estimates for alleged infringers are representative.

14For example, the estimate for single defendants was 0.608% (0.176%) for the full 200 day pre-event window and 0.609% (0.178%) for the truncated window.

3.2 Factors affecting Abnormal Returns

Tables 4 and 5 explore factors that might influence the magnitude of investors' reactions to lawsuit filings by comparing means of different sub-groups. We test differences in the means of different sub-groups using one-tailed *t*-tests, allowing unequal variances between the sub-groups and calculating the degrees of freedom using Satterthwaite's approximation (1946). We conduct these comparisons both for the subject firm's characteristics as well as characteristics of its opposing party in the lawsuit. We also ran regressions with various combinations of the variables in Table 4 (or continuous equivalents) on the right hand side. However, given the noisiness of our data, little conclusive could be drawn from these regressions and where significant results were found, they matched the results found with simple *t*-test comparisons of means.

For patentee litigants, we find that firms with high liabilities relative to assets (and to a lesser extent, firms with high current liabilities to current assets) have much more negative returns from initiating lawsuits. One explanation is provided by Haslem (2005), who observes that lawsuit settlements, including patent settlements, are associated with a decline in firm value, on average. Following Jensen and Meckling (1976), he argues that poorly governed firms will tend to settle lawsuits too soon (from the perspective of shareholders) because that allows managers to expend less effort. Firms with low debt have more leeway for managerial discretion. He finds that these firms experience greater declines in value on settlement. By similar logic, firms with low debt may have more discretion about which lawsuits to file. Therefore, they may choose to file just the most profitable lawsuits while managers in more debt-laden companies may be driven to file more marginal lawsuits, leading to relatively lower CARs.

Another explanation might arise if some industries have a "mutual forbearance" repeated game type equilibrium—firms mutually avoid suing each other because they recognize that if they initiate a suit, they may be punished in the future with retaliatory suits. However, a failing firm may have limited future prospects, hence little to fear from future retaliation. So failing firms, which have high liabilities, may be more likely to initiate suits, including less profitable suits.

For alleged infringers, we find five statistically significant differences. First, if the parties to the lawsuit are in different industries, then the alleged infringer suffers a substantially larger loss, which is statistically significant at the 1% level. Suits from outside the industry may be more of a surprise to investors and may be more indicative of inadvertent infringement.

Alternatively, when disputes occur within a narrow industry, the parties may have greater latitude to craft a settlement that benefits both jointly, including, perhaps, collusive settlements.

Second, if the patentee litigant is a newly public firm, the alleged infringer makes out better. This might be because newly public firms are less able to pursue sustained litigation, posing less of a threat to the alleged infringer. Or, perhaps, a suit by an entrant firm provides a signal that the technology may be more profitable than investors previously realized (see the discussion of signaling below).

The remaining three differences from the large sample, shown in Table 5, are statistically significant at the 5% level. First, small firms seem to have substantially more negative returns. This result appears robust to alternative cutoff points below 500 employees, but we found no significant variation in returns among firms larger than that. One explanation for this is that legal costs are relatively higher for small firms, creating a "floor" on the costs of litigation. Second, we find limited evidence that R&D intense firms suffer more negative returns, however, this result seems sensitive to the specific cutoff used. Finally, we also find some evidence that returns were worse during the 1990s compared to the 1980s. Note that the lower returns for alleged infringers do not appear to be matched by greater returns to patentee litigants (top of the table). In other words, this evidence of greater losses does not suggest a greater transfer of wealth to patent holders.

4 The Costs of Patent Litigation

4.1 Legal Costs

We first look at the legal costs of patent litigation using supplemental data we collected from legal records. We then estimate the total business costs of litigation to alleged infringers based on our event study estimates.

Certain U.S. patent lawsuits contain data on direct legal costs. American patent law provides that the winning party's attorneys' fees can be shifted to the losing party at the discretion of the judge in exceptional cases. Patentees usually get fee awards based on a finding of willful infringement, and alleged infringers usually get fee awards based on a finding the patent suit was frivolous or vexatious. When parties are awarded fees they are required to document their fees, and this information is often included in publicly available records. We searched Westlaw for all patent cases from 1985-2004 that discussed fee-shifting. We found 352 cases in which one of the

parties requested fees (about 100 patent cases go to trial per year). The request was granted in 137 (or 38.9%) of these cases. From this set of 137 cases we were able to determine the magnitude of the fees in 87 cases (63.5% of awards) from judicial opinions or from documents filed by the parties available through the PACER system.

Table 6 shows the median and mean amounts of the fee awards in millions of year 1992 dollars. Mean fees for cases that went through trial ranged were $1.04 million for patentee litigants and $2.46 million for alleged infringers. For cases that were decided prior to trial, the mean fees were $0.95 million for patentee litigants and $0.57 million for alleged infringers.[15] Median values tend to be smaller because the distribution is skewed. In the most extreme case, a $26 million fee was awarded to Bristol-Myers Squibb in conjunction with a successful defense against a pharmaceutical patent suit brought by Rhone-Poulenc. The next largest award was about $7 million.

Our fee-shifting data is in line with survey information collected by the American Intellectual Property Law Association (AIPLA). AIPLA asked patent litigators to estimate the fees associated with patent lawsuits under six different scenarios. Specifically, the survey question divided cases into three different intervals based on stakes, and asked for estimates for cases that concluded at the end of discovery, and cases that reached trial. Their 2001 report indicates the estimated cost through trial was $499,000 when the stakes are less than $1 million, $1.499 million when the stakes are between $1 million and $25 million, and $2.992 million when the stakes are over $25 million.[16] The estimated cost through discovery was $250,000 when the stakes are less than $1 million, $797,000 when the stakes are between $1 million and $25 million, and $1.508 million when the stakes are over $25 million.[17]

The expected legal cost associated with the filing of a patent lawsuit depends on the frequency of the different ways a lawsuit may be terminated. Kesan and Ball (2005) analyze patent lawsuit termination data available from the Administrative Office of the federal judiciary. Examining 5,207 lawsuits that were filed in 1995, 1997, and 2000, they find that most cases terminate short of trial, summary judgment, or other substantive court rulings. In particular, 4.6% of lawsuits reached trial, 8.5% of lawsuits terminated with a summary judgment, dismissal with

15 We included cases that ended in summary judgments, one case that settled, one case that was a default judgment, and one case that ended in a motion to dismiss.

16These amounts increased substantially in the 2003 and 2005 AIPLA reports.

17 The AIPLA estimate of costs through discovery should be larger than the fees shifted at the summary judgment stage to the extent that discovery continues after summary judgment.

prejudice, or confirmation of an arbitration decision, and the remaining 86.9% of cases terminated earlier in the process.

Kesan and Ball construct two proxies for legal fees in patent lawsuits: number of days until the suit terminates, and number of documents filed. Their data show that suits that go to trial last about 1.5 times as many days as suits that end with a summary judgment, and suits that end with a summary judgment last about 1.5 times as many days as all other suits. Further, their data shows that suits that go to trial generate about 2.5 times as many documents as suits that end with a summary judgment, and suits that end with a summary judgment generate about 2.5 times as many documents as all other suits.[18] If we assume that the expected legal cost in a suit that ends before summary judgment is one-half of the cost of suit that reaches summary judgment, then using our data in Table 6 we have estimates of $410,000 for the alleged infringer, and $624,00 for the patentee. A similar calculation using AIPLA data for stakes between $1 million and $25 million yields an estimate of $483,000.

4.2 Firm value and patent lawsuits

Using our CAR estimates, we can calculate the loss of wealth that occurs upon a lawsuit filing. From this, we can then infer a cost to alleged infringers. Multiplying the estimated CAR for each firm by the value of its outstanding shares of common stock immediately prior to the lawsuit filing, we obtain a mean loss of wealth in 1992 dollars of $83.7 million. This is an unbiased estimate of the mean loss of wealth, however, it is not the most efficient estimate. We can do better by multiplying the *mean* CAR by each firm's capitalization.[19]

Using means for three categories (suits with multiple defendants, those with single defendants with more than 500 employees and those with single defendants with 500 or fewer employees), we obtain a mean estimated loss of $52.4 million in 1992 dollars and a median loss

18 We derive these ratios from their Tables 10-12.

19The first estimator is $\frac{1}{N}\sum_{i=1}^{N}(r+e_i)\,x_i$ where N is the number of firms, r is the true CAR, e is the error in measuring the ith firm's CAR, and x is the ith firm's market capitalization. The second estimator is $\frac{1}{N}\left(r+\frac{\sum_{i=1}^{N}e_i}{N}\right)\sum_{i=1}^{N}x_i$. It is straightforward to show that both are unbiased but that the latter has smaller variance assuming that e and x are uncorrelated.

of $4.5 million.[20] These estimates are somewhat smaller than Lerner's estimate for biotech companies of a mean loss of $67.9 million and a median loss of $20.0 million.

This loss of wealth corresponds to the associated drop in investors' expected profits. But does this loss of wealth correspond to the cost of litigation? There are two reasons why it might not. First, the filing of a lawsuit might reveal information that causes investors to revalue the firm for reasons other than the direct and indirect costs of litigation. We explore these possibilities in this section. In the next section, we consider how much investment the firm must undertake in order to restore its investors' wealth—this might not equal the loss of wealth itself.

News of a lawsuit causes investors to re-evaluate their expectations of the discounted profit flow expected from the defendant firm for several different reasons. We assume that the Efficient Market Hypothesis holds, implying that investors incorporate all publicly available information into their valuation of the firm. Consider defendant firm i at time $t = 0$, before the lawsuit filing, and at $t = 1$, immediately after the news of the filing has been made public. At $t = 0$, investors' expected value of the firm based on publicly available information, V, is

$$V_i(0) = \pi_i(0) - p_i(0)\,C \tag{3}$$

where π represents the discounted expected profits of the firm (excluding litigation), p is the expected number of times the firm will be sued for patent infringement and C is the total expected cost to the firm of a patent lawsuit. This expected cost of litigation includes:

- Legal costs.
- Indirect costs, such as management distraction, loss of market share during the lawsuit, and loss of lead-time advantage.
- Financial costs arising from greater risk, including risk of bankruptcy. These include possibly higher costs of funds and also the loss of wealth associated with a higher risk-adjusted discount rate applied to the stream of future expected profits.[21]
- Costs of expected outcomes including those associated with a settlement agreement and trial outcome. Investors take expectations over all possible outcomes and also over the length of time and cost incurred before outcomes are reached.

20Specifically, we multiply the common stock capitalization by .00012 for firms in cases with multiple defendants, by .00564 for single defendants with more than 500 employees, and by .0208 for small single defendants.

21 Implying that π includes the discounted profit stream evaluated at the original discount rate. This interpretation is consistent with our definition of the cost of litigation being the level of investment necessary to restore the wealth of the firms' investors to the level just prior to the lawsuit.

Then at time $t = 1$,

$$V_i(1) = \pi_i(1) - p_i(1)\,C - C \tag{4}$$

Comparing (2) and (1) and taking expectations over all lawsuits, the mean CAR should equal

$$E[\Delta \mathrm{V}] = E[\Delta\pi] - E[\Delta \mathrm{p}]\,C - C \tag{5}$$

The first term represents the change in investors' expectations about the future profit stream based on new information made public by the lawsuit filing. The second term in (5) represents investors' re-assessment of the risk of future litigation. This occurs if the lawsuit provides information that the firm is somehow more prone to litigation than originally expected. Clearly, if the sum of the first two terms is non-zero, then the change in firm value provides a biased estimate of the cost of litigation.

There are two sources of information from the filing that might affect these two terms:

1. Information revealed by the filing documents themselves (and any associated press releases, etc.), and,
2. any information revealed by the event as a signal of the patentee's beliefs. For example, because litigation is costly, the lawsuit may signal that the patent holder believes that the opportunity at stake is particularly valuable—otherwise the suit might not be worth the cost. Note that the documents may reinforce this signal, e.g., the claim for damages may also be large, but with a signal the claim may become credible.

In order for either source to cause investors to revalue the firm, the lawsuit filing must somehow reveal information that was not previously public knowledge—under the Efficient Market Hypothesis we assume that investors correctly incorporate all public knowledge. In other words, the patent holder or the defendant firm must have some *private* knowledge that is revealed in the filing documents or by the signal generated by the filing.

Therefore, if the first two terms in (5) are to affect the mean CAR substantially, there must be a *systematic* reason for the patent holder or the alleged infringer to have private information that is revealed by the lawsuit filing. The documents in the lawsuit filing typically reveal relatively little hard information other than the fact of the filing, often exaggerated claims of damages, and possible allegations of bad behavior by the defendant (something we discuss

below). The patents themselves, of course, are necessarily public information before the suit is filed. But we can identify three reasons why the parties might have private information that is revealed by the filing:

1. Private information about the quality of the technology. For well-known reasons, managers have private information about the quality of their technology. A lawsuit may signal that the patent holder knows that the technology is of better quality than previously realized, hence the market potential is greater, hence a lawsuit may be more profitable. Note that in this case, $E[\Delta\pi] > 0$.
2. Private information about entry plans. If a patent holder plans on entering the defendant firm's market, then the lawsuit might reveal this knowledge, causing investors to revalue the defendant firm downwards because they expect greater competition for the firm. Note that in order for this factor to substantially affect our average CARs, such prospective entrants must initiate a substantial number of patent lawsuits. Also, the prospective entrants cannot have revealed any information about their entry plans prior to filing the lawsuit. This strikes us as a rather odd business strategy—one would think a superior strategy would be to enter the market *before* filing a lawsuit so as to capture market share from those customers who want to avoid the defendant firm. Nevertheless, we will look at empirical evidence regarding this story below. In this case, $E[\Delta\pi] < 0$.
3. Private information about managerial quality or level of effort. For well-known reasons, managers keep private information about their abilities and about the level of effort that they exert. Lawsuits might tend to indicate that managers at the defendant firm were not sufficiently diligent in clearing patent rights or, worse, that they copied technology rather than developing their own. If this tends to be true and if managers tend not to correct their behavior following a lawsuit, then investors might revalue future profits downwards. This occurs both because investors might expect more patent litigation in the future (the second term of (5)) and because poor managerial quality might also reduce profits generally (the first term in (5)).

Several empirical observations lead us to discount the second and third explanations, however. If lawsuit filings revealed news about previously unknown entrants, we might expect this to be particularly true for plaintiffs that had recently gone public. These plaintiffs might not be widely known and therefore, on average, defendant firms might lose greater value when sued

by newly public firms. However, we find that defendants' CARs are significantly more positive when the plaintiff is a newly public firm (see Table 4).[22]

In addition, if news about entry is a significant factor affecting average CARs, then we would expect to find that a significant portion of plaintiffs were: a.) not known as market rivals to the defendant firm prior to the lawsuit, but, b.) became market rivals subsequently. Using Compustat's market segment data, we found that this fact pattern is actually rather uncommon. Compustat reports SIC codes for each firm's major market segments. Of the plaintiffs who had no market segments in common with defendants prior to the lawsuit, we found that only 5% entered a market segment in common with the defendant during the three years following the lawsuit filing.[23] Thus it seems unlikely that a substantial part of defendants' CARs can be explained by revelation of previously unknown entrants.

Other evidence leads us to discount the significance of any news about managerial quality or effort revealed by the lawsuit. Managerial quality is less likely to be of significance in lawsuits that are filed the same year that the patent is granted—often these patents contain claims that were not previously publicly known, so there is less that managers could have done to avoid infringement and managerial quality is less of an issue. For this reason, lawsuits on these patents cannot reveal as much about managerial quality. If revelations about managerial quality explain a large portion of the defendants' CARs, we would expect the CARs to be more positive for patents issued the same year as the lawsuit. In fact, we find that the CARs are more negative for these patents, although the difference is not statistically significant.

Also, we would expect that the managerial quality explanation is much more significant the first time a firm is sued. That is, if a lawsuit reveals significant information about managerial quality, we would expect the second lawsuit to reveal less, and the fifth or tenth lawsuit to reveal even less. We would expect investors to learn and, for this reason, we would expect that, on average, CARs would reflect less revelation of information about managerial quality for, say, the fourth through tenth lawsuit than for the first three.[24] We compared defendant CARs depending on the number of lawsuits the firm had in our sample or on the sequence of the lawsuit. We

22The increase could be because startup firms are less able to pursue sustained litigation and therefore a lawsuit from a startup poses less of a threat. Alternatively, a lawsuit by an entrant may indicate that the technological opportunity is greater than investors previously realized.

23 This figure compares SIC market segments at the 4 digit level. A comparable calculation using 3-digit industry classifications finds a 6% entry rate. This comparison only concerns major market segments, so some entry is unrecorded in minor segments, however, rivalry in minor market segments is only likely to have a minor effect on firm value.

24 This assumes, of course, that management is not entirely replaced between lawsuits.

found no significant differences between CARs for a wide range of different comparisons. E.g., firms with only one lawsuit in our sample had CARs that were on average only .0008 (standard deviation of .0047) less than the CARs for firms sued multiple times. Thus revelations about managerial quality do not seem to explain much of the average loss in firm value from the filing of a lawsuit.

We have little empirical evidence bearing on the role of revelations about technological quality other than anecdote.[25] In Table 4, we saw that defendants do better when the lawsuit is filed by a newly public firm. One explanation is that suits by newly public firms reveal information about technological quality, but this is not the only possible explanation. However, as we noted above, for revelation about technological quality, $E[\Delta\pi] > 0$. Given this, we conclude that $E[\Delta\pi] \geq 0$ and $E[\Delta p] \approx 0$, so that $C \geq -E[\Delta V]$. That is, the cost of litigation is likely at least as large as the loss in firm market value.

4.3 Investment level costs

If we want to know how much litigation "taxes" investment in innovation, then we need to calculate something other than the loss of wealth. That is, we define the "cost of litigation" as the the amount that the firm has to invest in order to increase its value to the level it had just prior to the lawsuit, all else equal. This does not necessarily equal the amount of wealth the firm loses because firms are not necessarily operating at the long-run steady state. They may, instead, be undergoing dynamic adjustment. Then changes in investment will be larger or smaller than the associated changes in firm value. In particular, assuming constant returns to scale, an additional investment of one dollar should increase firm value by an amount equal to Tobin's Q.

Following this logic, to calculate the cost of litigation, we divide the estimated loss of wealth by Tobin's Q.[26] This gives us a mean cost of litigation to alleged infringers of $28.7 million and a median cost of $2.9 million in 1992 dollars.

These estimates are clearly much larger than the estimates of direct legal costs. Most of the cost of litigation to firms appears to arise from business costs such as loss of market share, management distraction, and increased financial costs from greater risk. These costs are incurred even if the suit does not proceed to trial, as happens most often.

25The tech industry joke on hearing that someone has been sued is, "Congratulations, you must be doing something right!"

26We calculate Tobin's Q as the aggregate value of firms divided by the inflation-adjusted value of the aggregate sum of accounting assets and R&D. For details on the computation of these quantities, see Bessen (2006b).

It is interesting to compare our estimate to data from cases that proceed to trial. For the small number of reported cases that go to trial, are won by the patentee, and which award damages to the patentee, we can compare the magnitude of these damages. Mean reported lawsuit damages from 1991-2005 are $10.7 million in 1992 dollars.[27] This number does not include the business cost of the injunction to the infringer, which is often much larger than the damages. For example, the court found damages of $53.7 million in NTP v. RIM, but because of the injunction, NTP eventually settled for $612 million. This mean also does not include the costs of pursuing the litigation, both direct payment of legal costs and indirect business costs. Nevertheless, it is reassuring that this figure is of the same order of magnitude as our mean estimate.

5 The Risk of Infringement for Public Firms

These cost estimates can be summed over all the observed lawsuit filings to obtain measures of firm risk. Table 7 shows three related measures.

The first column lists the annual cost of litigation obtained by summing the cost over all the events in our large sample in each year of the sample. During 1996-99, this averaged $14.9 billion in 1992 dollars. This number is large compared to estimates of patent value. Bessen (2006a), using renewal data to estimate patent value, reports the aggregate value of patents issued to *all* US patentees (not just public firms) in 1991was about $4.4 billion.

Moreover, this figure has varied considerably over time, increasing dramatically from $2.0 billion in 1984 to $16.1 billion in 1999. Figure 3 shows the annual time series. The rise began in the early 1990s and closely follows the increasing frequency of litigation (Bessen and Meurer 2005). But other factors contributed as well, including the increase in R&D spending. Below we look at infringement risk normalized by R&D. The absolute cost of litigation was borne almost entirely by large firms and nearly half by firms in the computer, electronics and software industries.

Note that this series may be substantially understated because, as is well-known, the Derwent Litalert data under-report lawsuits (Lanjouw and Schankerman 2004, Bessen and Meurer 2005). In our 2005 working paper using this sample, we find that only about 64% of lawsuits are reported in Derwent. We have left the first column uncorrected, since it reports a simple sum for our sample. However, the second and third columns compare litigation cost to

27This figure is the mean of deflated annual means reported in Pricewaterhouse Coopers (2006).

numbers of firms and to R&D spending, respectively, so to make the appropriate comparisons, we correct these for under-reporting by dividing by 0.64.[28]

On the other hand, this series may slightly overstate the aggregate cost of patent litigation *per se* because some of the suits listed involved more than just charges of patent infringement and validity. For example, sometimes patent owners will combine allegations of patent infringement with allegations that other rights (including other intellectual property rights) have been violated. Some of the suits of this sort might occur even if there were no patent infringement at issue, so it might not be appropriate to include all of the costs associated with these suits in an aggregate estimate of patent litigation costs. However, for two reasons we do not think this is a serious problem. First, searching published court decisions between 1991 and 1999, only 11% of patent infringement and validity suits also involved claims involving trade secrets, trademarks, copyright, false advertising, unfair competition or noncompete clauses.[29] Second, in Table 4 we observed that the alleged infringer's losses are much greater for inter-industry suits than for intra-industry suits. Since most of the cases involving these additional legal issues occur between rivals in the same industries, these suits do not contribute much to aggregate litigation costs. So it seems unlikely that our aggregate cost estimates overstate the costs of patent litigation by more than a few percent.

The second column displays the annual firm infringement risk. This is the mean expected cost of litigation for a firm from patent infringement lawsuits (or related declaratory actions). It averaged $4.5 million during 1996-99 and it shows a similar pattern of distribution.

The third column shows the ratio of annual litigation cost to annual aggregate R&D. This averaged 14.0% during 1996-99. This relative rate also increased from 1984 to 1999, more than tripling to 19.3% (roughly in line with the growth of the litigation hazard), but this increase was not as rapid as for the quantity in column 1. Note that relative to R&D, litigation risk is low for small firms and for firms outside of the chemical, pharmaceutical and tech industries.

It is tempting to compare this ratio with the "equivalent subsidy rate" for patents, that is, the aggregate value of patents divided by the value of the corresponding R&D. Schankerman (1998) suggests that this ratio represents an upper bound on the subsidy that patents provide to invest in innovation. But, as we argued above, this is clearly a gross subsidy that can be offset by litigation risk if innovators risk inadvertent infringement and by other costs. Several papers

28Lanjouw and Schankerman (2004) found no significant differences between the characteristics of the reported and unreported lawsuits.

29Based on a search of case synopses in the Westlaw FIP-CS database.

calculate this ratio by comparing the value of a nation's patents (estimated using patent renewal data) to R&D (calculated by allocating national R&D spending to the patents obtained in the subject country). Lanjouw et al. (1998) review this literature and report that most subsidy rates are on the order of 10-15%. Arora et al. (2005), use survey data to obtain a comparable estimate of 17%.

However, these numbers are not directly comparable to our estimates of relative litigation risk for at least three reasons. First, because of the way these studies allocate global R&D, they effectively report the subsidy provided by worldwide patents, not patents in a single country.[30] However, the litigation cost is only for US litigation and does not include the costs of litigation in other countries nor the costs of other dispute resolution such as opposition proceedings. An "apples-to-apples" comparison would include these costs as well.

Second, the subsidy rate calculations based on patent value use the value of all of the nation's patents, including patents from individual inventors and small firms. The litigation risk estimates are only for public firms; these are the firms that conduct the lion's share of R&D. A more appropriate comparison would calculate subsidy rates using patents values only for public firms (see Bessen 2006a for comparable figures). In any case, public firms may experience both different subsidy rates and different litigation costs than other firms.

Finally, the litigation costs are estimated for the current year, but the value of patents granted reflects a stream of profits in *future* years. Ideally, we would want to compare litigation costs to the profits from patents on the same cohort of technologies that were litigated. Some of these profits are realized prior to the time of litigation. Since both litigation costs and patent values are trending up, this use of current patent values understates the significance of litigation costs.

All three of these considerations suggest that a direct comparison of reported subsidy rates to US litigation risk overstates the relative positive value of patents. At the very least, these estimates suggest that litigation risk is quite large compared to the private benefits of patents, especially in recent years.

30That is, using trade data, they allocate a share of the R&D performed in every OECD country to, say, French patents when they calculate the subsidy rate using the value of French patents. The apparent assumption behind this allocation is that subsidy rates are the same across nations and that the share of trade is proportional to each nation's share of worldwide patent value. Then the calculated subsidy rate will represent the return from worldwide patents. Similarly, Arora et al. use US patents as a right hand variable, but this proxies for each firm's worldwide patents.

6 Conclusion

Using a large set of event studies, we estimate the total cost that patent litigation imposes on firms and we estimate the risk of infringement litigation. We find that, contrary to what is sometimes assumed, the business costs of litigation far exceed the direct legal costs. And we find that by the late 1990s, patent litigation risk was of the same order as, if not larger than, estimates of the private benefits firms receive from patents. Moreover, consistent with the previous literature, the losses to alleged infringers do not correspond to a transfer of wealth to patent holders; instead there is a substantial joint loss of wealth. Our estimates concern private costs rather than the social costs of litigation, nevertheless these estimates tell us something about the effectiveness of patents as a policy tool to encourage investment in innovation.

In the best case, this suggests that the patent system is at present an inefficient form of subsidy or regulation. Thomas Hopkins estimates the total 1992 cost of general regulatory compliance is $389,911 per firm (in 1995 dollars).[31] But the costs of complying with the patent system—annual infringement risk of $4.5 million—are much larger.

In the worst case, the net effect of patents today may be to reduce the profits of public firms and to possibly impose disincentives on innovative activity as well. Exploration of the possible causes and the significance of this for policy and for normative analysis are beyond the scope of this paper, however. Nevertheless, our analysis indicates that infringement risk should be an important consideration in the formulation of patent policy.

References

American Intellectual Property Law Association. 2001. "Report of the Economic Survey."

Arora, Ashish, Marco Ceccagnoli and Wesley Cohen. 2005. "R&D and the Patent Premium," working paper.

Barber, Brad M. and John D. Lyon. 1997. "Detecting Long-run Abnormal Stock Returns: The empirical power and specification of test statistics," *Journal of Financial Economics,* v. 43, pp. 341-72.

Bessen, James (2006a), "The Value of U.S. Patents by Owner and Patent Characteristics," Boston Univ. School of Law Working Paper No. 06-46.

Bessen, James (2006b), "Estimates of Firms' Patent Rents from Firm Market Value," Boston Univ. School of Law Working Paper No. 06-14.

31 Thomas D. Hopkins The Changing Burden of Regulation, Paperwork, and Tax Compliance on Small Business, http://www.sba.gov/gopher/Legislation-And-Regulations/Burden/burd7.txt

Bessen, James and Robert M. Hunt. 2007. "An Empirical Look at Software Patents," *Journal of Economics and Management Strategy* 16, no. 1, pp. 157-89.

Bessen, James and Michael J. Meurer. 2005. "The Patent Litigation Explosion," B.U.S. L. Law and Economics Working Paper Series, No. 05-18.

Bessen, James and Michael J. Meurer (2006), "Patent Litigation with Endogenous Disputes," *American Economic Review*, 96, no. 2, pp. 77-81.

Bessen, James and Michael J. Meurer (2008 forthcoming), *Do Patents Work?* manuscript.

Bhagat, Sanjai, John M. Bizjak, and Jeffrey L. Coles, 1998. "The shareholder wealth implications of corporate lawsuits," *Financial Management* 27, pp. 5–27.

Bhagat, Sanjai, James A. Brickley, & Jeffrey L. Coles. 1994. "The Costs of Inefficient Bargaining and Financial Distress: Evidence from Corporate Lawsuits," *Journal of Financial Economics*, v. 35, p. 221.

Brown, Stephen J. and Jerold B. Warner. 1985. "Using Daily Stock Returns: The case of event studies," *Journal of Financial Economics,* v. 14, pp. 3-31.

Cutler, David and Lawrence Summers. 1988. "The costs of conflict resolution and financial distress: Evidence from the Texaco-Pennzoil Litigation," RAND Journal of Economics, 19, pp. 157-72.

Dodd, Peter and Jerold B. Warner. 1983. "On Corporate Governance: A study of proxy contests," *Journal of Financial Economics*, v. 11, pp. 401-38.

Hall, Bronwyn H., Stuart Graham, Dietmar Harhoff, David C. Mowery, 2004. "Prospects for Improving U.S. Patent Quality via Postgrant Opposition," Innovation policy and the economy. Volume 4, Cambridge and London: MIT Press for the National Bureau of Economic Research, pp. 115-43.

Haslem, Bruce. 2005. "Managerial Opportunism during Corporate Litigation," Journal of Finance, 60, no. 4, pp. 2013-41.

Jensen, Michael C., and William H. Meckling, 1976, "Theory of the firm: Managerial behavior, agency costs and ownership structure," Journal of Financial Economics 3, 305–360.

Kesan, Jay P. and Gwendolyn G. Ball. 2005. "How Are Patent Cases Resolved? An Empirical Examination of the Adjudication and Settlement of Patent Disputes," U. Illinois Law & Economics Research Paper No. LE05-027.

Kothari, S. P. and Jerold B. Warner. 1997. "Measuring Long-horizon security price performance," *Journal of Financial Economics,* v. 43, pp. 301-39.

Kramer, Lisa A. 2001. "Alternative Methods for Robust Analysis in Event Study Applications," in Lee, C. F., ed., *Advances in Investment Analysis and Portfolio Management,* v. 8, pp. 109-32.

Lanjouw, Jean O., Ariel Pakes and Jonathan Putnam. 1998. "How to Count Patents and Value Intellectual Property: The uses of patent renewal and application data," *Journal of Industrial Economics* 46 405-32.

Lanjouw, Jean O. and Mark Schankerman. 2004. "Protecting Intellectual Property Rights: Are Small Firms Handicapped?," *Journal of Law and Economics*, 47, pp. 45-74.

Lerner, Josh. 1995. "Patenting in the Shadow of Competitors," *Journal of Law and Economics*, 38, no. 2, pp. 463-95.

Levin, R. C. and J. Levin. 2002. "Patent Oppositions." working paper.

Lunney, Glynn S. Jr. 2004. "Patent Law, the Federal Circuit, and the Supreme Court: A Quiet Revolution," Supreme Court Economic Review, 11, pg. 1.

Mackinlay, A. Craig. 1997. "Event Studies in Economics and Finance," *Journal of Economic Literature*, 35(1), pp. 13-39.

Meurer, Michael J. 1989. "The Settlement of Patent Litigation," *RAND Journal of Economics*, Spring 1989, v. 20, iss. 1, pp. 77-91.

Moore, Kimberly A. 2000. "Judges, Juries and Patent Cases – An Empirical Peek Inside the Black Box, *Michigan Law Review* 99, pp. 365-406.

Pakes, Ariel and Mark Schankerman. 1984. "The Rate of Obsolescence of Patents, Research Gestation Lags, and the Private Rate of Return to Research Resources," in Griliches, Zvi, ed., *R&D, Patents and Productivity*, chicago: University of Chicago Press for the NBER.

PricewaterhouseCoopers (2006), "2006 Patent and Trademark Damages Study."

Salinger, Michael. 1992. "Standard Errors in Event Studies," *The Journal of Financial and Quantitative Analysis* 27(1), pp. 39-53.

Satterthwaite, F. E. 1946. "An approximate distribution of estimates of variance components," Biometrics Bulletin, 2. pp. 110-14.

Schankerman, Mark. 1998. "How Valuable is Patent Protection: Estimates by Technology Fields," *RAND Journal of Economics*, 29, no. 1, pp. 77-107.

Ziedonis, Rosemarie Ham. 2004. "Patent Litigation in the US Semiconductor Industry," in Cohen, Wesley M. and Stephen Merrill, eds., *Patents in the Knowledge-Based Economy*, National Research Council.

Tables and Figures

Table 1. Summary Statistics

	Matched Sample				All Alleged Infringers	
	Patentee Litigants		Alleged Infringers			
	Mean	Median	Mean	Median	Mean	Median
Sales ($ million)	7,020.4	1,267.7	6,186.7	1,022.7	8,604.0	1,368.1
Employees (1000s)	40.2	9.2	36.1	6.7	46.3	9.3
R&D / Sales	9.4%	5.4%	18.9%	5.3%	13.9%	5.0%
No R&D reported	6.1%		9.0%		18.4%	
No. observations	771		720		2887	

Table 2. Cumulative Abnormal Returns from Suits Announced in Wall Street Journal, 1984-99

Event window	WSJ article T-2 to T+1	Suit filing T-1 to T+24	Bhagat et al. (1998)
Patentee Litigant (Plaintiff)			
mean	-0.3%	-0.1%	-0.31%
median	0.0%	0.9%	
no. of observations	86	86	
Alleged Infringer (Defendant)			
mean	-2.6%	-1.8%	-1.50%
median	-1.4%	-1.9%	
no. of observations	82	82	
Combined (matched parties)			
mean	-2.6%	-2.5%	
median	-1.8%	-0.5%	
no. of observations	80	80	
Addendum: mean combined abnormal returns			
Bhagat et al. (1994)	-3.13%		
Lerner (1995)	-2.0%		

Note: Events with possibly confounding news are excluded. Average cumulative abnormal returns are simple unweighted means.

Table 3. Cumulative Abnormal Returns

	Mean CAR	Median CAR	Robust *Z* Statistic	Observations
Sample: Matched Parties				
Patentee Litigants				
Base	-0.38% (0.30%)	0.00%	-1.51	667
Definite infringement suits	-0.63% (0.37%)*	-0.45%	-2.18*	412
Alleged Infringers				
Base	-0.62% (0.33%)*	-0.97%	-1.55	661
Definite infringement suits	-0.77% (0.42%)*	-0.83%	-1.70*	407
With possibly confounding events	-0.45% (0.31%)	-0.57%	-1.32	743
Sample: All alleged infringers				
Base	-0.50% (0.16%)**	-0.51%	-3.24**	2,887
Single defendants	-0.61% (0.18%)**	-0.54%	-2.94**	2,460
Multiple defendants	-0.01% (0.39%)	-0.39%	-1.38	427
Single defendants, definite infringement cases	-0.63% (0.27%)**	-0.42%	-2.37**	1,108

Note: Standard errors in parentheses. Single asterisk indicates statistical significance at the 5% level; double asterisk indicates 1% significance. Average cumulative abnormal returns (CARs) are weighted means, with weights proportional to the inverse of the estimated variance of each return. In matched sample, events with possibly confounding news are excluded, except where noted. Event window is 25 days (T-1 to T+24). Cumulative abnormal returns are estimated using OLS except for cases with multiple defendants (in large sample), which are estimated jointly. The robust *Z* statistic is a joint test of the individual firm *t* statistics (Kramer 2001).

Table 4. Differences in Mean CARs by Characteristics

Sample: Matched Parties

	Alleged Infringer	Patentee Litigant
Firm characteristic		
Employees < 500	-3.20% (2.32%)	-3.75% (2.42%)
R&D / Sales > .15	0.23% (1.62%)	-0.53% (1.22%)
Total liabilities / Total Assets > .5	1.40% (0.87%)	-2.35% (0.75%)**
Capital / Employee > $100,000	-0.02% (0.93%)	-1.02% (0.74%)
Current Assets / current liabilities < 1.5	0.94% (1.00%)	-1.91% (0.87%)*
Newly public firm	3.77% (1.51%)**	-1.92% (2.56%)
Rival characteristic		
Employees < 500	1.06% (1.19%)	-1.37% (1.07%)
R&D / Sales > .15	0.23% (1.62%)	0.81% (0.97%)
Total liabilities / Total Assets > .5	-0.15% (0.86%)	-0.35% (0.80%)
Capital / Employee > $100,000	-0.99% (0.95%)	1.02% (0.74%)
Current Assets / current liabilities < 1.5	1.69% (1.11%)	1.19% (0.86%)
Newly public firm	-0.94% (1.78%)	0.32% (1.05%)
Other Characteristics		
Year > 1989	-0.15% (0.82%)	0.09% (0.77)%
Firms in same SIC4 primary industry	2.67% (1.16%)**	-0.61% (0.89%)

Note: Standard errors in parentheses. Single asterisk indicates difference is statistically significant at the 5% level; double asterisk indicates 1% significance (one-tailed test allowing unequal variances and using Satterthwaite's calculation for degrees of freedom). Average cumulative abnormal returns are weighted means, with weights proportional to the inverse of the estimated variance of each return. Comparisons are for cases where infringement is known and no possibly confounding events have been found.

Table 5. Differences in Mean CARs by Firm Characteristics
Sample: All alleged infringers

Employees < 500	-1.70% (0.92%)*
R&D / Sales > .15	-1.79% (0.80%)*
Total liabilities / Total Assets > .5	0.05% (0.33%)
Capital / Employee > $100,000	-0.26% (0.44%)
Current Assets / current liabilities < 1.5	0.11% (0.34%)
Year > 1989	-0.56% (0.32%)*
Patentee is public firm	-0.12% (0.35%)
Industry	
SIC = 28 (chemicals, inc. pharma)	-0.41% (0.41%)
SIC = 35,36,73 (electronics, computer,sw)	0.06% (0.38%)
Other manufacturing	0.16% (0.33%)

Note: Standard errors in parentheses. Single asterisk indicates difference is statistically significant at the 5% level; double asterisk indicates 1% significance (one-tailed test allowing unequal variances and using Satterthwaite's calculation for degrees of freedom). Average cumulative abnormal returns are weighted means, with weights proportional to the inverse of the estimated variance of each return.

Table 6. Attorneys' Fees Awarded in Patent Lawsuits (in millions of year 1992 dollars)

	Mean	Median	Observations
Patentee Litigant			
Summary Judgment	.95	.40	8
Verdict	1.04	.78	51
Alleged Infringer			
Summary Judgment	.57	.30	10
Verdict	2.46	.98	18

Table 7. Measures of Infringement Risk, Public Firms

	Aggregate Annual Cost of Litigation to Alleged Infringers (billion $92)	Annual Firm Infringement Risk (million $92)	Aggregate Risk / R&D
1984	2.0	1.3	4.9%
1999	16.1	7.0	19.3%
1996 - 99			
All firms	14.9	4.5	14.0%
Small firms (employees <500)	0.1	0.1	1.3%
Large firms (employees>=500)	14.8	9.8	14.9%
SIC = 28 (chemicals, inc. pharma)	3.4	9.7	14.1%
SIC = 35,36,73 (electronics, computer, software)	6.8	5.7	14.8%
Other manufacturing	1.7	2.3	5.3%

Note: Annual cost of litigation is the mean CAR times the market capitalization of each firm's common stock divided by a GDP deflator and by the aggregate Tobin's *Q* (market value divided by replacement value of capital including R&D). Firm infringement risk is the expected annual cost of litigation. Column 1 includes all events in the large sample (2,887) with separate means for small firms and lawsuits with multiple defendants. Columns 2 and 3 have been adjusted for under-reporting of lawsuits (see Lanjouw and Schankerman 2004 and Bessen and Meurer 2005).

Figure 1. Frequency of Wall Street Journal Stories Relative to Court Filing Date

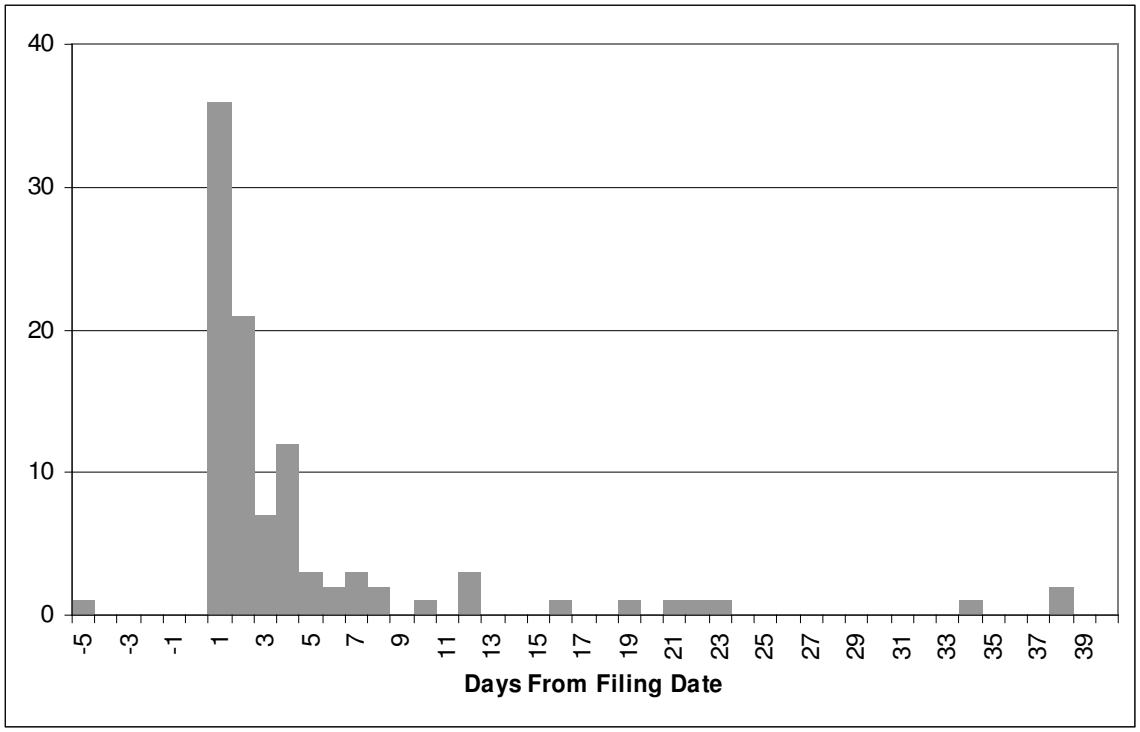


Figure 2. Histograms of Cumulative Abnormal Returns

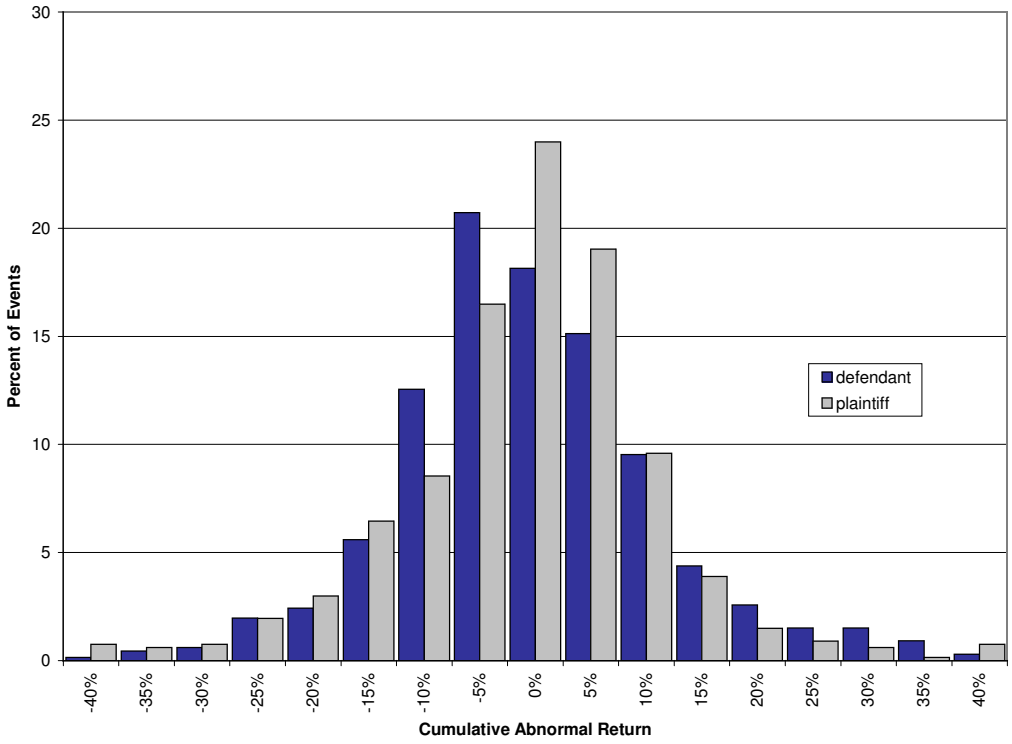
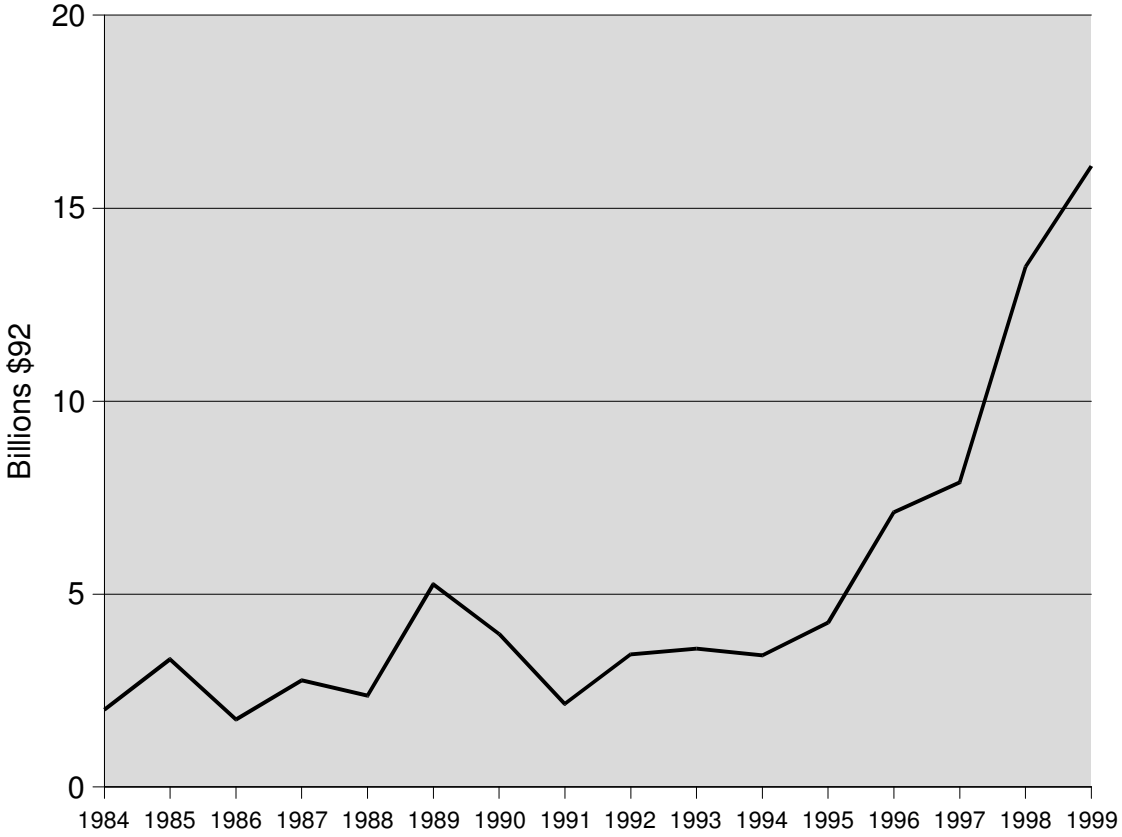


Figure 3. Aggregate annual cost of patent litigation to alleged infringers



7 Appendix

This appendix explores further our choice of a window around the lawsuit filing date rather than an announcement in a newspaper or wire service. First, we explore whether a sample based on Wall Street Journal articles is likely to suffer sample selection bias. Table A1 shows Probit regressions on whether a lawsuit in our matched sample received mention in the Wall Street Journal. The patentee litigant's capital intensity and the alleged infringer's stock beta are both highly significant (at the 1% level) predictors of a Wall Street Journal article. Because high beta stocks are likely to have a larger reaction to news of a lawsuit, this suggests that samples based on Wall Street Journal articles may have significant bias. We find, in fact, that our estimates from our sub-sample of lawsuits announced in the Wall Street Journal do have much more negative CARs.

Because we do not have information on whether a suit is an infringement suit or a declaratory action in all cases, we likely mis-identify some plaintiffs and defendants, possibly diluting our estimates for alleged infringers. One way to correct this would be to limit our sample to cases of definite infringement, but this might also introduce a selection bias. The last two columns of Table A1 explore characteristics that may affect whether the suit is an infringement suit or a declaratory action. It appears that newly public patentees may be a bit more aggressive in filing suits, while larger alleged infringers may be more likely to end up in an infringement suit. Large firms may avoid filing declaratory actions, waiting for evidence that the patent owner has the resources to conduct a lawsuit. Because there may be a selection bias, we report CARs both for the entire sample and also for cases that we know are infringement suits.

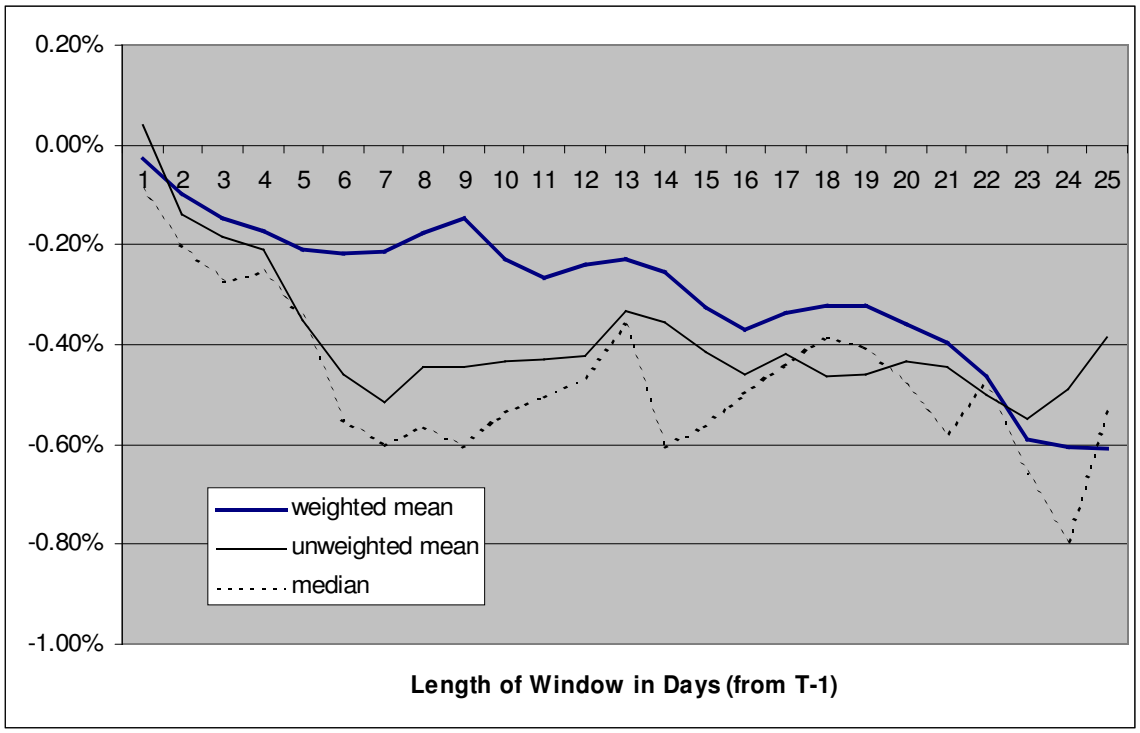
Finally, as discussed in the text, because news of a lawsuit filing leaks out more slowly than a newspaper announcement, we use a 25 day event window. Figure A1 shows the mean CARs we would obtain using shorter event windows. Note that the unweighted mean and the median CARs both react more sharply in the days after the filing. This is because high beta stocks respond more quickly after the filing (they are the ones where investors may have the greater incentive to obtain such news). Because the CARs for low beta stocks are estimated more precisely and their response is slower, the weighted mean responds more slowly. However, all three averages are roughly equal by the end of our 25 day window.

Table A1. Suit Announcement and Type

	Wall Street Journal Article		Infringement Suit	
	1	2	3	4
Plaintiff/patentee litigant				
Ln employment		0.05 (.03)	.02 (.03)	.01 (.04)
New firm		-.25 (.23)	**.63** (.29)	.62 (.33)
Stock Beta	.13 (.12)	.15 (.11)		.20 (.13)
Capital / employee	**1.01** (.38)	**1.12** (.40)		-.64 (.49)
Defendant/alleged infringer				
Ln employment		-.01 (.03)	.06 (.03)	**.07** (.03)
New firm		.28 (.20)	-.01 (.20)	-.05 (.22)
Stock Beta	**.35** (.13)	**.35** (.13)		.05 (.14)
Capital / employee	.05 (.36)	.11 (.36)		-.95 (.51)
No. of observations	637	637	507	475
Pseudo-R-squared	.049	.062	.023	.057

Note: Probit regressions. Robust standard errors in parentheses. Bold estimates are significant a the 5% level or better. Regressions include industry dummies (not shown).

Figure A1. Average Abnormal Cumulative Returns Over Time



Patent Failure: How Judges, Bureaucrats, and Lawyers Put Innovators at Risk

forthcoming Princeton University Press, 2008

Chapter 1

The Argument in Brief

Zimbabwe, a country once considered the "breadbasket of Africa," now suffers widespread starvation. Much of this decline can be attributed to the tyrannical policies of President Robert Mugabe — in particular, his disregard for property. In 2000, Mugabe's followers seized land on over one thousand farms owned by white farmers. But when Zimbabwe's Supreme Court ordered the squatters evicted, Mugabe forced the chief justice to resign and physically threatened the remaining justices, who relented. Owners abandoned their property, severely disrupting agricultural production, and within a few years Zimbabwe was wracked by famine.[1]

Even with the rule of law, property systems can fail. A successful property system also requires supportive institutions, and the technical details of property law must make sense. Consider one particular country where many property owners had a hard time enforcing their rights and were often forced to resort to expensive litigation. In one notorious case, a property owner had to assert its rights against more than one hundred parties, an ordeal that involved forty-three separate lawsuits. With so many ostensible trespassers, one might assume the property owner's claim was weak, but as the courts

found, this was not the case. Only one suit was dismissed on summary judgment, the owner's claims were largely upheld on appeal, and almost all the defendants settled.

This example does not come from a failed state or a "tinhorn dictatorship." The country in question is the United States; the property is patent 4,528,643, granted in 1985; the owner who initiated the lawsuits was a company called E-Data; and the alleged violators included a roster of technology companies as well as thousands of small businesses and individuals operating e-commerce websites.[2] This failure of property rights cannot be attributed to a breakdown of the rule of law. Rather it was caused by the failure of patent-related institutions and patent law generally to get the details right. This widespread pattern of alleged violation and litigation would surely be unusual in land, real estate, or personal property in the United States.

Such a rickety system of property rights seems unlikely to be an engine of growth. Burdensome means of enforcement lessen the value of property to its owners. Moreover, property disputes impose costs on other parties. Even though few are sympathetic to trespassers, squatters, and others who seek unjust enrichment, there is good reason to worry about costs imposed on innocent violators. In the case above, many of the defendants believed that they were not infringing upon the owner's rights, and they innocently made investments that turned out to be in violation. Those investments were exposed to unnecessary risk because of unclear property boundaries. A defective property system discourages trade and investment not just by property owners, but also by those who inadvertently face the threat of property-related lawsuits.

This book considers patents as a form of property right. If patents work as property, they should reward innovators and encourage investment in innovation. Below, we

explore how the laws and institutions of property, including patents, succeed and how they fail. The E-Data example suggests that even in an advanced society with well-developed legal institutions and strong respect for the law, property can fail. Yet this one example might not be entirely representative. We need to go further and ask whether patents work well as property overall.

This question is important because innovators have grown frustrated with the failings of the American patent system. Over the past several years, in newspaper articles and at hearings held by the Federal Trade Commission (FTC) and the National Academy of Science (NAS), industry executives have complained in growing numbers that the patent system is broken. In response, Congress has held its own hearings and debated reform. Critics argue that changes in patent law have created "a legal frenzy that's diverting scientists from doing science."[3] Some even believe that the patent system should be abolished.[4] Others say that the patent system can be fixed with some modest reforms.[5] Still others maintain that the patent system is not broken at all, and that current efforts to reform it are just an attempt to weaken the rights of small inventors.[6]

It is hard to tell who is right, however, because most evidence offered in support of these positions is anecdote, if not myth:

- Defenders of the current system tell stories about the role of patents in protecting small inventors from rapacious corporate giants. But most patents and most litigation do not come from independent inventors, so it is not clear how representative these stories are, or how important small inventors are overall.
- Critics of the system cite patents on a peanut-butter-and-jelly sandwich (6,004,596), a method of using a backyard swing (6,368,227), and a method of combing hair over a

bald spot (4,022,227) as evidence of poor patent "quality." Standing alone, however, these patents are not evidence that anything is seriously wrong. Silly patents and patents on unworkable inventions, such as perpetual-motion machines, have been around for at least two hundred years.[7]

- Critics also raise the issue of lawsuits initiated by "patent trolls"<m->people who obtain broad patents not for purposes of innovation, but solely to ensnare real innovators who might inadvertently cross the boundaries of the trolls patent. The label "troll" is potent rhetoric, but only a small percentage of patent litigation can be attributed to the most egregious trolls.

Stories about garage inventors are inspiring, while stories about frivolous patents and frivolous lawsuits are troubling, but better evidence is needed to guide patent reform. Without this evidence, it is hardly surprising that some reform proposals seem to be ad hoc.

This book moves beyond anecdote to provide the first comprehensive empirical evaluation of the patent system's performance. We measure patents against a simple, well-defined yardstick inspired by economic analysis of property rights. Our yardstick weighs the benefit of patents to an innovator against their cost, including the risk of inadvertent infringement. If the estimated costs of the patent system to an innovator exceed the estimated benefits, then patents fail as property.

Some readers might immediately find our objective to be somewhat oddly stated or, perhaps, overreaching. The key limiting qualifier here — the limitation that makes the

empirical exercise feasible — is "as property." At the risk of getting a bit pedantic, this phrase requires more careful discussion.

Many readers might think the phrase is redundant in the present context. Some are likely to assume that patents *are* property. What, then, could it mean to ask if they work "as property"? For example, in a paper comparing different ways of providing incentives to innovate, theoretical economists Gallini and Scotchmer (2004) tell us that patents are "intellectual property," which they define as "an exclusive right to market an invention for a fixed time period." As such, it might seem sensible to call patents "property" because exclusion is a hallmark of property. Property rights in land give a farmer exclusive rights to grow crops and bring them to market. If patents provide exclusive rights to market inventions, how could they not work "as property"?

It is important, however, to distinguish idealized depictions of patents from the actual workings of the patent system. Patents do not actually provide an affirmative right to market an invention; they provide only a right to exclude *others* from doing so. This might seem an inconsequential difference, but it has practical significance: other patent holders can block even a patented invention from coming to market. The power to block innovation is especially troublesome when property boundaries are not well identified.

Some of the troubling issues raised by the E-Data patent are explained by this difference. E-Data's patent was for a kiosk that produced digital audio tapes and the like in retail stores, but they interpreted this patent to cover a very broad swath of e-commerce. On the other hand, IBM holds hundreds of patents related to e-commerce, but this did not prevent E-Data from threatening to block IBM from marketing its own innovative e-commerce products. To market its own patented technologies, IBM was

forced to pay E-Data for a license. Similarly, Research in Motion (RIM) holds patents on its popular Blackberry personal-communication device, but this did not prevent NTP, a patent-licensing company, from famously suing RIM for patent infringement. To avoid being excluded from the market for its own patented invention, RIM paid NTP more than a half-billion dollars.

Gallini and Scotchmer present idealized notions of patents and property that might be useful for some theoretical inquiries.[8] Empirical investigation, however, requires us to be mindful of the ways in which real patents might fall short of such stylized concepts of property. If examples like E-Data and RIM are typical — and this is the kind of empirical question we explore in depth below — then patents will perform quite differently from the property ideal. Below, we estimate just how far actual patents fall short of that ideal. Patents work well as property for some kinds of technology and given the right institutional setting. Patents fail as property for other kinds of technology and given the wrong institutional setting.

Overall, the performance of the patent system has rapidly deteriorated in recent years. By the late 1990s, the costs that patents imposed on public firms outweighed the benefits. This provides clear empirical evidence that the patent system is broken. Both our empirical analysis and our comparative institutional analysis provide clues about the causes of this deterioration<m->and about what might be done to fix it.

Our focus is on the American patent system, and some of the problems we identify with the American patent system are unique. Nevertheless, there are two reasons that our analysis has relevance to innovation in other countries. First, many other patent systems are under some pressure to become more like the American patent system. For example,

Japan and Europe have loosened restrictions against software patents. Second, patent rights and patent litigation are global matters. Important inventions are usually patented in all major markets. This means that patent holders can choose where to litigate. Increasingly, patent disputes are being litigated in the United States, usually resulting in worldwide settlement agreements. Indeed, European inventors file more lawsuits in the United States than they file in any European country other than Germany. This means that the United States patent system directly affects firms and innovators in Europe, Japan, and elsewhere.

Patents as Property

We begin by comparing patents to tangible property. Lawyers and legal scholars<m->perhaps because they have endured at least a semester of training in property law and are therefore aware that things might not be so neat — tend to speak of patents not as a form of property, but as *analogous to* other forms of property. Some argue that the analogy might not be appropriate (Lemley 2005), others that the analogy is long-standing (Mosoff 2007), but most recognize that the law and institutions of property systems are complicated and patent law necessarily diverges from the law of tangible property.

In chapter 2 we begin our inquiry by looking at the appropriateness of this analogy, comparing the property-like features of patent law to features of the law of tangible property. We have already noted one important difference: patents do not provide an affirmative right to use an invention. More than one person can use an invention at a time and more than one inventor can claim rights over an invention. Many people can even

invent the same technology independently at the same time. In contrast, tangible property is a "rival" good — that is, only one person can use it at a time. This means that the *right to exclude* others more or less conveys an affirmative *right to use* tangible property. As we shall see, this difference between inventions and tangible property is important.

In many other ways, however, patent law shares essential doctrinal features with the law of tangible property. Specifically, patents provide partial rights to exclude others from using an invention as well as rights to transfer ownership. Just as property rights provide incentives to invest in the acquisition, development, and maintenance of tangible property, patents potentially provide incentives to conceive a new technology ("invention"), develop it into a commercial product or process ("commercialization"), and put it to use ("innovation"). Such "innovation incentives" are central to the Constitutional mandate to "promote the progress of . . . the Useful Arts," which the framers set out when empowering Congress to devise a patent system.

But property and patents only *potentially* provide these incentives. Our review finds well-known evidence that property systems sometimes fail to provide such incentives efficiently:

- Property rights can fail when their validity is uncertain. Such was the case when the transition from Mexican to American rule in California during the nineteenth century clouded the validity of land titles granted under Spanish and Mexican rule. This uncertainty led to squatting and a decline in agricultural productivity.
- Property rights can fail when rights are so highly fragmented that the costs of negotiating the rights needed to make an investment become prohibitive. Such was the case with Russian retail establishments in that country's transition to a private

economy. Ownership rights to stores were granted to large numbers of parties, making it too difficult for any one group to obtain the required permissions to operate each store. The stores were often shuttered while street vendors conducted a busy trade nearby.

- Property can fail when boundary information is not publicly accessible. In many less-developed nations, cumbersome regulations discourage impoverished people from recording property boundaries. This limits their ability to trade that property or to use it as collateral for obtaining loans.
- Finally, property rights can fail when the boundaries of the rights are not clear and predictable. This problem sometimes arises with property extracted from nature, such as mineral rights. For example, mineral veins beneath the surface of the earth twist and intersect in unpredictable ways. Such a boundary-related failure in the copper mines of Butte, Montana, led to a violent struggle between rival claimants.

These failures emphasize the importance of implementation in property rights systems. The economic effectiveness of any property system depends not just on what it sets out to do, but also on the laws, regulations, institutions, and norms that implement the system. Consequently, the doctrinal similarity between patent law and the law of tangible property can obscure important differences in economic performance that arise because these doctrines are implemented differently. Patents might not work well as property if patent law is not implemented effectively; the messy details of how patents work matter.

The Notice Function: If You Can't Tell the Boundaries, It Ain't Property

We can identify one very important difference between the way property rights and patent rights are implemented. This difference concerns the "notice function" of property. An efficient property system notifies non-owners of property boundaries. For example, land rights have a well-developed and efficient system to notify third parties of boundaries. Because of this, only rarely does someone invest millions of dollars constructing a building that encroaches on someone else's land without permission. Far more typically, would-be investors "clear" the necessary rights before investing. They locate markers, check land deeds, conduct surveys, and so forth, in order to determine the adjacent boundaries. They then either negotiate rights to the needed land or design the building to avoid encroachment.

The notice function does not always work so well with patents. For example, the E-Data dispute arose because hundreds of parties, including some very large companies, ignored, did not see, or misunderstood the boundaries created by the patent in question. That patent, awarded to Charles Freeny Jr., was entitled a "System for Reproducing Information in Material Objects at a Point of Sale Location." Its unhelpful title obscures the fact that Freeny actually invented a kiosk for producing music tapes or other products in retail stores using digital information. But, as we have seen, the patent was asserted against thousands of companies doing e-commerce, a rather different technology.

Why did notice fail so completely in this case? For one thing, a prospective technology investor needs to check a very large number of patents. According to David M. Martin, CEO of a patent risk-management firm, "if you're selling online, at the most

recent count there are 4,319 patents you could be violating. If you also planned to advertise, receive payments for, or plan shipments of your goods, you would need to be concerned with approximately 11,000."[9]

But even if a website developer could check all these patents, it would be very difficult to know what their boundaries are. The boundaries of the E-Data patent depend on the meaning of abstract phrases such as "point of sale location," "information manufacturing machine," and "material object." Consider, for example, the meaning of "point of sale location." This is a bit of computer and retail-industry jargon first used when electronic terminals replaced cash registers. It refers to the location within a store where items are checked out and transactions take place. Did this term in the patent claim limit the patent to transactions in retail stores, or did it cover all e-commerce, including transactions that might take place in buyers' offices or even in their bedrooms? The district court limited the patent right to retail locations. In 2001, however, the Court of Appeals for the Federal Circuit, using legal rules that place little weight on actual industry usage or on dictionary definitions, concluded that the "point of sale location" included bedrooms, offices, and anywhere else with an Internet connection. Thus, sixteen years after the patent was granted, it was given boundaries that many people, including a district court judge, would find surprisingly broad. In the interim, the correct boundaries of this patent were essentially unknown. The patent offered poor notice.

Poor notice causes harm because it subjects technology investors to an unavoidable risk of disputes and litigation. The expected cost of inadvertent infringement imposes a *dis*incentive on technology investors. Potential innovators consider not only the reward that they might reap from owning patents, but also the risk of being sued for infringing

upon the patents of others. Clearly, if the risk of inadvertent infringement is too great, the net incentives provided by the patent system will be negative, and patents will fail as a property system. This is similar to the failures that occurred with Mexican land grants in California, with Russian retail store ownership, and with the copper-mine war in Butte, Montana, noted above.

Establishing notice is often inherently easier for tangible property because, as opposed to patents, tangible property is a rival good. This means that active possession of tangible property is often sufficient to inform the world about what is owned and who owns it<m->consider, for example, the shirt on one's back. For nonrival inventions, such as RIM's technology, however, the fact that RIM independently developed and actively possesses the technology does not help clarify the relevant patent boundaries and ownership.

In addition, the implementation of patent notice suffers important deficiencies. In chapter 3 we explore several institutional differences between patents and the property system for land that might make the former particularly prone to notice problems. These institutional features affect patent notice and are thus central to our analysis:

1. *Fuzzy or unpredictable boundaries*: Surveying land is inexpensive, and the survey boundaries carry legal weight. While surveyors can plainly map the words in a deed to a physical boundary, it is much harder to map the words in a patent to technologies, as the E-Data patent dispute illustrates. Not only are the words that lawyers use sometimes vague, but the rules for interpreting the words are also sometimes unpredictable. Although innovators can obtain expensive legal opinions about the

boundaries of patents, these opinions are unreliable. There is thus no reliable way of determining patent boundaries short of litigation.

2. *Public access to boundary information*: The documents used to determine boundaries for both land and patents are eventually available to the public. It is possible, however, for patent owners to hide the claim language that defines patent boundaries from public view for many years, a practice that is becoming increasingly frequent.

3. *Possession and the scope of rights*: Generally, tangible property rights are linked closely to possession, hence the well-known expression, "possession is nine points of the law." Patent law also requires possession of an invention, but often this requirement is not rigorously enforced. Courts sometimes grant patent owners rights to technology that is new, different, and distant from anything they actually made or possessed. Not surprisingly, this practice makes patent boundaries especially unclear in fast-paced fields such as biotechnology and computer software development. It certainly seems that E-Data was granted ownership of technology that was far removed from what Charles Freeny Jr. actually invented.

4. *The patent flood*: Clearance costs are affected by the number of prospective rights that must be checked for possible infringement. Investments in land or structures rarely involve many parcels of land, and property law discourages fragmentation of land rights. In contrast, investments in new technology often need to be checked against many patents<m->even thousands, in the case of e-commerce. Although the patent system has features that discourage patent proliferation<m->notably the requirement that an invention not be obvious<m->empirical evidence suggests these are not working well.

These differences mean that patents might diverge significantly from an ideal property system that grants an inventor a well-defined, exclusive right to develop a technology and bring it to market. Because of such differences, patents might not work well as property. Whether or not they do is the empirical question at the heart of this book.

Empirical Evidence: Do Patents Work As Property?

Do patents give inventors positive net incentives to invest in innovation? An answer requires careful attention to the details of the patent system and the markets for innovation. The empirical evidence must account for incentives from many sources, including exclusion of competitors, licensing, and sale of the patent. We must be careful to distinguish between patent-based incentives and other incentives to invest in innovation. We must also account for disincentives that arise indirectly from the threat of litigation. We study how the pattern of incentives varies over time, and across industries, technologies, and types of inventor.

Our question is simpler and more basic than the questions economists often ask when evaluating policy. Economists like to ask whether policies increase "net social welfare," a generalized measure of the overall well-being of society. Short of that, economists like to ask whether innovation policies increase innovation or R&D spending. But these are even more difficult and complicated questions to investigate empirically. Many interrelated factors can influence R&D spending, innovation, and the resulting social welfare, so it is difficult to disentangle these to determine the independent

influence of patents. Not surprisingly, economic studies that attempt to answer these more difficult questions typically have arrived at inconclusive results.

Our approach, instead, is to ask a more limited question. We can determine, with reasonable accuracy, whether or not patents provide net positive incentives for a given group of inventors. This does not tell us whether patent policy is optimal or not. To the extent that incentives are positive, we cannot tell whether they are too big or too small relative to the social optimum. There are many factors we cannot measure that go into a calculation of the optimal incentive. On the other hand, if patent incentives are negative, then they fail as property in a basic sense. In this case, patents do not do what they are supposed to do, and it is not likely that they will spur innovation and increase social welfare.[10] Even though society might receive benefits far beyond the innovator□s profit, if patents discourage innovators on net, then patents will not help realize these benefits.

We begin our empirical analysis in chapter 4 by reviewing the literature. It has been almost fifty years since the empirical evidence on patents was last comprehensively reviewed. The reviewer, Fritz Machlup (1958), concluded that it was not possible to decide whether patents were good or bad policy instruments. In the interim, a wide variety of research has looked at the performance of patent systems.

We review this scholarship not to determine whether patents are good or bad policy instruments in general — the discussion above suggests that firm conclusions of this sort might be very difficult to reach. Rather, we simply ask whether a nation's patents seem to have a similar effect on its economic performance as do other property rights in that nation. If they do not, this suggests that the implementation differences between patents and other property rights might be significant. Even though economic

performance is ultimately influenced by *global* property rights, the contrast between a nation's patents and its other property should reveal important differences or similarities.

Specifically, the research we review includes:

- *Historical research on the Industrial Revolution*. Although property rights and markets fostered economic growth and innovation throughout Europe and the United States, patents played a much more limited role. In Britain, few major inventors received much benefit from patents, although in the United States more did benefit. More generally, however, countries without patents were just as innovative during this period as those that had patents.
- *Statistical studies that compare the performance of countries over time*. These studies use indices of the strength of property rights or the strength of patent rights to explain each country's economic-growth rate. Although general measures of property rights exhibit robust correlations with economic growth, measures of patent rights do not. Patents might still play an indirect role, however. Patent rights appear to be somewhat correlated with R&D spending, although this relationship exists only among more-developed countries, and it is not clear whether patents cause R&D or vice versa.
- *Studies of economic experiments*. These studies explore what happens when legal rights change. Some researchers have explored the role of property rights in the transition of former Soviet-bloc countries to market-based economies. Those countries that developed property institutions to support a robust market economy have experienced strong economic growth after an initial period of sharp decline. This success apparently depends, however, on the establishment of specific supportive institutions, including market-oriented legal systems, commercial banking,

regulatory infrastructure, and labor-market regulation. Countries that introduced private property and markets without developing these institutions have experienced persistently declining per capita income. By contrast, economic experiments that extended or strengthened *patent* rights do not seem to show clear evidence of increased innovation, except, perhaps, to a limited degree among the wealthiest nations.

- *Miscellaneous research*. Case studies present a convincing argument that patents are critical for investment in R&D in the pharmaceutical industry. On the other hand, survey evidence suggests that in most other industries, patents do not pose much of a barrier to imitation, and firms rely mainly on other means, such as lead-time advantages and trade secrecy, to obtain returns on their R&D investments. Moreover, several studies suggest that a moderate degree of competition might actually spur innovation.

In summary, patents do not work "just like property." While they do play some role in promoting innovation and economic growth, that role is limited and highly contingent compared to the role property rights normally play in promoting economic growth. The laws and institutions that implement property rights might be harder to get right for patents than for tangible property rights.

Nevertheless, patents might still work effectively, even if they have a more limited impact on economic growth than property rights for other assets. To arrive at a more definitive evaluation, we need to perform a careful accounting of the incentives and

disincentives for investing in innovation that patents provide. We do this by drawing on estimates found in the literature, and on some of our own estimates.

Figure 1.1, copied from chapter 6, conveys the basic calculations we make for United States public firms. Because chemical patents, especially on pharmaceuticals, are much more valuable and much less likely to be litigated, we display the chemical and pharmaceutical industries separately from other industries. The heavy, solid lines show the annual aggregate costs to these firms of defending against patent litigation. This estimate includes not only direct legal costs of litigation, but also business costs such as loss of market share or the costs of management distraction. The dashed lines represent an estimate of the incremental annual profit flow from all patents worldwide associated with inventions patented in the United States. We derive these estimates, based on a review of over twenty research papers, in chapters 5 and 6.[11]

{Figure 1.1 in back}

The incremental profit flows represent the gross positive incentives provided to innovators by worldwide patents above and beyond the profits that could be earned without patents. Litigation costs represent an important disincentive to innovation. A firm looking to invest in innovation will consider the risk that the innovation will inadvertently expose it to a patent-infringement lawsuit. Since infringement lawsuits are usually filed against firms exploiting new technologies, development of a new technology exposes the innovator to risk of inadvertent infringement if patent boundaries are hidden, unclear or unpredictable.[12] That risk weighs against the profits that can be made from innovation. Of course, firms are both patent holders and potential defendants, so a

comparison of profit flows and litigation costs for a group of firms should reveal the sign of net incentives.[13]

The results in figure 1.1 show that chemical and pharmaceutical firms earn far more from their patents than they lose to litigation. But for other firms, figure 1.1B tells a simple but dramatic story: during the 1980s, these firms might have, at best, broken even from patents, but in the mid-1990s litigation costs exploded. By almost any interpretation, the United States patent system could not be providing overall positive incentives for these United States public firms by the end of the 1990s. The risk of patent litigation that firms faced in their capacity as technology adopters simply outstripped the profits that they made by virtue of owning patents. A firm looking to invest in an innovative technology during the late 1990s, taking this risk into account, would expect the net impact of patents to reduce the profits from innovation rather than to increase them. Moreover, preliminary data for more recent years suggest that this problem has gotten worse since 1999.

Note that patents *do* provide profits for their owners, so it makes sense for firms to get them. But taking the effect of *other* owners' patents into account, including the risk of litigation, the average public firm outside the chemical and pharmaceutical industries would be better off if patents did not exist.

Moreover, figure 1.1 understates the extent to which costs exceeded benefits for several reasons: disputes settled before a lawsuit was filed are not counted, nor are foreign disputes; this comparison ignores the costs of obtaining patents and clearance; and for a variety of reasons, the estimates of worldwide patent profits are biased upwards, while the estimates of litigation costs are biased downwards.

The patent system clearly provides large positive incentives for innovators in the chemical and pharmaceutical industries. Also, small firms generally receive benefits that exceed costs, but the net incentives for these patentees are not large. We will comment more on these exceptions below.

Interpreting the Results

To understand the meaning of the evidence in figure 1.1, we explore several related issues. First, what is driving this surge in the cost of litigation? In chapter 7 we look at a variety of alternative explanations. The increase in aggregate litigation cost is mainly driven by the increasing frequency of litigation, which has roughly tripled since the 1980s.[14] Yet when we look in detail at what determines the rate of litigation, we find that only a small part of this increase can be explained by measurable factors, such as how much the parties to a lawsuit spend on R&D or how many patents they have. This suggests that most of the increase arises from unmeasured factors that might include legal, institutional, and technological changes. We explore several possible factors, including deterioration of patent notice, industry-specific factors, greater rewards for litigation, a general increase in litigiousness, the rise of patent "trolls," and the declining quality of patent examination.

We can directly rule out several of these explanations. All industries appear to have experienced a rapid increase in patent litigation, although the increase seems somewhat more rapid in software-related industries. This means that industry-specific factors are unlikely to explain most of what is happening. Also, business-to-business litigation has not been increasing in general, so we cannot attribute the increase in patent litigation to

an overall rise in litigiousness. In addition, we find no evidence to suggest that the rewards that patent holders gain from litigation increased in the 1990s, although they might have increased during the 1980s.[15]

We also considered the role of patent "trolls," which we define narrowly as individual inventors who do not commercialize or manufacture their inventions. One story claims that the increasing availability of patent litigators willing to work on contingency fees has spurred lawsuits by such trolls, who might otherwise be unable to afford litigation. The share of lawsuits initiated by public firms has not declined, however, nor has the share of lawsuits involving patents awarded to independent inventors increased. This suggests that the increase in litigation cannot be mainly attributed to patent "trolls," at least through 1999. Of course, if we use a broader definition of "troll" that includes all sorts of patentees who opportunistically take advantage of poor patent notice to assert patents against unsuspecting firms, then troll-like behavior might be a more important explanation. Indeed, if patent notice is poor, then all sorts of patent owners might quite reasonably assert patents more broadly than they deserve. But then it is more appropriate to attribute the surge in litigation to poor patent notice, not to trolls per se.

In fact, the distinctive pattern of litigation over time and across technologies does provide support for an explanation based on the deterioration of patent notice. Several changes to patent notice occurred during the mid-1990s, including the way that courts interpret patent claims, increased "hiding" of patent claims while applications are under review at the Patent Office, problematic legal decisions in software and biotechnology that extended the reach of patents to technologies far beyond what was actually invented,

and the growing flood of patents that began during the mid-1980s and gathered strength during the 1990s. Many of these changes are the specific work of the Court of Appeals for the Federal Circuit, the specialized appeals court for patents that was established in 1982. In any case, these changes provide a natural explanation for the concurrent increase in litigation.

In addition, the pattern of litigation costs across technologies is consistent with differences in patent notice. Litigation costs are particularly low for patents on chemical compounds, including pharmaceuticals. At the same time, the value of these patents is much higher than the value of other patents<m->perhaps, in part, because litigation costs are low and enforcement is effective. Economists have long recognized that patents on chemicals work particularly well because these patents have very well-defined boundaries.[16] In contrast, economists recognize that complex technologies, such as electronics and computers, might have relatively poorly defined patent boundaries. Patents on complex technologies have higher litigation rates and lower values than chemical patents. By the late 1990s, these patents generated more litigation cost than profit.

Software patents, in particular, often have boundaries that are especially difficult to determine, for reasons we explore further in chapter 9. Software patents have even higher litigation rates and a high frequency of appeals over the meaning of patent claims. Not surprisingly, the costs of litigation for software patents far exceed the profits. The distinctive pattern of litigation rates across technologies thus supports the notion that patent notice might explain differences in patent value.

The deterioration of patent notice might also be roughly associated with a decline in patent quality, broadly conceived. Many critics equate low patent quality with frequent issuance of invalid patents. These critics contend that poor examination allows invalid patents to be issued for inventions that are obvious or lack novelty. Specifically, they assert that inadequate search of previous patents and publications causes examiners to overlook novelty and obviousness problems. Other critics attribute patent-quality decline in part to the Federal Circuit's proclivity to weaken the legal test of obviousness. These two sources of patent-quality decline contribute to the patent flood and make clearance difficult and costly, leading indirectly to litigation. Our broader conception of patent quality acknowledges problems with novelty and obviousness, but our evidence shows that quality problems are more fundamentally connected to problematic boundaries associated with patents that are vaguely worded, overly abstract, of uncertain scope, or that contain strategically hidden claims.

The narrow conception of patent-quality decline does not explain the surge of patent litigation or the pattern of litigation across technologies. Perhaps there has been a recent surge of invalid patents granted, but no such surge appears in the data on litigation outcomes. Similarly, invalidation rates are not higher for technologies featuring higher litigation rates. This suggests that patent-examination search quality is not primarily responsible for the increase in costly litigation by itself, although it might well be a contributing factor and it might also be a problem for other reasons.

This analysis leads us to the conclusion that during the late 1990s, the American patent system failed as a system of property rights for public firms. It did so because it failed to provide clear and efficient notice of the boundaries of the rights granted.

Small Inventors

But this evidence of failure applies only to one group of inventors — namely, those at public firms. Now this is a large and very important group of inventors, especially if we assume that the main purpose of patents is to provide incentives to invest in R&D — this group of firms is responsible for about 90 percent of R&D spending. Nevertheless, some important inventions are made by small inventors, including independent individuals and small nonpublic firms. Perhaps the patent system works sufficiently well for these small inventors to offset its other failures. Indeed, some people claim that almost all "breakthrough" inventions come from small inventors, and their interests should be paramount in debates about patent reform.

Chapter 8 explores several issues regarding small-inventor patents. There are good reasons to think that small inventors make important inventions. This is not true of all types of small inventors, of course; many small inventors patent games, simple machines, and other low-tech inventions. Nevertheless, many small inventors do make important high-tech inventions. But there is no evidence to suggest that *most* breakthrough inventions come from small inventors. What limited evidence exists — for example, the characteristics of inventors nominated to the National Inventors Hall of Fame — suggests that most recent major inventions originated in large organizations, although a significant minority of important inventions are developed by independent inventors or inventors working in small firms.

How does the patent system perform for small inventors? In our analysis of public firms, we find that small public firms enjoy positive incentives from patents — their

litigation costs are lower than the profits they receive from patents, although their absolute level of profits from patents is not large. Other small inventors are also likely to enjoy a positive incentive, but we lack the data to estimate their litigation costs. Certainly, the many small inventors who do not commercialize any technology have little to fear as defendants in patent lawsuits. Even so, we find troubling evidence that patent notice adversely affects small inventors, too. Patent-notice problems impair the market for technology and rob many small inventors of the larger reward they could get from licensing or selling their patents in a world with good patent notice.

The troubling evidence is this: all types of small inventors, including small firms, realize substantially less value from their patents than do large firms. This is true for the independent inventors who work in low-tech fields, as well as for small public firms in many high-tech industries. Indeed, relative to large firms, many small inventors, even small high-tech firms that go public, forgo patents entirely, relying instead on trade secrecy and other means of protecting their profits deriving from innovations. The patent system works for small inventors, but does so only weakly. Why? In part because small inventors do not have access to the resources needed to commercialize inventions. They cannot quickly ramp up manufacturing and marketing, they do not have established distribution channels, and they cannot easily finance acquisition of these assets.

Lack of such complementary assets would not be a problem if markets for technology worked better. Small inventors who lack resources should be able to sell or license their technology to large firms who have those resources. Indeed, technology markets are often the best means that small inventors can employ to capture value from their inventions. In a world with competitive and efficient technology markets, licensing

royalties and sales contracts would deliver value to small inventors comparable to the value that large firms gain from their own patents. The fact that small inventors actually gain much less from their patents, however, indicates that these markets do not always work very well.

Better patent notice would improve technology markets in two ways. A direct improvement flows from clearer property rights. Unclear property rights increase bargaining costs and the probability of bargaining breakdown. Better patent notice makes technology markets more efficient, and hence more attractive to small inventors.

An indirect and possibly larger benefit flows from the impact of notice on buyers in technology markets. When potential buyers face substantial risk of patent litigation, they cannot profit as much from the technology they seek to exploit and are therefore unwilling to pay as much for the technology. Better notice would reduce the risk of inadvertent infringement and any ensuing litigation, increase the willingness of buyers to pay for technology, and increase the value of patents to small firms who sell in technology markets.

Small inventors and large firms alike suffer from poor patent notice.[17] The positive incentives that small inventors receive from patents give us no reason to be sanguine about the current state of the patent system.

The Particular Problem of Software Patents

As we noted above, the patent system performs particularly poorly for software patents. Software is a vital technology and, as we shall see, software patents contribute

substantially to the overall failure of the patent system for public firms. We explore the reasons for this in chapter 9.

Software patents have been controversial in part because the software-publishing industry grew up largely without patents and most computer professionals oppose patenting software. But judicial decisions during the 1990s eliminated certain obstacles to software patents, and now close to 200,000 software patents have been granted.

Some argue that there is nothing different about patents on software, and if there are any problems, these will be resolved as the Patent Office adapts to this new technology. Some say that because the software-publishing industry remains innovative, patents cannot be hurting innovation. But evidence about the software-publishing industry is not definitive; the majority of software firms still do not obtain patents, and most software patents are awarded to firms in other industries, chiefly the manufacture of computers, semiconductors, and electronics.

Critically, software patents *do* seem to exhibit some marked differences from other patents when it comes to litigation costs. Software patents are more than twice as likely to be litigated as other patents; patents on methods of doing business, which are largely software patents, are nearly *seven* times more likely to be litigated. And, despite being a relatively new area for patenting, software patents accounted for 38 percent of the total cost of patent litigation to public firms during the late 1990s. This does not appear to be a temporary problem that is dissipating as the Patent Office adapts<m->the probability that a software patent will be litigated has been *increasing* substantially rather than decreasing.

Why are software patents more frequently litigated? In a word, abstraction. In chapter 9, we will elaborate upon what we mean by "abstraction" and how it affects patent notice, but for the present consider that software is an abstract technology, and this sometimes makes it more difficult, if not impossible, to relate the words that describe patent boundaries to actual technologies. In this context, it is helpful to recall the abstract concepts described in the claims of the E-Data patent — "point of sale location," "information manufacturing machine," "material object," and so on. These words reach far beyond the actual kiosk technology that Charles Freeny Jr. invented, yet during the course of litigation it was not clear what, precisely, they covered. In other cases, the words in some broad software patents seem clear enough, but because the patents claim technology far beyond what was actually invented, judges will sometimes interpret the claims narrowly (for example, see the discussion of Wang's patent 4,751,669 in chapter 9). But it is hard to predict which broad claims will be narrowed. This becomes another cause of boundary unpredictability that contributes to inadvertent infringement and, ultimately, to litigation.

Of course, software is not the only technology that can be described in abstract terms. Indeed, the problem of abstraction in patents has been recognized at least since the eighteenth century, when British law attempted to exclude patents on general "principles of manufacture" as opposed to specific inventions. In the United States, judges also developed doctrines to exclude patents with abstract claims during the nineteenth century.

Nevertheless, there are two major reasons why abstraction poses a particular problem for software. First, as we will discuss in chapter 9, the Court of Appeals for the Federal Circuit has tolerated more abstraction in software patents than seems warranted

by these patent doctrines. Second, software is inherently more abstract than other technologies. Indeed, it is well known among computer scientists that software technologies (algorithms, system structures) can be represented in many different ways, and it might be difficult to know when alternative representations are equivalent. This means that the technology claimed in a patent can be difficult to distinguish from alternatives; it might be hard to know whether a given patent protects an invention that is different from previous inventions, or whether an allegedly infringing program is different from the claimed technology. If computer scientists cannot unambiguously make these distinctions, there is little hope that judges and juries can do better.

Although not all software patents suffer from abstract or overly broad claims, software technology is especially prone to these problems. Indeed, software patents are much more likely than other patents to have their claim construction reviewed on appeal<m->an implicit indication that parties to lawsuits have fundamental uncertainty over the boundaries of these patents. This uncertainty leads to more frequent litigation and substantially higher litigation costs.

Software patents are not just like other patents. The problems of software patents — problems arising partly from the nature of the technology and partly from the way the courts have treated this technology — are a substantial factor in the overall poor performance of the patent system. The problem of implementing patent law to deal with abstract patents appears to be particularly stubborn and is unlikely to go away unless it is addressed directly.

Making Patents Work Like Property

What, then, will it take to fix the failure of patent notice and make the patent system an effective tool for encouraging innovation in all industries? At first glance, this might not seem too difficult a task, given that patents seem to have performed reasonably well as recently as the 1980s. Indeed, many people have been quite optimistic that the current round of draft legislation and recently renewed attention from the Supreme Court will soon lead to a rebirth of effective patent policy.

In chapter 10, however, we suggest that effective reform might well prove surprisingly difficult to achieve. Many reform advocates believe that the poor performance of the patent system flows from deterioration of patent "quality" (narrowly defined) that can be fixed by improving the patent-examination process. We agree that invalid patents are a problem, and that patent examination can be improved; however, we see this as only *part* of the problem. We suspect that many people focus on patent quality because there has been so much publicity about bad patents on inventions that lack novelty or seem obvious, such as those mentioned previously involving the peanut butter and jelly sandwich or the backyard swing. Patents of doubtful validity create social costs, but our evidence suggests that concerns about validity are not the main drivers of the patent-litigation "explosion."

Moreover, we think that attempts to improve patent quality, including review procedures involving third parties, will not be very effective unless there are broader improvements in patent notice. This is because patent examination *depends* on clear, predictable patent boundaries. For example, critics of the E-Data ruling contend that e-commerce had been discussed and practiced before Freeny's invention. Under patent law this "prior art" should have invalidated the E-Data patent. But if the patent examiner, like

almost everyone else, interpreted the patent narrowly in 1983 as claiming only in-store kiosks and vending machines, not e-commerce, then that prior art would not have seemed relevant. Thus, patent quality depends on well-defined patent notice, which involves much more than simply improving the examiners' access to prior art.

Finally, improving patent notice will be challenging because it cannot be achieved merely by a few court decisions or statutory changes; rather, it requires changing institutions. As we discussed above, the institutions of the patent system fail to perform basic functions required for notice — functions other property systems perform smoothly. Indeed, the institutions of the patent system actually seem to have contributed to the deterioration of notice over the last two decades.

In particular, the structure of the courts — specifically, the designation of a single court for patent appeals — appears to have undermined notice in at least two ways. First, a specialized court is more likely than a typical appellate court to take actions to expand its influence. This seems to have been the case with the changes in the interpretation of patent claims. The Federal Circuit downgraded the role of the Patent Office and the district courts in claim interpretation while increasing its own. We will show that this shift has decreased the predictability of patent boundaries. The Federal Circuit has also increased its influence by expanding the range of patentable subject matter to include software, business methods, early-stage inventions, and more.[18] Increased patenting of these new technologies might have created problems because of a second institutional shortcoming: a single appellate court might not be well suited for developing new law. Because power is concentrated in the Federal Circuit, patent law misses the benefits of the inter-circuit competition that exists in most other areas of federal law.

We thus think it likely that effective reform will require structural changes, including, possibly, multiple appellate courts, specialized district courts, and greater deference to fact-finders. What other changes might improve patent notice? In Chapter 11 we consider many reforms, most of which have also been advanced by others who have preceded us. These include:

- *Make patent claims transparent*. We recommend changes in the way patent claims are defined, published, recorded in the application process, and used for subsequent determinations so that innovators have clear, accessible, and predictable information on patent boundaries. This includes strong limits on patent "continuations," a procedure used to keep patent claims hidden from the public for extended periods. We also consider a new role for the Patent Office where, for a fee, innovators can obtain opinion letters on whether their technology infringes a patent.
- *Make claims clear and unambiguous by enforcing strong limits against vague or overly abstract claims*. This includes a robust "indefiniteness" standard that invalidates patent claims that can be plausibly interpreted in multiple, fundamentally different ways. Also, we recommend reforms to limit overly abstract patents in software and other technologies. At the very least, patent law should prevent software patents from claiming technologies far beyond what was actually disclosed as the patented invention. If this proves inadequate, then we suggest subject-matter tests to limit the range of software inventions that can be patented, tests similar to those used during the 1970s and 1980s.
- *Make patent search feasible by reducing the flood of patents*. This includes a strong requirement that patents should not be granted on obvious inventions, coupled with

substantially higher renewal fees. Ideally, patent-renewal fees should be set by a quasi-independent agency and should be based on empirical economic research. These reforms will help stem the patent flood by screening out unwarranted patents and discouraging renewal of low-value patents. Reducing the number of such patents should help notice by reducing the cost of clearance search.

- Besides improving notice, we also favor reforms to mitigate the harm caused by poor notice. These include an exemption from penalties when the infringing technology was independently invented, as well as changes in patent remedies that might discourage opportunistic lawsuits.

In presenting these policy suggestions, we admit that we cannot know with any certainty what it will take to substantially improve patent notice. These policy reforms move us in the direction of an effective patent system, but we cannot as yet tell whether they are sufficient to get us there.

Some have argued strongly that our policy prescriptions will not suffice. Economists Boldrin and Levine (2007) argue that the patent system does not work at all and should be abolished. We doubt that such an extreme move is warranted. Our evidence suggests that the patent system *does* provide positive innovation incentives for small inventors, and, on a larger scale, for chemical and pharmaceutical inventions. It seems likely that reform can improve notice and overall patent performance in some areas, especially since the patent system did provide positive innovation incentives as recently as the 1980s.[19]

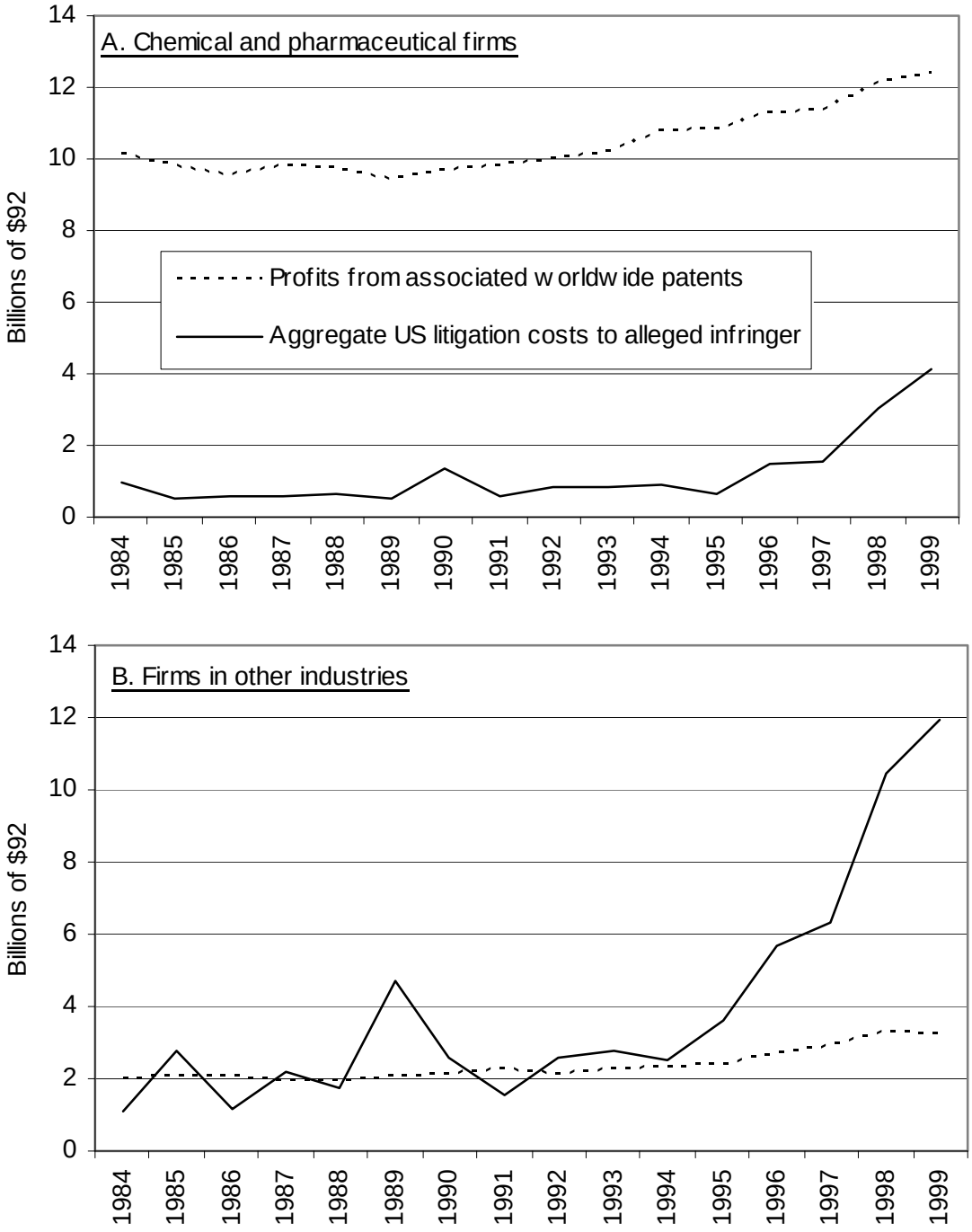
On the other hand, we are troubled by the expansionist view of the courts that "everything under the sun made by man" should be patentable, including software, business methods, and even mental correlations. Tangible property systems are not so expansive. They restrict property to assets that can be clearly defined with unambiguous boundaries. A landowner gets no rights to untapped oil flowing beneath her land, or to migratory ducks that set down on it, or to the airplanes that fly over it. Similarly, we doubt that all types of inventive ideas can have clear boundaries, and our empirical evidence shows that many software and business-method patents fail to provide efficient notice. We are quite sure that the patent system needs to recognize the limits to its grasp, even if we are not sure of the best way to implement those limits.

Perhaps many of these reforms are not politically feasible today. Perhaps the political will to thoroughly fix patent notice does not yet exist. The patent bar has long dominated patent policymaking, and its interests — at least in the short run — do not always coincide with improved notice.

Yet there is some reason to think that this impasse is temporary and that some of the patent bar's opposition to improved notice will prove to be shortsighted. Our estimates suggest that the litigation burden imposed by patents is growing, and the performance of the patent system will continue to deteriorate. Moreover, the trends suggest that the deterioration might be particularly bad for software patents and other patents used in information technology (IT) industries — not only is the rate of litigation per software patent rising, but the share of software patents out of total patents continues to grow rapidly. It is no accident that computer, semiconductor, software, Internet, and finance companies have begun to lobby for patent reform.

If this prognosis is correct, then the political landscape will continue to change. In the end, the survival of the patent system will require major improvements in the notice function. Despite all the rhetoric calling for "protection of inventors' property," today's patents fail as property, and tomorrow's might yet do even worse. Too often, such rhetoric is used to justify policies that actually undermine the property nature of patents. We hope our message and our empirical evidence succeeds in distinguishing actual patent performance from rhetoric. But in the long run, the pressures of market competition will determine the fate of the patent system based on its performance. If patents fail to work as property, over time they will make the United States economically less competitive, and industry will demand change.

Figures 1.1A and 1.1B



1. BBC News, "Starvation strikes Zimbabwe," May 3, 2002, available online at <URL>http://news.bbc.co.uk/1/hi/world/africa/1966365.stm</URL>; The Heritage Foundation, Index of Economic Freedom, "Zimbabwe," available online at <URL>http://www.heritage.org/index/country.cfm?id=Zimbabwe</URL>.

[2]. E-Data sent out 75,000 letters to website developers offering "amnesty" from lawsuits for a fee.

3. Clifton Leaf and Doris Burke, "The Law of Unintended Consequences," *Fortune* 152, no. 6 (September 19, 2005).

4. Boldrin and Levine (2007)

5. Jaffe and Lerner (2004).

6. Frank Hayes, "Patents Pending," *Computerworld*, May 2, 2005, available online at <URL>http://www.computerworld.com/governmenttopics/government/legalissues/story/0,10801,101434,00.html</URL>.

7. Macleod et al. (2002) find that a large portion of nineteenth-century steam engine patents were technically unviable. They go on to quote inventor Richard Roberts: "Our patent list now contains a great number of very silly things, which no man, who had been long in a workshop, would ever think of patenting; and the reason is that the patentee has money, though deficient in experience and mechanical talent; probably he thinks he cuts a figure by being in the patent list."

8. In fact, we suspect that the issues we raise in this book about patent notice are highly relevant to the comparison of patents to other forms of innovation incentives and to many other theoretical inquiries as well.

[9]. David Streitfeld, *Los Angeles Times*, February 8, 2003.

10. Some people argue that a major benefit of patents is that they disseminate information. The limited evidence available makes us skeptical of this claim, and it certainly seems unlikely this benefit could be large enough to justify a patent system that imposes a net tax on innovators.

11. The profit estimates are based on estimates of patent value (chapter 5) multiplied by a rate of return. The estimated profits are the value of the aggregate stock of patents held by public firms multiplied by the annual discount rate (the rate-of-return "hurdle" required to justify an investment compared to alternative investments). We use a discount rate of 15 percent, net of depreciation. We use estimates of the value of United States patents that economists have obtained using well-established techniques based on patent-renewal behavior (decisions to pay maintenance fees reveal the actual value patentees place on patents). We also draw on several studies of the stock market value of firms to obtain estimates of their worldwide patent values (investors valuations of firms reveal the value of firm assets including patents). In chapter 6, we use stock market–event studies to estimate the total business cost of litigation. This, too, is an established technique that we have employed on a large scale<m->some 2,460 filings of lawsuits<m->to obtain an aggregate cost of litigation for public firms.

[12]. For this reason, patent-infringement risk is not a general cost of doing business, but is specifically related to innovative activity. In fact, the risk of being sued increases

with a firm's R&D spending. Of course, some lawsuits are filed against copyists, not inadvertent infringers. In chapter 6 we argue that most costly litigation is associated with inadvertent infringement rather than piracy.

[13]. Note that some, but not all, of the costs of litigation show up as profits for the firm holding the patent. To the extent that litigation costs represent a transfer to the patent holder<m->as we shall see, this is not largely the case — our calculation already includes these profits in the profit flow from patents.

14. There is some evidence of a modest increase in the cost per lawsuit during the 1990s. Because we have used conservative assumptions, however, our calculations in figure 1 do not factor in this increase.

15. Although the cost of patent litigation to alleged infringers might have increased modestly during the 1990s, we do not have evidence that there was a corresponding increase in the rewards to patentee litigants. In general, we find that what alleged infringers lose from patent litigation does not substantially accrue to patentee litigants.

16. The specific nature of the pharmaceutical industry might also play a role, but inorganic chemicals also have low litigation rates, while biotechnology patents that are not simple chemical entities have high litigation rates.

17. Trolls and their patent suppliers do profit from poor notice, but we doubt that they account for a large share of small inventors.

18. To be sure, the Supreme Court also contributed to this expansion during the early 1980s.

19. Of course, even if the patent system provides positive incentives, some people argue that it should be replaced by alternative incentives for other reasons. See, for

instance, Hubbard and Love (2004a, 2004b) on pharmaceutical innovation. See also Wright (1983), Kremer (1998), and Shavell and van Ypersele (2001).