2012

Patent Litigation and the Internet

Samantha Zyontz

John R. Allison

Emerson H. Tiller

Tristan Bligh

Stanford Technology Law Review

Patent Litigation and the Internet*

JOHN R. ALLISON,† EMERSON H. TILLER,‡ SAMANTHA ZYONTZ,§
TRISTAN BLIGH**

CITE AS: 2012 STAN. TECH. L. REV. 3
http://stlr.stanford.edu/pdf/allison-patent-litigation.pdf

I. INTRODUCTION

¶1 Patent infringement litigation has not only increased dramatically in frequency over the past few decades,[1] but also has also seen striking growth in both stakes and cost.[2] Although a relatively rich literature has added much to our understanding of the nature, causes, and consequences of patent litigation during the past two decades,[3] many interesting questions remain inadequately addressed.

 We wish to acknowledge the helpful comments on our preliminary data by attendees at the 2011 University of San Diego Law School conference on empirical patent research, particularly Stuart Graham, Ted Sichelman, and Polk Wagner, as well as the various kinds of data assistance received from Mohammad Morovati, Ph.D. candidate in economics, University of Texas at Austin.

† Spence Centennial Professor of Business Administration, and Professor of Intellectual Property, McCombs Graduate School of Business, University of Texas at Austin.

‡ J. Landis Martin Professor of Law and Business, Northwestern University School of Law.

§ Senior Research Associate, Searle Civil Justice Institute, Law and Economics Center, George Mason University School of Law.

** JD 2010 Northwestern University Law School.

[1] *See, e.g.*, Alan C. Marcos & Ted M. Sichelman, *Do Economic Downturns Dampen Patent Litigation?* 7 (5th Annual Conference on Legal Studies Working Paper, July 16, 2010), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1641425 (showing the dramatic increase in such litigation since 1970 in the context of examining the effects of economic downturns on patent infringement litigation, with the exception of the most recent recession).

[2] In its 2001 economic survey, the American Intellectual Property Law Association (AIPLA) reported that 251 law firms had responded to its request for data on patent infringement litigation costs for cases with stakes exceeding $25 million, and that the median cost per party through the end of discovery in such cases was $1.5 million. AMERICAN INTELLECTUAL PROPERTY LAW ASSOCIATION, REPORT OF THE ECONOMIC SURVEY 2001, at 85. Eight years later, AIPLA reported that 398 law firms had responded to its inquiry about patent cases with stakes exceeding $25 million, and that the median cost per party through the end of discovery in such cases was $3 million. AMERICAN INTELLECTUAL PROPERTY LAW ASSOCIATION, REPORT OF THE ECONOMIC SURVEY 2009, at I-129.

[3] In its 2001 economic survey, the American Intellectual Property Law Association (AIPLA) reported that 251 law firms had responded to its request for data on patent infringement litigation costs for cases with stakes exceeding $25 million, and that the median cost per party through the end of discovery in such cases was $1.5 million. AMERICAN INTELLECTUAL PROPERTY LAW ASSOCIATION, REPORT OF THE ECONOMIC SURVEY 2001, at 85. Eight years later, AIPLA reported that 398 law firms had responded to its inquiry about patent cases with stakes exceeding $25 million, and that the median cost per party through the end of discovery in such cases was $3 million. AMERICAN INTELLECTUAL PROPERTY LAW ASSOCIATION, REPORT OF THE ECONOMIC SURVEY 2009, at I-129.

[3] *See generally* Jay P. Choi, *Patent Pools and Cross-Licensing in the Shadow of Patent Litigation*, 51 INT'L ECON. REV. 441 (2010) (analyzing the incentives to form a patent pool or engage in cross-licensing arrangements in the presence of uncertainty about the validity and coverage of patents); John R. Allison & Mark A. Lemley, *Empirical Evidence on the Validity of Litigated Patents*, 26 AIPLA Q.J. 185 (1998) (hereinafter Allison & Lemley, *Empirical Evidence on the Validity of Litigated Patents*) (analyzing findings of patent validity and invalidity in litigation); John R. Allison, Mark A. Lemley & Joshua W. Walker, *Patent Quality and Settlement Among Repeat*

The nuances of and trends in patent litigation in different technology fields and industries, for example, are still understudied.[4] Litigation of patents on new technologies has likewise received a dearth of attention. Here we seek to help begin filling these gaps by empirically analyzing the phenomenon in a very particular context: the litigation of Internet patents. In particular, we study litigation of patents on Internet business processes issued during the first few years in which such patents were granted, and determine whether it differs in meaningful ways from litigation of patents in other fields.

¶2 Patents on methods of doing business on the Internet (Internet patents) have been the subject of intense debate and criticism for a number of years.[5] Indeed, since 1998, when the Court of Appeals for the Federal Circuit held that there was no per se exclusion of these Internet-implemented methods from the realm of patentable subject matter,[6] many have questioned the wisdom of the decision and sought to have its result altered.[7] In addition to the issue of subject matter eligibility for patenting, critics have questioned whether the U.S. Patent and Trademark Office (PTO) has improvidently granted patents on Internet processes that appear at first glance to be obvious, thus failing one of the key requirements for patentability.[8] The importance of the Internet as a rapidly growing commercial platform combined with concerns for an open and free Internet added to the intensity of these debates.

¶3 Despite the importance of these debates, there has been relatively little empirical study of Internet patents, and no study of the increasing litigation in which they have been involved. In one of the few studies of the patents themselves, Allison and Tiller analyzed the quality of Internet patents

---

Patent Litigants, 99 GEO. L.J. 677 (2011) (hereinafter Allison, Lemley & Walker, *Repeat Patent Litigants*) (empirically studying patent characteristics and litigation outcomes when patent owners are repeat players in litigation); John R. Allison, Mark A. Lemley & Joshua Walker, *Extreme Value or Trolls on Top: Characteristics of the Most-Litigated Patents*, 158 U. PA. L. REV. 1, 20-22 (2010) (hereinafter Allison, Lemley & Walker, *Trolls on Top*) (empirically finding that patents litigated 8 or more times during a 9-year period from Jan. 1, 2000 to Feb. 28, 2009 possessed much greater value-indicating characteristics than a sample of patents litigated only once during the same period); John R. Allison, Mark A. Lemley, Kimberly A. Moore & Derek Trunkey, *Valuable Patents*, 92 GEO. L.J. 435 (2004) (hereinafter ALMT) (finding empirical evidence that litigated patents are marked by several important signs of private economic value); John R. Allison & Thomas W. Sager, *Valuable Patents Redux: On the Enduring Merit of Using Patent Characteristics to Identify Valuable Patents*, 85 TEX. L. REV. 1769 (2007) (hereinafter Allison & Sager) (defending the statistical power of the results in ALMT); Jean O. Lanjouw & Mark Schankerman, *Characteristics of Patent Litigation: A Window on Competition*, 32 RAND J. ECON. 129 (2001) (demonstrating substantial variation among patents in exposure to litigation risk); Jean O. Lanjouw & Mark Schankerman, *Stylized Facts of Patent Litigation: Value, Scope and Ownership* (Nat'l Bur. Econ. Res. Working Paper No. W6297, 1997, rev. 2000) (empirically analyzing factors contributing to patent infringement litigation); Jean O. Lanjouw & Josh Lerner, *The Enforcement of Intellectual Property Rights: A Survey of the Empirical Literature*, 49/50 ANNALES D'ECONOMIE ET DE STATISTIQUE 223-46 (Jan./June 1998) (surveying the empirical literature on the link between patent value and litigation); Michael J. Meurer, *The Settlement of Patent Litigation*, 20 RAND J. ECON. 77 (1989) (modeling the role of litigation in inducing patent infringement dispute settlement); Kimberly A. Moore, *Forum Shopping in Patent Cases: Does Geographic Choice Affect Innovation?*, 79 N.C. L. REV. 889 (2001) (studying the nature and effects of forum shopping in patent infringement litigation); Kimberly A. Moore, *Judges, Juries, and Patent Cases—An Empirical Peek Inside the Black Box*, 99 MICH. L. REV. 365 (2000) (hereinafter Moore, *Inside the Black Box*) (empirically analyzing the performance of judges and juries in patent infringement cases); Kimberly A. Moore, *Jury Demands: Who's Asking?*, 17 BERKELEY TECH. L.J. 847 (2002) (empirically analyzing jury demands in patent cases); Kimberly A. Moore, *Populism and Patents*, 82 N.Y.U. L. REV. 69 (2007) (hereinafter Moore, *Populism and Patents*) (empirically studying the relative success of individual and corporate patent infringement plaintiffs and finding significant jury bias in favor of individual plaintiff-inventors); Kimberly A. Moore, *Xenophobia in American Courts*, 97 NW. U. L. REV. 1497 (2003) (hereinafter Moore, *Xenophobia*) (finding empirical evidence of anti-foreign bias in patent infringement litigation); Deepak Somaya, *Strategic Determinants of Decisions not to Settle Patent Litigation*, 24 STRAT. MGMT. J. 17 (2003) (analyzing firms' decisions whether to settle patent litigation from strategic management perspective).

[4] However, Ziedonis has analyzed the trend of litigation rates in the semiconductor industry. Rosemarie H. Ziedonis, *Patent Litigation in the U.S. Semiconductor Industry*, *in* PATENTS IN THE KNOWLEDGE-BASED ECONOMY 180 (Wesley M. Cohen & Stephen A. Merrill eds. 2003).

[5] For a more detailed discussion of the criticism and debate, see John R. Allison & Emerson H. Tiller, *The Business Method Patent Myth*, 18 BERKELEY TECH. L.J. 989 (2003) (hereinafter Allison & Tiller) (after detailing the debate, empirically comparing characteristics of a large set of Internet business method patents with random sample of contemporaneously issued patents from the general population).

[6] State St. Bank v. Signature Fin. Grp., Inc., 149 F.3d 1368, 1375-76 (Fed. Cir. 1998). The court's decision applied to business processes whether Internet-implemented or not, but the most important and controversial of these have involved Internet implementations. In *Bilski v. Kappos*, 130 S. Ct. 3218 (2010), the Supreme Court agreed that there was no patentability exclusion for such processes. The Court disagreed with the legal standard employed by the Federal Circuit in *State Street*, but upheld the absence of a patentability exclusion. *Id.* at 3231.

[7] See Allison & Tiller, *supra* note 5, for a detailed discussion of the criticisms of the *State Street* decision and efforts to alter it.

[8] 35 U.S.C. § 103 (2006); *see, e.g.*, Rochelle Cooper Dreyfuss, *Are Business Method Patents Bad for Business*, 16 Santa CLARA COMPUTER & HIGH TECH L.J. 263, 269-70 (2000).

by comparing them to other kinds of patents.[9] They built a data set of 1,093 Internet-implemented process patents and compared many of their characteristics with those of a randomly selected set of 1,000 contemporaneously issued patents from the general population of patents (non-Internet patents, or NIPs).[10] Their purpose was to test empirically the merits of the many criticisms of these patents, all of which had been made without the support of any data. Measuring a number of characteristics that previous research had associated with patent quality and private economic value, such as the total number of claims and prior art references, as well as several other characteristics first employed by the authors, the study found that Internet patents appeared to be of higher average quality and value than the average patent.[11]

¶4 Researchers in economics and law have compiled evidence to support a link between the economic value of patents to their owners (private patent value) and litigation propensity.[12] In 2004 Allison, Lemley, Moore, and Trunkey (ALMT) conducted the most comprehensive comparison yet made of litigated and unlitigated patents.[13] Measuring a number of patent characteristics that had been linked to litigation propensity and private economic value, as well as some characteristics not previously considered but that logically might suggest value and a greater likelihood of litigation, the authors found that litigated patents were a completely different breed than those that had not been involved in litigation. With a high degree of significance, the authors found that litigated patents, compared with unlitigated ones, contained more claims and more references to prior U.S. patents, foreign patents, and other kinds of publications ("nonpatent prior art"), and were cited more often as prior art by subsequent patents (i.e., they had more "forward citations"). They were disproportionately represented in some technology areas and in some industries. They were also much more likely to have originally been issued to individuals and small businesses, and to be owned by domestic rather than foreign entities. Further, they had spent much more time in "prosecution" (examination within the PTO) from their original filing dates than unlitigated patents, primarily as a result of patent applicants having invested more in continuing applications leading to the generation of multiple patents on closely related inventions—a patent portfolio that can have greater value than the sum of its parts.[14]

¶5 Allison and Tiller's findings revealed that Internet patents, at least those issued during the formative years of digital commerce, possessed value-indicating characteristics very similar to those of the litigated patents subsequently studied by ALMT. Given that these characteristics suggest not only private value but also litigation propensity, the next logical questions are whether these Internet patents later experienced unusually high rates of litigation, and how these patents fared in court compared with other patents. The current study contributes to the empirical literature on patent infringement litigation by comparing litigation rates and outcomes for early Internet patents with those for a large comparison set of contemporaneously issued NIPs. Delving more deeply, we further investigate litigation rates and outcomes for two subgroups of Internet patents identified by Allison and Tiller—those covering relatively broad Internet business "models"[15] and those covering

[9] Allison & Tiller, *supra* note 5.

[10] *Id.*

[11] *Id.* at 1003. Allison & Tiller also found that these early Internet patents appeared to be of greater quality and value than patents in many individual technology fields. *Id.*

[12] See sources cited in ALMT, *supra* note 3, at 448-51.

[13] *Id.* The authors employed a data set of approximately 7,000 patents that were the subject of patent infringement lawsuits that terminated in 1999 and 2000. *Id.* at 444 n.33. The authors did not identify how many times those patents had been litigated, and thus it can only be said that they were litigated at least once. The sample of litigated patents was compared with a randomly selected set of unlitigated patents from the general patent population.

[14] *Id.* at 438. Either counting the number of continuing applications or counting the number of patents issuing from the same original application allows us to view decisions of patent applicants to invest more in creating families of patents on related technologies and more closely tailoring their patents to potentially infringing activities in an industry. Each of these techniques picks up similar information. *Id.* at 457-60. On the value of patent portfolios, see, e.g., Gideon Parchomvosky & R. Polk Wagner, *Patent Portfolios*, 154 U. PA. L. REV. 1 (2005).

[15] Such a patent contains claims of a sufficiently broad nature that the invention could form the foundation for an entire business, or line of business, on the Internet. The Priceline.com reverse auction model for purchasing airline tickets was an early example of this type of patent on a broad business model. Method, Apparatus, & Program for Pricing, Selling, & Exercising Options to Purchase Airline Tickets, U.S. Patent No. 5,797,127 (filed Dec. 31, 1996). Another example of a patent of this kind is

narrower Internet business "techniques."[16] The important difference between these subgroups is that the claim language in those Internet patents classified as business models is usually broader—that is, more general—than in the other subgroup. More general claim language tends to increase the universe of potential infringers, thus creating at least a possibility of higher litigation rates and perhaps greater win rates.[17]

¶6 Using both univariate[18] comparisons and multiple regression techniques, we find primarily that: (1) Internet patents and their two subtypes were litigated at a far higher rate than NIPs—they were between 7.5 and 9.5 times more likely to end up in infringement litigation, depending on the model we used. (2) Within the category of Internet patents, those on business models were litigated at a significantly higher rate than those on business techniques. (3) Across both Internet patents and NIPs, patents issued to small entities, especially individuals and small businesses, were much more likely to be litigated than those issued to large entities; (4) Patents of all kinds with more independent claims were significantly more likely to be litigated than those with fewer independent claims. (5) Including both Internet patents and NIPs, litigated patents received many more forward citations—citations received from later patents—than did unlitigated patents. (6) Patents issued to foreign entities were significantly less likely to be litigated than patents issued to U.S. entities. (7) The more time that an application for an Internet patent or NIP had spent in the PTO prior to issuance, the more likely it was that the patent granted from that application was to be involved in infringement litigation. (8) There was no difference in the ages of Internet patents and NIPs when they became the subject of litigation—both kinds were about 4.5 years old; (9) Once patent infringement litigation was initiated, the owners of litigated Internet patents were significantly more likely to settle before judgment than the owners of litigated NIPs (especially when *probable* settlements were taken into account along with *obvious* settlements, which we believe is the more accurate metric). (10) Across both sets of patents, the larger the number of potential infringers involved in a case (defendants in infringement actions and plaintiffs in declaratory judgment actions), the less likely the case was to settle. (11) Internet patents and NIPs went to trial at about the same rate. (12) When failing to settle, the owners of NIPs won on the merits at a significantly higher rate than did owners of Internet patents—although the win rate for NIP owners was quite low at around 16%, the win rate of Internet patents was even lower by a substantial margin. This finding did not hold up in regression analysis, however; when the effects of other variables were taken into account in a logistic regression analysis, there was no significant difference in the win rate for accused infringers between Internet patents and NIPs. Accused infringers did win more often when Internet patents were asserted against them than when they defended against NIP complaints, but the relatively small number of observations prevented the difference from being statistically significant. (13) Surprisingly, owners of both kinds of patents were significantly more likely to win as the number of inventors on the patents increased. (14) The longer that applications for Internet patents and NIPs had spent in the PTO before issuance, the less likely accused infringers were to win. (15) Accused infringers were less likely

---

Hyperlinks Resolution at & by a Spec. Network Server in order to Enable Diverse Sophisticated Hyperlinking upon a Digital Network, U.S. Patent No. 6,006,265 (filed Apr. 2, 1998), which is a method for distributing hypervideo, *i.e.*, digital video broadcasting incorporating hyperlinks.

[16] Such a patent covers an invention narrower in scope and intended not as a business model, but only as a means of solving a specific business problem. Amazon.com's patent on the "one-click" technique for merchandise ordering was an early prototype of the narrower business technique type of patent. Method & Sys. for Placing a Purchase Order via a Commc'ns Network, U.S. Patent No. 5,960,411 (filed Sept. 12, 1997). Another example of such a patent is Rights Mgmt. Sys. for Digital Media, U.S. Patent No. 6,006,332 (filed Sept. 12, 1997), which provides a method for controlling unauthorized access to copyrighted material distributed by content owners over the Internet.

[17] The breadth of patent claims is a function of relative language generality or specificity (fewer or more "limitations" in patent law jargon), and is only determined once a court has construed the language of disputed terms in patent claims. Claim construction (interpretation) occurs prior to trial in an infringement case, and is a function to be performed solely by a judge, not a jury. Hearings in which claim construction takes place are called "Markman hearings," after Markman v. Westview Instruments, 517 U.S. 370 (1996), in which the Supreme Court held that claim construction is a question of law that should only be determined by a judge. Subsequently, the U.S. Court of Appeals for the Federal Circuit concluded that a trial court's claim construction decisions are reviewable under a de novo standard. Cybor Corp. v. FAS Technologies, Inc., 138 F.3d 1448 (Fed. Cir. 1998) (en banc). To our knowledge, no one has yet devised a means to empirically estimate the breadth of patent claims.

[18] Univariate statistics are the results of the comparison of the same variable between two or more data sets. Some statisticians use the term "bivariate" statistics for the same procedure.

to win on the merits when the Internet patents or NIPs asserted against them had been litigated previously. (16) Across both sets of patents, the larger the number of potential infringers involved in a case, the more likely these potential infringers were to win a judgment on the merits. That is, the more infringement defendants per case, the more likely these defendants were to win. (17) There was no difference between the different types of patents in the percentage of cases that were terminated for procedural reasons. We also discuss a few other findings of interest.

¶7 These findings raise some intriguing questions for further research: (1) How does litigation of patents on other relatively new technology fields compares with litigation of patents from the general population, the latter consisting predominantly of patents on more mature technologies? (2) How might litigation of patents on different technologies early in their maturation periods compare to each other? (3) Are patents on young technologies likely to be stronger and more valuable on average because there is less relevant prior art to require the narrowing of patent claim language or because applicants perceive greater potential innovation importance and are willing to invest more in the patenting enterprise, thus leading to more litigation? (4) Do patents on young technologies generate more uncertainty because of their newness, contributing to more contention? (5) The Internet patents in our data set, whether litigated or not, showed many of the same internal characteristics as patents in all fields that wind up in litigation.[19] Is the same true of patents in other emerging fields of patenting activity such as flash memory, smart phones, nanotechnology, and others?

¶8 We also posit other questions for future research that are not limited to the context of new patent fields. For instance, our current study found that the average number of defendants per case significantly decreased the odds of settlement and increased the odds of a win on the merits by defendants (the term "defendants" here referring to "potential infringers" whether they are actually defendants in infringement cases or plaintiffs in declaratory judgment actions who then become infringement counterclaim defendants). This confirms recent findings by Allison, Lemley, & Walker, and may indicate that such a result is generalizable.[20] Thus, an investigation empirically probing the wisdom of patent owners' litigation strategies when deciding whether to sue multiple alleged infringers in a single lawsuit or in several different ones may bear interesting fruit in other contexts.[21] In addition, given our finding of significantly different litigation rates for patents on online business models and online business techniques, the former clearly appearing to have broader claims, another research question is whether one might devise methods to create empirical estimates of patent claims' breadth for use in better predicting which patents companies should fear when they assess how much freedom of action they have to innovate in a given field. The role of so-called "non-practicing entities" (NPEs) also deserves more research attention. These are companies that do not make or sell products and thus are not vulnerable to patent infringement counterclaims, as are product companies that sue for infringement. NPEs consequently may be less reluctant to sue.[22] There is a significant but imperfect correlation between the fact that a patent was originally issued to a small entity and the identity of a patent infringement plaintiff as an NPE, and in the current study one of our findings is that patents granted to small entities, especially to individuals and small businesses, are much more likely to be litigated.[23] Thus it may be worthwhile to investigate the NPE/product company question in future studies of patent litigation over other types of patents.

---

[19] Allison & Tiller, *supra* note 5.

[20] *See* Allison, Lemley & Walker, *Repeat Patent Litigants*, *supra* note 3, at 699.

[21] The Leahy-Smith America Invents Act will itself have a meaningful effect on the litigation strategies of patent owners in that one may no longer name large numbers of unrelated defendants based on the allegation that they all have infringed the same patent. Patent owners can sue two or more defendants in the same infringement action only when the defendants' alleged infringements arise out of "the same transaction, occurrence, or series of transactions or occurrences" and questions of fact common to all defendants will arise in the case. Pub. L. 112–29, 125 Stat. 33 § 299(a) (Sept. 16, 2011). There nevertheless will be situations in which a patent owner may sue multiple defendants for infringement of the same patent, such as an action against a manufacturer, distributors, and retailers of the same allegedly infringing product, and perhaps an action against the maker of a patented component and makers and sellers of end products that incorporate the component. In such cases, patent owners will continue to face choices about the best litigation strategy.

[22] *See, e.g.*, Robert A. Matthews, Jr., *Legal Nuances When A Patent-Holding Company Seeks To Enforce A U.S. Patent*, 49 IDEA 549, 551 (2009).

[23] *See infra* at Tables 2, 3 and accompanying text.

¶9 In Part II we describe our data. In Part III we first make univariate statistical comparisons of litigation rates between Internet patents and NIPs, between the two subgroups of Internet patents and NIPs, and between the Internet patent subgroups themselves. We then use multivariate logistic regression to determine whether being an Internet patent or being within one of the Internet patent subgroups still contributes to the likelihood of being litigated after we controlled for other patent characteristics found to be positively or negatively associated with litigation by previous studies. Part IV expands our study by analyzing more closely a number of other case-specific litigation variables for both Internet patents and NIPs. We conclude and revisit potential future research questions in Part V.

II. DESCRIPTION OF THE DATA

¶10 Among other data, we employ the data set of Internet patents from authors Allison and Tiller.[24] That database included 1,093 patents drawn from PTO "data processing" classifications 705, 707, and 709[25] that had issued through the end of 1999, most having 1998 and 1999 issue dates.[26] These patents covered business processes clearly intended for use on the Internet. We then employed the Derwent LitAlert database ("LitAlert") to identify patents from that data set that had been litigated through April 2009. We also made use of Stanford's Intellectual Property Litigation Clearinghouse (IPLC) to verify and find additional cases filed after January 1, 2000, (when IPLC data begins) through April 2009.[27] The earliest filed lawsuit alleging infringement of an Internet patent from our data set was initiated in 1998. We thus captured litigation during a period of approximately ten years after the last patents in our data set were issued and longer for some of them.[28] Most patent litigation occurs when patents are relatively young.[29] Thus, if a patent is going to end up in infringement litigation at all, it usually does so relatively early in its life. The reason is that, with the exception of pharmaceutical patents, those patents that do have any private economic value typically experience a diminution of that value long before the expiration of the term of protection. At the date of lawsuit filing, the average age of all litigated patents in our data set, both Internet and NIPs, is 4.52 years. Of the 1,093 Internet patents, 111 were litigated and 982 were unlitigated.

¶11 The first portion of the data set of NIPs came from ALMT's 2004 study of litigated patents, which included a random sample of 1,000 unlitigated NIPs that were issued between mid-1996 and mid-1998 and had *not* been the subject of litigation, plus a random sample of 300 NIPs that were issued between mid-1996 and mid-1998 and *had* been the subject of infringement litigation that terminated during 1999-2000 (litigated NIPs).[30] For purposes of the current analysis, some modifications to the original data set were made from updated information. Ten of the original 1,000 unlitigated NIPs were deleted because they had since been litigated. Of the original 300 litigated NIPs, one was dropped because it appeared twice in our dataset. Thus, 990 unlitigated NIPs and 299 litigated NIPs remained from the original ALMT dataset. Both of these sets included patents issued

[24] The laborious process by which this data set was built is described in detail in Allison & Tiller, *supra* note 5, at 1032-36.

[25] *See* U.S.P.T.O. Classification 705, Data Processing: Fin., Bus. Practice, Mgmt., or Cost/Price Determination (Feb. 2011), http://www.uspto.gov/web/patents/classification/uspc705/defs705.htm; U.S.P.T.O. Classification 707, Data Processing: Database, Data Mining, and File Mgmt. or Data Structures (Feb. 2011), http://www.uspto.gov/web/patents/classification/uspc707/defs707.htm; U.S.P.T.O. Classification 709, Elec. Computers and Digital Processing Sys.: Multicomputer Data Transferring (Jan. 2011), http://www.uspto.gov/web/patents/classification/uspc709/defs709.htm.

[26] The earliest issue date of patents in the data set was Apr. 16, 1996.

[27] Stanford Law School recently spun off the IPLC to a private company, Lex Machina, in which Stanford is a large stakeholder. The web site for the company is http://lexmachina.com.

[28] Depending on when the patent in our data set was issued, the period between issuance and our cut-off date ranged from 9.3 to 12 years.

[29] *See, e.g.*, ALMT, *supra* note 3, at 460. Pharmaceutical patents appear to be an exception because FDA approval requires years, they are likely to have value until the very end of their terms of protection, and on average are litigated later in their lives than other patents. *Id.* at 475. *See also* Allison & Tiller, *supra* note 5, at 1066.

[30] ALMT, *supra* note 3, at 445-48.

between June 1996 and May 1998, and the litigated patents were involved in cases that terminated in 1999 or 2000.

¶12 However, since we included every Internet patent case we could find that was the subject of an infringement lawsuit filed through April 2009, for comparability we also needed a set of litigated NIPs in cases that were restricted to those in which litigated terminated in 1999-2000 as were the 299 from ALMT. To identify a set of litigated NIPs similar to the 299 from ALMT, we identified the first patent number issued in the beginning of June 1996 (5,522,091) and searched through the last patent number issued in May 1998 (5,758,361). We then randomly generated 5,000 patent numbers from between the two end-points using a uniform distribution. To identify which of these NIPs had been litigated, we searched LitAlert in the same manner as for the Internet patents. The search generated 55 litigated NIPs, or about a 1 percent NIP litigation rate. Because we wanted a similar number of litigated NIPs and litigated Internet patents, we searched an additional 5,000 randomly selected NIPs assuming that the 1 percent litigation rate would be constant. This second set of 5,000 was generated by adding 10 to each patent number in the original set of 5,000 patents, *e.g.*, 5,289,210 became 5,289,220. The additional set did not contain duplicates of any patents in the first list of 5,000 patents.[31] We checked this final list against the NIPs used from ALMT and against the Internet patents to ensure that the new chosen NIPs were unique. To find the litigated NIPs from the 10,000 patents we again relied on the information available in the LitAlert database and in the IPLC. These searches resulted in a total 136 litigated NIPs.

¶13 For each patent we then searched for the specific cases in which it was the subject of litigation using LitAlert and the IPLC. We dropped a small number of cases because of incorrect associations in LitAlert between cases and patent numbers that we were unable to correct. We also filtered out cases that had not yet terminated so that we could examine outcomes. The result was 365 unique terminated cases, of which 196 involved at least one NIP, 167 involved at least one Internet patent, and two involved both kinds of patents. Because some patents were litigated in more than one distinct case, and because some cases involved the litigation of more than one patent, we treated each assertion of a distinct patent in a distinct lawsuit as an observation. Our final data set includes 453 assertions, or "case-patent pairs."[32] To prevent having duplicate cases in our data, we used only the docket number from the court in which the case closed as an identifier.

¶14 We used the IPLC, Westlaw's Patent-Docket database,[33] and Public Access to Court Electronic Records (PACER)[34] to identify a number of case-specific variables from the dockets. These variables included the total number of defendants, whether the patent owner or the accused infringer initiated the case,[35] whether the case had been transferred from another federal district, the year the case was filed, the regional circuit in which the deciding U.S. district court was located, the age of the patent at the time of the litigation, and whether the patent had been litigated previously.

---

[31] Adding the same number to a set of random numbers, called a monotonic transformation (i.e., using a monotonic function) produces another set of random numbers.

[32] An allegation of patent infringement is an assertion that one or more claims within the patent have been infringed. Assertions that more than one claim in a patent have been infringed are common, and it is possible for a court to reach different outcomes for different allegedly infringed claims within one patent. If we had found in our research of outcomes that at least one case resulted in different outcomes on the merits for different claims in the same patent, we would have treated each merits outcome on a particular claim as an observation, but we did not encounter such an occurrence.

[33] This database is called DOCK-PATENT in Westlaw.

[34] PACER, http://www.pacer.gov/ (last visited Jan. 18, 2012).

[35] In the remainder of the article, we often use the term "plaintiff" to refer to the patent owner and "defendant" to refer to the accused infringer. Even though patent litigation is sometimes initiated by the party that does not own the patent, by far the most common scenario has the patent owner filing as a plaintiff. When a patent owner sues for infringement, the accused infringer always files a counterclaim seeking a declaratory judgment of noninfringement and invalidity. When, because of a dispute with the patent owner, an accused infringer initiates the case by seeking a declaratory judgment of noninfringement (and usually invalidity, as well), the patent owner will file a counterclaim for infringement. The case involves the same issues regardless of which party is the initiator, but being able to initially select the venue is sometimes viewed as an advantage. *See* Kimberly A. Moore, *Forum Shopping in Patent Cases: Does Geographic Choice Affect Innovation?*, 79 N.C. L. REV. 889 (2001) (finding substantive and procedural differences among federal districts).

III. INTERNET PATENTS AND THE LIKELIHOOD OF LITIGATION

A. Descriptive Results

¶15 As shown in Table 1, the Internet patents in our data set were litigated at a far greater rate than the comparison set of NIPs. Moreover, broader Internet business model patents were litigated at a much greater rate than narrower business technique patents, possibly supporting the idea that broader claim language increases the universe of potential infringers and promotes litigation.[36] Given that there is no good way to measure the breadth of claim language empirically, however, we can only surmise that such a phenomenon contributed to these results.[37]

Table 1: Litigation Rates for Internet Patents and NIPs

	Total Patents in Data Set	Total Patents Litigated	Patent Litigation Rate	P-Value Univariate Comparison w/ NIPs
Internet patents-all	1,093	111	10.16%	p = 0.000
Internet patents-model	345	52	15.10%	p = 0.000
Internet patents-technique	748	59	7.89%	p = 0.000
NIPs	10,000	136	1.36%	

B. Logistic Regression Design

¶16 We then determined whether the differences in litigation rates reported in Table 1 continued to be significant while taking into account the effects of several patent characteristics shown by prior research to be positively or negatively associated with litigation propensity and value. This is a particularly interesting question in our current study because Allison and Tiller found that the average Internet patent issued during the first few years in which such patents began being issued in substantial numbers possessed many of the same characteristics as litigated patents more generally, regardless of whether the Internet patents had been litigated. Some patent characteristics suggesting litigation propensity and private economic value can be disaggregated into sub-characteristics which themselves may be differently associated with litigation likelihood than the aggregated measure of which they form a part. Because it is statistically incorrect to include both an aggregate measure and its disaggregated subparts as independent variables in the same regression model, we run different regressions for aggregates and their constituents.

¶17 Recall that, in our effort to identify litigation rates, we found it necessary to acquire a random sample of 10,000 NIPs issued during the same time period as our data set of Internet patents to discover a large enough number of litigated NIPs for statistical analysis. Compared with the total number of issued patents, litigated ones represent rare events, thus requiring this kind of

[36] If we compare the two subgroups of Internet patents (models & techniques) only with each other, the litigation rate for the broader Internet models exceeded the rate for narrower Internet techniques with an extremely high degree of statistical significance (p = 0.0003).

[37] In our assessment of Internet model patents as having broader claims than Internet technique patents, we borrowed from Justice Potter Stewart:

> I shall not today attempt further to define the kinds of material I understand to be embraced within that shorthand description [of hard-core pornography]; and perhaps I could never succeed in intelligibly doing so. But I know it when I see it, and the motion picture involved in this case is not that.

Jacobellis v. Ohio, 378 U.S. 184, 197 (1964) (Stewart, J., concurring) (emphasis added).

oversampling. Unlike the univariate comparisons of litigation rates, there is a need in our regression models for the data sets to be of sizes that are not wildly incomparable. This is the situation that would confront us if we used the sample of 10,000 NIPs in the regressions, unless we also drew a comparably sized sample of Internet patents. The latter is an impossibility because (1) a relatively small number of Internet patents had been granted by the late 1990s when we collected our population of them,[38] and (2) we were constrained to include only Internet patents and NIPs that had been issued a long enough time ago that enough of them would have been litigated for us to have sufficient numbers for statistical analysis. It would have been impossible, therefore, for us to have Internet patents in numbers comparable to the set of NIPs from which we derived our NIP litigation rate.

¶18 As a consequence, we use only the 2,382 Internet patents and NIPs from the Allison & Tiller and ALMT studies in our regressions: 982 unlitigated and 111 litigated Internet patents from the Allison & Tiller set of 1,093, plus 990 unlitigated and 299 litigated NIPs from ALMT. If we used these 2,382 patents without adjustment, however, they would reveal a litigation rate among NIPs (i.e., patents in general) of approximately 23% (299/(299 + 990), or 299/1,289), which of course we know to be breathtakingly wrong based on both common knowledge and the obviously more accurate rate of approximately 1% derived from our examination of litigation in a random sample of 10,000 NIPs—with a random sample of this magnitude, we can infer with great confidence that this is very close to the actual rate of litigation in the general population of patents (excluding certain small subsets such as early Internet patents). We thus had to accomplish an adjustment by specifying a weighted logistic regression model derived from the rare events logistic regression technique proposed by King and Zeng.[39] The model is weighted so that the proportion of litigated NIPs within our sample is the same as the inferred proportion of litigated NIPs in the general population. The binary dependent variable (litigated or not) is identified as 1 if a patent was litigated and 0 if the patent was not litigated.

¶19 The weights are denoted as

$$w_1 = \frac{\tau}{\hat{y}} \quad \text{and} \quad w_0 = \frac{1-\tau}{1-\hat{y}},$$

where τ is defined as the 1% rate of NIP litigation in the general population and $\hat{y}$ is the rate of NIP litigation implied by the sample (the 23% resulting from our deliberate oversampling).[40] Furthermore, w_1 is the weight assigned to all records containing litigated NIPs and w_0 is the weight assigned to all records containing unlitigated NIPs. The Internet patents, both litigated and unlitigated, received a weight of only one because the data set already reflects the appropriate proportion of litigated Internet patents. Using this paradigm, the following weights were applied to the NIPs in our data set:

$$w_1 = \frac{1}{\frac{\frac{100}{299}}{1289}} \quad \text{and} \quad w_0 = \frac{\frac{99}{100}}{\frac{990}{1289}}$$

Our binary dependent variable is whether a patent was litigated or not. As independent variables, we used: (1) whether the patent was classified as an Internet patent (and also whether it was an Internet patent-model or Internet patent-technique); as well as (2) several patent characteristics previously found to be either positively or negatively associated with litigation, including (a) the time the patent

[38] The data set of 1,093 Internet patents was the product of an effort to collect all of the patents meeting our definition issued through the end of 1999. Thus, it is a population (or "census") rather than a sample. We do not claim to have identified 100% of such patents issued during that time frame, because we undoubtedly missed some, but the set is nevertheless a substantially complete population.

[39] *See* Gary King & Langche Zeng, *Logistic Regression in Rare Events Data*, 9 POL. ANALYSIS 137 (2001). The concept is similar to that of a simple weighted average, used in the logistic regression context to help solve the problem of logistic regression with rare events.

[40] *Id.* at 145.

spent in the PTO from original application filing,[41] (b) the number of inventors,[42] (c) the number of patent claims,[43] (d) the number of prior art references in the patent,[44] (e) whether the patent was originally issued to a small entity owner,[45] and (f) whether the patent was originally issued to a foreign owner.[46] Because some of our independent variables are aggregates of others (i.e., *Total Number of Claims* is the aggregation of *Independent Claims* and *Dependent Claims*; *Total Number of Prior Art References* is the aggregation of references in a patent to prior U.S. patents, foreign patents, and other publications ("nonpatent references"); and *Small Entity* ownership status is the aggregation of individual, small business, and nonprofit ownership status), we cannot place all covariates in the same regression model. Thus, we constructed two different models, one with all the aggregated independent variables and the other with the disaggregated ones.

C. Logistic Regression Results

1. Using Aggregated Independent Variables

¶20 First, we examine the relationship between the probability of a patent being litigated and whether or not it is characterized as an Internet patent (or Internet patent-model or Internet patent-technique), while controlling for all of the aggregated independent variables. The results of these regressions are reported in Table 2. Logit (1) describes the Logit model for conducting logistic regression (used when the dependent variable is binary) with all Internet patents as an independent variable, and Logit (2) describes the same model with, instead, each of the two Internet patent subgroups as independent variables.

---

[41] *See, e.g.*, ALMT, *supra* note 3, at 459-60 (finding that applications for patents that were later litigated spent far more time in prosecution from original filing than did applications for patents that were not litigated).

[42] *Id.* at 478-479 (finding in both a population study and a more granular sample study that litigated patents covered inventions with slightly fewer inventors than unlitigated ones, although in the sample study for which the data correspond most closely to those in the present study, the difference was too minor to be significant). Allison and Tiller, *supra* note 5, at 1058-63, argued that a larger number of inventors might be indicative of greater patent value, and found empirically that Internet patents did have significantly more inventors than the average NIP. However, Allison and Tiller's finding largely disappeared when Internet patents were compared with NIPs in individual technology areas because the finding pertaining to the average NIP was substantially driven by the relatively large number of patents in a few technology fields such as mechanics that were characterized by small numbers of inventors. In any event, there is enough evidence that the number of inventors may be associated with litigation propensity to include it in the model, especially since it does not appear to be significantly correlated with the other independent variables we use.

[43] Most researchers have found that the total number of claims within patents is strongly associated with greater litigation propensity. *See, e.g.*, Jean O. Lanjouw & Mark Schankerman, *Characteristics of Patent Litigation: A Window on Competition*, 32 RAND J. ECON. 129 (2001); Jean O. Lanjouw & Mark Schankerman, *Stylized Facts of Patent Litigation: Value, Scope and Ownership*, (Nat'l Bureau of Econ. Research, Working Paper No. 6297, 1997). For this reason, Allison and Tiller used this metric as an independent variable in their study. However, ALMT broke down total claims into independent and dependent claims, finding that most of the association between total claims and litigation likelihood was accounted for by the number independent claims. Because of the finding by ALMT, in the current study we have used the textual analysis capability of Adobe Professional 9.0 software to separate independent and dependent claims in our data sets.

[44] The total number of prior art references was found in prior research to be associated with the likelihood of litigation with a high level of significance. *See* ALMT, *supra* note 3, at 453. Also, Internet patents contained many more total prior art references than did NIPs. *See* Allison & Tiller, *supra* note 5, at 1040-45.

[45] Litigated patents were found in prior research to have been issued disproportionately to small entity owners rather than large entities. Segregating small entities into subgroups, it was found that patents granted to individual and small business owners were much more likely to be asserted in litigation, but this was not true of nonprofits (such as universities). *See* ALMT, *supra* note 3, at 479 tbl.5. It is not uncommon for the ownership of a patent to change after issuance and before litigation, however, so the fact that litigated patents were significantly more likely to have been issued originally to small entities does not automatically mean that small entities are more likely than large entities to assert patents in litigation. *See* Moore, Populism and Patents, *supra* note 3, at 96–97 (documenting a high rate of pre-litigation transfer of patents); *see also* Allison, Lemley & Walker, *Trolls on Top*, *supra* note 3, at 20-22 (in a study comparing characteristics of patents that had been litigated eight or more times during a recent nine-year period with those that had been litigated once during that period, finding evidence that a substantial portion of patents in both data sets of litigated patents had been assigned after issuance and before litigation).

[46] Patents issued to foreign owners are much less likely to be litigated than those issued to U.S. owners. *See* ALMT, *supra* note 3, at 478-79 tbls. 4 & 5 ($p < 0.0001$ for both foreign ownership in the population study and for foreign inventorship in the more finely graded sample study). This finding is unsurprising given the significant jury bias against foreign patent owners in American patent infringement litigation. *See* Moore, *Xenophobia*, *supra* note 3, at 1548-50. Jury bias does not explain all of the clearly lower win rates in patent infringement litigation for foreign-owned U.S. patents, because foreign entities disproportionately acquire U.S. patents in technology areas where there is less patent litigation overall. *Id.* at 1533-37. However, Moore's data make it clear that jury bias provides a substantial part of the explanation. *Id.* at 1548-49. The evidence shows, inter alia, that foreign-owned patents are not more likely to be invalid that those with domestic owners. *Id.* at 1533-35, 1542.

Table 2: Litigation Likelihood: Logistic Regression with Aggregated Variables

Dependent Variable = Litigated or Not	Logit (1)	Logit (2)
Independent variables	**Odds ratios; Standard errors in parenthesis**	
Internet patent-all	7.497 (1.185)***	
Internet-model		9.073 (1.962) ***
Internet-technique		6.646 (1.192) ***
Years in PTO	1.146 (0.062) **	1.146 (0.062) **
Number of inventors	1.019 (0.046)	1.017 (0.046)
Total # of claims	1.005 (0.003) *	1.005 (0.003)
Total # of prior art references	0.999 (0.001)	0.999 (0.001)
Small entity	3.035 (0.584) ***	2.792 (0.584) ***
Foreign owner	0.424 (0.167) **	0.427 (0.169) **
Adjusted # of Forward Citations	1.153 (0.043) ***	1.145 (0.041) ***
Standard Errors[47]	Robust	Robust
Pseudo R-squared[48]	0.182	0.184
N = 2382 *** p < 0.01; ** p < 0.05; * p < 0.10		

¶21 We reported results as odds ratios for ease of interpretation.[49] These results show that, even after controlling for the other factors itemized in Table 2 that are either known or suspected to influence litigation likelihood, being an Internet patent greatly increased the likelihood of a lawsuit ($p < 0.01$).[50] A patent issued during the late 1990s on an Internet business process was about 7.5 times more likely to be litigated than a non-Internet patent issued during the same timeframe, holding other factors constant. Separating Internet patents into broader online business models and narrower online business techniques, we see that business models were nine times more likely to be litigated than NIPs, while business techniques were 6.6 times more likely to be litigated. These findings were also significant at the 1% level.

[47] When we say "robust" in connection with standard errors, we mean that we ran our logistic regression models in a way that accounts for the fact that the variances in the distributions of each of our independent variables might not be equal. That is, we ran our models to account for any possible heteroskedasticity in the error terms.

[48] The pseudo R-squared is a measure of the goodness-of-fit for a logistic regression model, the model being the selected group of independent variables. *See* JOSEPH F. HAIR ET AL., MULTIVARIATE DATA ANALYSIS 318-20 (5th ed. 1998). It is not a true R-squared as is found in OLS regression, but it may be interpreted similarly. We used McFadden's pseudo R-squared, a commonly employed pseudo R-squared method, which runs from 0 to 1 in the same manner as a true R-squared. The closer the measure is to 1, the better fit the model is to the actual data. There are a multitude of variables that affect whether various patents are asserted in litigation, most of them idiosyncratic and undiscoverable. When, as here, there is so much noise in the data, a pseudo R-squared for even a good-fitting regression model necessarily will be relatively low on a scale of 0 to 1, as illustrated by the 0.167 and 0.170 shown in Table 2.

[49] If the odds ratio is exactly 1 for an independent variable, the independent variable has no effect on the likelihood of the patent being litigated. When the odds ratio is greater than 1 for an independent variable, the variable has a positive effect on the likelihood of the patent being litigated. When the odds ratio is less than 1 for an independent variable, the variable has a negative effect on the likelihood of the patent being litigated. The amount by which the odds ratio is more or less than 1 reveals the magnitude of the effect on likelihood.

[50] Our logistic regression results from using Logit models are robust to Probit and Linear Probability (OLS) models. In other words, we also tried other generally accepted models for regressing on binary dependent variables and did not get results that would contradict those from our Logit models.

¶22 The regression models also reveal that the number of years a patent spends in the PTO from original filing[51] and initial ownership of the patent by a small entity both have a positive and statistically significant effect on a patent's likelihood of being litigated. Indeed, initial issuance of a patent to a small entity made the patent about three times more likely to be litigated, controlling for the effects of the other variables. These patents were thus either much more likely to be asserted in infringement litigation by their small entity owners or sold to others who were far more likely to assert them.[52] Our results also confirmed previous findings that the number of forward citations—citations to patents by later patents—is highly associated with litigation propensity. An additional forward citation made litigation about 15% more likely, a finding that was statistically significant at the 1% level.[53] A caveat must accompany this result, however. A portion of the forward citations received by litigated patents in our dataset occurred after litigation was initiated. To the extent that forward citations occurred after the initiation of litigation of any given patent, the variable is exogenous and cannot be said to actually "predict" the occurrence of litigation. The finding is nevertheless an interesting one and again confirms the previous results of prior research, and the magnitude and high level of statistical significance of our finding warrant reporting.

¶23 In addition, and as expected, whether the patent had a foreign owner when it was granted significantly decreased the likelihood that it would be litigated, a finding that we discuss in more detail below.[54] In the model using all Internet patents but not their subtypes as an independent variable, the total number of claims had a small positive effect on litigation likelihood that was significant at the 10% level, but not at 5% or below.[55] Even that level of significance was missing when the Internet patent subtypes were the independent variables, further highlighting the weakness of the relationship. Also of note is that the total number of prior art references had no effect on the likelihood of litigation.

¶24 Three observations about odds ratios may be helpful. We included both categorical variables denoting states of being (here, they are the binary type of categorical variable) and quantitative variables reflecting measurements of values (either "count" variables that must be whole numbers or "continuous" variables that do not have to be whole numbers). Odds ratios are a bit different for categorical and quantitative variables. One may properly say that, for each additional X of a quantitative independent variable, the odds of the dependent variable occurring increase Y amount. Thus, it could be that, for each additional prior art reference in a patent or for each additional unit of

---

[51] When the time spent in prosecution at the PTO is measured from the original filing date, the metric captures the amount of time devoted by the applicant to the filing of various types of continuing applications, something that requires significant investment and suggests that the applicant is seeking to build a portfolio of patents on closely related technologies. This is probably an indicator of perceived patent importance by the applicant and greater willingness to invest more in the acquiring the patent. *See supra* note 15.

[52] There is an active market for subsequently litigated patents. *See, e.g.*, Moore, Populism and Patents, *supra* note 3, at 96–97; Allison, Lemley & Walker, *Trolls on Top*, *supra* note 3, at 20-22. Thus, even though some small entity owners may have sold their patents prior to litigation, small entities nevertheless can be very active themselves as patent infringement plaintiffs. *See, e.g.*, Allison, Lemley & Walker, *Repeat Patent Litigants*, *supra* note 3, at 689-90 (small entities extremely over-represented as plaintiffs for patents litigated 8 or more times during the period 1-1-2000 to 2-28-2009). Within the category of small entities, individuals do litigate at a slightly lower rate than corporations (large and small combined), but are still well represented as plaintiffs. *See* Moore, *Populism and Patents*, *supra* note 3, at 94 ("It appears that individuals litigate slightly less often than corporations—individuals receive 14.9% of patents and initiate only 11.9% of patent suits."). But it is unclear what the litigation rate is for small businesses (corporations with fewer than 500 employees) within the overall small entity category.

On average, small entity patent owners are far less likely than large entity owners to make and sell products. *See* ALMT, *supra* note 3, at 468-69. On average, small entity owners may thus have less to lose from filing infringement lawsuits than large entities. *Id.* The reason is that a patent owner that does not make and sell products (a so-called non-practicing entity, or NPE) need not fear a counterclaim for patent infringement because an NPE does not do anything that could constitute infringement. *Id.*

[53] *See, e.g.*, ALMT, *supra* note 3, at 435 (an additional citation received—a forward citation—increased the odds of litigation by 16%; there, too, however, that proportion of forward citations received after the first instance of litigation was exogenous).

[54] The Internet patents from the Allison & Tiller, *supra* note 5, study had few foreign owners relative to NIPs, thus creating a possible biasing effect in the regression models for our current study of Internet patent litigation. For the current study, however, we ran a correlation table for Internet patents and foreign owners and found that the negative correlation was not high enough to affect our logistic regression results in a meaningful way.

[55] Indeed, the magnitude of the positive effect on litigation likelihood was very small, but the total number of claims in all of the patents in our data set represents an extremely large number of observations, thus increasing the chances that our finding was not due to random chance (less than a 10% chance in this case) and leading to a finding of minor statistical significance.

time the patent application was pending in the PTO, the odds of it being litigated increased Y%. One cannot properly make such a statement about a categorical independent variable—if the independent variable is, say, whether the patent was initially issued to a small business, one could say that this fact increased the odds of litigation by Y%. But there are no degrees of being a small business owner. The owner either falls within the PTO definition of a small business or it does not.

¶25 Another interesting fact about odds ratios in logistic regression is that each increase in the amount of a quantitative independent variable produces an effect on the dependent variable (if there is any positive or negative effect) that is multiplicative rather than additive. Suppose that one of our findings is that the number of children that inventors had at the time they filed their patent application increased the odds of litigation by 4%, and that this was statistically significant. If the inventors had a total of five children at the time of filing, the chances of litigation would not increase by 20% (.04 times 5 = 0.2). Instead, the odds of litigation would increase by almost 22% (1.04^5 = 1.216). The above clearly does not apply to a categorical independent variable.

¶26 Third, there sometimes are occasions when one has a choice about the unit of measure to use for a quantitative independent variable. Thus, for the amount of time a patent application spent in the PTO before issuance of the patent (pendency time), one could choose days, weeks, months, years, 300-day periods, etc. (in recent times, some applicants might say that millennia would be a plausible choice). The choice clearly has an effect on what the odds ratio will look like. Even if the amount of time in the PTO has what most experienced observers would view as a very meaningful positive effect on the likelihood of litigation, the effect would appear to be very small if we chose a day as our unit of measure. Thus, if we were to say that, for each day of additional time the application spent during examination in the PTO, any effect on the odds of litigation would necessarily be miniscule. This is one of countless situations in which experienced judgment is essential in empirical research. We happen to know that patent issuance requires years, so that is what we chose. However, it does not normally take 50 or 100 years, but rather about three years or so on average, depending on the technology. Thus, one year is a large increment relative to the mean, and one must keep this in mind when assessing the magnitude of any effect on odds.

2. Using Disaggregated Independent Variables

¶27 Before decisively drawing the conclusions made above about Internet patents and litigation, we performed a sensitivity analysis to ascertain whether separating the aggregated variables into their components affects the statistical significance of the variables of interest. More specifically, we broke the total claims, total prior art references, and small entity owner variables into their constituent elements, and report these results in Table 3.

Table 3: Litigation Likelihood: Logistic Regression with Disaggregated Variables

Dependent Variable = Litigated or Not	Logit (1)	Logit (2)
Independent variables	**Odds ratios; Standard errors in parenthesis**	
Internet patent-all	7.969 (1.310) ***	
Internet-model		9.470 (2.082) ***
Internet-technique		7.114 (1.322) ***
Years in PTO	1.123 (0.068) *	1.122 (0.068) *
Number of inventors	1.025 (0.048)	1.021 (0.047)
Number of independent claims	1.063 (0.025) ***	1.062 (0.025) **
Number of dependent claims	1.000 (0.004)	1.000 (0.004)
Number of U.S. patent references	1.008 (0.005)	1.008 (0.005)
Number of foreign patent references	1.030 (0.018) *	1.030 (0.017) *
Number of nonpatent references	0.987 (0.008)	0.988 (0.008)
Individual owner at issuance	3.023 (0.716) ***	2.748 (0.695) ***
Nonprofit owner at issuance	2.478 (1.948)	2.243 (1.802)
Small business owner at issuance	3.182 (0.719) ***	3.002 (0.712) ***
Foreign owner at issuance	0.424 (0.165) **	0.427 (0.167) **
Adjusted Forward Citations	1.155 (0.043) ***	1.148 (0.042) ***
Standard Errors	Robust	Robust
Pseudo R-Squared	0.192	0.194
N = 2382 *** p < 0.01; ** p < 0.05; * p < 0.10		

¶28 As shown in Table 3, the Internet patent variable in the disaggregated-variables model was also a predictor of litigation at a 1% significance level. The substitution of the disaggregated variables for their aggregates increased the odds ratios for Internet patents to where they were about eight times more likely to be litigated than NIPs. When we separated Internet patents into subgroups, we found that online business models and online business techniques were respectively 9.5 and 7.1 times more

likely to be litigated than NIPs. These findings were also significant at the 1% level. Thus, the increased sensitivity created by drilling down more deeply to disaggregated independent variables led to findings of somewhat greater litigation likelihood for Internet patents as a whole, as well as for the subgroups. As seen by the pseudo R-squared measures, disaggregating some of the independent variables slightly increased the models' goodness-of-fit—their overall explanatory power.

¶29 While taking into account the cumulative effects of all other independent variables in the disaggregated models, including the type of patent, we found that the number of independent claims, the status of the patent owner as an individual at the time of patent issuance, and the status of the patent owner as a small business all had highly significant positive effects on a patent's likelihood of being litigated.

¶30 Thus, our results confirm the finding of ALMT in 2004 that it is the number of *independent* claims, not dependent claims, that accounts for the previously observed effect of *total* numbers of claims on litigation likelihood. Moreover, the disaggregated analysis teases out the fact that the large effect on litigation propensity of small entity status is mostly accounted for by patents originally issued to either individuals or small businesses and not by nonprofits such as universities or research foundations. The nonprofit status of the owner at the time the patent was issued did have a large odds ratio, but the number of patents having such owners in our data sets was much too small for the large odds ratio to be statistically significant. We can only speculate about whether patents initially owned by universities and other nonprofits would have been as statistically prone to end up in litigation as individuals and small businesses had their representation been greater in the data.[56]

¶31 As with the aggregated variables models reported in Table 2, here too we see that initial issuance of a patent to a foreign owner significantly reduces the probability of the patent being litigated. Again, this is unsurprising given Kimberly Moore's clear findings of jury bias against foreign patent owners.[57] The finding also presents a potential contradiction in our regression results. What we have is this: (1) Foreign owners of U.S. patents are quite significantly *less likely* to litigate;[58] (2) Foreign owners of U.S. patents cite *significantly more* references to foreign patents as prior art than do domestic owners of U.S. patents;[59] and (3) More citations to foreign patents in U.S. patents overall *increases* the likelihood of litigation (only at a 10% level of significance, however, and not at 5% or less).[60]

¶32 Table 4 shows a matrix of U.S.-foreign ownership of U.S. patents and the citation of foreign patents as prior art references. One can see rather striking, but logically intuitive differences.

---

[56] It bears repeating that a patent's original ownership status does not necessarily mean that entities having the same ownership status later file an infringement lawsuit because there is a relatively active market for patents that end up in litigation.

[57] *See* Moore, Xenophobia, *supra* note 3, at 1504.

[58] *See supra* Tables 2 & 3; ALMT, *supra* note 3, at 454.

[59] *See* Moore, Xenophobia, *supra* note 3, at 1537-42.

[60] This positive effect on litigation likelihood is of modest magnitude (odds ratio of 1.034) but is statistically significant at $p < 0.10$ because the number of observations was relatively large. The effect that the number of foreign patent references has on litigation likelihood is probably due to the fact that it is more time-consuming and expensive for U.S. patent applicants to find and analyze foreign patents than it is for them to find and analyze prior U.S. patents. This suggests that, when a patent applicant perceives that its invention is likely to be relatively important, and perhaps more likely to be sold to someone else or litigated by the applicant itself, the applicant probably has a greater willingness to invest more in acquiring the patent. It should not be as difficult or expensive for a foreign applicant to acquire knowledge of and to analyze prior foreign patents within its own area of localized news dissemination and language facility, and litigated U.S. patents with foreign ownership do cite significantly more foreign patents than do unlitigated U.S. patents with foreign ownership. Moreover, domestic owners of litigated U.S. patents who likely have to invest more to cite prior foreign patents than they have to invest to cite prior U.S. patents are represented in much greater numbers in our data set than are foreign owners of litigated patents.

Table 4: Matrix of U.S.-Foreign Patent Ownership and Citation of Foreign Patents as Prior Art References

		Foreign Owned	
		No	**Yes**
Litigated	**No**	0.759	2.426
		1413	559
	Yes	2.505	3.361
		374	36
Percent Litigated		26.47%	6.44%
In each cell, the top number is the average number of foreign patent references and the bottom number is the total number of patents.			

¶33 Rounding off the numbers reported in the table, we see the following: In our data set of 2,382 U.S. patents, foreign owners of *litigated* patents cited an average of 3.361 foreign patent references, while foreign owners of *unlitigated* patents cited an average of 2.426 foreign patent references. On the other hand, U.S. owners of *litigated* patents cited an average of 2.505 foreign patent references, while owners of *unlitigated* patents cited an average of cited an average of 0.76 foreign patent references.

¶34 Going back to the apparent conundrum presented by the facts that, on average, foreign owned patents are much less likely to be litigated, foreign owned patents have significantly more foreign patent references, and litigated patents have significantly more foreign patent references, we have a classic econometric question about the difference between individual effects and cross effects of variables in a model. The answer is that the previous statements look at one variable at a time, whereas the overall model does not do that. Instead, it considers any variable in light of the effects of all other variables, and is also affected by attributes of a data set that are not even included explicitly as regression variables. For example, foreign ownership is clearly under-represented in our set of Internet patents because U.S. law and U.S. inventors were the innovation forerunners in this field.[61] Another example is that, even though U.S. owners cited many fewer foreign patent references *overall* than did foreign owners, they cited more such references in *litigated* patents than they did in *unlitigated* ones, and the number of U.S. owners of litigated patents in our data set overwhelmed the number of foreign owners of litigated patents. The situation also presents an opportunity to recognize the very important distinction between statistical significance and magnitude.[62] Here, for instance, even though we found that the number of foreign patent references as a predictor of litigation was

[61] Moreover, foreign owners still would be underrepresented in a data set of recent Internet patents because of U.S. dominance of Internet innovation and patenting, and because the laws of some countries make it much more difficult to get this sort of patent.

[62] Although an independent variable in a logistic regression model can have a positive or negative magnitude of meaningful size but not be statistically significant, and a variable can be statistically significant while having an effect of very small magnitude, significance and magnitude clearly are not independent of one another. *See, e.g.*, Allison & Sager, *supra* note 3, at 1792 (observing the interrelationship between the statistical significance and the magnitude of effect of an independent variable's effect). Not only magnitude of difference but also sample size has an effect on statistical significance levels, and in multiple regression, the presence of other variables also has an impact.

statistically significant at $p < 0.10$, the magnitude of the effect was modest as shown by the odds ratio of 1.030 (addition of a reference to a prior foreign patent increases the odds of litigation by 3%).[63] On the other hand, the negative effect of foreign ownership on litigation likelihood was much greater in magnitude (odds ratio of 0.391—odds of litigation *decreased* by 61% if the patent had a foreign owner) despite having the same level of statistical significance.

¶35 We return now to the other results reported for the two disaggregated models shown in Table 3. The amount of time a patent application spent in the PTO from its original filing was significant only at the marginal 10% level, whereas it had been significant at $p < 0.05$ in the aggregated variables models results reported in Table 2. The magnitude of this variable's effect on litigation odds also decreased somewhat when some of the independent variables were disaggregated. This decrease in the effect on litigation likelihood of how long patent applications spent in the PTO is probably the result of changes in the correlations between time-in-prosecution at the PTO and a couple of other variables, such as the different types of prior art references.[64] It is inevitable that, when working statistically with patent characteristics, some of them are correlated with each other, a fact that can complicate multiple regression analysis. The idea of investment-willingness in the face of greater perceived importance applies to more than one patent characteristic and thus to more than one of our independent variables.[65] Thus, a patent applicant holding such a perception may be likely to make investments that cause an increase in multiple metrics within a patent that increase time spent in the PTO, such as numbers of claims, numbers of prior art references, and numbers of continuing applications.

¶36 Recall that, in the aggregated variables models for which results were reported in Table 2, the total number of prior art references had no effect at all on litigation likelihood. Separating the aggregated variable into three types of prior art references yielded one result that was interesting, one that was uninteresting, and one that might be termed a meaningful insight. The *interesting* result was that a "nonpatent reference" had a trivial negative effect. Nonpatent references, sometimes called "other publications," are references in patents to prior articles in scientific journals, theses and dissertations, books and book chapters, government and industry group reports, company press releases and popular press articles, web pages, sales brochures, and so on. It is interesting because it fails to comport with earlier research. We have no ready explanation for the fact that our finding concerning the number of nonpatent references was quite contrary to the significant positive effect found in ALMT.[66] Early Internet patents overall did include much larger numbers of nonpatent references than did contemporaneous NIPs.[67] Because Internet patents were litigated at a much greater rate than NIPs, one might expect that the number of nonpatent prior art references would have a significant positive effect on the odds of litigation, but such was not the case. The result that was *uninteresting*, given the insignificance of total prior art references discovered in our aggregated models, was that a reference to a prior U.S. patent had a trivial positive effect on the odds of litigation to which we can attach no meaning. The *insight* was that that the presence of a reference to a prior foreign patent had a positive effect on the odds of litigation of a magnitude that warrants notice—3% for one foreign patent reference in a U.S. patent. And remember that additional such references have a multiplicative effect on litigation propensity. This result was significant only at the 10% level, which does not inspire great confidence, but this is likely because of the fact that the total number of observations was almost certainly quite small. The smallness of the absolute numbers of

[63] Although an odds ratio of 1.030 reveals a modest 3% increase in the odds of litigation when a patent includes a foreign patent reference, the effect is much more than modest as additional foreign patent references are added. Recall that the odds ratio is multiplicative rather than additive. Thus, one cannot just add 3% for each reference. For example, suppose that a patent has 10 foreign references. The odds of litigation are not increased by 30%, but by 1.03^10 = 1.3439, or nearly 35%.

[64] For instance, the independent variable "total prior art references" was correlated with the "years in the PTO" independent variable but two of the constituent elements of total references—number of U.S. patent references and foreign references accounted for the largest part of that correlation.

[65] *See supra* note 14 and accompanying text.

[66] ALMT, *supra* note 3, at 453-54.

[67] Allison & Tiller, *supra* note 5, at 1040 (finding that Internet patents included an average of 10 nonpatent references, while NIPs included an average of 2.37).

references to foreign patents, especially in such largely home-grown phenomenon the Internet business process patents, is probably attributable not only to the general rarity of patents issued in foreign countries that are relevant to applications for U.S. Internet patents, but also to the fact that it historically has been more difficult and costly for those seeking U.S. patents on any sort of technology to find relevant foreign patents. That problem has been easing as searchable databases of foreign patents have become more available and extensive, but they are typically rather expensive.

¶37 We should not overlook the most important result of disaggregating certain independent variables into their constituents, however—being an Internet patent generally or being one of the two subtypes continued to have strikingly large effects on the odds of litigation. No matter how one slices it, patents issued on Internet business methods during the early years of their existence were the subject of infringement litigation at a far greater rate than other patents granted during the same time frame.

D. Limitations of Our Data

¶38 The data used in the analyses just explained have at least one unavoidable limitation. In their study of litigated patents, ALMT controlled for technology field, and to a lesser degree, for industry. Prior research had shown differential litigation propensities among technologies and industries,[68] which ALMT confirmed.[69] We also wished to refine our regression analyses further by controlling for technology field but were unable to do so. Initially, we conceived of accounting for technology differences by using the PTO classifications. This did not work. Neither the PTO nor the international patent classification (IPCs) system is suitable for delineating technology categories or industries at a conceptual level.[70] The main problem, though, is that our data set of Internet patents all fall within the same field of technology and the same industry (computer), and consequently we were not able to effectuate a technology-industry control despite the fact that a number of different fields were represented in our set of NIPs.[71]

IV. INTERNET PATENTS IN LITIGATION

¶39 In Part III we showed that, among other things, Internet patents issued during the formative years of such patenting were about eight times more likely than non-Internet patents to be litigated after controlling for other factors, the rate being even higher for patents on more broadly claimed Internet models. We next analyze the cases themselves to see if there are differences in how Internet

---

[68] *See, e.g.*, Moore, Xenophobia, *supra* note 3, at 1533-35.

[69] ALMT, *supra* note 3, at 438, 471-73.

[70] *See* Allison & Tiller, *supra* note 3, at 1027-28. The PTO classification system identifies most classes and subclasses at a very low level of functional abstraction and is not especially helpful in conceptually defining technologies in a science or engineering sense. PTO classes and subclasses also are frequently entwined in rather cryptic ways. For example, Class 345 is defined as "Computer Graphics Processing, Operator Interface Processing, and Selective Visual Display Systems." The class definition states: "This class provides for processes and apparatus for selective electrical control of two or more light-generating or light-controlling display elements in accordance with a received or stored image data signal. The image data includes character, graphical information or display attribute data. The image data may include, for example, information data from a peripheral input device, from the reception of a television signal, from the recognition of image data, or from the generation or creation of image data by a computer. This class also provides for digital data processing systems or methods for data processing for visual presentation, wherein the processing of data includes the creation or manipulation of graphic objects (e.g., artificial images), text or use of an operator interface by a digital data processing system prior to use by or within a specific display system."

Not only does the class combine different technology areas, such as computer hardware, software, and optics, but it also does so in a very functional way rather than a conceptual way. To further illustrate, Subclass 74 within Class 395 is "Machine Element or Mechanism, particularly subclass 471 for control elements which move in two planes," showing the system's focus on low-abstraction functionality.

The system is designed to assist patent examiners in searching for prior art according to functions achieved by inventions, not by technological concepts. Although the World Intellectual Property Organization's system of International Patent Classifications may be better suited than the PTO's classifications for identifying technology areas, it also was designed for purposes quite different than ours and similarly operates too often at too low a functionally oriented level to serve our definitional objectives.

[71] Another unavoidable limitation is that, as stated before, the litigation of a patent is a rare event and to collect sufficient data for these events, the litigated NIPs were disproportionately sampled. Therefore, even with the weighting we performed, the standard errors and the coefficients may be biased since the rare events are coming from a finite sample. For a more in-depth discussion about the possible biases in the standard errors and the coefficients, see King and Zeng, *supra* note 39, at 142-43.

patents and NIPs fared in court after the litigation decision was made. In several sections of Part IV, we examine the way Internet patents and NIPs fare in courts, and in Part V we present the conclusions from our analysis.

¶40 To examine any possible differences between outcomes in Internet patent and NIP cases, we grouped the cases into several categories: clear settlement, likely settlement, judgment on the merits (with subcategories for the stage of litigation at which the judgment was rendered), and procedural dispositions (such as dismissals on jurisdictional grounds). None of the cases was still pending at the time of the time of the final draft of this paper, so all 453 are accounted for in the following categories.

- **Clear Settlement:** We placed a case into the *clearly* settled category if there was a specific docket entry stating that the case had settled, or a statement in a dismissal order or other court document that the case had settled.
- **Likely Settlement:** We placed a case into the *likely* settled category if there was evidence in docket entries or court documents creating a strong inference that the case had settled, but no specific statement that a settlement had occurred. Such evidence included judicial settlements based on stipulated (or agreed, or joint) motions to dismiss. We also placed cases in this category when they terminated as the result of a plaintiff's voluntary dismissal if there was no evidence of some reason other than settlement for the dismissal by plaintiff.
- **Judgments on the Merits:**[72] We treated a case as a win for one party (on a patent-by-patent basis) if there was a default judgment for the patent owner, or a judgment for either the patent owner or accused infringer at summary judgment, JMOL, or trial (we did not differentiate between jury and bench trials because of small numbers) on validity, enforceability, or infringement that clearly led to termination of the case. In addition, we treated a consent judgment as a win for one party despite the fact that such a judgment incorporates a settlement between the parties if the consent judgment followed closely after an event that could be characterized as a clear win for one of the parties. For example, if a consent judgment followed closely on the heels of a summary judgment or jury verdict for the patent owner on infringement or validity, this was treated as a win for the patent owner incorporated into a consent judgment. Our reasoning was that the agreed terms of the judgment almost certainly favored in a substantial way the party that had prevailed on a critical merits issue just prior to consent judgment.
- **Procedural Disposition:** We placed a case in this category if it was dismissed on personal or subject matter jurisdictional grounds, ordered by the judge to be closed administratively (put on long-term "hold" without removal from the docket), terminated because of a party's bankruptcy, or terminated for other reasons when there was no suggestion of settlement and no decision on the merits.

A. Settlement—Internet Patents and NIPs

¶41 We first analyzed the differences in settlement rates between the different types of litigated patents, considering only clear settlements.

[72] It is worth observing that, with many cases involving multiple patents and multiple claims within each patent being asserted, saying who "won" can be dicey. Janicke and Ren sought to identify the "main winner," which can be rather subjective and fraught with peril. Paul M. Janicke & LiLan Ren, *Who Wins Patent Infringement Cases?*, 34 AIPLA Q.J. 1, 8-9 (2006) (hereinafter Janicke & Ren). The more precise and accurate way of trying to determine who wins a patent infringement case is to treat each patent asserted within a case as a separate assertion with a separate outcome as we did in our study by using case-patent pairs as the unit of observation. There can even be situations in which outcomes on different claims within a patent can be so different as to make separate claims within a patent the appropriate unit of observation, but we did not encounter such a situation in our data set.

Table 5: (Clear) Settlement Rates for Internet and NIP Cases

Patent Type	All Cases	Internet Patents-All	Internet Model	Internet Technique	NIPs
Settled Cases	107	49	24	25	58
Total Cases	453	251	103	148	202
Settlement Rate	23.62%	19.52%	23.30%	16.89%	28.71%
P-Value Chi-squared Comparison w/ NIPs		0.022	0.313	0.010	

¶42 Table 5 shows that, just considering clear settlements, cases involving NIPs settled at a significantly higher rate than those involving Internet business patents as a whole ($p = 0.022$), and at a far higher rate than those involving Internet business techniques ($p < 0.010$). Although NIP cases settled more often than Internet business model cases, the difference was not statistically significant. Not shown in the table, but to be expected from glancing at the rates, is that, within the category of all Internet patents, cases involving patents on Internet models and Internet techniques settled at rates that were not significantly different.

¶43 Any number of factors may affect settlement rates, and some of these factors may differ depending on the type of case. For example, risk aversion and strategic behavior may vary from one kind of case to another as a result of asymmetry of stakes—the negative consequences of a loss at trial may be more severe for one party than another. Allison, Lemley, and Walker have observed that, in a patent case, the patent owner may be more willing to settle if it expects to be a repeat player by later suing at least one other defendant for infringement of the same patent, given the existence of non-mutual collateral estoppel.[73]

> Patent owners who file lawsuits put their underlying patents at risk. A significant percentage of litigated patents are held invalid, [74] and a finding of invalidity is the death knell for a patent. Because of the arcane civil procedure doctrine of offensive nonmutual collateral estoppel, the consequences of validity and invalidity holdings are highly asymmetric.[75] A patentee who wins a suit against defendant *A*, having proven the patent infringed and fought off a validity challenge, gets no credit for the win in a subsequent suit against defendant *B*.[76] Because *B* was not a party to the first suit, it is entitled to once again challenge the validity of the patent, even on the very same grounds rejected in the first lawsuit.[77] The same is true in subsequent suits against defendants *C*, *D*, *E*, and so on.[78] Indeed, for this reason Federal Circuit Judge Rich used to insist that patents were not held valid, but merely held "not invalid."[79] By contrast, should *A* succeed in proving the patent invalid, the game is up. The

[73] Allison, Lemley & Walker, *Repeat Patent Litigants*, *supra* note 3, at 678-79. The U.S. Supreme Court recognized the doctrine of non-mutual collateral estoppel patent infringement cases in Blonder-Tongue Labs., Inc. v. University of Illinois Found., 402 U.S. 313 (1971).

[74] *See, e.g.*, Allison & Lemley, *Empirical Evidence on the Validity of Litigated Patents*, *supra* note 3, at 205 (finding that 46% of patents litigated to judgment are held invalid); Janicke & Ren, *supra* note 70, at 5-6 (finding that patentees win only 25% of cases litigated to judgment, in part because of invalidity and in part because of noninfringement).

[75] *E.g.*, Blonder-Tongue Labs, Inc. v. Univ. of Ill. Found., 402 U.S. 313, 325 (1971) (discussing consequences and fairness of nonmutuality of estoppel in patent litigation).

[76] Indeed, the Federal Circuit has even refused to allow the prior judgment to be considered by the jury in a subsequent lawsuit. *See* Mendenhall v. Cedarapids, Inc., 5 F.3d 1557, 1575 (Fed. Cir. 1993).

[77] *Id.*

[78] *Id.*

[79] *See, e.g.*, Thomson, S.A. v. Quixote Corp., 166 F.3d 1172 (Fed. Cir. 1999).

> doctrine of collateral estoppel will prevent the patentee from enforcing the patent against *B, C, D*, or *E*; each of those defendants is entitled to rely on the patentee's prior loss to defeat the lawsuit.[80] And even existing licensees will be permitted to stop paying royalties and file their own challenge to the patent.[81] [footnotes in the original]

¶44 A patent owner's propensity to settle or not could also be affected by its concern about establishing a reputation for "going to the mat" for the purpose of getting better settlements from future defendants. Also, a repeat player in patent litigation, asserting multiple patents and/or the same patent against many accused infringers may refrain from settlement and take cases to judgment more frequently than an objective observer might think rational could be simply playing a lottery in the hope of one gigantic payoff.[82] Moreover, either party's settlement-trial mentality can sometimes be affected by factors the same as or analogous to factors in other kinds of cases. These could include vindication, revenge, and various other kinds of reputational interests.

¶45 Settlements are by far the most common litigation outcome in general, not just in patent cases.[83] Settlement cannot necessarily be considered either a success or a failure for the plaintiff in any kind of case, including a patent one. On the one hand, anecdotal evidence reveals that at least some form of payment usually accompanies the settlement of a patent infringement case.[84] On the other hand, a plaintiff receiving a damages payment that does not even cover its litigation expenses is unlikely to view the outcome as a success, and a plaintiff receiving a payment exceeding its expenses is still likely to take that view if far more was genuinely expected in order to be viewed as compensable.

¶46 However, before drawing any conclusions about settlement rates, we also examined rates after combining clear and likely settlements. Because we did not characterize a case outcome as a likely settlement unless the evidence created a fairly strong inference of such, the results of combining clear and probable settlements reported in Table 6 may actually be the most reliable.

---

[80] In patent law, this is true even if the patentee has already litigated and won one or more cases before its first loss. *See, e.g.*, Mendenhall v. Barber-Greene Co., 26 F.3d 1573, 1577 (Fed. Cir. 1994); Miss. Chem. Corp. v. Swift Agric. Chems. Corp., 717 F.2d 1374, 1379 (Fed. Cir. 1983); Stevenson v. Sears, Roebuck & Co., 713 F.2d 705, 709 (Fed. Cir. 1983).

[81] MedImmune, Inc. v. Genentech, Inc., 549 U.S. 118, 137 (2007); Lear, Inc. v. Adkins, 395 U.S. 653, 673–74 (1969).

[82] Allison, Lemley & Walker, *Repeat Patent Litigants*, *supra* note 3, at 710. Such a plaintiff still faces the risk created by nonmutual collateral estoppel, however.

[83] *See, e.g.*, Theodore Eisenberg & Charlotte Lanvers, *What is the Settlement Rate and Why Should We Care?*, 6 J. EMPIRICAL LEG. STUD. 111, 115 (2009) (in a 2009 study of 3,300 cases of various kinds from the Eastern District of Pennsylvania and the Northern District of Georgia decided in 2001-2002, finding a settlement rate in the two districts combined of 66.9%, leading them to challenge the oft-asserted proposition that at least 90-95% of civil cases end with a settlement); Jason Scott Johnston & Joel Waldfogel, *Does Repeat Play Elicit Cooperation? Evidence from Federal Civil Litigation*, 31 J. Legal Stud. 39, 40 (2002) ("settlement rates for some type of cases—such as torts—exceeding 90 percent"); Jay. P. Kesan & Gwendolyn G. Ball, *How Are Patent Cases Resolved? An Empirical Examination of the Adjudication and Settlement of Patent Disputes*, 84 WASH. U. L. REV. 237, 265 (2006) (in a study of patent cases finding likely settlement rates ranging from 65 percent to 68 percent for the years 1995, 1997, and 2000). Whatever the settlement rate in any category of case, including patent infringement, it is without doubt the most frequent result.

[84] Email correspondence with Mark A. Lemley, a noted patent scholar at Stanford Law School who also litigates patent infringement cases (Feb. 19, 2011) (on file with authors). Hard statistics of any quality are very difficult to generate because the terms of settlement agreements are almost always made confidential by terms of the agreements.

Table 6: (Clear + Likely) Settlement Rates for Internet and NIP Cases

Patent Type	All Cases	Internet Patents-All	Internet Model	Internet Technique	NIPs
Settled Cases (Clear plus Likely)	334	201	72	129	133
Total Cases	453	251	103	148	202
Settlement Rate	73.73%	80.08%	69.90%	87.16%	65.84%
P-Value Chi-squared Comparison w/ NIPs		0.001	0.475	0.000	

¶47 Our results were the same in kind but dramatically different in degree after combining the two groups: (1) As one would expect, settlement rates rose very substantially to levels closer to, but somewhat larger than those observed in patent cases by Kesan & Ball.[85] (2) The settlement-rate differences between all Internet patents and NIPs, and those between Internet business techniques and NIPs are far more significant. (3) The settlement rates of Internet business models and NIPs move much closer to one another; the rate is higher for Internet business models but for statistical purposes is essentially the same. (4) Internet technique cases settle at a rate that is higher than that of Internet business model cases with an exceptionally high degree of significance.

¶48 The much higher settlement rate for cases involving Internet patents was driven primarily by those cases in which narrower Internet business technique patents were being litigated. Internet business technique patents are more likely to have been issued to large entities, and Internet business model patents were more likely to have been issued to small entities, but prior research reveals no reason why this fact would cause a major difference in settlement rates. Allison, Lemley, and Walker found no statistically significant difference in settlement rates between large and small entities—the rates were essentially identical.[86] The data set used in that study did contain a large proportion of anomalous patents, however, because the authors were comparing patents that had been litigated eight or more times during a nine-year period with those that had been litigated only once during that period. Moreover, the data in our current study includes a large portion of patents of a very specific type—Internet business processes—that were not at issue in the Allison, Lemley, and Walker study. Consequently, we cannot draw any reliable conclusions about whether the characteristics of the patent owners contributed to the higher settlement rate for Internet patents.

¶49 The much higher settlement rates for Internet patents overall and for Internet business technique patents particularly defies ready explanation when we examine the rates without taking into account the effects of other variables. Thus, we constructed a logistic regression model in hopes of gaining more insight into the cause of the observed difference. Although we ran a regression using only clear settlements, and another combing clear and probable settlements, we report only the latter because we are rather confident that the combination produces a fairly accurate portrait of reality and

[85] Kesan & Ball, *supra* note 83, at 265 (finding likely settlement rates ranging from 65-68% for the years 1995, 1997, and 2000); *see also* Allison, Lemley & Walker, *Repeat Patent Litigants*, *supra* note 3, at 688-89 (in an empirical study of outcomes in patent cases filed during 2000-2008, finding a settlement rate of almost 90% in patent cases, but one of the two data sets used in that study was quite different than what we used in our study—one set consisted of 106 patents that had been litigated 8 or more times during 1-1-2000 to 2-28-2009, and these repeat players in patent litigation probably were more risk averse about allowing a case to go to judgment).

[86] Allison, Lemley & Walker, *Repeat Patent Litigants*, *supra* note 3, at 690.

because the combination produces larger numbers of observations that should enhance reliability. We also sought to discover other variables that may have affected these settlement rates.

¶50 Because this and the other regressions that follow use a database of only litigated patents, whereas previous ones also included unlitigated ones, we included litigation-related controls for (1) "litigation year" (the year in which the case was filed), and (2) "regional circuit" (the regional federal circuit in which the district court was located). The first control accounts for any possible temporal differences in patent law or litigation. The second accounts for any possible effects caused by different circuit court rules; it also provides a rough control for any geographic differences in jury or judge attitudes toward this type of litigation and of the parties' knowledge of these attitudes on settlement negotiation strategies—doing so at the district court level would have been more precise, but we believed that controlling for all 53 districts in which these cases were litigated would have provided more complication than value. An example of how the "regional circuit" control would help the explanatory power of these regressions is that an inordinate percentage of early Internet patents had California inventors, and cases involving these patents were disproportionately litigated in California federal districts where juries might favor these inventors.[87] We report these results in Table 7.

[87] Of the 453 case-patent pairs across both Internet patents and NIPs, 99 were litigated in the three California federal districts, which accounted for most of the cases within the geographic area of a single circuit: the Ninth. Again, Internet patents were disproportionately of California origin.

Table 7: Logistic Regression with Clear and Likely Settlements Included

Dependent Variable = (Clear + Likely) Settlement or Not	Logit (1) Case-Patent Pairs	Logit (2) Case-Patent Pairs
Independent variables	*Odds ratios; Standard errors in parenthesis*	
Internet patent-all	5.853 (3.193) ***	
Internet-model		4.069 (2.350) **
Internet-technique		9.178 (5.759) ***
Total Defendants	0.876 (0.041) ***	0.882 (0.037) ***
Patent owner filed case (not declaratory judgment action by accused infringer)	2.047 (0.867) *	2.263 (1.022) **
Transferred from other ct.	0.485 (0.234)	0.539 (0.257)
No. of Independent Claims	0.98 (0.044)	0.999 (0.047)
No. of Dependent Claims	1.007 (0.007)	1.006 (0.007)
No. of U.S. Pat Reference	0.982 (0.008) **	0.981 (0.008) **
No. of Inventors	0.883 (0.087)	0.874 (0.083)
Small Entity	0.665 (0.219)	0.707 (0.231)
Years in PTO	1.019 (0.058)	1.017 (0.057)
Age of Patent at Litigation	1.507 (0.339) *	1.528 (0.345) *
Patent Previously Litigated	1.135 (0.349)	1.104 (0.341)
Adjusted Forward Citations	0.998 (0.069)	0.989 (0.072)
Additional Controls	year and circuit	year and circuit
Robust Standard Errors	Clustered on patent number	Clustered on patent number
Sample Size	453	453
Pseudo R-Squared	0.160	0.167
*** p < 0.01; ** p < 0.05 * p < 0.10		

¶51 Regarding the results reported above in Table 7, the fact that we are examining an outcome (settlement or not) in a data set characterized by the appearance of some patents in more than one case warrants a small observation. Because there are some patents the infringement of which was asserted in more than one of our cases, settlement or any other outcome in a particular case may not be completely independent of outcomes in some other cases. Regression analysis is more reliable when each observation is independent of other observations. Fortunately, standard error clustering can be used to substantially ameliorate the potential problems caused by incomplete independence.

We note in Table 7 that the regression analyses were conducted with clustering on the standard errors of individual patents. We do the same in other regressions on outcomes.

¶52 We can see several rather striking results from our regression models on likelihood of settlement. First, when we take into account the cumulative effects on settlement of the other variables, Internet patent cases settled at a far greater rate than NIP cases—they were almost six times more likely to settle overall, and cases involving Internet technique patents were nine times more likely to settle than those involving NIPs. And we were surprised to find that Internet model cases were four times more likely to settle than NIP cases when the effects of other factors were considered. These findings were either significant or highly significant.

¶53 We also discovered that adding an additional defendant to a case made it about 13% *less likely* that a case would settle, whether it involved an Internet patent or an NIP. This finding, which was highly significant statistically, confirms earlier research revealing that the more defendants there are, the lower is the probability of settlement. [88] Recall that we are using the term "defendant" to refer not only to actual defendants in patent infringement suits, but also to plaintiffs in declaratory judgment actions that then become counterclaim defendants. The possible reasons that larger numbers of defendants seem to make settlement in patent cases less much more difficult to achieve were examined by Allison, Lemley, and Walker.[89] Recall, too, that the new America Invents Act places some meaningful constraints on the ability of patent owners to name large numbers of defendants in infringement litigation, although the phenomenon is likely to continue on a somewhat smaller scale.[90]

¶54 Some other findings from our regressions on settlement likelihood are worth mentioning, but we haven't a clue about why they turned out this way. For example, a case begun in the more common manner, the patent owner suing for infringement, was more than twice as likely to settle as an action instituted by a potential infringer seeking a declaratory judgment of noninfringement and invalidity (and then becoming a counterclaim defendant). This finding was statistically significant at $p < 0.05$ in one of the above models (separating Internet patents into subtypes) and at the marginally significant level of < 0.10 in the other model (not separating Internet patents into subtypes). There could be something about the mindsets of the average litigants in the two situations that makes them more or less intransigent, a possibility that may be worth investigating in future research. Also, more references to U.S. patents as prior art decreased the odds of settlement slightly, but enough so that the result was significant at $p < 0.05$. This finding, which defies explanation, must be a statistical artifact of some interaction among certain other independent variables. Finally, the age of a patent at the time litigation was initially instituted was weekly predictive of greater settlement likelihood: the magnitude was quite substantial in that each year of additional patent age increased settlement odds by 50%, but the result was only marginally significant statistically ($p < 0.10$). As with the "years in PTO" variable, one year is a large portion of the average total time that an application spends in the PTO before issuance, so we probably should not make much out of this.[91] Also worth nothing, although not part of this regression analysis, is that there was no difference in the average ages of Internet patents and NIPs at the time litigation was first instituted—both kinds of patent averaged 4.5 years of time elapsed since the patent grant.

¶55 As with the decision to institute patent litigation in the first place, the decision to settle is a complex one affected by countless factors, most of which are undiscoverable or at least unmeasurable.[92] Thus, we should not feel particularly bad (and we don't, as a matter of fact) about the fact that our regression models explained only about 1/6 of the occurrence of settlement (pseudo r-squared values of 0.160 and 0.167).

---

[88] Allison, Lemley & Walker, *Repeat Patent Litigants*, *supra* note 3, at 699-700.

[89] *Id.*

[90] *See supra* note 21.

[91] It is possible, of course, that several very old or very young patents all settled or all did not settle. Given that the average patent in our data set was only about ¼ of the way through its life when litigation was instituted, and that

[92] *See supra* note 48.

B. Trial Rates—Internet Patents and NIPs

¶56 The rate at which litigated Internet patents and NIPs went to trial presents another relevant dimension on which to make comparisons. This category *excludes* those cases that clearly or probably settled before trial (but not those that settled after trial) and those that were disposed of by the judge's ruling on a pretrial motion. The resulting measures, presented in Table 8, reveal a bit more evidence about the parties' attitudes toward risk after having allowed pretrial discovery and *Markman* hearings[93] to have concluded and having both survived summary judgment motions without bargaining to an outcome. We present only univariate comparisons of trial rates, as the numbers of observations are clearly too small for regression.

Table 8: Rate at Which Cases Went to Trial

Patent Type	All Cases	Internet Patents-All	Internet Model	Internet Technique	NIPs
Went to Trial	18	8	5	3	10
Total Cases	453	251	103	148	202
Trial Rate	3.97%	3.19%	4.85%	2.03%	4.95%
P-Value Fisher's Exact Comparison w/ NIPs		0.347	1.000	0.252	

¶57 In absolute terms, NIP cases went to trial more frequently than Internet patents overall and Internet technique patents. The rates for NIPs and Internet models were not only very close to each other but also very close to the 5% rate found by Kesan and Ball in the 1990's, although their results were based on a larger number of observations. The data set employed by Kesan and Ball resembled our data set of NIPs in that it involved a cross-section of technology areas.

¶58 The number of cases that made it to trial was not only too small for us to conduct regression analysis, but also was not large enough for any differences to be statistically significant in the univariate comparisons. We can take comfort, however, in the fact that our results were not out of line with those of prior research.

C. Judicial Dispositions—Internet Patents and NIPs

1. Patent Owner Wins

¶59 In our study of outcomes on the merits, we first looked at the rate at which patent owners won those cases in which there was a formal judicial disposition, and then at the composition of those wins.

[93] So named after Markman v. Westview Instruments, Inc., 517 U.S. 370 (1996), which held that claim construction—interpretation of disputed language in the claims within a patent that are being asserted against an accused infringer—is a question of law solely for judicial determination and is not a jury question. A trial judge engages in claim construction after having held a *Markman* hearing, which is a prerequisite to any finding of infringement or invalidity because of the need to first determine the precise contours of the claimed invention. Among other things, the *Markman* case has resulted in the parties to a patent case having to invest more at an earlier stage of the proceeding and increasing the allotment of resources to the pretrial stage. All of the cases in our study were decided well after the Markman ruling, so we do not have to account for any effect it may have had on settlement rates.

Table 9: Rate at Which Patent Owners Won

Patent Type	All Cases	Internet Patents-All	Internet Model	Internet Technique	NIPs
Patent Owner Win	41	8	5	3	33
Total Cases	453	251	103	148	202
Win Rate	9.05%	3.19%	4.85%	2.03%	16.34%
P-Value Fisher's Exact Comparison w/ NIPs		0.000	0.003	0.000	

¶60 Our examination of patent owner win rates necessitated that we work with a relatively small number of observations because most cases, whether patent infringement or other kind, do not make it to a formal judicial disposition on the merits. Given the small number of data points, it is quite remarkable the owners of NIPs won at a rate so much greater than Internet patents and their subtypes that the differences were statistically significant at extremely high levels ($p = 0.000$ when NIPs were compared with all Internet patents and with Internet technique patents, and $p = 0.003$ when NIPs were compared with Internet model patents).[94] One would need a much larger data set than ours with much larger numbers of merits outcomes to be completely sure that owners of Internet patents win at rates so much lower than other patents, but we have reasons to believe our results. Internet patents are but a subset of patents on software inventions more generally, and prior research disclosed that owners of software patents do not win their infringement cases very often. A recent study showed that, across two data sets of litigated patents—one including patents that had been the subject of multiple lawsuits during a nine-year period and the other including patents that had been litigated only once during that period—owners of software patents won only 12.9% of the cases.[95] This is a considerably larger win rate than we find with our Internet patents, but it is still small and in rough accord with the rate for Internet patents. Moreover, there was a substantial representation of NPE's—non-practicing entities that own patents but that do not make and sell products—among our group of Internet patents. There were product-producing companies among this group, but NPEs accounted for a share that was disproportionate when compared with their share of all kinds of patents. Prior research also has shown that NPEs win a very small portion of the patent cases they file—9.2% across the multiply- and once-litigated data sets referred to above.[96]

¶61 Almost as remarkable, though, is that owners of NIPs won only 1/6 of the time. Most of our NIPs were not software, and most of the owners were product companies—not all in either case, but a clear majority. Given this information, one would be justified in expecting a higher win rate. Prior research showed product companies winning 40% of the time, and owners of non-software patents winning 51% (both including default judgment wins as in the current study). [97]

¶62 Next we examination the composition of patent owner wins, i.e., their stage and nature.

---

[94] The difference between win rates for the two subtypes of Internet patents was not statistically significant (using Fischer's exact test for the comparison because of the small numbers of observations, $p = 0.279$).

[95] Allison, Lemley & Walker, *Repeat Patent Litigants*, *supra* note 3, at 686.

[96] *Id.* at 694. One should not interpret these statements as implying any kind of value judgment about the enforcement of patents by NPEs. The pejorative term "patent troll" that is often applied to NPEs indiscriminately fails to recognize the many types of patent owners who do not make products and engage in perfectly legitimate activities, such as universities, independent inventors and the small businesses they create, and so on. Some such patent owners probably engage in abusive litigation tactics, and some do not. The subject of NPEs is far beyond the scope of this article.

[97] *Id.* at 696.

Table 10: Composition of Patent Owner Wins

	Type of Patent Owner Win	Percent of Total
Default Judgment	4	9.76%
Summary Judgment	1	2.44%
Consent Judgment	24	58.54%
Trial	12	29.27%
Total	41	100.00%

¶63 By far the greatest portion of patent owner wins took the form of consent judgments that followed soon after a significant owner win (such as a summary judgment that did not fully resolve the case) on infringement or validity. Others could reasonably disagree with our characterization of these events as patent owner wins rather than settlements. We did not take this decision lightly and required substantial evidence in the docket entries or court documents of a very meaningful patent owner win not long before entry of the CJ. If such evidence was not present, we called the CJ a clear settlement. We believe, therefore, that those situations in which we treated a CJ as a patent owner win showed a major qualitative difference when compared with those we classified as settlements.

¶64 There was no surprise in our finding that patent owners won only rarely at the summary judgment stage, given that a patent owner must prove infringement of at least one claim—an intensely fact-driven issue—to such a degree that there is no genuine issue of disputed fact and that the accused infringer's evidence of invalidity must be so deficient as to similarly not create a genuine fact issue.

¶65 That patent owner wins at trial constituted less than 30% of overall wins does not seem out of line with reality and warrants no comment.

¶66 Results of our logistic regression on patent owner wins follow in Table 11.

Table 11: Logistic Regression with Patent Owner Wins

Dependent Variable = Patent Owner Wins	Logit (1) Case-Patent Pairs	Logit (2) Case-Patent Pairs
Independent variables	Odds ratios; Standard errors in parenthesis	
Internet patent-all	0.149 (0.101) ***	
Internet-model		0.337 (0.231)
Internet-technique		0.062 (0.053) ***
Total No. of Accused Infringers	1.030 (0.045)	1.016 (0.043)
Transferred from other ct.	2.128 (2.530)	1.824 (2.03)
No. of Independent Claims	0.839 (0.092)	0.814 (0.091) *
No. of Dependent Claims	0.992 (0.010)	0.992 (0.01)
No. of U.S. Pat Reference	1.008 (0.010)	1.01 (0.011)
No. of Inventors	1.331 (0.180) **	1.348 (0.188) **
Small Entity	1.594 (0.811)	1.352 (0.689)
Years in PTO	1.003 (0.063)	1.001 (0.065)
Age of Patent at Litigation	1.009 (0.195)	0.989 (0.196)
Patent Previously Litigated	1.708 (0.773)	1.78 (0.782)
Adjusted Forward Citations	1.061 (0.075)	1.104 (0.09)
Additional Controls	year and circuit	year and circuit
Standard Errors	Clustered on patent number	Clustered on patent number
Sample Size	453	453
Pseudo R-Squared	0.236	0.251
*** p < 0.01; ** p < 0.05 * p < 0.10		

¶67 The regression results strongly confirm what we found in the univariate comparisons; namely, that owners of Internet patents won at a strikingly smaller rate than did owners of NIPs. The odds

ratios were a small fraction of 1, and the differences were significant at $p < 0.01$ for all Internet patents and Internet technique patents. Owners of Internet model patents also won at a small fraction of the rate of NIP owners, but the small number of observations prevented a finding of statistical significance. These findings, of course, take into account the cumulative effects of all of the other independent variables we included, as well as controls for the year of litigation and the regional circuit in which the district court was located.

¶68 An unexpected finding was that the number of inventors had a strong positive effect on the likelihood that owners of both Internet patents and NIPs would win—the effect of an additional inventor on the odds of a patent owner win was over 30% with statistical significance at $p < 0.05$. To our knowledge, this is not an effect that has been found elsewhere, and should be viewed with caution unless replicated in other contexts with larger sample sizes.

¶69 What would have been a notable finding had it been statistically significant was that, if a patent had been litigated previously (the *fact* that it *had been*, not the number of times that it had been), the patent owner was more likely to win. The magnitude of the odds increase was high, but not statistically significant, presumably because the number of patents that had been previously involved in litigation was rather low, again presenting the problem of small sample size.

¶70 The regression models for patent owner wins demonstrated greater goodness-of-fit than the settlement models, our independent variables explaining about 25% of the outcomes, possibly reflecting the fact that fewer imponderables affect win rates. As with settlements, though, there are many unknown or unmeasurable factors that influence whether a patent owner wins an infringement case.

2. Accused Infringer Wins

¶71 We next compare accused infringer wins among the patent types. We do so because these are not just (1 minus patent owner wins). We look at win rates using the total number of 453 patent assertions, or case-patent pairs, as the denominator, and most of these cases do not result in a win for either party—out of the 453, there also are settlements and procedural dispositions. As seen below, accused infringers did not win very many cases either.[98]

Table 12: Rate at Which Accused Infringers Won

Patent Type	All Cases	Internet Patents-All	Internet Model	Internet Technique	NIPs
Accused Infringer Win	29	18	12	6	11
Total Cases	453	251	103	148	202
Win Rate	6.40%	7.17%	11.65%	4.05%	5.45%
P-Value Fisher's Exact Comparison w/ NIPs		0.563	0.066	0.622	

¶72 The only notable thing about these findings is that the rates are so low. There is little of interest in the comparisons among patent types. Although Fischer's Exact test was able to work with these

[98] We should always remember, however, that a settlement could easily be viewed as being predominantly a win for one side or the other, depending on its terms. The terms of settlements remain, and probably will always remain, a black box for those of us who would like to know because one of these terms is usually a confidentiality clause. It is not unheard of for one party to a settlement that feels as though it has received vindication and has the upper hand to successfully insist on publicity, but even then it is unusual for the precise terms to be made publicly known.

small numbers of observations, no finding was statistically significant at $p < 0.05$. The win rate of accused infringers in cases involving broad Internet model patents was sufficiently larger than their win rate in NIP cases for marginal significance at $p < 0.10$ (0.066).

¶73 As we did with patent owner wins, we then examined the composition of wins by accused infringers.

Table 13: Composition of Accused Infringer Wins

	Type of Accused Infringer Win	Percent of Total
Consent Judgment	2	6.90%
Summary Judgment	21	72.41%
Trial	6	20.69%
Total	29	100.00%

¶74 That the greatest portion of accused infringer wins came at the summary judgment stage is entirely expected. The reason is that the patent owner bears the burden of proving infringement, and for a win on summary judgment the accused infringer need only show that the owner has failed to create a genuine fact issue on one of its elements of proof. The trend of accused infringer wins on summary judgment of noninfringement has been an upward one in recent years.[99] The experience of accused infringers with consent judgments shortly following a significant victory on infringement or validity was very different than that of patent owners. Whereas this kind of outcome predominated in the experience of patent owners, it was almost nonexistent for accused infringers with only two such outcomes—less than 7% of the total.

¶75 We regressed the same independent variables, and used the same controls, on accused infringer wins as we did on patent owner wins. The results are below in Table 14.

[99] *See generally* John R. Allison & Mark A. Lemley, *The (Unnoticed) Demise of the Doctrine of Equivalents*, 59 STAN. L. REV. 955 (2007) (observing throughout the increasing willingness of federal district judges to grant summary judgments of noninfringement to accused infringers on the Doctrine of Equivalents, which traditionally had been viewed as an issue especially suited for jury determination, as well as the increasing tendency of judges to grant summary judgments to accused infringers on the more fundamental question of literal infringement after the Supreme Court's decision in *Markman v. Westview Instruments, Inc.*).

Table 14: Logistic Regression with Accused Infringer Wins

Dependent Variable = Accused Infringer Wins	Logit (1) Case-Patent Pairs	Logit (2) Case-Patent Pairs
Independent variables	Odds ratios; Standard errors in parenthesis	
Internet patent-all	1.841 (1.332)	
Internet-model		2.897 (2.265)
Internet-technique		1.054 (0.978)
Total No. of Accused Infringers	1.184 (0.065) ***	1.175 (0.056) ***
Transferred from other ct.	0.846 (0.931)	0.678 (0.694)
No. of Independent Claims	0.964 (0.064)	0.95 (0.063)
No. of Dependent Claims	1.026 (0.009) ***	1.026 (0.009) ***
No. of U.S. Pat Reference	1.011 (0.017)	1.013 (0.017)
No. of Inventors	0.928 (0.187)	0.923 (0.176)
Small Entity	0.750 (0.383)	0.667 (0.341)
Years in PTO	0.764 (0.105) **	0.768 (0.105) *
Age of Patent at Litigation	1.358 (0.311)	1.386 (0.335)
Patent Previously Litigated	0.389 (0.191) *	0.414 (0.206) *
Adjusted Forward Citations	0.894 (0.099)	0.871 (0.122)
Additional Controls	year and circuit	year and circuit
Standard Errors	Clustered on patent number	Clustered on patent number
Sample Size	453	453
Pseudo R-Squared	0.313	0.322
*** $p < 0.01$; ** $p < 0.05$ * $p < 0.10$		

¶76 The regression results are in accord with those of the univariate comparisons that the odds of accused infringers winning were greater by a substantial degree when Internet patents were asserted against them than when NIPs were asserted. Despite the large magnitude of the odds increase, the

findings are not statistically significant, again because of a relatively small number of total observations.

¶77 The most notable finding from these models, however, is that the larger the number of accused infringers in a case, the more likely it was that those accused infringers would win. Recall that we use the term "accused infringer" instead of "defendant" because several of our cases were instituted by potential infringers as declaratory judgment plaintiffs, who then wound up as counterclaim defendants. It is really all the same, and to put into terms that are more common and easily understood, the more defendants there were, the greater were the odds that those defendants would win on the merits: the odds of a defendant win increased by approximately 18% with the addition of another defendant in the case, an increase that was statistically significant at $p < 0.01$. And recall that the effect is even greater as more and more accused infringers are involved because of the multiplicative effect.

¶78 This finding is in accord with that of Allison, Lemley, and Walker, who also discovered that the presence of more defendants substantially increased the odds that defendants would win.[100] This result augments the finding that, the more accused infringers there were, the lower were the odds of settlement by a substantial margin, also a finding that comported with prior research.

¶79 A finding that almost certainly should be ignored despite its high level of statistical significance is that accused infringers were more likely to win as the number of dependent claims increased in the patents asserted against them. The magnitude of the odds increase was relatively small, and statistical significance at a high level was due to the extremely large numbers of dependent claims in these patents.

¶80 Accused infringers were about 25% less likely to win with each additional year that the patent owners' applications had spent during prosecution in the PTO. The significance level was at $p < 0.05$ in one of the models and $p < 0.10$ in the other. The reverse was not true in our regression on patent owner wins—additional application pendency time had essentially no effect on patent owners' odds of winning. In the case of accused infringers, it is hard to know what to make of the finding. One could speculate that, when patent applicants invested more time and money in filing more continuing applications, which accounts for much of the average additional pendency time, they made their patents stronger in other ways and were more formidable opponents when they asserted them. Although there could be some truth in that proposition, we will not pursue it further.

¶81 When we regressed on patent owner wins, we found that a patent's having been litigated previously *increased* the patent owner's odds of winning by a degree that was large in magnitude but without statistical significance. Looking at the same variable when regressing on wins by accused infringers reveals the symmetry that the odds of an accused infringer win *decreased* by a dramatic margin when the patent asserted. The magnitude of the negative effect was large enough to overcome the small numbers problem to the point that it was marginally significant statistically at $p < 0.10$.

3. Procedural Dispositions

¶82 Finally, we ran comparisons of procedural dispositions among the different patent types. Unsurprisingly, we found no meaningful differences in the rate at which cases involving Internet patents and NIPs were terminated for procedural reasons.

[100] Allison, Lemley & Walker, *Repeat Patent Litigants*, *supra* note 3, at 699.

Table 15: Rate at Which Cases Ended on a Procedural Disposition

Patent Type	All Cases	Internet Patents-All	Internet Model	Internet Technique	NIPs
Procedural Dispositions	49	24	14	10	25
Total Cases	453	251	103	148	202
Rate	10.82%	9.56%	13.59%	6.76%	12.38%
P-Value Chi-squared Comparison w/ NIPs		0.338	0.764	0.083	

V. CONCLUSION AND SUGGESTIONS FOR FURTHER RESEARCH

¶83 Using both univariate comparisons and multiple regression techniques, we found the following: (1) Internet patents and their two subtypes were litigated at a far higher rate than NIPs—they were between 7.5 and 9.5 times more likely to end up in infringement litigation, depending on the model we used. (2) Within the category of Internet patents, those on business models were litigated at a significantly higher rate than those on business techniques. (3) Across both Internet patents and NIPs, patents issued to small entities, especially individuals and small businesses, were much more likely to be litigated than those issued to large entities. (4) Patents of all kinds with more independent claims were significantly more likely to be litigated than those with fewer independent claims. (5) Including both Internet patents and NIPs, litigated patents received many more forward citations—citations received from later patents—than did unlitigated patents. (6) Patents issued to foreign entities were significantly less likely to be litigated than patents issued to U.S. entities. (7) The more time that an application for an Internet patent or NIP had spent in the PTO prior to issuance, the more likely it was that the patent granted from that application was to be involved in infringement litigation. (8) There was no difference in the ages of Internet patents and NIPs when they became the subject of litigation—both kinds were about 4.5 years old. (9) Once patent infringement litigation was initiated, the owners of litigated Internet patents were significantly more likely to settle before judgment than the owners of litigated NIPs (especially when *probable* settlements were taken into account along with *obvious* settlements, which we believe is the more accurate metric). (10) Across both sets of patents, the larger the number of potential infringers involved in a case (defendants in infringement actions and plaintiffs in declaratory judgment actions), the less likely the case was to settle. (11) Internet patents and NIPs went to trial at about the same rate. (12) When failing to settle, the owners of NIPs won on the merits at a significantly higher rate than did owners of Internet patents—although the win rate for NIP owners was quite low at around 16%, the win rate of Internet patents was even lower by a substantial margin. This finding did not hold up in regression analysis, however; when the effects of other variables were taken into account in a logistic regression analysis, there was no significant difference in the win rate for accused infringers between Internet patents and NIPs. Accused infringers did win more often when Internet patents were asserted against them than win they defended against NIP complaints, but the relatively small number of observations prevented the difference from being statistically significant. (13) Surprisingly, owners of both kinds of patents were significantly more likely to win as the number of inventors on the patents increased. (14) The longer that applications for Internet patents and NIPs had spent in the PTO before issuance, the less likely accused infringers were to win. (15) Accused infringers were less likely to win on the merits when the Internet patents or NIPs asserted against them had been litigated previously. (16) Across both sets of patents, the larger the number of potential infringers involved in

a case, the more likely these potential infringers were to win a judgment on the merits. That is, the more infringement defendants per case, they more likely these defendants were to win. (17) There was no difference between the different types of patents in the percentage of cases that were terminated for procedural reasons. We also discussed several other findings of interest.

¶84 These findings raise some intriguing questions for further research: (1) How does litigation of patents on other relatively new technology fields compares with litigation of patents from the general population, the latter consisting predominantly of patents on more mature technologies? (2) How might litigation of patents on different technologies early in their maturation periods compare to each other? (3) Are patents on young technologies likely to be stronger and more valuable on average because there is less relevant prior art to require the narrowing of patent claim language or because applicants perceive greater potential innovation importance and are willing to invest more in the patenting enterprise, thus leading to more litigation? (4) Do patents on young technologies generate more uncertainty because of their newness, contributing to more contention? (5) The Internet patents in our data set, whether litigated or not, showed many of the same internal characteristics as patents in all fields that wind up in litigation; is the same true of patents in other emerging fields of patenting activity such as flash memory, smart phones, nanotechnology, and others?

¶85 There also are other questions for future research that are not limited to the context of new patent fields. For instance, our current study found that the average number of defendants per case significantly decreased the odds of settlement and increased the odds of a win on the merits by defendants (the term "defendants" here referring to "potential infringers" whether they are actually defendants in infringement cases or plaintiffs in declaratory judgment actions who then become infringement counterclaim defendants). This confirms recent findings by Allison, Lemley, and Walker, and may indicate that such a result is generalizable. Thus, an investigation empirically probing the wisdom of patent owners' litigation strategies when deciding whether to sue multiple alleged infringers in a single lawsuit or in several different ones may bear interesting fruit in other contexts. Although the new America Invents Act places meaningful constraints on the ability of patent owners to name multiple unrelated defendants, there nevertheless will still be circumstances in which patentees have a choice about whether to name multiple defendants in the same suit. Other things relevant to the decision being equal, might there be an optimal number of defendants per case? In addition, given our finding of significantly different litigation rates for patents on online business models and online business techniques, the former clearly appearing to have broader claims, another research question is whether one might devise methods to create empirical estimates of patent claims' breadth for use in better predicting which patents companies should fear when they assess how much freedom of action they have to innovate in a given field. The role of so-called "non-practicing entities" (NPEs) also deserves more research attention. These are companies that do not make or sell products and thus are not vulnerable to patent infringement counterclaims, as are product companies that sue for infringement. NPEs consequently may be less reluctant to sue. There is a significant but imperfect correlation between the fact that a patent was originally issued to a small entity and the identity of a patent infringement plaintiff as an NPE, and in the current study one of our findings is that patents granted to small entities, especially to individuals and small businesses, are much more likely to be litigated. Thus, it may be worthwhile to investigate the NPE/product company question in future studies of patent litigation involving other types of patents.