Empirical Evidence on the Validity of Litigated Patents¹

John R. Allison² & Mark A. Lemley³

The lack of empirical evidence on the function and impact of the patent system has long been lamented.⁴ In recent years, a number of scholars have begun to address this deficiency in two primary ways. First, several researchers have focused attention on how patents are perceived and used by firms in various economic sectors. The 1987 study by Levin, *et al.* is a prominent example.⁵ There, the authors surveyed a large number of high-level R & D executives in over one hundred industries to identify preferences among patents, secrecy, lead time, and other methods of protecting the competitive advantages of important new processes and products. Although significant inter-industry variances were found, across industries patents were not viewed as the most effective means of encouraging innovation. Indeed, in some industries patents were considered the least effective contributor to innovation.⁶

¹ Copyright 1998 John R. Allison & Mark A. Lemley.

² The Spence Centennial Professor of Business Administration, University of Texas Graduate School of Business.

³ Professor of Law, University of Texas School of Law; of counsel, Fish & Richardson P.C., Austin, Texas.

The authors would like to thank John Barton, Rose Hagan, Rose Marie Ham, Weston Headley, Paul Janicke, Josh Lerner, David McGowan, Peter Menell, Rob Merges, F.M. Scherer, and participants in the Stanford Workshop on Intellectual Property and Industry Structure for comments on an earlier draft; Ryan Garcia, Shari Heino, Brian Spross, and Wade Williams for research assistance; and Dr. Alex Chien for statistical expertise.

⁴ George Priest complained ten years ago that there was virtually no useful economic evidence on the impact of intellectual property. George Priest, *What Economists Can Tell Lawyers About Intellectual Property*, 8 RES. L. & ECON. 19 (1986). Fritz Machlup told Congress that economists had essentially no useful conclusions to draw on the nature of the patent system. See SENATE SUBCOMM. ON PATENTS, TRADEMARKS, AND COPYRIGHTS, SENATE COMM. ON THE JUDICIARY, 85TH CONG., AN ECONOMIC REVIEW OF THE PATENT SYSTEM 55 (Comm. Print 1958) (prepared by Fritz Machlup). For some of the disagreements among historians over the impact of the patent system on innovation, see ROBERT P. MERGES ET AL, INTELLECTUAL PROPERTY IN THE NEW TECHNOLOGICAL AGE 125-27 (1997).

These complaints may be unfair. As noted below, there is increasing attention in academic circles to the relationship between patents and innovation.

⁵ Richard C. Levin et al, *Appropriating the Returns from Industrial Research and Development*, 1987 BROOKINGS PAPERS ON ECON. ACTIVITY 783.

⁶ *Id.* at 798. Not surprisingly, patents were viewed as much less effective for processes than for products. *Id.* at

Several other authors have evaluated patent acquisition and licensing strategies in various industries through case studies.⁷ One especially interesting study of licensing by Josh Lerner in 1995 empirically examined the patenting behavior of 419 new biotechnology firms with varying litigation costs.⁸ One of Lerner's key findings was that firms with relatively higher litigation costs are less likely to seek patents in those subclasses⁹ in which there had been many patent

794-95. Among the most important reasons found by the authors for the perceived limitations on the effectiveness of patents for were the ease of inventing around both process and product patents and doubts about patentability in the case of processes. *Id.* at 803. Given their findings, the authors were led to question why firms patent so much, and at an increasing rate. Their research did not explore this question. However, a recent study by Cohen, *et al.*, directly addresses this question. See Wesley M. Cohen, *et al.*, *Appropriability Conditions and Why Firms Patent and Why They Do Not in the American Manufacturing Sector*, presented at the Stanford Workshop on Intellectual Property and Industry Competitive Standards, Stanford Law School, April 17-18, 1998. Before seeking to answer this question, the Cohen study updates that of Levin and finds that, across many manufacturing sectors, patents are viewed as substantially less effective for appropriating the value of product innovations than all other alternatives, secrecy and lead time being the most preferred alternatives. The study finds a number of reasons why firms nonetheless seek patents. Unsurprisingly, the most important reason given by respondents was to prevent others from copying. The authors recognize, however, that the importance of this reason could have been exaggerated because many respondents may have viewed this as the most "socially desirable response." The second most important reason was "blocking," or preventing other firms from patenting related technology. Blocking and related defensive motives may help explain the finding in our study that patent litigation commonly occurs long after the issuance of a patent. See *infra* notes ___-___ and accompanying text. For another test of possible explanations why firms increasingly seek patents, see Samuel Kortum & Josh Lerner, *Stronger Protection or Technological Revolution: What is Behind the Recent Surge in Patenting?* (N.B.E.R. working paper, Sept. 1997). Kortum & Lerner reject the explanation that changes in patent law have prompted the surge in patenting.

For an interesting variant on the industry patent study, which accounts for the extent firms value patents by examining renewal data and multi-country filings, see Jean O. Lanjouw *et al.*, *How to Count Patents and Value Intellectual Property: Uses of Patent Renewal and Application Data* (N.B.E.R. working paper, Sept. 1996). Cf. Mark Schankerman & Ariel Pakes, *Estimates of the Value of Patent Rights in European Countries During the Post-1950 Period*, 96 ECON. J. 1052 (1986) (attempting to value patents in Europe).

⁷ Rebecca S. Eisenberg, *Public Research and Private Development: Patents and Technology Transfer in Government-Sponsored Research*, 82 VA. L. REV. 1663 (1996); Josh Lerner & Robert P. Merges, *The Control of Technology Alliances: An Empirical Analysis of the Biotechnology Industry*, __ J. Indus. Econ. __ (forthcoming 1998); Robert P. Merges & Richard R. Nelson, *On the Complex Economics of Patent Scope*, 90 COLUM. L. REV. 839 (1990); Robert P. Merges, *Contracting Into Liability Rules: Institutions Supporting Transactions in Intellectual Property Rights*, 86 CAL. L. REV. 1293 (1996); David J. Teece, *Profiting From Technological Innovation: Implications for Integration, Collaboration, Licensing and Public Policy*, 15 RES. POL. 285 (1986).

⁸ Josh Lerner, *Patenting in the Shadow of Competition*, 38 J.L. & ECON. 463 (1995).

⁹ The reference to subclasses is to sublevels within the patent classification system maintained by the United States Patent & Trademark Office (PTO). There are over 120,000 subclasses. Lerner's finding means that firms with relatively high litigation costs are more likely to use litigation-avoidance patenting strategies.

damage awards to rivals, especially to those with lower litigation costs.¹⁰

A smaller number of empirical studies have been conducted on patent litigation. Two of these studies cover relatively early periods. The first, by P.J. Federico, provided validity and infringement data for litigated patents reported in the *United States Patent Quarterly (U.S.P.Q.)* during the years 1925-1954, with more in-depth study of patents litigated during the years 1948-1954. Although Federico did not attempt to examine a large number of variables, he did examine overall validity rates in a relatively thorough manner, and in the 1948-1954 portion of the study he explored the courts' treatment of uncited and cited prior art.¹¹ The other, first published by Koenig in 1974 and then updated through 1980, constitutes the most extensive set of data ever gathered on patent litigation.¹² Koenig collected all patent cases reported in the *U.S.P.Q.* in the years 1953-1978 to produce an array of descriptive statistics. She also selected a random sample of 150 patents from the years 1953-1967 for more in-depth study. In addition to finding that most courts held patents invalid, and noting the wide disparity of validity rates across regional circuits, she also found that obviousness (or "lack of invention") was the most frequently used basis for judicial invalidation of patents.¹³

¹⁰ Lerner, *supra* note <8>, at 478-79. Using a number of ingenious data collection and testing methods, particularly in estimating relative litigation costs, Lerner contributes not only to the literature on patenting strategy but also to the literature on the various effects that litigation and other dispute resolution costs have on firms' behavior.

¹¹ P.J. Federico, *Adjudicated Patents, 1948-54*, 38 J. PAT. OFF. SOC'Y 233 (1956). Federico found that courts upheld the validity of patents in only about 30-40% of the cases in which validity was an issue. *Id.* at 236. He also concluded that the prior art before the courts was often better than that used by the PTO in issuing the patent, based on his observation that accused infringers were generally more successful in convincing courts to invalidate patents on the basis of uncited prior art than on the basis of cited prior art. *Id.* at 249. Our data confirm this observation, though the explanation Federico offers is by no means the only one. See *infra* notes ___-___ and accompanying text.

¹² GLORIA K. KOENIG, *PATENT INVALIDITY: A STATISTICAL AND SUBSTANTIVE ANALYSIS* (rev. ed. 1980). Koenig also notes other early studies, but they appear to not add anything meaningful to the work of Federico and Koenig. KOENIG, *id.* at 3-4 to 3-10. Like Federico's data for 1925-1954, Koenig's data for 1953-1978 revealed that district and circuit courts found patents valid only about 35% of the time.

¹³ *Id.* at 5-70 to 5-78. Koenig also studied the many kinds of prior art relied on by courts, and the ways in which uncited prior art played a role in the courts' decisions. *Id.* at 5-25 to 5-69.

Of the relatively few recent contributions to the empirical literature on patent litigation, the work by Lanjouw and Lerner on injunctive relief in patent cases is notable.¹⁴ They evaluate a sample of 252 patent suits, testing the hypothesis that preliminary injunctive relief in patent litigation is used to impose costs on rivals.¹⁵ Coolley has also produced a useful empirical study of a purely descriptive nature on patent infringement damages.¹⁶ Lanjouw and Schankerman evaluate data provided by the PTO about litigated patents to determine the ways in which litigated patents differ from the general patent pool.¹⁷ At least two other studies¹⁸ have attempted, with mixed success, to empirically analyze the decision-making behavior of the

¹⁴ Jean O. Lanjouw & Josh Lerner, *Preliminary Injunctive Relief: Theory and Evidence From Patent Litigation* (N.B.E.R. working paper, July 1996).

¹⁵ *Id.* Lanjouw and Lerner find that their data is consistent with the hypothesis that preliminary injunctive relief is a predatory weapon in patent cases.

¹⁶ Ronald B. Coolley, *Overview and Statistical Study of the Law on Patent Damages*, 75 J. PAT. & TRADEMARK OFF. SOC'Y 515 (1993). This study analyzed several factors from 152 decisions between 1982-1992 in which the amount of damages was reported. Although unstated in the article, it appears that both district court and Federal Circuit decisions were included. The article also did not define the source of its data set, but apparently included decisions reported in West reporters, U.S.P.Q., and Lexis.

¹⁷ Jean O. Lanjouw & Mark Schankerman, *Stylized Facts of Patent Litigation: Value, Scope and Ownership* (working paper 1997).

¹⁸ Ronald B. Coolley, *What the Federal Circuit Has Done and How Often: Statistical Study of the CAFC Patent Decisions—1982-1988*, 71 J. PAT. & TRADEMARK OFF. SOC'Y 385 (1989); Donald R. Dunner, *et al.*, *A Statistical Look at the Federal Circuit's Patent Decisions: 1982-1994*, 5 FED. CIR. B.J. 151 (1995).

The Coolley study of Federal Circuit decision making is quite difficult to use as a basis for any type of conclusion. It did not identify the source of its data, did not attempt any kind of precise definition of its data set, and has a number of data comparability problems. Some of these problems stem from the inclusion of design patent decisions, decisions on appeal from all lower tribunals over which the Federal Circuit has appellate jurisdiction, and inclusion of all subjects of Federal Circuit decision and all types Federal Circuit judgments.

The study by Dunner, *et al.*, on the other hand, provides much more useful descriptive statistics. This research had the avowed objective of determining whether the Federal Circuit was “biased” in favor of patents. Specifically, Dunner examines whether the Federal Circuit is generally more pro-patent than its predecessor patent appeals courts, namely, the regional circuits and the Court of Customs and Patent Appeals (C.C.P.A.). Dunner's study was based on 1302 Federal Circuit decisions of all kinds, many unreported; the source of the data set is not clear. Although the study was based on a very large data set that may present data comparability problems, one portion of it did segregate Federal Circuit decisions on patent validity. Like other studies of the Federal Circuit, the Dunner study found a much higher validity rate than had been found in district court and regional court of appeals decisions prior to the Federal Circuit's creation. This was found to be true both overall and with respect to the individual grounds of novelty/statutory bars, obviousness, and description and claim adequacy. *Id.* at 154.

United States Court of Appeals for the Federal Circuit.¹⁹

Given that patent litigation tends to be exceptionally costly, with legal expenses often exceeding \$1 million per party, lawyers and clients should be eager for more systematic data and fewer anecdotes about how patents actually fare in the courts and why.²⁰ In the present study, our hope is to contribute both to the scholarly empirical literature on the patent system at work and to offer patent practitioners harder evidence to either confirm or contradict their assumptions about the courts' decision making in patent cases. To do this, we have produced a database of all written, final validity decisions by either district courts or the Federal Circuit reported in the *U.S.P.Q.* during an almost eight-year period from early 1989 through 1996.²¹ We use this database to develop many descriptive statistics about these patent validity decisions, test a number of hypotheses about what accounts for the success or failure of challenges to patent validity, and produce a regression model to determine the predictive power of our findings. Throughout the study we examine a greater number of variables than in any modern study of patent litigation.

¹⁹ Congress created the Federal Circuit in 1982 to serve, *inter alia*, as the only United States court of appeals to review district court patent cases. Although not relevant to our study, the same legislation also gave the newly created Federal Circuit appellate jurisdiction over appeals from decisions of the Patent and Trademark Office's Board of Patent Appeals and Interferences in instances where the Board had affirmed the patent examiner's rejection of a patent application. This latter form of appellate jurisdiction had previously been within the province of the C.C.P.A, the existence of which was extinguished by the 1982 legislation. Federal Courts Improvement Act of 1982, P.L. 97-164, 96 Stat. 25, 37 (codified as amended at 28 U.S.C. § 1295 (1994)).

²⁰ One illustration of the generally recognized high costs of patent litigation is found in *Blonder-Tongue Lab. v. University of Ill. Found.*, 402 U.S. 313 (1971). In *Blonder-Tongue* the Supreme Court adopted the doctrine of collateral estoppel for judicial decisions of patent invalidity, so that a finding of invalidity in one infringement case renders the patent invalid for all purposes and thus useless by the patent owner against other alleged infringers. *Id.* at 350. One of the Court's reasons for applying collateral estoppel in the stead of mutuality of estoppel was the extreme expense of patent litigation. *Id.* at 334-36. It bears mentioning that the Court's observation was made before the dramatic rise in such costs during recent years.

See also Lerner, *supra* note 7 at 470-71 (discussing abnormal direct and indirect costs of patent litigation).

²¹ The relevant volumes of *U.S.P.Q.* are 10 *U.S.P.Q.*2d through 40 *U.S.P.Q.*2d, inclusive. Although we did not extend the full study through 1997, we did include any 1997 Federal Circuit case reported in volumes 41-44 of *U.S.P.Q.*2d producing a final judgment on validity in a decision that had been in our population as a district court decision. Thus, in a small number of instances, we substituted 1997 Federal Circuit decisions for earlier district court decisions in the population.

Our study is structured as follows: Part I describes the data set, explains what information we collected and the methodology we have used, and identifies some of the limitations inherent in this data set. Part II contains descriptive statistics about the data set, our findings on each of the hypotheses we have tested, and the results of our regression model that attempts to predict the outcome of patent litigation based on the factors we have tested. Finally, in Part III, we draw some general conclusions about the universe of patent validity decisions, and suggest avenues for further research.

. Description of the Study

. **The Population**²²

Our defined population contains 299 patents litigated in 239 different cases. These cases represent of all written, final validity decisions by either district courts or the Federal Circuit reported in the *U.S.P.Q.* during an almost eight-year period from early 1989 through 1996. Several more detailed observations about this population and its limitations are discussed below.²³

- We have studied only validity decisions on issued United States patents. Our study does not include decisions on unenforceability due to misuse, inequitable conduct, estoppel or the like.

²² This is a population study rather than a sample study. The general parameters of that population are noted above, and we have included every case within this definition. There are several reasons for this decision. First, because we have defined a population in relatively narrow terms to ensure data comparability, the size of our data set is small enough that sampling is unnecessary. Second, when using reported cases as data sources, there are intractable problems with treating the grouping of cases as a representative, random sample regardless of how carefully one has defined the grouping. Although it is self-evident that any grouping of cases represents a subset of something larger (a population of something), it is practically impossible to assure one's self that the grouping is a representative sample, much less a random one.

Despite the use of a population, we were able to not only generate a large number of descriptive statistics but also to perform statistical testing. As discussed in more detail in Part I.C., hypothesis testing and prediction from a population can be done by treating it as a subset of a “superpopulation” -- in this case, the hypothetical population of all past and future validity decisions -- without any pretense that the data set is a representative sample of that superpopulation. See M.E. Thompson, *Superpopulation Models*, 9 **Encyclopedia Stat. Sci.** 93 (1988) (discussing the concept of a “superpopulation”).

²³ In this regard, it should be borne in mind that, although a few of our decisions about the population definition were due to practical constraints on data collection, most were made to ensure data comparability.

Nor does it include infringement decisions.²⁴ Because the population is limited to issued United States patents, it does not include appeals from the rejection of a patent application by the Board of Patent Appeals and Interferences of the Patent and Trademark Office (PTO).²⁵ Because it is limited to judicial decisions on the validity of U.S. patents, it does not include either the decisions of foreign courts or decisions of or appeals from the United States International Trade Commission (ITC).²⁶

- The population consists only of final decisions that resulted in written opinions. Where there is more than one decision in a case, we have reported the last final decision ruling on the validity of that patent. For example, final Federal Circuit decisions supersede trial court decisions in the population; in such a case, the "final" validity decision we have reported is the Federal Circuit ruling.²⁷ Where the Federal Circuit remands a case, and there is then a subsequent final, trial court decision on the validity of the patent reported in the specified volumes of the *U.S.P.Q.*, we have included the trial court rather than the Federal Circuit decision in the population. Where there is no final decision on a patent—for example, in

²⁴ We suspect most litigated patent cases include at least one validity argument, however.

²⁵ The inclusion of appeals from the PTO would introduce data into the population that are not comparable with data derived from infringement actions in district courts or the Federal Circuit. The proceedings, the burdens of proof, and the nature of the parties are entirely different. Three of the more important examples of data comparability problems are the purely *ex parte* nature of PTO Board proceedings, the absence of juries, and the absence of the strong presumption of validity that is present when a court considers a patent that has finally issued from the PTO.

²⁶ For many purposes one plausibly could characterize a Federal Circuit decision reviewing a patent validity determination in a United States International Trade Commission exclusion order as a judicial validity decision. We did not do so, however, because there are enough differences between ITC adjudication and district court litigation to have introduced data comparability problems into the study. The most important examples are the absence of jury trials in ITC adjudications and the extra requirements for relief (such as proof of domestic industry) in ITC actions. In any event, the number of these decisions found in, and then excluded from, our population was extremely small.

²⁷ We have also collected information on the trial court ruling, and use it to test a number of variables in the study. The case is only reported once in the population, however, even if it produced two or more opinions. We must note, however, that even though only the highest level final validity decision was included in the population, we did find it necessary in many instances to study earlier decisions in the same case to obtain more detail on matters such as the particular prior art references relied on in the final decision.

rulings denying summary judgment or motions to dismiss, or appeals that result in a remand for further decision making that is not subsequently reported—we *have not included the case in the population*. Similarly, if a case settles before a final decision on validity, it is not included (although cases that settle after a final decision that is not vacated remain in the population). Finally, while the population is defined to include final decision rendered from early 1989 through 1996, we have replaced district court rulings in this time frame with later final Federal Circuit decisions where appropriate.²⁸

- The population is defined in terms of patents, not cases. Many of the cases we studied produced final decisions on a number of different patents. Each of those patents is treated as a separate unit for purposes of our analysis, just as it would be under patent law.²⁹
- Our population consists of reported written decisions. This is broader than "published" decisions as that term is used in local court rules. The population includes cases denominated "not for publication" by the Federal Circuit, and district court opinions that are not included in the Federal Supplement. At the same time, it is not as comprehensive as "all decisions" in patent validity cases. Some decisions consist only of jury verdicts, or unwritten conclusions on the validity of a patent by a district court judge. We have excluded these decisions from the population, both because they do not produce intelligible data for most of the hypotheses we test, and because any effort to include all such cases necessarily would involve very unsystematic collection, producing only a haphazard subset of the cases and almost certainly introducing various instances of data incomparability.
- If we did our study several years from now, the population of validity decisions almost certainly would be much larger, although there is nothing to indicate that our results would

²⁸ See *supra* note --- for the details of this extension of our data set.

²⁹ Indeed, in one case, different *claims* of the patent produced different reported final decisions on validity. In that case, we have treated the two claim sets as two different patents for purposes of some of the hypotheses. That fact will be noted where appropriate.

differ.³⁰ The reason is that prior to the Supreme Court's 1993 decision in *Cardinal Chemical Co. v. Morton International, Inc.*,³¹ the Federal Circuit followed the practice of refusing to entertain validity issues once it had determined that no infringement had occurred.³² Moreover, in such a case, the Federal Circuit also vacated any decision on validity that had been rendered by the district court.³³ The Supreme Court overturned the practice, in part because the Federal Circuit is not a court of last resort and thus a finding of noninfringement does not moot validity issues that have been properly raised.³⁴ Since the Court's decision in *Cardinal Chemical*, therefore, validity should be an issue in practically every litigated patent dispute, and we would expect to see more final decisions on validity.³⁵

The Data Collected

For each patent in the population, we collected the following information, to the extent it was relevant and could be discerned from the court's opinion:³⁶

³⁰ The data set we use is more than large enough for valid statistical analysis.

³¹ 508 U.S. 83 (1993).

³² The Court specifically concluded that the Federal Circuit's practice had become very firmly engrained. *Id.* at 92 n.12.

³³ *Id.* at 91-92.

³⁴ *Id.* at 97.

³⁵ Technically, the Supreme Court held that the Federal Circuit's practice was improper only when the accused infringer sought a declaratory judgment of invalidity, either as a counterclaim to the patentee's infringement suit or as a declaratory judgment action initiated by the accused. The Court's ruling does not apply to a patent infringement suit in which the defendant raises invalidity solely as an affirmative defense. *Id.* at 93-94. As a practical matter, however, it seems that virtually all accused infringers will desire a finding of invalidity rather than only a finding of noninfringement, and will not limit themselves to raising invalidity only as a defense. The latter proposition is intuitive; although it is an empirical question, we have not investigated it empirically.

³⁶ In a small number of cases this was impossible. Some court opinions simply did not contain enough information to enable us to understand what issue was being litigated or on what basis. In those few cases, we have included as much information as we could discern. As a result, a few cases in the database include information about validity but not about the particular grounds, or a ruling on obviousness but not the nature of the prior art that was at issue.

In most cases in which deficiencies of this nature were found, however, we could solve the problem by

- Case name, citation and patent number.
- Whether the patent was finally held valid or invalid.³⁷
- Whether the finder of fact held the patent valid or invalid.
- Whether the finder of fact was a jury, a judge in a bench trial, a judge ruling on a pretrial motion, or a judge granting JMOL.³⁸
- Whether there was an appeal, and if so whether the fact finder's decision was affirmed or reversed.³⁹
- Whether the different claims in the patent were “disaggregated” (treated differently for validity purposes) by the court—either because the court found one set valid and the other invalid, or because it rendered a final ruling on one set but remanded the other set for a further decision.
- Whether the court decided litigated validity issues on each of the following grounds: enablement/written description,⁴⁰ claim indefiniteness, best mode, patentable subject

gathering more detail from earlier decisions in the same case. Thus, when we encountered a Federal Circuit case with inadequate detail on matters such as the precise pieces of prior art that were relied upon in a making a validity/invalidity decision, we usually were able to find the necessary detail in the district court’s opinion (or sometimes even in a preliminary decision by the district court). In the few instances in which a district court decision contained insufficient detail about its validity or invalidity decision, a preliminary decision by the court in the same case sometimes provided more detail.

As an aside, we are disturbed by the lack of information in some reported decisions in the federal courts. While the federal judiciary certainly owes no sensitivity to the needs of social science researchers, we would expect that Fed. R. Civ. P. 52(a) would compel a more detailed set of findings than is provided in some of these decisions.

³⁷ Strictly speaking, courts generally refer to patents as “invalid” or “not invalid” because an issued patent carries a strong presumption of validity that can be overcome by the challenger’s demonstration of invalidity by clear and convincing evidence. See, e.g., *Applied Materials, Inc. v. Advanced Semiconductor Materials Am., Inc.*, 98 F.3d 1563, 1569 (Fed. Cir. 1996). For convenience, however, we use the terms “invalid” and “valid” in the study.

³⁸ Where a judge *denies* a JMOL motion after a jury trial, we have treated the case as involving a decision by a jury.

³⁹ As noted above, if a Federal Circuit case produced no final decision on any aspect of validity but was remanded for further proceedings, the case is excluded from the population.

⁴⁰ We have treated these two grounds for decision together in the database.

matter, utility, section 102 prior art,⁴¹ section 102 non-art, obviousness, double-patenting, incorrect inventorship, and broadening of claims during reexamination. For each of these grounds of decision, we have indicated whether the finder of fact finally held the patent valid or invalid *on this ground*,⁴² and whether the final decision in the case held the patent valid or invalid on this ground.

- For obviousness and section 102 prior art cases, whether the decision was made primarily on the basis of cited prior art, uncited prior art, or both. In the case of obviousness, whether secondary considerations played a significant role in the decision.
- Whether the subject matter of the patented invention was general (also referred to as “mechanical”), electrical, chemical, pharmaceutical, biotechnology, computer-related, or software.⁴³
- Whether the inventor or the assignee of the patent reside outside the United States.
- The total number of prior art references cited on the face or inside the patent, and the breakdown between patent and non-patent references.
- For cases litigated on obviousness or section 102 prior art grounds, the number of references relied upon by the defendant, and the breakdown between cited and uncited prior art references.

⁴¹ We have divided section 102 decisions into two categories. The first, which we call “section 102 prior art” decisions, includes section 102(a) novelty determinations, section 102(b) determinations that depend on the prior art of another rather than the patentee’s own actions, and section 102(e) prior art. The second category, denominated “section 102 non-art” decisions, includes the statutory bars of section 102(b) to the extent they are triggered by the patentee’s own actions, abandonment under section 102(c), derivation under section 102(f), and priority decisions under section 102(g).

⁴² Because a patent need only be invalid on one ground to be invalidated, some courts may hold a patent invalid on, say, obviousness grounds while rejecting challenges under section 102 prior art and enablement. In such cases, we have listed the patent as finally invalid, but have listed recorded each of the subject matter decisions: e.g., invalid on obviousness grounds, valid on section 102 prior art grounds, and valid on enablement grounds. For this reason, the data set contains some rulings of “validity” on particular grounds even in invalid patents (though not the reverse). For the same reason, the number of validity decisions on each particular ground will not necessarily total to the number of final validity decisions.

⁴³ Note that a patent can fall into more than one category. For example, most software patents are also classed as “electrical” and “computer-related”.

- The amount of time the patent spent in prosecution before the PTO from first filing to issuance ("time in prosecution"), the amount of time elapsed between issuance and the final decision on validity ("delay"), and the total time from first filing to final decision on validity ("lag").

Methodology

We have used three different statistical approaches to evaluate the data we have produced. Because we have defined a population and included all the members of that population in our data set, the normal statistical tests designed to evaluate the statistical significance of the data do not apply. *Within the population*, all the numbers we reproduce below are by definition "statistically significant."⁴⁴ Thus, one type of data we produce are simply descriptive statistics about the population: what percentage of patents were invalid, how many patents were challenged on enablement grounds, etc. These descriptive statistics also include data relating one variable to another: comparing the percentage of jury decisions finding a patent valid with the percentage of bench trial decision finding a patent valid, for example. These statistics are interesting for what they reveal about the population of judicial validity decisions during a recent eight-year period. As a matter of pure statistics, however, they do not *predict* anything about future litigation.

Because one of our interests is in the predictive significance of the data we have collected, we also have evaluated the data set in a second way, as noted earlier. This approach defines our population as a subset of an indeterminate "superpopulation" consisting of final reported validity decisions across a range of time. We then apply the techniques of statistical inference to the population to test a number of hypotheses about the relationship between validity and other factors in the superpopulation. We detail the results in Part II; the full list of

⁴⁴ Unlike a sample study, all differences found in a population study are statistically significant. Such differences may nevertheless be small enough to have no practical significance to lawyers.

hypotheses is contained in Appendix A.

Finally, we combine a number of these hypotheses into a logistic regression model designed to determine how well the factors we have identified predict the outcome of patent cases.⁴⁵ We report the results of this predictive model in Part II as well; the full model is described in Appendix B.

Limitations of the Study

The study contains a number of limitations and potential biases. Some of these relate to our definition of the population, while others are inherent in the project of categorizing and interpreting data about litigation.

1. Population Biases

First, one might quibble with the population we have chosen. It would probably be desirable to evaluate *all* judicial patent validity decisions, whether or not they result in reported written opinions. Some might even suggest that written decisions are likely to have biases in their results, though it is not obvious to us what those biases would be. However, we doubt whether most of the data we collected for this study is in fact available for unwritten decisions. Further, we know of no reliable way to identify all final patent validity decisions in U.S. courts, and we suspect that the haphazard collection of some but not all decisions—or decisions from certain districts but not others, depending on the state of record-keeping—is at least as likely to produce biased results as the approach we take.⁴⁶

⁴⁵ We use logistic regression here because virtually all of the variables in the model are binary, and the rest can be converted to binary data. *See, e.g.*, AMIR D. ACZEL, COMPLETE BUSINESS STATISTICS 558-59 (1989) (discussing use of logistic regression model when the response to a set of independent variables is in the binary form, *i.e.*, “yes/no,” “success/failure,” etc.).

⁴⁶ Some limited data, such as fact finder and final validity decision, could perhaps be collected for such a study. If someone conducts such a study, it would be useful to compare those limited results with ours to determine whether such a bias in fact exists.

Although we explained the necessity of excluding from our population Federal Circuit appeals from PTO and ITC decisions, we do note that some types of cases are more likely to arise in appeals from the PTO than in normal litigation, and therefore may be underrepresented in our population. For example, the numerous patentable subject matter decisions involving software patents are largely absent from our population, because they usually arose in appeals from the PTO's refusal to issue a patent on section 101 grounds.

Finally, the fact that we have included only final decisions on validity has introduced some bias into our population. In particular, the percentage of decisions that are affirmed on appeal in this population is extraordinarily high. We suspect this is because we have excluded remands in cases that later settle or that are later decided without written opinion. Because many reversals by the Federal Circuit result in a remand, not an outright reversal, the "affirmed/reversed" statistic may be unreliable.⁴⁷

This does not mean that excluding remands was a mistake, however. Allowing nonfinal decisions to be included in the database would have introduced a host of other problems. It would have required us to further fragment the data set by measuring each case according to the procedural posture in which it arose. It would also have made the calculation of hard numbers of validity and invalidity impossible, and would have required the exercise of more judgment in categorizing cases and hence more possibility for error.

0. **Inherent Limitations**

We have already noted some of the problems that inhere in converting written opinions into hard numbers. In a few cases, it was impossible to discern enough information even to identify the issues being decided in a case. In other cases, determining how a case should be categorized required the exercise of judgment. For example, in characterizing the subject matter of an invention as "pharmaceutical" or "software," we necessarily had to make a judgment as to

⁴⁷ We discuss the ways we have sought to deal with this problem *infra* notes ___-___ and accompanying text.

what those categories should include. Similarly, some of the categories we have established—such as reliance by the courts on particular pieces of prior art, or on secondary considerations of nonobviousness—require analysis of the court's opinion. We are well aware that others might disagree with our judgment in any given case. However, we do not believe there is any reason to think that our evaluation of these cases is biased in any systematic way.⁴⁸ We have, of course, retained our complete data set and will make it available to those wishing to conduct similar research.

We are also limited in what we can test by the inherent nature of the litigation process. We suspect that there are numerous variables that affect patent validity decisions that are not included in this study for the simple reason that there is no obvious, rigorous way to test them. The skill of the lawyers on each side, the composition of the jury, the interests and experience of the judges, the demeanor of the witnesses, the financial resources of both parties, and the quality of the patent itself are all likely determinants of validity in at least some cases. They are not considered here, because there simply is not data available that would allow us to test these things. This does not mean that the data we have are problematic in any way; but it does mean that they cannot possibly tell the whole story.

Finally, a word of caution is in order regarding our hypothesis testing on the population to make predictions about the superpopulation. We conducted these tests because we suspect that practitioners and scholars are at least as interested in what is likely to happen prospectively over time than in what happened in a circumscribed historical time period, no matter how recent that historical period. However, even the best predictive efforts in this area encounter the fundamental limitation imposed by the fact that law and the litigation process change over time. A study that did not foresee the creation of the Federal Circuit in 1982 would not have given

⁴⁸ The authors worked closely together to personally study every decision and every patent; although we received various kinds of help from our research assistants and statistician, we did not delegate any of these tasks involving the exercise of judgment. Many other cases and patents were studied before being ultimately discarded for not meeting the strict parameters of our population. Moreover, in order to gather richer detail about matters such as the specific prior art relied upon by a court, it was sometimes necessary to study earlier decisions in the same case.

accurate predictions about patent validity. Similarly, efforts to predict the future from our data assume a static model of legal rules and systems which simply does not exist. The recent changes in the law of software patents may serve as a good example. There are remarkably few software patents in the population, but case law changes in this field beginning in 1994 are just beginning to have an effect on patent litigation, and the numbers may be much higher in the future. By its very nature, our population study cannot take these changes into account. All we can do is warn the reader to be aware of them.

The Results

We have evaluated a number of characteristics of these cases in order to identify possible predictors of patent validity. We will discuss each of these factors separately, and then in the context of the logistic regression model.

Overall Validity

Of the 300 final validity decisions in the data set,⁴⁹ 54% (or 162) found the patent valid, and 46% (or 138) found the patent invalid. These results are summarized in Figure 1.

⁴⁹ While there were 299 litigated patents reported, one of those patents produced two different final validity results for two different sets of claims. We have therefore counted that patent twice, as both a final valid and a final invalid result. In other cases where n=300, it is for the same reason.

This result is broadly consistent with other work on overall patent validity, which has generally found that courts litigating the validity of patents since creation of the Federal Circuit adjudge approximately 55% of them to be valid.⁵⁰

For predictive purposes, we tested
Hypothesis 1: Issued patents are not more likely to be held valid than invalid.

We set $H_0: q=0.5$ where q is the probability of final invalidity. The observed probability was $q=0.46$. The G-square p-value for the test was greater than 0.1, however, indicating that we cannot predict with confidence that patents in general are more likely to be held valid than invalid.⁵¹ From the perspective of an outsider to the patent system, it may be surprising that

⁵⁰ Donald R. Dunner, *The United States Court of Appeals for the Federal Circuit: Its First Three Years*, 13 AIPLA Q.J. 185, 186-87 (1985); Robert P. Merges, *Commercial Success and Patent Standards: Economic Perspectives on Innovation*, 76 CAL. L. REV. 803, 820-21 (1988); Mark A. Lemley, *An Empirical Study of the Twenty-Year Patent Term*, 22 AIPLA Q.J. 369 (1994) (finding 56% of all litigated patents held valid during the period 1989-1994). By contrast, before creation of the Federal Circuit, studies had found that only about 35% of litigated patents were held valid on average. The percentages were similar in the district courts and the courts of appeal. See KOENIG, *supra* note 11, at 4-18 to 4-19, 4-22 to 4-23. At the court of appeals level, validity rates varied widely among circuits, ranging from around 10% to over 55%. *Id.* at 4-32. Koenig's data covered the 1953-1978 period. See also Federico, *supra* note 10, at 244. Federico's study presented validity and infringement data for the years 1925-1954, although his more in-depth study covered only the 1948-1954 period. Koenig's validity data for 1953-1978 closely track those of Federico for 1925-1954.

⁵¹ The p-value is a measure of the confidence with which a hypothesis can be rejected. Hypotheses in our study are generally stated in the null form. The null hypothesis posits "no difference," or "no relationship." If the null hypothesis is rejected, then one cannot rule out the fact that there is a difference or a relationship. A rejection of the null hypothesis with a p-value of .01 means that such rejection can be made with 99% confidence; a rejection with a p-value of .05 or less means that such rejection can be made with 95% confidence; a rejection with a p-value of .10 or less means that such rejection can be made with 90% confidence. P-values less than .05 are viewed as an indication that the null hypothesis can be rejected with sufficient confidence and that any differences or relationships are statistically significant. Depending on several factors, p-values up to .10 may be viewed as supporting rejection with statistical significance. However, p-values greater than .10 for the null hypothesis are always viewed as showing that any differences or relationships are not statistically significant.

We have tested our hypotheses using both chi-square p-values (the "Pearson statistic") and G-square p-values (the "likelihood-ratio statistic"). Although both are reported in the results of our logistic regression analysis (Appendix B), in our discussion we refer only to the G-square p-value. The two are very similar and for most purposes interchangeable. Although there can occasionally be reasons for choosing between the two, the choice is more typically a matter of personal preference, as it was here. The important thing is to use either one consistently throughout a study. See THOMAS D. WICKENS, *MULTIWAY CONTINGENCY TABLES ANALYSIS FOR THE SOCIAL SCIENCES* 17-50 (1989) (thoroughly discussing tests of relatedness). For a discussion of the choice between the chi-square and G-square p-values, see *id.* at 38-39.

once a patent has been issued, the chance that a court will hold it valid is only slightly better than even.⁵²

Grounds for Attacking Validity

Of the 138 patents held invalid in the population, the majority of the grounds for invalidity were rooted in the prior art. Table 1A shows the breakdown of the findings of invalidity by grounds of invalidity.⁵³

⁵² This fact, while true, is somewhat misleading. First, it may still be the case that the PTO is effective in weeding out some invalid patents. For example, if only 25% of the patent applications objectively *should* result in valid patents, and if the PTO threw out 2/3 of the invalid applications, the validity rate of those patents that were issued would still be only 50%. Second, it may be that cases select themselves for litigation based on uncertainty as to outcome, and that cases involving obviously valid or obviously invalid patents are likely to settle. See ROBERT COOTER & THOMAS ULEN, LAW AND ECONOMICS __ (2d ed. 1996). This latter explanation has trouble accounting for the change in validity rates over time, however.

⁵³ Because a patent may be invalid on more than one ground, the percentages in Table 1A exceed 100%.

Table 1A

patentable matter	subject utility	enablement/ written description	indefiniteness	best mode	102 prior art	102-non- prior art	obviousness
0.72%	0.72%	15.22%	5.80%	11.59%	26.81%	31.16%	42.03%
1	1	21	8	16	37	43	58
double patenting	reexamination	incorrect inventorship					
3.62%	0.72%	0.72%					
5	1	1					

By far the largest number of invalidity determinations were made on the basis of obviousness (58, or 42% of all cases finding invalidity), comporting with the results of earlier studies. The second largest number were made on the basis of section 102 prior art (37, or 26.8%). Together, sections 102 and 103 accounted for 138 out of 191 total determinations of invalidity. Section 112 accounts for virtually all of the remaining grounds: 45 out of 191 determinations were made on the basis of enablement, written description, claim indefiniteness, or best mode. The remaining grounds for attacking validity resulted in very few invalidity findings.

We also tested the number of times particular theories were asserted in the population, as described in the court opinion, and the success rates of each theory discussed by the court.⁵⁴ The results for both inquiries are presented in Table 1B and Figure 1B.

⁵⁴ It is likely that in some cases, some of the invalidity arguments made by a party were not discussed at all by the court in rendering its decision. Those arguments are not included in the numbers that follow.

Table 1B

Likelihood of Success

Patentable Subject Matter

Decisions	Invalid
	100.00%
1	1
Utility	Invalid
Decisions	20.00%
5	1
Enablement	Invalid
Decisions	36.11%
36	13
Claim Indefiniteness	Invalid
Decisions	34.78%
23	8
Best Mode	Invalid
Decisions	35.56%
45	16
102 PA	Invalid
Decisions	40.66%
91	37
102 Non-PA	Invalid
Decisions	60.56%
71	43
Obviousness	Invalid
Decisions	32.50%
160	52
Double Patenting	Invalid
Decisions	71.43%
7	5
Reexamination	Invalid
Decisions	100.00%
1	1
Incorrect Inventor	Invalid
Decisions	10.00%
10	1

Figure 1B

The five most popular theories of invalidity asserted by defendants are obviousness (asserted in 160 out of 300 cases); 102 prior art (asserted in 91 out of 300 cases); 102 non-prior art (71 out of 300 cases); best mode (45 out of 300 cases); and enablement/written description (36 out of 300 cases).

The most successful theories on percentage terms are not necessarily those most often asserted. Some high success rates are simply due to the small number of instances—witness the 100% success rates of objections to patentable subject matter and broadening claims during reexamination, both of which result from having only one case in the population.⁵⁵ Among theories that were tried more often, the most successful objections to validity were double-patenting (successful 5 times out of 7, or 71.4% of the time), 102 non-prior art (successful 43 times out of 71, or 60.6% of the time), and 102 prior art (successful 37 times out of 91, or 40.7%

⁵⁵ Obviously, one should not conclude from this data that asserting such claims as a defendant is a sure route to invalidity.

of the time). The *least* successful theories were incorrect inventorship (successful 1 time out of 10, or 10% of the time), utility (successful 1 time out of 5, or 20% of the time), and obviousness (successful 52 times out of 160, or 32.5% of the time). It is perhaps instructive that the theory most often asserted against patents is also one of the least likely to succeed.⁵⁶

Judge vs. Jury

The role of jury trials in patent cases has increased dramatically over the past twenty years. Federal Judicial Center statistics indicate that in 1978, only 8.3% of all patent cases were tried to a jury, while in 1994 (the last year for which statistics are available) fully 70% of all patent trials were held before juries.⁵⁷ This represents a fundamental change in the nature of patent litigation. Apparently, somebody (presumably patentees) think trial by jury will benefit them.

The data bear out this assumption. There are significant differences between the validity rates reported by different triers of fact. Of the 300 patents litigated in the population, the validity of 73 (or 24.3%) was decided by a jury, 143 (or 47.7%) by a judge during a bench trial, and 82 (or 27.3%) during pretrial motions.⁵⁸ This distribution is depicted in Figure 2. It is worth noting that because of the population definition, these numbers may actually underestimate the number of patent validity cases tried to a jury. Jury verdicts are less likely than bench trials to result in reported written opinions, unless the opinion is written in the context of denying a post-trial motion.

The differences between the pretrial, bench and jury decisions is striking. Table 2 reflects the number and percentages of patents held valid by each trier of fact.

⁵⁶ On the other hand, perhaps the fact that it is so often asserted is responsible for the low success rate.

⁵⁷ HERBERT F. SCHWARTZ, PATENT LAW & PRACTICE 130 (2d ed. 1995).

⁵⁸ We have grouped together summary judgments and directed verdicts as "pretrial motions." Also, the validity of one case was decided on motion for JMOL.

Figure 2**Table 2**

	valid	Invalid
--	-------	---------

jury	valid	Invalid
------	-------	---------

24.33%	67.12%	32.88%
73	49	24

court bench trial		
-------------------	--	--

47.67%	57.34%	42.66%
143	82	61

court before trial		
--------------------	--	--

27.33%	28.05%	71.95%
82	23	59

More than two-thirds of the patents tried to juries were held valid (49 of 73, or 67.1%). By contrast, just over one quarter of the cases decided on pretrial motion were decided in favor of the patentee (23 of 82, or 28.1%). While some of this reflects the procedural nature of pretrial motions—the better cases are more likely to make it to trial, and the weaker cases to be dismissed early—that is not the whole explanation. Given the strong presumption of validity

that patents are given,⁵⁹ it is reasonable to expect that patentees would be entitled to pretrial judgment that their patent was valid with some frequency. Instead, it appears that pretrial rulings disposing of validity issues largely go in favor of the defendant, and jury verdicts in favor of the patentee. There is also a lesser, but still notable, difference between the 57.1% validity rate in bench trials and the 67.1% validity rate in jury trials.

Our test of this data for prediction about the superpopulation reveals the strong predictive significance of these findings. We tested *Hypothesis 2*: There is no difference between the likelihood that a patent tried to a jury, a patent tried to a court, and a patent resolved before trial will be held invalid *by the finder of fact*.

The hypothesis was rejected with great confidence (G-square p-value < 0.001). These numbers confirm the conventional wisdom that juries are likely to favor patentees, and unlikely to second-guess the decision of the PTO.

The differences between pretrial, bench and jury decisions also carry over to the *type* of invalidity argument each is likely to find persuasive. As Table 2A demonstrates, juries are largely receptive to arguments based on prior art: 62.5% of the jury findings of invalidity were premised on obviousness, and 37.5% were premised on section 102 prior art arguments. Only in one case (4.2%) did a jury find invalidity on the basis of enablement or written description violations.⁶⁰ By contrast, judges in bench trials were much more receptive to enablement and written description arguments: 19.7% of all invalidity determinations in bench trials were made on this basis. Judges in bench trials were somewhat less likely to find invalidity on the basis of prior art, however: only 42.6% of judicial findings of invalidity were made on the basis of obviousness, and 19.7% on the basis of section 102 prior art.

⁵⁹ 35 U.S.C. § 282 (1994) (presumption of validity). This presumption can be overcome only by clear and convincing evidence of invalidity. *See, e.g.,* Kahn v. GMC, 135 F.3d 1472, 1480, 45 U.S.P.Q.2d 1608 (Fed. Cir. 1998); Glaverbel Societe Anonyme v. Northlake Mktg. & Supply, Inc., 45 F.3d 1550, 1555, 33 U.S.P.Q.2d 1496, 1499 (Fed. Cir. 1995).

⁶⁰ As noted above, the percentages are larger than 100% total because many fact finders found invalidity on more than one ground.

Table 2A

Jury: Enablement/written description	claim indefinitenes s	Best mode	102 prior art	102 non- prior art	Obviousness
4.17%	8.33%	12.50%	37.50%	16.67%	62.50%
1	2	3	9	4	15
Bench Trial: Enablement/written description	claim indefinitenes s	Best mode	102 prior art	102 non- prior art	Obviousness
19.67%	8.20%	14.75%	19.67%	22.95%	42.62%
12	5	9	12	14	26
Court Pretrial: Enablement/written description	claim indefinitenes s	Best mode	102 prior art	102 non- prior art	obviousness
0.00%	1.69%	6.78%	25.42%	38.98%	22.03%
0	1	4	15	23	13

The rulings on pretrial motions differ significantly from both jury and bench trials. No court found a patent invalid on the basis of enablement on a pretrial motion, and only 22.0% of the courts' findings of invalidity pretrial were made on the basis of obviousness. The largest single ground of invalidity in pretrial decisions was on the basis of section 102 *non*-prior art: on sale bars and the like.

The predictive significance of these data for the superpopulation was uneven. In only two cases were the differences in findings by subject matter statistically significant: in the case of section 102 prior art, where

Hypothesis 2c: There is no difference between the likelihood that a patent tried to a jury, a patent tried to a court, and a patent resolved before trial will be held invalid on section 102 prior art grounds *by the finder of fact*.

was rejected (G-square p-value = 0.042); and in the case of section 102 non-prior art decisions, where

Hypothesis 2f: There is no difference between the likelihood that a patent tried to a jury, a patent tried to a court, and a patent resolved before trial will be held invalid on section 102 non-prior art grounds *by the finder of fact*.

was rejected (G-square p-value < 0.001).⁶¹

These differences have some important implications for patent validity litigation. Beyond the obvious—patentees generally want to request jury trials, and defendants want to try very hard to prevail before trial—there are also some indications of the types of arguments that are likely to be most persuasive in each forum. Juries are most receptive to prior art arguments, and respond to very little else. Judges in bench trials are somewhat less impressed by prior art arguments, but are more willing to hear enablement, written description, and a variety of other, more "technical" defenses. Finally, judges in pretrial motions are unlikely to invalidate a patent on the basis of obviousness or enablement, presumably because both generally require resolution of significant factual questions. They are more receptive to arguments of statutory bar.

Subject Matter of the Invention

We also tested the inventions litigated in the population by subject matter. We divided them into several categories. The first three categories—general, chemical, and electrical patents—are mutually exclusive, and track the classification scheme used by the PTO.⁶² In addition, we categorized certain patents as (1) pharmaceutical or (2) biotechnology-related (both overlapping subsets of the chemical group), or (3) software or (4) computer-related (both overlapping subsets of the electrical group) based on our own evaluation of the subject matter. The numbers and percentages of each type of patent in the population, as well as the validity statistics for each group, are reprinted in Table 3.

⁶¹ In one case—section 102 cited prior art—the hypothesis

Hypothesis 2d: There is no difference between the likelihood that a patent tried to a jury, a patent tried to a court, and a patent resolved before trial will be held invalid on section 102 cited prior art grounds *by the finder of fact*.

could be rejected with 90% confidence but not with 95% confidence (G-square p-value = 0.082).

⁶² The "general" category tracks the "mechanical" classification used by the PTO.

Table 3

Chemical	Valid	Invalid
23.00%	55.07%	46.38%
69	38	32
Electrical		
19.00%	52.63%	47.37%
57	30	27
General		
57.67%	54.34%	45.66%
173	94	79
Pharmaceutical	Valid	Invalid
3.67%	72.73%	27.27%
11	8	3
Biotech	Valid	Invalid
3.00%	55.56%	44.44%
9	5	4
Computer-related	Valid	Invalid
8.67%	65.38%	34.62%
26	17	9
Software	Valid	Invalid
1.00%	33.33%	66.67%
3	1	2

The first thing that is striking about these statistics is the nature of the patents that were litigated during this period. The majority of the patents litigated (173, or 57.7%) were classed as "general" inventions; only a much smaller number were chemical (69, or 23.0%) or electrical (57, or 19.0%) inventions. Contrast this with the number of patents issued. There, the evidence indicates that general patents are significantly less than half of the total number of patents issued.⁶³ It is evident that run-of-the-mill mechanical inventions, not chemical and electrical

⁶³ See Lemley, *supra* note ___, at 394 Table 4 (874 out of 2,081 patents in sample were classified as "general," 604

inventions, are the ones most commonly litigated.⁶⁴ This conclusion is borne out when one looks at the numbers of specialized inventions we have identified. Biotechnology, pharmaceuticals, computer and software patents are "hot" areas in patent law; firms, practitioners, and scholars spend a great deal of time thinking about these types of patents. Our population includes relatively few litigated patents in these areas: only 11 pharmaceutical patents (3.7% of the total), 9 biotech patents (3.0% of the total), 26 computer-related patents (8.67% of the total), and 3 software patents (1% of the total).⁶⁵ Even these small numbers are overstated because many of these categories overlap. One may suggest several possible explanations for this phenomenon. First, it may be that the new and dynamic nature of technology in areas such as biotechnology and software leads firms to settle disputes through cross-licensing or other means, because litigation is time-consuming and affects a firm's ability to move quickly in further developing the technology and fully appropriating its value. Second, in some of these areas there may be the type of litigation cost asymmetries identified by Lerner for new biotechnology firms, leading those firms to pursue a patent strategy that avoids litigation risk.⁶⁶ Third, the data may reflect reluctance to litigate early in an uncertain field. If this is correct, recent developments in the law governing biotechnology and software patents may very well lead to more patent litigation in the future.⁶⁷ Fourth, and perhaps most straightforward, at least a partial explanation

as chemical, and 603 as electrical).

⁶⁴ While it is conceivable that there is a large number of chemical and electrical patents on which suit is brought, but which do not face an invalidity defense, we find it extremely unlikely that that is the case.

⁶⁵ By contrast, Lanjouw and Schankerman conclude that "litigation is particularly frequent in new technology areas." Lanjouw & Schankerman, *supra* note __, at __. It is not clear what explains these different results; it may be that high-technology patents are commonly litigated but only rarely litigated to final judgment.

⁶⁶ Lerner, *supra* note 7.

⁶⁷ Because of the recent development of both biotechnology and software patent law, and in particular because of the changes to software patent law made in 1994, *see* *In re Alappat*, 35 F.3d 1526, 31 U.S.P.Q.2d 11545 (Fed. Cir. 1994) (en banc), it may be that these numbers will grow in the future. For example, a number of software patent validity decisions were handed down in 1997, after the closing date for the population. *See, e.g.,* *Lockwood v. American Airlines*, __ F.3d __ (Fed. Cir. 1997); *Fonar v. General Electric Co.*, __ F.3d __ (Fed. Cir. 1997); *In re Zurko*, __ F.3d __ (Fed. Cir. 1997); *Robotic Vision Sys., Inc. v. View Engineering*, __ F.3d __ (Fed. Cir. 1997); *In*

of the dearth of validity litigation in these areas may be found in the natural lag time between filing a patent and getting a final validity decision.⁶⁸

The comparison between the validity of different types of inventions was less illuminating. The large categories have strikingly similar validity rates: 55.1%, 52.6%, and 54.3% for chemical, electrical, and general patents, respectively. While the other groups have some wider variation in validity numbers (pharmaceutical patents were 72.7% valid, software patents only 33.3% valid), the small size of the populations in those groups renders ineffectual any attempt to predict a sustained difference in validity with any accuracy. Indeed, our test of *Hypothesis 3*: There is no difference between the likelihood that a chemical invention, an electrical invention, a general invention, a pharmaceutical invention, a biotechnological invention, a computer-related invention, and a software invention will be held invalid.

demonstrates that while there are differences in the numbers, one cannot predict with confidence that any group of patents is either more or less likely to be found valid.⁶⁹

re Dossel, ___ F.3d ___ (Fed. Cir. 1997); IGT v. Global Gaming Tech., 42 U.S.P.Q.2d 1144 (D. Nev. 1997).

⁶⁸ See *infra* notes ___-___ and accompanying text.

⁶⁹ For greater statistical accuracy, we divided the subject matter classifications into nine mutually exclusive categories, as noted in Table 3A.

Table 3A

Invention	Category	Probability of final invalid	Count
1	General	0.457	173
2	chemical but not pharm or biotech	0.472	53
3	chem and pharm but not biotech	0.375	8
4	chem and biotech but not pharm	0.667	6
5	chem and pharm and biotech	0	3
6	electrical but not software or computer	0.559	34
7	electrical and software but not computer	1	1
8	electrical and computer but not software	0.3	20
9	electrical and software and computer	0.5	2

The Chi-square test for independence produced a p-value of 0.3916, indicating that the null hypothesis 3 could not be rejected. [On the choice between Chi-square and G-square tests in this situation, see WICKENS, *supra* note ___, at 38-39.]

We also tested the grounds for invalidity of each of these types of inventions. These findings were somewhat more interesting. We present the data in Table 3B.

Table 3B

Chemical	enablement/written description	Claim indefiniteness	best mode	102 prior art	102 prior art	non-obviousness
23.00%	25.00%	9.38%	18.75%	37.50%	18.75%	37.50%
69	8	3	6	12	6	12
Electrical	enablement/written description	Claim indefiniteness	best mode	102 prior art	102 prior art	non-obviousness
19.00%	3.70%	7.41%	0.00%	25.93%	25.93%	51.85%
57	1	2	0	7	7	14
General	enablement/written description	Claim indefiniteness	best mode	102 prior art	102 prior art	non-obviousness
57.67%	5.06%	3.80%	12.66%	22.78%	37.97%	40.51%
173	4	3	10	18	30	32
Pharmaceutical	enablement/written description	Claim indefiniteness	best mode	102 prior art	102 prior art	non-obviousness
3.67%	33.33%	0.00%	33.33%	33.33%	0.00%	33.33%
11	1	0	1	1	0	1
Biotech	enablement/written description	Claim indefiniteness	best mode	102 prior art	102 prior art	non-obviousness
3.00%	100.00%	25.00%	0.00%	0.00%	0.00%	0.00%
9	4	1	0	0	0	0
Computer-related	enablement/written description	Claim indefiniteness	best mode	102 prior art	102 prior art	non-obviousness
8.67%	11.11%	22.22%	0.00%	22.22%	33.33%	33.33%
26	1	2	0	2	3	3
Software	enablement/written description	Claim indefiniteness	Best mode	102 prior art	102 prior art	non-obviousness
1.00%	0.00%	0.00%	0.00%	0.00%	0.00%	50.00%
3	0	0	0	0	0	1

Some of the differences are apparent. Compare the treatment of enablement and best mode in

chemical and electrical cases, for example. In chemical cases, 25% of all invalidity findings are made on the basis of enablement or written description, and another 18% are on the basis of best mode. In electrical cases, by contrast, only 3.7% of the invalidity findings result from enablement or written description and *none* from best mode.

We ran two tests of these relationships. The first tested for independence in the invalidity probabilities between general, electrical, and chemical patents. The test found no statistically significant difference between the overall probabilities of invalidity of each of the three types of patents (G-square p-value = 0.97).⁷⁰ The second test was of independence in the invalidity probabilities between the nine mutually exclusive categories of patents defined above. The test also found no statistically significant differences (G-square p-value = 0.26). That is, *Hypothesis 3*: There is no difference between the likelihood that a chemical invention, an electrical invention, and a general invention will be held invalid.

could not be rejected.

We tested these same groups and hypotheses for each ground of invalidity as well. Most of these tests also demonstrated no statistically significant relationship.⁷¹ However, a few of the relationships met the standard threshold of significance ($p < 0.05$). In particular,

Hypothesis 3h: There is no difference between the likelihood that a chemical invention, an electrical invention, and a general invention will be held invalid on obviousness grounds due to cited prior art.

was rejected (G-square p-value = 0.041). In this test, 28% of the general inventions, 23% of the chemical inventions, and 60% of the electrical inventions in which cited prior art was argued as a grounds for obviousness were found invalid on that ground.

Other differences of statistical significance were double patenting and best mode:

Hypothesis 3j: There is no difference between the likelihood that a chemical invention,

⁷⁰ The full data for each of these tests is reported in Appendix B.

⁷¹ The full set of test results is reprinted in Appendix B.

an electrical invention, and a general invention will be held invalid for double patenting.

was rejected (G-square p-value = 0.034). In this test, 100% of the general inventions, 33% of the chemical inventions, and none of the electrical inventions in which double-patenting was argued were held invalid on that ground.⁷²

Hypothesis 3l: There is no difference between the likelihood that a chemical invention, an electrical invention, and a general invention will be held invalid for failure to disclose best mode.

was also rejected (G-square p-value = 0.026). In this test, 39% of the general inventions, 50% of the chemical inventions, and none of the electrical inventions in which best mode violations were argued were found invalid on this ground.⁷³

Despite these significant results, the overall validity rates and the general grounds of invalidity are remarkably congruent across different types of inventions. This result may prove surprising to some—including us—who expected a greater divergence in patent law treatment of different fields of invention.

Foreign vs. Domestic Inventors

We also collected data from the population on the number of cases in which either the inventor or the assignee named on the patent resided outside the United States. An invention was classified as "foreign" for our purposes if *either* the inventor or the assignee was foreign.⁷⁴ The results of this analysis are presented in Table 4, and in graphical form in Figure 4.

Table 4

Foreign inventors/assigns	valid	Invalid
---------------------------	-------	---------

⁷² The number of such cases was very small, however: 4 general, 3 chemical, and 1 electrical case.

⁷³ Hypotheses 3h and 3j were also rejected in the parallel analysis of nine mutually independent groups of inventions; hypothesis 3l was not.

⁷⁴ If anything, this definition is overinclusive. One could argue that a patent with a foreign inventor but a U.S. owner was "domestic" in a meaningful sense.

13.71%	58.54%	41.46%
41	24	17
American inventors/assigns	valid	Invalid
86.29%	53.49%	46.90%
258	138	121

The first thing to note about this data is that the overwhelming majority of patents actually litigated in the United States have neither foreign inventors nor foreign owners. The data released by the PTO on the origin of all patents obtained in the U.S. show a consistent pattern: between 42% and 48% of all patents issued in the U.S. are of foreign origin.⁷⁵ Yet only 14% of the litigated patents are of foreign ownership. For whatever reason, foreign owners

⁷⁵ For recent data in representative years through 1995, found in the Additional Information section of the PTO's 1996 Annual Report, see <http://www.uspto.gov/web/offices/com/annual> (visited Nov. 21, 1997). Data for those representative years are:

Year	US Pats	Foreign Pats	Total Pats	% Foreign
1983	32871	23989	56860	42.2%
1986	38126	32734	70860	46.2%
1989	50185	45354	95539	47.5%
1992	59760	49968	109728	45.5%
1995	64562	49679	114241	43.5%

obtain U.S. patents, but do not sue to enforce them very often.⁷⁶

One reason foreign owners might not sue might be a fear that they would have a hard time persuading an American jury to grant them relief.⁷⁷ In fact, however, the population data on validity do not bear out that fear. Table 4 demonstrates that foreign patent owners actually do better than their American counterparts when they do enforce their patents. These data show that 58.5% of the foreign-owned or invented patents litigated were held valid in the population, while only 53.5% of the U.S.-owned and invented patents were held valid. These numbers do not suffice to show a statistically significant difference, however;
Hypothesis 4: There is no difference between the likelihood that a foreign invented or owned patent and a domestic invented and owned patent will be held invalid.

cannot be rejected (G-square p-value = 0.529). Still, the intuitive sense of many patent lawyers that foreign inventors or companies will not appear sympathetic to the trier of fact does not seem to be borne out in practice.⁷⁸

Claim Disaggregation

Most patents have multiple claims -- both "dependent" claims that narrow the scope of independent claims, and in many cases different sets of independent claims. Writing multiple claims has real costs, both in attorney time and in payments to the PTO, which charges its fees on a per-claim basis. The goal of multiple independent claims may be either to sweep the net of the invention more broadly -- catching more potential infringers -- or to help preserve the

⁷⁶ This is consistent with the findings of Lanjouw and Schankerman, who find that domestic patent owners are considerably more likely than foreign patent owners to be involved in litigation. Lanjouw & Schankerman, *supra* note ____.

⁷⁷ See Toshiko Takenaka, *The Role of the Japanese Patent System in Japanese Industry*, 13 **UCLA Pac. Basin L.J.** 25 (1994) (arguing that patent systems are biased towards nationals).

⁷⁸ It is possible that there is a relationship between these two findings; that the reason so few foreign-owned or invented patents are litigated is that foreign companies expect resistance, and only the "best" foreign-owned patents ever make it to court. The data we have collected do not provide a way to test this hypothesis.

validity of the patent by formulating the claim in alternative ways. The goal of dependent claims, which virtually every patent includes, is solely to preserve the validity of the patent by giving patentees a narrower alternative claim in case their broad independent claims are held invalid.

We have tested how often these multiple claims in fact make a difference in validity litigation -- that is, how often courts actually "disaggregate" their decision on the validity of a patent by claim, producing one validity holding on one set of claims and a second holding on another.⁷⁹ The results are striking. Courts actually disaggregated claims in only 7 of the 299 patents litigated (2.34%). Further, most of these were not instances of "true" claim disaggregation, because they did not produce two different *final* decisions on patent validity. In six of the seven cases, the court produced a final holding on validity with respect to one set of claims, and a *non-final* decision (such as a remand or denial of pretrial) with respect to another set. In only one case in the population (0.34%) did a court actually hold that one set of claims was valid and the other set invalid.

Some litigators have indicated to us that they are not terribly surprised by this result, because in their experience claim disaggregation is rarely even argued by a patentee. While that may be true, it is clear that patent prosecutors are spending a great deal of time, effort and money writing long strings of dependent claims that almost never make a difference in litigation.

Prior Art Citations

Patent applicants are required to disclose to the PTO during the application process all material prior art of which they are aware.⁸⁰ In addition, the Examiner conducts a search of the relevant classes for prior art. Prior art discovered through either method is listed in the final

⁷⁹ Our use of the term "claim disaggregation" should be distinguished from the doctrine of "claim differentiation," which we do not discuss.

⁸⁰ 37 C.F.R. § 1.56.

patent, either on the cover or in the text of the patent application itself. Lawyers frequently speculate on the relationship between prior art cited in the patent and the patent's strength; a common theory is that patents in which the Examiner has considered a great deal of prior art are stronger than other patents (because attacking the patent on the basis of this prior art would mean revisiting the Examiner's conclusions). Further, some have suggested that patents are stronger if they cite a great deal of "nonpatent" prior art -- prior art in the form of journal articles, commercial products, and the like -- because that art really does represent the state of the technology. We tested the relationship between cited and uncited prior art and patent validity.

We first determined the distribution of prior art citations in the population. That distribution is reflected in Table 5.

Table 5

	Total Cited PA	Cited Patent PA	Cited nonpatent PA
Mean	16.6	12.23	4.37
Median	10	9	0
Standard Deviation	32.89	12.3	28.61

Table 5 demonstrates both that there is tremendous variation in the number of prior art references cited in the patents in the population, and that the great majority of the prior art cited to patent Examiners consists of other patents. Indeed, the median number of non-patent art citations is zero, meaning that only a minority of patents in the population cite *any* nonpatent art.

We examined the relationship between total cited prior art, cited patent prior art, and cited nonpatent prior art and the validity of the patent.⁸¹ The results are presented in Table 5A.

Table 5A

Total PA in Patents	Cited Valid Patents	Total Cited PA Invalid on 102 PA or Obv.	Cited PA in Valid Patents	Cited Patent PA Invalid on 102 PA or Obv.	Cited PA in Patents Invalid on 102 PA or Obv.	Cited nonpatent PA in Valid Patents	Cited nonpatent PA in Patents Invalid on 102 PA or Obv.
Mean	18.85	15.03	12.52	13.62		6.33	1.41
Median	12	9	10.5	0	0		

⁸¹ For this test, we used a subset of the general population: those validity decisions made on the basis of prior art (obviousness and section 102 prior art cases). There were 202 patents in this group.

	10				
Std. Dev.	13.71	13.56	11.68	38.53	3.22
	43.13				

The data do not demonstrate a simple relationship between cited prior art and patent validity. While patents held invalid on prior art grounds tended to have cited less nonpatent prior art than valid patents did (1.4 nonpatent prior art references cited on average in invalid patents, compared with 6.3 in valid patents), the reverse is true for patent prior art (13.6 in invalid patents, and 12.5 in valid patents). The end result is that none of the relationships proves statistically significant.

None of the hypotheses:

Hypothesis 5: There is no relationship between the number of cited prior art references in a patent and its invalidity on section 102 prior art or section 103 obviousness grounds.

Hypothesis 5a: There is no relationship between the number of cited patent prior art references in a patent and its invalidity on section 102 prior art or section 103 obviousness grounds.

Hypothesis 5b: There is no relationship between the number of cited non-patent prior art references in a patent and its invalidity on section 102 prior art or section 103 obviousness grounds.

can be rejected.⁸² Thus, while it is true that not much nonpatent prior art is cited in the patents in the population, we cannot draw robust conclusions about the relationship between cited prior art and patent validity.

Cited vs. Uncited Art

Another issue of significance relates not to prior art cited in the patent, but to prior art relied upon by defendants in litigation. It is received wisdom among patent lawyers that it is much easier to invalidate a patent on the basis of "uncited" prior art -- that is, art that was not considered by the Examiner during prosecution. The rationale is that the trier of fact will be

⁸² We used the log odds of final invalidity based on cited art, calculated in the logistic regression model described below. The p-values for hypotheses 5, 5a, and 5b were 0.543, 0.353, and 0.389 respectively.

reluctant to second-guess the Examiner about an art reference that she has already considered, but that the trier of fact may be willing to invalidate a patent based on information that the Examiner didn't have available to her.

To test this assumption, we compared the validity determinations made primarily on the basis of cited art and those made primarily on the basis of uncited art.⁸³ The data on which these tests were conducted are reported in Table 6.

⁸³ Once again, the population for this test is the 202 patents in which prior art issues were at stake.

The determination of which prior art references were "primarily" relied upon by the courts necessarily involved some very difficult judgment calls in interpreting the courts' opinions. We have excluded cases from this test where it was not possible to make that judgment with any certainty.

Table 6

Cases where 102 PA or obviousness at issue

	PA Argued	Cited PA Argued	Uncited PA argued
Mean	2.78	1.16	1.62
Median	2	0	1
standard deviation	2.9	2	2.09

Valid under 102 PA or Obviousness

	PA Argued	Cited PA Argued	Uncited PA argued
Mean	2.88	1.36	1.52
Median	2	0	1
standard deviation	3.09	2.19	2.15

Invalid under 102 PA or Obviousness

	PA Argued	Cited PA Argued	Uncited PA argued
Mean	2.78	0.89	1.89
Median	2	0	1
standard deviation	2.64	1.61	2.08

The first thing worth noting is that most prior art attacks rely on a relatively small number of references: 2.8, on average. Further, most of the references that are argued in court are uncited references; indeed, in the majority of prior art-related cases, no cited art is relied upon at all. In the cases where patents were actually held invalid, defendants disproportionately relied upon uncited prior art: 1.9 uncited references on average, compared with 0.9 cited references.

Further, statistical tests in the superpopulation indicate with a fair degree of confidence that reliance on uncited art was more likely to lead to a finding of invalidity than reliance on cited art. We tested

Hypothesis 6: There is no difference between the likelihood that a patent will be held invalid on section 102 prior art or section 103 obviousness grounds because of cited prior art and the likelihood that it will be held invalid on section 102 prior art or section 103 obviousness grounds because of uncited prior art.

The probability of invalidity based on uncited art was 40.8%, while the probability of invalidity

based on cited art was 29.6%.⁸⁴ Based on these numbers, the G-square p-value for the hypothesis was 0.057. The null hypothesis can be rejected with 90% confidence, but cannot be rejected with 95% confidence. The likely result therefore confirms the conventional wisdom and the results of two early studies⁸⁵: uncited prior art is a more effective tool for invalidating patents than cited prior art.

Elapsed Time to Final Judgment

We evaluated several measures of elapsed time before the final judgment in the case, in an effort to determine whether any of these measures had any relationship to validity. Our collection of these data was motivated by several factors. First, the length and importance of the time a patent spends in prosecution has been hotly debated because beginning in 1995 the term of a patent is measured from the date the application is filed, not the date the patent issues.⁸⁶ Some have accused the PTO of "selling out" inventors, while those on the other side complain of "submarine patents" that languished in prosecution for decades at the behest of the inventor, only to "surface" and torpedo a mature industry.⁸⁷ Second, information about a technology changes over time, making it reasonable to assume that factfinders might look less favorably upon old patents than upon newer ones.

To test various hypotheses about elapsed time, we have defined three measures. The first is "time in prosecution," by which we mean the total time elapsed from the earliest filing of a patent application (or parent or grandparent application upon which the current application relies for priority) to the issuance of the patent. The second is "delay," by which we mean the total

⁸⁴ $n=267$ for this test, because in some cases defendants relied on *both* cited and uncited art. Where that occurred, we listed the case in both categories.

⁸⁵ Koenig, *supra* note __, at __.

⁸⁶ 35 U.S.C. § 154(a).

⁸⁷ For a discussion of this debate, see Lemley, *supra* note __, at 369.

elapsed time between the date the patent issues and the date the validity of the patent is finally decided in litigation.⁸⁸ The final measure is "lag," by which we mean the total elapsed time between the filing of a patent application and the final validity decision. Lag is simply "time in prosecution" plus "delay."

The descriptive statistics for all three measures are presented in Tables 7 (lag), 7A (delay), and 7B (time in prosecution). For each measure, we have presented the mean, median and distribution over the entire population, and for four subsets of the population: valid patents, invalid patents, patents held valid over a prior art claim, and patents held invalid on the basis of a prior art claim.

⁸⁸ Delay is therefore a function of two periods of time: the amount of time between issuance and the filing of a lawsuit, and the amount of time it takes to resolve the lawsuit, including appeal if any. The first period is generally within the control of the patentee, though it may depend on the patentee's knowledge of infringing activity. The second period is only partially and indirectly within the patentee's control.

We emphasize that by using the term "delay," we do not intend to imply anything about either the propriety of refraining from filing suit or the state of mind of the patent owner.

Table 7

Lag Between Filing and Resolution

	All	Valid	Invalid	Valid 102PA/103	Invalid 102PA/103
mean	12.26	12.14	12.36	12.42	12.24
Median	11.3	11.05	11.5	11.8	11.5
Standard deviation	5.68	5.29	6.12	5.09	5.64

Table 7A

Delay between Issuance and Resolution

	All	Valid	Invalid	Valid 102PA/103	Invalid 102PA/103
Mean	8.61	8.69	8.49	8.81	8.62
Median	7.8	8	7.5	8.2	7.75
standard deviation	5.08	5.02	5.16	4.84	5.17

Table 7B

Time in Prosecution

	All	Valid	Invalid	Valid 102PA/103	Invalid 102PA/103
mean	3.64	3.45	3.87	3.62	3.61
Median	2.7	2.65	2.75	2.8	2.75
standard deviation	2.98	2.56	3.39	2.69	2.64

The average time in prosecution for litigated patents in the population is 3.64 years, though the lower median (2.7 years) indicates that a few patents spent a great deal of time in prosecution, driving the mean up. The average delay after issuance is 8.61 years, and the average lag between filing and resolution is 12.26 years.

The lag and delay statistics are quite interesting, because it appears that most patents litigated to judgment involve fairly old technology. This may be of some significance in the debate over what the patent system contributes to innovation: litigation does not appear to provide early, certain protection to inventors. Of course, the patent system may still provide ex ante incentives to innovate, and up front licensing and preliminary injunctions both provide significant relief to patent owners in many cases. It appears that most patent litigation in fact involves inventions that not only are fairly old, but that have been patented for several years

before the patents have been enforced.⁸⁹ One possible explanation for this finding is that many firms may build patent portfolios over time to give them more freedom of action in their markets, resorting to litigation only when that freedom is subsequently challenged, either by a patent suit by a rival or by a shifting marketplace.⁹⁰ In other words, many firms patent with no immediate purpose or early need to enforce them, but rather to fence out competitors and potential competitors. Alternatively, the rapid developmental pace of some cutting-edge technologies may militate against investment in lengthy and expensive patent litigation.

The large lags and delays we found also may explain another mildly surprising finding. Given the increasing complexity, cost, and interdisciplinary nature of research and development that today characterize a number of cutting-edge technologies, one might expect that the number of inventors per patent would have increased dramatically. There is reason to believe that this is happening as we write. In our population, however, only 138 out of 299 patents listed more than one inventor, compared to 161 listing a single inventor. Although the highest number of inventors was 8, the mean number was only 1.82.

The amount of time the patents spent in prosecution is also significant. The findings here are broadly consistent with Lemley's prior work on the subject, which tested a smaller population and found the average time in prosecution to be 1274 days (approximately 3.5 years) and the median time in prosecution to be 1012 days (approximately 2.8 years) for litigated cases.⁹¹ By contrast, that study found a much shorter time in prosecution in a sample of *all*

⁸⁹ This prevalence of old technology may help to explain why so few "hot" fields of invention are represented in the litigation population. It may be several years before we start to see significant numbers of software or biotechnology patents litigated to judgment.

⁹⁰ This is implicitly supported by a recent study. See Cohen, *supra* note __. In the study, Cohen, *et al.* found that firms generally view patents as less effective than other means of obtaining value from new product innovations, secrecy and lead time being the most important. One of the most important motives for patenting was defensive—preventing other firms from patenting related innovations, perhaps providing a partial explanation for the fact that most patent validity decisions occur a lengthy time after patent issuance.

⁹¹ See Lemley, *supra* note __, at 420 & Table 8.

issued patents (not just those that were litigated).⁹² As noted in that study, this discrepancy may result either from the lag (if patents now are issued more quickly than they were in the past) or from the fact that "important" patents (of which litigation to judgment might be a predictor) actually do spend longer in prosecution.⁹³

The data also demonstrate no significant relationship whatever between lag, delay or time in prosecution and patent validity. The lag, delay and time in prosecution statistics for valid patents are all quite close to those for invalid patents, under either population subset. As one might expect from looking at this data, each of the hypotheses we tested from this data:

Hypothesis 7: Delay between issuance of a patent and decision in the case has no effect on the validity of the patent.

Hypothesis 7a: Delay between issuance of a patent and decision in the case has no effect on the validity of the patent on 102 prior art or section 103 obviousness grounds.

Hypothesis 8: Lag between original application for a patent and decision in the case has no effect on the validity of the patent.

Hypothesis 8a: Lag between original application for a patent and decision in the case has no effect on the validity of the patent on 102 prior art or section 103 obviousness grounds.

Hypothesis 9: Time in prosecution (measured from original application to issue date of current patent) has no effect on the validity of the patent.

Hypothesis 9a: Time in prosecution (measured from original application to issue date of current patent) has no effect on the validity of the patent on 102 prior art or section 103 obviousness grounds.

cannot be rejected.⁹⁴ There is no demonstrable relationship between the validity of a patent and

⁹² *Id.* at 388 & Table 2.

⁹³ *See id.* at 421-22 (noting both hypotheses).

⁹⁴ The p-values for the predictors in the logistic regression model are:

Hypothesis 7 p=0.731

Hypothesis 7a p=0.818

Hypothesis 8 p=0.733

Hypothesis 8a p=0.821

Hypothesis 9 p=0.220

Hypothesis 9a p=0.983

the time it spends in prosecution, delay in filing suit or bringing the case to judgment, and the total lag between invention and judgment.

Appeals

In addition to measuring the final validity of patents, we also tested how various trial court decisions fared on appeal. Table 8 presents affirmance/reversal data for Federal Circuit decisions in the population.

Table 8

Patents	Appealed	Affirmed	Reversed	Not Appealed
	48.83%	86.30%	17.12%	51.17%
299	146	126	25	153

Approximately half of the cases in the population are appellate decisions (146 out of 299). The high percentage of appellate decisions reflects two facts. First, we replaced district court decisions with appellate court decisions wherever there were such decisions. Second, appellate courts are more likely than district courts to produce *reported* decisions.

Of the 146 appellate decisions, the overwhelming majority (126, or 86.3%) affirmed the district court judgment; only 25 (17.1%) were reversed by the district court.⁹⁵ On the basis of these numbers, we can clearly reject *Hypothesis 10*: The decision of the finder of fact in patent cases is not more likely to be affirmed than reversed on appeal.

(G-square p-value < 0.01). However, we think that this result is at least in significant part an artifact of our population choice. Because we have included only final reported decisions in the population, appellate decisions which *remand* a case but do not reverse the validity finding

respectively.

⁹⁵ The numbers and percentages do not total perfectly because in some cases, the district court's decision was *both* affirmed and reversed.

outright are not included in the population. We suspect that if remands were included in our population definition, the percentage of affirmances would drop sharply.

The population definition does allow us to test how various trial court decisions fare on appeal, however. We first tested whether trial court decisions holding a patent valid were more likely to be affirmed than decisions holding a patent invalid. The results are presented in Table 8A.

Table 8A

Valid	Valid Appealed	Valid Affirmed	Valid Reversed	Valid Not Appealed
	43.87%	89.71%	10.29%	56.13%
155	68	61	7	87

Invalid	Invalid Appealed	Invalid Affirmed	Invalid Reversed	Invalid Not Appealed
	54.17%	83.33%	23.08%	45.83%
144	78	65	18	66

While the affirmance rates for both trial court decisions of validity and of invalidity are quite high, there is a noticeable discrepancy: trial court findings of invalidity are more commonly reversed by the Federal Circuit. The statistical tests in the population bear out this observation.

We tested twin hypotheses

Hypothesis 10a: The decision of the finder of fact holding a patent valid is not more likely to be affirmed on appeal than the decision of the finder of fact holding a patent invalid.

Hypothesis 10b: The decision of the finder of fact holding a patent valid is not more likely to be reversed on appeal than the decision of the finder of fact holding a patent invalid.

The result was a G-square p-value = 0.056, meaning that the hypothesis could be rejected with greater than 90% confidence, but not with 95% confidence.⁹⁶ This finding bears out the received

⁹⁶ The test of hypotheses 10a and 10b characterized cases that included both affirmances and reversals as two separate decisions for purposes of testing. The result was that n(trial valid)=68 and n(trial invalid)=83 in the test,

wisdom that the Federal Circuit is more likely to favor patentees on appeal, and confirms the findings of previous studies that the court is substantially more likely to affirm a district court's finding of validity than invalidity.⁹⁷ This finding is in polar opposition to that of studies using data from the regional circuits prior to creation of the Federal Circuit.⁹⁸

We also broke down appellate decisions into categories based on who the trier of fact was. The results of that division are reported in Table 8B.

and the respective percentages affirmed on appeal were 89.7% for trial validity decisions and 78.3% for trial invalidity decisions in the test.

⁹⁷ See Dunner, et al., *supra* note __, at 154 (found this to be true both overall and within the separate invalidity categories of novelty/statutory bars, 35 U.S.C. § 102, obviousness, 35 U.S.C. § 103, and disclosure/claim adequacy, 35 U.S.C. § 112).

⁹⁸ See Koenig, *supra* note __, at 4-39 to 4-40.

Table 8B

Jury Trial	Appealed	Affirmed	Reversed	Not Appealed
	61.64%	86.67%	13.33%	38.36%
73	45	39	6	28
Bench Trial	Appealed	Affirmed	Reversed	Not Appealed
	47.55%	89.71%	16.18%	52.45%
143	68	61	11	75
Pretrial	Appealed	Affirmed	Reversed	Not Appealed
	39.02%	78.13%	25.00%	60.98%
82	32	25	8	50
JMOL	Appealed	Affirmed	Reversed	Not Appealed
	100.00%	100.00%	0.00%	0.00%
1	1	1	0	0

The results show that district court decisions on pretrial motion are the least likely to be appealed, but that when they are appealed they are the most likely to be reversed (reversal in 25% of all appealed cases). By contrast, bench and jury trials show about the same reversal rate (13.3% for juries, and 16.1% for bench trials).

Table 8C shows the results of considering both variables together: trier of fact, and result below.

Table 8C

Jury Trial	Valid	Valid Appealed	Valid Affirmed	Valid Reversed	Valid Not Appealed
		57.14%	82.14%	17.86%	42.86%
73	49	28	23	5	21
Bench Trial	Valid	Valid Appealed	Valid Affirmed	Valid Reversed	Valid Not Appealed
		39.02%	93.75%	6.25%	60.98%
143	82	32	30	2	50
Pretrial	Valid	Valid Appealed	Valid Affirmed	Valid Reversed	Valid Not Appealed
		30.43%	100.00%	0.00%	69.57%
82	23	7	7	0	16
JMOL	Valid	Valid Appealed	Valid Affirmed	Valid Reversed	Valid Not Appealed
		100.00%	100.00%	0.00%	0.00%
1	1	1	1	0	0
Jury Trial	Invalid	Invalid Appealed	Invalid Affirmed	Invalid Reversed	Invalid Not Appealed
		70.83%	94.12%	5.88%	29.17%
73	24	17	16	1	7
Bench Trial	Invalid	Invalid Appealed	Invalid Affirmed	Invalid Reversed	Invalid Not Appealed
		59.02%	86.11%	25.00%	40.98%
143	61	36	31	9	25
Pretrial	Invalid	Invalid Appealed	Invalid Affirmed	Invalid Reversed	Invalid Not Appealed
		42.37%	72.00%	32.00%	57.63%
82	59	25	18	8	34
JMOL	Invalid	Invalid Appealed	Invalid Affirmed	Invalid Reversed	Invalid Not Appealed
		0.00%	0.00%	0.00%	0.00%
1	0	0	0	0	0

The data suggest some notable differences in the treatment of findings of validity and invalidity. No pretrial decisions holding a patent valid were reversed on appeal; by contrast, 32% of the pretrial decisions holding a patent *invalid* were reversed on appeal. Similarly, only 6.25% of the

bench trial decisions holding a patent valid were reversed on appeal, compared with 25% of the bench trial decisions holding a patent invalid. (The same cannot be said of jury verdicts, however; on appeal from jury verdicts, 17.9% of the juries' validity decisions were reversed, compared with only 5.9% of the juries' invalidity decisions).

The small population size of some of these groups makes it harder to find statistical significance in the superpopulation. None of the following hypotheses

Hypothesis 10c: The decision of the court in a bench trial in a patent case is not more likely to be affirmed on appeal than the decision of a jury or the decision of a court on pretrial motion holding a patent invalid.

Hypothesis 10d: The decision of the court in a bench trial in a patent case holding a patent valid is not more likely to be affirmed on appeal than the decision of a jury or the decision of a court on pretrial motion holding a patent invalid.

Hypothesis 10e: The decision of the court in a bench trial in a patent case holding a patent invalid is not more likely to be affirmed on appeal than the decision of a jury or the decision of a court on pretrial motion holding a patent invalid.

could be rejected with confidence based on this data (G-square p-values were 0.425, 0.163, and 0.108 respectively).

In short, it appears that the Federal Circuit is somewhat more hospitable to patentees than to accused infringers. The Federal Circuit gives somewhat more deference to decisions at trial than to pretrial rulings. But interestingly, it does not significantly distinguish between bench trials and jury trials in reviewing trial decisions.

Multiple Patents in Suit

The 299 patents in the population were litigated in 239 suits. In several cases, the validity of two or more patents was decided in the same suit. These patents were often but not always related to each other. We set out to determine whether different patents litigated together in the same lawsuit were more likely than chance would predict to stand or fall together. The data regarding multiple patents in suit are presented in Table 9.

Table 9

Cases	Validity of More than One	All Valid	All Invalid	Mixed
	18.83%	53.33%	33.33%	13.33%
239	45	24	15	6

These data show remarkable agreement. In 86.7% of the cases with multiple patents in suit, the patents in the case were either all held valid together or all held invalid together. Courts produced "mixed" results in only 6 of 45 cases, or 13.3%.

Using this data, we tested twin hypotheses

Hypothesis 11: In cases in which more than one patent is litigated, there is no relationship between the validity of one patent in the case and the validity of remaining patents in the case.

Hypothesis 12: In cases in which more than one patent is litigated, there is no relationship between the invalidity of one patent in the case and the invalidity of remaining patents in the case.

These hypotheses were rejected with great confidence (G-square p-value < 0.001). Whether it is because the patents themselves are closely related, or because triers of fact tend to make decisions as a whole, it clearly appears that asserting multiple patents in the same suit is an all-or-nothing proposition.

Where Patent Cases Are Litigated

We also examined the districts in which patents were most likely to be litigated. Of the 108 federal district courts in the country, just over half (56) issued one or more final patent validity decisions during the period of our study. Twenty-five of these districts litigated the validity of more than five patents during this period. The districts in which patents were most commonly litigated were the Southern District of New York, the Northern District of Illinois, the Eastern District of Michigan, the Central District of California, and the District of Delaware. The list of the twenty-five most active jurisdictions is presented in Table 10.⁹⁹

⁹⁹ While we report the number of litigated patents, not the number of cases, we have evaluated the number of cases litigated in these districts, and the results are not significantly different. The District of New Jersey and the Eastern District of Pennsylvania replace the Western District of Wisconsin and the Southern District of Florida on the top

Table 10

Court	Patents
S.D. N.Y.	23
N.D. Ill.	20
E.D. Mich.	17
C.D. Cal.	17
D. Del.	17
D. Minn.	15
N.D. Cal.	10
E.D. Va.	9
W.D. Wisc.	8
S.D. Fla.	8
N.D. Tex.	8
E.D. N.Y.	7
S.D. Cal.	7
E.D. Pa.	7
D. N.J.	7
W.D. Tex.	6
S.D. Tex.	6
D. Or.	6
E.D. Mo.	6
D. Mass.	6
W.D. N.C.	6
M.D. Fla.	5
S.D. Ohio	5
D. Kan.	5
W.D. Pa.	5

Patterns are hard to discern in these figures. Patent litigation occurs across the geographic spectrum. However, the ten most active districts are the ones patent litigators might expect: centers of industry (the Southern District of New York, the Northern District of Illinois, the Eastern District of Michigan, the Central District of California), centers of technology (the Northern District of California, the District of Minnesota), or courts known to be patent-friendly for legal or procedural reasons (the District of Delaware, the Eastern District of Virginia). There are a few surprises, however: we would not have expected the Western District of Wisconsin to appear on the top-ten list, nor the District of Kansas to appear in the top 25. By contrast, we would have predicted that the District of Massachusetts would have more patent cases than appears to be the case.

ten list.

We also examined the validity rates by district for all those districts that decided the validity of five or more patents during the period of the study. The results are presented in Table 11.¹⁰⁰

¹⁰⁰ The data we report here are the percentage of patents held valid *in the district court* in each district, not the final determination made by any court. Since our goal is to identify "patent-friendly" and "patent-unfriendly" districts, the district court numbers are of more interest than numbers that reflect Federal Circuit decisions.

Table 11

District Court Valid/Invalid

Court	Patents	Valid	% Valid	Invalid	% Invalid
W.D. Wisc.	8	8	100.00%	0	0.00%
S.D. Fla.	8	7	87.50%	1	12.50%
E.D. N.Y.	7	6	85.71%	1	14.29%
W.D. Tex.	6	5	83.33%	1	16.67%
M.D. Fla.	5	4	80.00%	1	20.00%
S.D. Tex.	6	4	66.67%	2	33.33%
D. Or.	6	4	66.67%	2	33.33%
E.D. Mich.	17	11	64.71%	6	35.29%
D. Minn.	15	9	60.00%	6	40.00%
N.D. Cal.	10	6	60.00%	4	40.00%
S.D. Ohio	5	3	60.00%	2	40.00%
D. Kan.	5	3	60.00%	2	40.00%
S.D. Cal.	7	4	57.14%	3	42.86%
E.D. Pa.	7	4	57.14%	3	42.86%
E.D. Va.	9	5	55.56%	4	44.44%
C.D. Cal.	17	9	52.94%	8	47.06%
N.D. Tex.	8	4	50.00%	4	50.00%
E.D. Mo.	6	3	50.00%	3	50.00%
D. Del.	17	7	41.18%	10	58.82%
S.D. N.Y.	23	8	34.78%	15	65.22%
N.D. Ill.	20	5	25.00%	15	75.00%
W.D. Pa.	5	1	20.00%	4	80.00%
D. Mass.	6	1	16.67%	5	83.33%
D. N.J.	7	0	0.00%	7	100.00%
W.D. N.C.	6	0	0.00%	6	100.00%

Based on this data, patent plaintiffs would do well to eschew some of the districts historically considered "patent-friendly," such as the District of Delaware and the Eastern District of Virginia. Both were more likely than average to hold patents invalid.

We should emphasize that the numbers of cases in each district are quite small, and that one should not draw robust conclusions from these data. In particular, we have not attempted any statistical test of the differences in validity rates between the districts.

The Model

We developed a logistic regression model using most of the variables we have tested here in an effort to find reliable predictors of patent invalidity over time. The factors included in the model were

trial to jury, trial to bench, pretrial disposition, JMOL disposition, chemical invention, electrical invention, general invention, pharmaceutical invention, biotechnology invention, computer-related invention, software invention, foreign inventorship/assignee, total number of cited prior art references, number of cited patent prior art references, number of cited nonpatent prior art references, total number of references argued in court, number of cited references argued in court, number of uncited references argued in court, delay, lag, time in prosecution, appeal.

The results were disappointing, although perhaps not surprising. Of all the factors we tested as possible predictors of invalidity, only one set showed any significant predictive value. That one set was the choice of finder of fact. Compared to pretrial disposition, trial to either a jury or the bench were significant predictors of the validity of a patent.¹⁰¹ The exponent coefficients for jury trial and bench trial were 0.359 and 0.389 respectively.¹⁰²

The failure of most other factors as predictors of invalidity suggests that most of the variables that determine invalidity are in fact things that our study cannot measure. Some obvious but untested variables include the skill of the lawyers, the amount of money each side is willing to spend on the litigation, the "hometown advantage," the skill, character and demeanor of the fact and expert witnesses, and the quality of the patent. There are clearly relationships of interest to lawyers between validity and many of the factors we have described; but it is also evident that we have uncovered at most only a small part of the story.

Conclusions

Some of the more interesting findings from the data are summarized below:

- Approximately 54% of all litigated patents are held valid.
- Section 102 arguments (both those based on prior art and those based on the statutory bars) generally fare better in the courts than do arguments regarding obviousness.
- Defendants' best hope of invalidating a patent lies in obtaining a pretrial ruling from the

¹⁰¹ Strictly speaking, trial to bench and trial to jury were negative predictors of invalidity (coefficients -1.02 for jury and -0.94 for bench trial), which amounts to the same thing. Pretrial disposition is a positive predictor of invalidity (coefficient 0.55).

¹⁰² For further detail on the model, see Appendix B.

court. Bench trials are more favorable to patentees, and juries are extremely favorable to patentees.

- There is virtually no difference between the validity rates of patents in different fields of invention.
- Very few software, biotechnology and pharmaceutical patents have actually been litigated to final judgment in the recent eight-year period. The majority of the patents that are litigated are general (“mechanical”) inventions.
- Foreign inventors and owners bring very few patent cases, but when they do bring suit, they do not face discrimination by fact finders.
- In virtually no cases does the result actually turn on different treatment of multiple claims in the patent.
- Uncited prior art is more likely to invalidate a patent than previously cited prior art.
- Litigated patents are generally not newly obtained ones; on average, well over a decade elapses between the filing of a patent application and a court's ruling on the validity of the patent.
- The Federal Circuit is more likely to affirm trial courts that find patents valid than courts that find patents invalid, and is more likely to affirm trial decisions than pretrial decisions.
- Patents that are raised in the same suit almost always stand or fall together.

Some of these findings support the conventional wisdom—that juries and the Federal Circuit favor patentees, for example. But other findings ran counter to our expectations, notably the rarity of claim disaggregation and the overwhelming number of “traditional” mechanical inventions in the population. In other instances, there were simply no significant relationships to be found, though that fact might itself be considered important. We hope that all of these findings will be useful to scholars seeking to better understand patent cases and to practitioners who must try them.

APPENDIX A

Hypotheses Tested in the Superpopulation

Hypothesis 1: Issued patents are not more likely to be held valid than invalid.

Hypothesis 2: There is no difference between the likelihood that a patent tried to a jury, a patent tried to a court, and a patent resolved before trial will be held invalid *by the finder of fact*.

Hypothesis 2a: There is no difference between the likelihood that a patent tried to a jury, a patent tried to a court, and a patent resolved before trial will be held invalid on utility grounds *by the finder of fact*.

Hypothesis 2b: There is no difference between the likelihood that a patent tried to a jury, a patent tried to a court, and a patent resolved before trial will be held invalid on enablement or written description grounds *by the finder of fact*.

Hypothesis 2c: There is no difference between the likelihood that a patent tried to a jury, a patent tried to a court, and a patent resolved before trial will be held invalid on section 102 prior art grounds *by the finder of fact*.

Hypothesis 2d: There is no difference between the likelihood that a patent tried to a jury, a patent tried to a court, and a patent resolved before trial will be held invalid on section 102 cited prior art grounds *by the finder of fact*.

Hypothesis 2e: There is no difference between the likelihood that a patent tried to a jury, a patent tried to a court, and a patent resolved before trial will be held invalid on section 102 uncited prior art grounds *by the finder of fact*.

Hypothesis 2f: There is no difference between the likelihood that a patent tried to a jury, a patent tried to a court, and a patent resolved before trial will be held invalid on section 102 non-prior art grounds *by the finder of fact*.

Hypothesis 2g: There is no difference between the likelihood that a patent tried to a jury, a patent tried to a court, and a patent resolved before trial will be held invalid on obviousness grounds *by the finder of fact*.

Hypothesis 2h: There is no difference between the likelihood that a patent tried to a jury, a patent tried to a court, and a patent resolved before trial will be held invalid on obviousness grounds due to cited prior art *by the finder of fact*.

Hypothesis 2i: There is no difference between the likelihood that a patent tried to a jury, a patent tried to a court, and a patent resolved before trial will be held invalid on obviousness grounds due to uncited prior art *by the finder of fact*.

Hypothesis 2j: There is no difference between the likelihood that a patent tried to a jury, a patent tried to a court, and a patent resolved before trial will be held invalid for double-patenting *by the finder of fact*.

Hypothesis 2k: There is no difference between the likelihood that a patent tried to a jury, a patent tried to a court, and a patent resolved before trial will be held invalid for incorrect inventorship *by the finder of fact*.

Hypothesis 2l: There is no difference between the likelihood that a patent tried to a jury, a patent tried to a court, and a patent resolved before trial will be held invalid for failure to disclose best mode *by the finder of fact*.

Hypothesis 2m: There is no difference between the likelihood that a patent tried to a jury, a patent tried to a court, and a patent resolved before trial will be held invalid for lack of patentable subject matter *by the finder of fact*.

Hypothesis 2n: There is no difference between the likelihood that a patent tried to a jury, a patent tried to a court, and a patent resolved before trial will be held invalid for claim indefiniteness *by the finder of fact*.

Hypothesis 3: There is no difference between the likelihood that a chemical invention, an electrical invention, a general invention, a pharmaceutical invention, a biotechnological invention, a computer-related invention, and a software invention will be held invalid.

Hypothesis 3a: There is no difference between the likelihood that a chemical invention, an electrical invention, a general invention, a pharmaceutical invention, a biotechnological invention, a computer-related invention, and a software invention will be held invalid on utility grounds.

Hypothesis 3b: There is no difference between the likelihood that a chemical invention, an electrical invention, a general invention, a pharmaceutical invention, a biotechnological invention, a computer-related invention, and a software invention will be held invalid on enablement or written description grounds.

Hypothesis 3c: There is no difference between the likelihood that a chemical invention, an electrical invention, a general invention, a pharmaceutical invention, a biotechnological invention, a computer-related invention, and a software invention will be held invalid on section 102 prior art grounds.

Hypothesis 3d: There is no difference between the likelihood that a chemical invention, an electrical invention, a general invention, a pharmaceutical invention, a biotechnological invention, a computer-related invention, and a software invention will be held invalid on section 102 cited prior art grounds.

Hypothesis 3e: There is no difference between the likelihood that a chemical invention, an electrical invention, a general invention, a pharmaceutical invention, a

biotechnological invention, a computer-related invention, and a software invention will be held invalid on section 102 uncited prior art grounds.

Hypothesis 3f: There is no difference between the likelihood that a chemical invention, an electrical invention, a general invention, a pharmaceutical invention, a biotechnological invention, a computer-related invention, and a software invention will be held invalid on section 102 non-prior art grounds.

Hypothesis 3g: There is no difference between the likelihood that a chemical invention, an electrical invention, a general invention, a pharmaceutical invention, a biotechnological invention, a computer-related invention, and a software invention will be held invalid on obviousness grounds.

Hypothesis 3h: There is no difference between the likelihood that a chemical invention, an electrical invention, a general invention, a pharmaceutical invention, a biotechnological invention, a computer-related invention, and a software invention will be held invalid on obviousness grounds due to cited prior art.

Hypothesis 3i: There is no difference between the likelihood that a chemical invention, an electrical invention, a general invention, a pharmaceutical invention, a biotechnological invention, a computer-related invention, and a software invention will be held invalid on obviousness grounds due to uncited prior art.

Hypothesis 3j: There is no difference between the likelihood that a chemical invention, an electrical invention, a general invention, a pharmaceutical invention, a biotechnological invention, a computer-related invention, and a software invention will be held invalid for double patenting.

Hypothesis 3k: There is no difference between the likelihood that a chemical invention, an electrical invention, a general invention, a pharmaceutical invention, a biotechnological invention, a computer-related invention, and a software invention will be held invalid for incorrect inventorship.

Hypothesis 3l: There is no difference between the likelihood that a chemical invention, an electrical invention, a general invention, a pharmaceutical invention, a biotechnological invention, a computer-related invention, and a software invention will be held invalid for failure to disclose best mode.

Hypothesis 3m: There is no difference between the likelihood that a chemical invention, an electrical invention, a general invention, a pharmaceutical invention, a biotechnological invention, a computer-related invention, and a software invention will be held invalid for lack of patentable subject matter.

Hypothesis 3n: There is no difference between the likelihood that a chemical invention, an electrical invention, a general invention, a pharmaceutical invention, a biotechnological invention, a computer-related invention, and a software invention will

be held invalid for claim indefiniteness.

Hypothesis 4: There is no difference between the likelihood that a foreign invented or owned patent and a domestic invented and owned patent will be held invalid.

Hypothesis 5: There is no relationship between the number of cited prior art references in a patent and its invalidity on section 102 prior art or section 103 obviousness grounds.

Hypothesis 5a: There is no relationship between the number of cited patent prior art references in a patent and its invalidity on section 102 prior art or section 103 obviousness grounds.

Hypothesis 5b: There is no relationship between the number of cited non-patent prior art references in a patent and its invalidity on section 102 prior art or section 103 obviousness grounds.

Hypothesis 6: There is no difference between the likelihood that a patent will be held invalid on section 102 prior art or section 103 obviousness grounds because of cited prior art and the likelihood that it will be held invalid on section 102 prior art or section 103 obviousness grounds because of uncited prior art.

Hypothesis 7: Delay between issuance of a patent and decision in the case has no effect on the validity of the patent.

Hypothesis 7a: Delay between issuance of a patent and decision in the case has no effect on the validity of the patent on 102 prior art or section 103 obviousness grounds.

Hypothesis 8: Lag between original application for a patent and decision in the case has no effect on the validity of the patent.

Hypothesis 8a: Lag between original application for a patent and decision in the case has no effect on the validity of the patent on 102 prior art or section 103 obviousness grounds.

Hypothesis 9: Time in prosecution (measured from original application to issue date of current patent) has no effect on the validity of the patent.

Hypothesis 9a: Time in prosecution (measured from original application to issue date of current patent) has no effect on the validity of the patent on 102 prior art or section 103 obviousness grounds.

Hypothesis 10: The decision of the finder of fact in patent cases is not more likely to be affirmed than reversed on appeal.

Hypothesis 10a: The decision of the finder of fact holding a patent valid is not more likely to be affirmed on appeal than the decision of the finder of fact holding a patent

invalid.

Hypothesis 10b: The decision of the finder of fact holding a patent valid is not more likely to be reversed on appeal than the decision of the finder of fact holding a patent invalid.

Hypothesis 10c: The decision of the court in a bench trial in a patent case is not more likely to be affirmed on appeal than the decision of a jury or the decision of a court on pretrial motion holding a patent invalid.

Hypothesis 10d: The decision of the court in a bench trial in a patent case holding a patent valid is not more likely to be affirmed on appeal than the decision of a jury or the decision of a court on pretrial motion holding a patent invalid.

Hypothesis 10e: The decision of the court in a bench trial in a patent case holding a patent invalid is not more likely to be affirmed on appeal than the decision of a jury or the decision of a court on pretrial motion holding a patent invalid.

Hypothesis 11: In cases in which more than one patent is litigated, there is no relationship between the validity of one patent in the case and the validity of remaining patents in the case.

Hypothesis 12: In cases in which more than one patent is litigated, there is no relationship between the invalidity of one patent in the case and the invalidity of remaining patents in the case.

APPENDIX B**Hypothesis 1:**

Let p be the probability of final invalid
 $H_0: p=0.5$

Observed Probability	Count	Z-value	p-value
0.46	300	-1.386	>0.1

Hypothesis 2:

Trial	Category	Probability of Trial Valid	Count
1	Jury	0.671	73
2	Bench	0.573	143
3	Pretrial	0.28	82

Total cases 298

Test of independence:

Statistic	p-value
G-square	< 0.001
Chi-square	< 0.001

Hypothesis 2a (utility):

Trial	Probability of trial valid	Count
1	0.5	2
2	0	2
3	0	1

Total cases 5

Test of independence:

Statistic	p-value
G-square	0.3277
Chi-square	0.3916

Hypothesis 2b (enablement):

Trial	Probability of trial valid	Count
1	0.778	9
2	0.462	26
3	0	1

Total cases 36

Test of independence:

Statistic	p-value
G-square	0.1124
Chi-square	0.1472

Hypothesis 2c (102 prior art):

Trial	Probability of trial valid	Count
1	0.625	24
2	0.659	44
3	0.348	23

Total cases 91

Test of independence:

Statistic	p-value
G-square	0.0418
Chi-square	0.0416

Hypothesis 2d (102 cited prior art):

Trial	Probability of trial valid	Count
1	0.778	9
2	0.824	17
3	0.333	6

Total cases 32

Test of independence:

Statistic	p-value
G-square	0.0818
Chi-square	0.0643

Hypothesis 2e (102 uncited prior art):

Trial	Probability of trial valid	Count
1	0.5	16
2	0.625	32
3	0.353	17

Total cases 65

Test of independence:

Statistic	p-value
G-square	0.1849
Chi-square	0.1883

Hypothesis 2f (102 non-prior art):

Trial	Probability of trial valid	Count
1	0.556	9
2	0.611	36
3	0.04	25

Total cases 70

Test of independence:

Statistic	p-value
G-square	< 0.001
Chi-square	< 0.001

Hypothesis 2g (obviousness):

Trial	Probability of trial valid	Count
1	0.651	43
2	0.622	82
3	0.543	35

Total cases 160

Test of independence:

Statistic	p-value
G-square	0.6042
Chi-square	0.6014

Hypothesis 2h (obviousness due to cited prior art):

Trial	Probability of trial valid	Count
1	0.696	23
2	0.712	52
3	0.5	20

Total cases 95

Test of independence:

Statistic	p-value
G-square	0.2321
Chi-square	0.2190

Hypothesis 2i (obviousness due to uncited prior art):

Trial	Probability of trial valid	Count
1	0.613	31
2	0.578	64
3	0.533	30

Total cases 125

Test of independence:

Statistic	p-value
G-square	0.8198
Chi-square	0.8197

Hypothesis 2j (double patenting):

Trial	Probability of trial valid	Count
1	0	1
2	0.6	5
3	0	1

Total cases 7

Test of independence:

Statistic	p-value
-----------	---------

G-square	0.2429
Chi-square	0.3499

Hypothesis 2k (incorrect inventorship):

Trial	Probability of trial valid	Count
1	1	2
2	0.75	4
3	0.75	4

Total cases 10

Test of independence:

Statistic	p-value
G-square	0.6033
Chi-square	0.7316

Hypothesis 2l (best mode):

Trial	Probability of trial valid	Count
1	0.636	11
2	0.458	24
3	0.2	10

Total cases 45

Test of independence:

Statistic	p-value
G-square	0.1169
Chi-square	0.1301

Hypothesis 2m (patentable subject matter):

Trial	Probability of trial valid	Count
3	0	1

Total cases 1

No test is available

Hypothesis 2n (indefiniteness):

Trial	Probability of trial valid	Count
1	0.25	4
2	0.333	12
3	0.714	7

Total cases 23

Test of independence:

Statistic	p-value
G-square	0.1881
Chi-square	0.1936

Hypothesis 3:

Invention	Category	Probability of final invalid	Count
1	General	0.457	173
2	chemical but not pharm or biotech	0.472	53
3	chem and pharm but not biotech	0.375	8
4	chem and biotech but not pharm	0.667	6
5	chem and pharm and biotech	0	3
6	electrical but not software or computer	0.559	34
7	electrical and software but not computer	1	1
8	electrical and computer but not software	0.3	20
9	electrical and software and computer	0.5	2

Total cases 300

Test of independence:

Statistic	p-value
G-square	0.2620
Chi-square	0.3916

Hypothesis 3a (utility):

Invention	Probability of final invalid	Count
1	0.333	3
8	0	2

Total cases 5

Test of independence:

Statistic	p-value
G-square	0.2764
Chi-square	0.3613

Hypothesis 3b (enablement):

Invention	Probability of final invalid	Count
1	0.267	15
2	0.375	8
3	0.5	2
4	0.8	5
6	0	2
8	0.2	5

Total cases 37

Test of independence:

Statistic	p-value
G-square	0.2055
Chi-square	0.2449

Hypothesis 3c (102 prior art):

Invention	Probability of final invalid	Count
1	0.391	46
2	0.478	23
3	0.5	2
5	0	3
6	0.714	7
8	0.222	9

Total cases 91

Test of independence:

Statistic	p-value
G-square	0.1671
Chi-square	0.2463

Hypothesis 3d (102 cited prior art):

Invention	Probability of final invalid	Count
1	0.143	14
2	0.4	10
3	0.25	4
5	0	3
8	1	1

Total cases 32

Test of independence:

Statistic	p-value
G-square	0.3480
Chi-square	0.4288

Hypothesis 3e (102 uncited prior art):

Invention	Probability of final invalid	Count
1	0.471	34
2	0.467	15
6	0.714	7
8	0.222	9

Total cases 65

Test of independence:

Statistic	p-value
G-square	0.2540
Chi-square	0.2740

Hypothesis 3f (102 non-prior art):

Invention	Probability of final invalid	Count
1	0.566	53
2	0.75	8
4	0	1
6	0.667	9
8	1	1

Total cases 72

Test of independence:

Statistic	p-value
G-square	0.3963
Chi-square	0.5045

Hypothesis 3g (obviousness):

Invention	Probability of final invalid	Count
1	0.33	97
2	0.333	33
3	0.333	3
6	0.526	19
7	1	1
8	0.5	6
9	0	1

Total cases 160

Test of independence:

Statistic	p-value
G-square	0.4121
Chi-square	0.4691

Hypothesis 3h (obviousness due to cited prior art):

Invention	Probability of final invalid	Count
1	0.276	58
2	0.25	20
3	0	2
6	0.538	13
8	1	2

Total cases 95

Test of independence:

Statistic	p-value
G-square	0.0452
Chi-square	0.0594

Hypothesis 3I (obviousness due to uncited prior art):

Invention	Probability of final invalid	Count
1	0.372	78
2	0.348	23
3	0.5	2
6	0.615	13
7	1	1
8	0.5	6
9	0	1

Total cases 125

Test of independence:

Statistic	p-value
G-square	0.4090
Chi-square	0.4809

Hypothesis 3j (double patenting):

Invention	Probability of final invalid	Count
1	1	4
2	0.5	2
3	0	1
6	0	1

Total cases 8

Test of independence:

Statistic	p-value
G-square	0.0501
Chi-square	0.1183

Hypothesis 3k (incorrect inventorship):

Invention	Probability of final invalid	Count
1	0.2	5
2	0	2
3	0	2
8	0	1

Total cases 10

Test of independence:

Statistic	p-value
G-square	0.6828
Chi-square	0.7744

Hypothesis 3l (best mode):

Invention	Probability of final invalid	Count
1	0.385	26
2	0.625	8
3	0.5	2
4	0	2
6	0	4
8	0	3

Total cases 45

Test of independence:

Statistic	p-value
G-square	0.0606
Chi-square	0.1689

Hypothesis 3m (patentable subject matter):

Invention	Probability of final invalid	Count
9	1	1

Total cases 1

No test is available

Hypothesis 3n (indefiniteness):

Invention	Probability of final invalid	Count
1	0.231	13
2	0.4	5
3	0	1
4	1	1
6	0	1

8	1	2
---	---	---

Total cases 23

Test of independence:

Statistic	p-value
G-square	0.1113
Chi-square	0.1837

Hypothesis 3' (with respect to general, chemical, electrical categories only):

Invention	Category	Probability of final invalid	Count
1	general	0.457	173
2	chemical	0.457	70
3	electrical	0.473	57

Total cases 300

Test of independence:

Statistic	p-value
G-square	0.9738
Chi-square	0.9738

Hypothesis 3a (utility):

Invention	Probability of final invalid	Count
1	0.333	3
3	0	2

Total cases 5

Test of independence:

Statistic	p-value
G-square	0.2764
Chi-square	0.3613

Hypothesis 3b (enablement):

Invention	Probability of final invalid	Count
1	0.267	15

2	0.533	15
3	0.143	7

Total cases 37

Test of independence:

Statistic	p-value
G-square	0.1284
Chi-square	0.1362

Hypothesis 3c (102 prior art):

Invention	Probability of final invalid	Count
1	0.391	46
2	0.414	29
3	0.438	16

Total cases 91

Test of independence:

Statistic	p-value
G-square	0.9447
Chi-square	0.9445

Hypothesis 3d (102 cited prior art):

Invention	Probability of final invalid	Count
1	0.143	14
2	0.313	16
3	0.5	2

Total cases 32

Test of independence:

Statistic	p-value
G-square	0.3948
Chi-square	0.3951

Hypothesis 3e (102 uncited prior art):

Invention	Probability of final invalid	Count
1	0.471	34
2	0.467	15
3	0.438	16

Total cases 65

Test of independence:

Statistic	p-value
G-square	0.9753
Chi-square	0.9753

Hypothesis 3f (102 non-prior art):

Invention	Probability of final invalid	Count
1	0.566	53
2	0.667	9
3	0.7	10

Total cases 72

Test of independence:

Statistic	p-value
G-square	0.6529
Chi-square	0.6591

Hypothesis 3g (obviousness):

Invention	Probability of final invalid	Count
1	0.33	97
2	0.333	36
3	0.519	27

Total cases 160

Test of independence:

Statistic	p-value
G-square	0.1910
Chi-square	0.1806

Hypothesis 3h (obviousness due to cited prior art):

Invention	Probability of final invalid	Count
1	0.276	58
2	0.227	22
3	0.6	15

Total cases 95

Test of independence:

Statistic	p-value
G-square	0.0408
Chi-square	0.0328

Hypothesis 3I (obviousness due to uncited prior art):

Invention	Probability of final invalid	Count
1	0.372	78
2	0.346	26
3	0.571	21

Total cases 125

Test of independence:

Statistic	p-value
G-square	0.2140
Chi-square	0.2076

Hypothesis 3j (double patenting):

Invention	Probability of final invalid	Count
1	1	4
2	0.333	3
3	0	1

Total cases 8

Test of independence:

Statistic	p-value
G-square	0.0339
Chi-square	0.0759

Hypothesis 3k (incorrect inventorship):

Invention	Probability of final invalid	Count
1	0.2	5
2	0	4
3	0	1

Total cases 10

Test of independence:

Statistic	p-value
G-square	0.4729
Chi-square	0.5738

Hypothesis 3l (best mode):

Invention	Probability of final invalid	Count
1	0.385	26
2	0.5	12
3	0	7

Total cases 45

Test of independence:

Statistic	p-value
G-square	0.0261
Chi-square	0.0800

Hypothesis 3m (patentable subject matter):

Invention	Probability of final invalid	Count
3	1	1

Total cases 1

No test is available

Hypothesis 3n (indefiniteness):

Invention	Probability of final invalid	Count
1	0.231	13
2	0.429	7
3	0.667	3

Total cases 23

Test of independence:

Statistic	p-value
G-square	0.3174
Chi-square	0.3118

Hypothesis 4:

Foreign	Probability of final invalid	Count
0	0.467	259
1	0.415	41

Total cases 300

Test of independence:

Statistic	p-value
G-square	0.5294
Chi-square	0.5305

Hypothesis 5:

The analysis is based on the logistic regression model on the log odds of final art invalid.

Total cases 202

Predictor: total number of cited prior art

p-value for the predictor is 0.5431

Hypothesis 5a:

Predictor: total number of cited patent prior art

p-value for the predictor is 0.3532

Hypothesis 5b:

Predictor: total number of cited non-patent prior art

p-value for the predictor is 0.3891

Hypothesis 6:

Cited art	Probability of final invalid	Count
0	0.408	152
1	0.296	115

Total cases 267

Test of independence:

Statistic	p-value
G-square	0.0572
Chi-square	0.0584

Hypothesis 7:

The analysis is based on the logistic regression model on the log odds of final invalid.

Total cases 300

Predictor: delay

p-value for the predictor is 0.7309

Hypothesis 7a:

The analysis is based on the logistic regression model on the log odds of final art invalid.

Total cases 202

Predictor: delay

p-value for the predictor is 0.8175

Hypothesis 8:

The analysis is based on the logistic regression model on the log odds of final invalid.

Total cases 300

Predictor: lag

p-value for the predictor is 0.7325

Hypothesis 8a:

The analysis is based on the logistic regression model on the log odds of final art invalid.

Total cases 202

Predictor: lag

p-value for the predictor is 0.8206

Hypothesis 9:

The analysis is based on the logistic regression model on the log odds of final invalid.

Total cases 300

Predictor: Time in prosecution

p-value for the predictor is 0.2199

Hypothesis 9a:

The analysis is based on the logistic regression model on the log odds of final art invalid.

Total cases 202

Predictor: time in prosecution

p-value for the predictor is 0.9832

Hypothesis 10:

Let p be the probability of affirmed

$H_0: p=0.5$

Observed Probability	Count	Z-value	p-value
0.834	151	8.219	<0.01

Hypothesis 10a & b:

Trial invalid	Probability of affirmed	Count
0	0.897	68
1	0.783	83

Total cases 151

Test of independence:

Statistic	p-value
G-square	0.0563
Chi-square	0.0609

Hypothesis 10c:

Factor Level	Category
1	Bench
2	Jury
3	Pretrial

Level	Probability of affirmed	Count
-------	-------------------------	-------

1	0.847	72
2	0.867	45
3	0.758	33

Total cases 150

Test of independence:

Statistic	p-value
G-square	0.4246
Chi-square	0.4018

Hypothesis 10d (given trial valid):

Level	Probability of affirmed	Count
1	0.938	32
2	0.821	28
3	1	7

Total cases 67

Test of independence:

Statistic	p-value
G-square	0.1632
Chi-square	0.2163

Hypothesis 10e (given trial invalid):

Level	Probability of affirmed	Count
1	0.775	40
2	0.941	17
3	0.692	26

Total cases 83

Test of independence:

Statistic	p-value
G-square	0.1082
Chi-square	0.1512

Hypothesis 11 & 12

Let p be the probability of agreement
 $H_0: p=0.5$

Observed Probability	Count	Z-value	p-value
0.864	44	4.824	<0.001

The logistic regression:

Modeled Response: The log odds of invalidity (i.e.)

$$\log \frac{\Pr(\text{invalid})}{1 - \Pr(\text{invalid})}$$

Only the trial factor (trial to jury, trial to bench, pretrial disposition) is significant.

Trial Factor	Definition
TRIAL_F(1)	trial to jury
TRIAL_F(2)	trial to bench
BASELINE	pretrial disposition

Trial Factor	Coefficients
TRIAL_F(1)	-1.0245
TRIAL_F(2)	-0.9439
Constant	0.5500

Trial Factor	EXP(Coefficient)	Interpretation
TRIAL_F(1)	0.3590	The odds that a jury decision in a patent holding the patent final invalid is 0.3590 times the odds that a pretrial disposition in a patent holding the patent final invalid
TRIAL_F(2)	0.3891	The odds that a bench trial in a patent holding the patent final invalid is 0.3891 times the odds that a pretrial disposition in a patent holding the patent final invalid

----- Variables in the Equation -----

Variable	B	S.E.	Wald	df	Sig	R	Exp(B)
TRIAL_F			13.1926	2	.0014	.1492	
TRIAL_F(1)	-1.0245	.3324	9.4981	1	.0021	-.1348	.3590
TRIAL_F(2)	-.9439	.2854	10.9414	1	.0009	-.1472	.3891
Constant	.5500	.2293	5.7554	1	.0164		

Classification Table for F_IVD

		Predicted				Percent Correct
		.00	1.00			
Observed		0	1			
.00	0	I 131	I 30			81.37%
	1	I 86	I 52			37.68%
						Overall 61.20%