

Who's Suing Us? Decoding Patent Plaintiffs since 2000 with the Stanford NPE Litigation Dataset

Shawn P. Miller*

Ashwin Aravind, J.D. '18
Bethany Bengfort, J.D. '16
Clarisse De La Cerda, LL.M. '17
Matteo Dragoni, LL.M. '16
Kevin Gibson, J.D. '17
Amit Itai, J.S.D. '20
Charles Johnson, J.D. '17

Deepa Kannappan, J.D. '16
Emily Kehoe, B.S. '17
Hyosang Kim, J.D. '16
Katherine Mladinich, J.D. '18
Roberto Pinho, LL.M. '17
John Polansky, J.D. '18
Brian Weissenberg, J.D. '16

21 STAN. TECH. L. REV. 235 (2018)

ABSTRACT

Despite widespread interest in the impact of patent assertion entities (“PAEs”) on the U.S. patent system, there has been no publicly available dataset that categorizes more than a fraction of lawsuits as involving practicing entities, non-practicing entities (“NPEs”), or PAEs. To address this knowledge gap, Stanford Law School student researchers, led by Mark Lemley and Shawn Miller, have created the Stanford NPE Litigation Dataset (“the Dataset”).[†] The Dataset is the first comprehensive patent litigation dataset to categorize patent asserters and will do so for every U.S. patent lawsuit filed since 2000. With over 80% of total cases categorized as including practicing entities or one of eleven types of NPE patent asserters, the Dataset is nearing completion of the 63,000 lawsuits filed between 2000 and 2017. Thereafter, we will continue to update the Dataset with more recently filed lawsuits. The Dataset provides an invaluable tool to help policy makers craft effective rules, and help judges, litigators, and scholars better understand the nature of the entities filing patent suits. This is especially true because the Dataset captures how patent litigation patterns

* Lecturer in Law and IP Research Fellow, Stanford Law School. Ph.D., George Mason University; D.D., Notre Dame Law School; B.S., University of Notre Dame.

[†] We thank Lex Machina for extensive technical support and Colleen Chien, Christopher Cotropia, Jay Kesan, Mark Lemley, Shawn Miller, Michael Risch and David Schwartz for their contribution to this project of patent owner datasets they developed for prior analysis of various subsets of patent litigation.

evolve over an era of heightened activity and policy reform. The Dataset will reveal trends before and after passage of the America Invents Act, key Supreme Court patent cases including eBay and Alice, and various executive orders focused on increasing transparency and reducing costs in patent suits. The first half of this paper explains the motivation for creating the Dataset and details the methodology used to create it. At present, we have completed and made public a random sample of 20% of the lawsuits filed from 2000 through 2015 (10,812 lawsuits).¹ We utilize this sample in the second half of the paper to reveal, for the first time ever, trends in the share of patent disputes attributed to different types of patent asserters over a span of sixteen years. These trends show that while practicing entities dominated patent litigation in the first half of the 2000s, NPEs and PAEs now assert patents in most lawsuits. Further, the trends show that the rise of NPEs and PAEs began before—and thus is not attributable to—the 2011 change in joinder rules. We hope this data will be used to further policy discussions and therefore conclude this paper with examples of how the Dataset can be used in future research on the impact of different types of patent asserters on the patent system.

1. STANFORD NPE LITIGATION DATASET (Dec. 16, 2017), <https://law.stanford.edu/projects/stanford-npe-litigation-dataset/#slnav-brief-dataset-methodology> [<https://perma.cc/96Y4-8TP9>]. Starting Summer 2018 the full Dataset will be accessible at <https://npe.law.stanford.edu>.

TABLE OF CONTENTS

I. INTRODUCTION.....	237
II. CREATING THE STANFORD NPE LITIGATION DATASET.....	243
A. <i>Taxonomy</i>	244
B. <i>Coding Methodology</i>	246
III. DESCRIPTION OF THE DATA: DESCRIPTIVE STATISTICS, TIME TRENDS, AND PRELIMINARY FINDINGS.....	252
A. <i>Stanford NPE Litigation Dataset Descriptive Statistics and Time Trends</i>	253
1. <i>Counts of Lawsuits by Patent Asserter Type</i>	253
2. <i>Trends in Lawsuits by Patent Asserter Type</i>	257
B. <i>Findings Combining Stanford NPE Litigation Dataset with other Data Sources</i>	262
1. <i>Technology</i>	263
2. <i>Lawsuit Duration and Outcome</i>	266
IV. IMPACT ON AND SUGGESTIONS FOR FUTURE RESEARCH.....	270
V. CONCLUSION.....	274

I. INTRODUCTION

The term “patent troll” has become part of the American legal lexicon, though few can agree on a precise definition. The concept of the patent troll first appeared in the 1994 short film, “The Patents Video,” which featured a troll demanding fees from would-be bridge-crossers as a satirical illustration of the modern patent system.² The origin of the term in common usage, however, is most often credited to Peter Detkin, counsel to Intel in the late 1990s.³ At the time, Intel was hit by a swath of what it believed to be unmeritorious lawsuits by businesses that did not manufacture products. One lawsuit in particular infuriated Detkin, prompting him to designate the lawyer who filed it a “patent extortionist.”⁴ In retaliation, the lawyer filed suit for libel.⁵ Detkin realized that he would need to create a term that was similarly evocative yet appreciably less literal. He conducted an internal contest at Intel to create a new moniker and the term “patent troll” was born.⁶

2. See Paula Natasha Chavez, *The Original Patent Troll*, YOUTUBE, (Jan. 28, 2007), <https://www.youtube.com/watch?v=LOGoZFzHkhs> [<https://perma.cc/96AV-TXSW>].

3. See Ira Glass, *When Patents Attack!*, THIS AMERICAN LIFE 4 (July 22, 2011), http://www.thisamericanlife.org/sites/default/files/TAL441_transcript.pdf [<https://perma.cc/5JQ8-UF8N>].

4. *Id.*

5. *Id.*

6. *Id.*

To date, interest groups and scholars that use the term disagree about which types of patent owners are and are not patent trolls, while many others refuse to use the term at all. Among those who use it, almost all agree that a patent troll must be a non-practicing entity (“NPE”), or an entity that owns patents but does not create or sell products or services. Most definitions further specify that a patent troll must be an entity that exists to assert patents against other actors, birthing the term “patent assertion entity” (“PAE”).

While the terms PAE and NPE are often used interchangeably, the difference in definition is important. For example, universities and technology development companies are NPEs, but most commentators would not consider these PAEs. Individual inventors who assert to collect on others’ use of their innovations are NPEs, but there is disagreement on whether they are PAEs. Still other NPEs are startups that failed to commercialize a patented idea that may still be financially worthwhile. At a minimum, patent policymakers should consider the different reasons why various types of NPEs exist, as well as differences in their assertion behavior generally and their litigation activity in particular. Indeed, research, including ours, reveals a complex ecosystem of patent litigation, filled with various practicing entities and NPEs employing diverse litigation strategies.

The principal motivation for this project, however, is the perceived prevalence of the PAE business model and we define PAEs as entities that employ patents primarily to obtain license fees, rather than to support the transfer or commercialization of technology. Over the last decade, a growing debate among scholars, the bar, industry, and public agencies about the impact of PAEs has sparked government scrutiny. Critics have come to believe that the steady escalation of PAE enforcement activity, including litigation, is harming innovation and acting as a tax on producers and consumers. While there are a variety of theoretical business models for PAEs, with different predicted effects, critics essentially argue that PAE enforcement imposes costs on business, consumers, and the courts that exceed the financial benefits to PAEs themselves and the incentives their gains might generate for inventors.

Some critics have gone so far as to argue abusive litigation tactics conducted by PAEs, coupled with excessive verdicts, cost the U.S. economy \$1.5 billion per week.⁷ PAE activities have been described as taxes on innovation and colossal drains on manufacturing, with PAEs asserting patents of dubious validity and value in order to extract licensing fees from the entities that are actually making products available to the

7. Larry Downes, *Everyone Hates Patent Trolls, But Here’s the Root Problem with Our Broken System*, WASH. POST (May 4, 2015), <https://www.washingtonpost.com/news/innovations/wp/2015/05/04/everyone-hates-patent-trolls-but-heres-the-root-problem-with-our-broken-system> [https://perma.cc/YP7F-RC2K].

public.⁸ Similarly, many scholars and defense attorneys argue that PAEs are a corrupt byproduct of existing law, which enables PAEs to operate under the veil of shell companies to hide the true nature of their litigation activities and rent empty office spaces to take advantage of favorable federal jurisdictions for filing suit.⁹ Finally, critics believe—and at least some empirical evidence supports the idea—that in many cases only a tiny fraction of the value of the patent extracted by a PAE ever reaches the original inventor, thus distorting the incentive structure designed by the patent system to encourage innovation.¹⁰ This final criticism of PAEs is a response to the principal argument of PAE defenders—that PAEs serve as efficiency-enhancing business intermediaries between inventors and commercializers.¹¹

In response to PAE critics, public officials have begun to explore the nature and impact of these business entities. For example, in December 2012 the Federal Trade Commission (“FTC”) and the Department of Justice (“DOJ”) held a joint workshop on the behavior of PAEs, followed a month later by a Patent and Trademark Office (“PTO”) workshop on proposed “sunshine rules” that would provide more ownership transparency.¹² In October 2016, the FTC released an in-depth study of the

8. See James Bessen, *The Evidence Is In: Patent Trolls Do Hurt Innovation*, HARV. BUS. R. (Nov. 2014), <https://hbr.org/2014/07/the-evidence-is-in-patent-trolls-do-hurt-innovation> [<https://perma.cc/FSD5-BUSG>] (finding the more R&D a firm performs, the more likely it is to be sued for patent infringement).

9. See *Judge Takes on Patent Troll with ‘Sham Employee’; Forces Troll to Defend Practice Before a Jury*, TECHDIRT, (Sept. 26, 2013), <https://www.techdirt.com/articles/20130926/02410524662/judge-takes-patent-troll-with-sham-employee-forces-troll-to-defend-practice-before-jury.shtml> [<https://perma.cc/3BTE-6TK2>] (finding PAE created a shell company, rented empty office space in East Texas, and hired sham employees). However, the Supreme Court’s decision in *TC Heartland LLC v. Kraft Foods Group Brands LLC*, 137 S. Ct. 1514 (2017), may significantly limit patent-litigation forum shopping. See, e.g., Colleen V. Chien & Michael Risch, *Recalibrating Patent Venue*, 77 MD. L. REV. 47 (2017) (estimating that 58% of cases filed in 2015 would have been filed in a different venue had the more restrictive venue rule adopted in *TC Heartland* been in effect).

10. See James Bessen & Michael J. Meurer, *The Direct Costs from NPE Disputes*, 99 CORNELL L. REV. 387, 411 (2014) (finding that payments to individual inventors come to only 7% of PAE licensing revenues and that less than one-quarter of these PAEs’ revenues flow to innovative activity, which Bessen and Meurer define as purchases of other patents or direct R&D expenses).

11. See, e.g., Daniel F. Spulber, *Patent Licensing and Bargaining with Innovative Complements and Substitutes*, (Northwestern Law & Econ Research Paper No. 16-12, Aug. 3, 2016); Michael Risch, *Licensing Acquired Patents*, 21 GEO. MASON L. REV. 979 (2014); Ryan Holte, *Trolls or Great inventors: Case Studies of Patent Assertion Entities*, 59 ST. LOUIS. U. L.J. 1 (2014); James F. McDonough III, *The Myth of the Patent Troll: An Alternative View of the Function of Patent Dealers in an Idea Economy*, 56 EMORY L.J. 189 (2006); Ashish Arora & Robert P. Merges, *Specialized Supply Firms, Property Rights, and Firm Boundaries*, 13 INDUS. & CORP. CHANGE 451 (2004).

12. See *List of Public Workshops*, U.S. DEP’T OF JUSTICE, ANTITRUST DIV. (May 12, 2016), <http://www.justice.gov/atr/public/workshops/pae/index.html> [<https://perma.cc/Q5NB-UJV8>]; *Roundtable on Real Party in Interest Information*, U.S. PATENT

behavior and organization of a select group of PAEs.¹³ Further, as part of the America Invents Act (“AIA”), Congress directed the Government Accountability Office (“GAO”) to conduct a study of the consequences of patent litigation brought by NPEs.¹⁴ Additionally, President Obama took executive action aimed at curtailing PAE litigation, and issued an executive order requesting the PTO initiate a rulemaking process requiring patent applicants and owners to regularly update ownership information.¹⁵ Congress also has considered—though not enacted—several substantive reform bills since the AIA. The Innovation Act, for example, would allow manufacturers to step in and defend end users in litigation and offer measures to reduce the often enormous costs of discovery.¹⁶

The goal of this paper is to contribute quantitative evidence to the ongoing policy debate surrounding PAEs. We aim to accomplish this goal by introducing and describing the first comprehensive public dataset to classify the type of entities asserting patents in every patent lawsuit filed in U.S. district court from 2000 to present. With the release of the Stanford NPE Litigation Dataset (“the Dataset”), researchers will have access to data that can help inform patent policy related to PAEs and other types of entities.¹⁷ We have made sufficient progress in amassing the data that we contribute early analysis in this paper, as well as provide a full description of the Dataset and our underlying methodology. In addition, we invite comment from public agencies, policymakers, scholars, and practitioners, on the methodology used to create the underlying dataset, our initial findings on the scope of PAE and NPE litigation, and the direction that future research should take to help inform national legislation and PTO policy and regulation.

Roadmap

We describe the Dataset and the methodology utilized to produce it in Part II. Briefly, we use thirteen different classifications to describe the

& TRADEMARK OFFICE (May 12, 2016), <http://www.uspto.gov/ip/officechiefecon/roundtable-RPI-agenda.1.pdf> [https://perma.cc/6RGX-SND8].

13. See *Patent Assertion Entities (PAE) Study*, U.S. FED. TRADE COMM’N (May 12, 2016), <https://www.ftc.gov/policy/studies/patent-assertion-entities-pae-study> [https://perma.cc/UY43-V4KF].

14. See *Intellectual Property: Assessing Factors that Affect Patent Infringement Litigation Could Help Improve Patent Quality*, GOV’T ACCOUNTABILITY OFFICE 17 (Aug. 22, 2013), http://www.uspto.gov/sites/default/files/aia_implementation/GAO-12-465_Final_Report_on_Patent_Litigation.pdf [https://perma.cc/6SKH-479F].

15. *Attributable Ownership*, U.S. PATENT & TRADEMARK OFFICE (May 12, 2016), <http://www.uspto.gov/patent/initiatives/attribution-ownership> [https://perma.cc/E3AJ-JB2U].

16. See generally *Innovation Act*, H.R. 9, 114th Cong. (2015).

17. See *infra* notes 19 and 20 and accompanying text.

different types of entities asserting patents in litigation. These classifications include companies manufacturing products or selling services (a.k.a., “product companies” or “operating companies”), PAEs started by individual inventors, and PAEs that acquired patents from others. Additionally, we employ various methods of quality control to ensure the accuracy of the data. To date, we have completed review of over 80% of the 63,000 lawsuits filed between 2000 and 2017, including a random sample of 20% of all lawsuits filed between 2000 and 2015. The random sample is available for download on the project website¹⁸ and starting Summer 2018 the full Dataset will be accessible to scholars at <https://npe.law.stanford.edu>.¹⁹

In Part III, we present descriptive statistics and basic time trends revealed by the Dataset and also a preliminary demonstration of how the Dataset can be combined with other data sources to analyze patent litigation. Our general findings indicate that litigation growth in the past fifteen years is due largely to the rise of PAE-initiated lawsuits. These lawsuits tend to involve software patents and resolve much more quickly than other lawsuits, lending credence to the theory that PAEs often assert with the goal of extracting quick “nuisance value” money from defendants.²⁰ When PAEs assert their patents through judgment on the merits, they fare far worse than practicing entities asserting patents.

Finally, in Part IV, we explain the Dataset’s potential in future research and present suggestions on how we hope it will be used. At its core, our research fills a knowledge gap by categorizing every plaintiff in every patent lawsuit over many years. As a result the Dataset enables scholars to conduct more complete analysis of the U.S. patent system. We encourage the Patent Office and other patent policymakers to use the Dataset and future research that utilizes it to craft targeted data-driven policy.

Key Findings from Part III

Preliminary results utilizing our 20% random sample reveal:

18. STANFORD NPE LITIGATION DATASET (Dec. 16, 2017), <https://law.stanford.edu/projects/stanford-npe-litigation-dataset/#slnav-brief-dataset-methodology> [<https://perma.cc/96Y4-8TP9>].

19. Contact Shawn Miller, Research Lead, Stanford NPE Litigation Dataset (smiller@law.stanford.edu) to request access.

20. Mark Lemley & A. Douglas Melamed, *Missing the Forest for the Trolls*, 113 COLUM. L. REV. 2117, 2170 (2013).

- **Litigation Growth:** Using the total number of defendants rather than lawsuits (hereinafter “defendant-lawsuit pairs”) to account for the AIA joinder rule, we find that the number of distinct patent disputes nearly doubled between 2000 and 2015.
- **Rise of PAEs:** As shown by defendant-lawsuit pairs, the share of litigation attributable to PAEs (our Categories 1, 4 and 5)²¹ rose from about 15% during the early 2000s to about 45% since 2010. Licensing firms that acquire their patents from third parties (our Category 1) are responsible for most of the growth of PAE litigation, increasing from less than 5% of all pairs in 2000 to over 30% since 2010.
- **Decrease in Share of Litigation Attributable to Practicing Entities:** Conversely, since about 2006 there has been a marked decrease in the share of all patent litigation attributable to practicing entities (our Category 8). Prior to 2006, about 70% of all lawsuits and 60% of all defendant-lawsuit pairs asserted practicing-entity patents; since then, those two percentages have dropped to around 45% each.
- **Software Patent Assertion:** In lawsuits filed in 2014, about 50% of PAEs (our Categories 1, 4 and 5) asserted a patent that meets a conservative definition of covering software. In contrast, only 11% of Category 8 practicing entity lawsuits that year involved a patent that covered software.
- **Time to Termination:** Across our entire 20% random sample, suits involving only Category 1 patent asserters terminated most rapidly (average of 327 days). Suits involving only practicing entities took significantly longer to end (average of 443 days).
- **Settlement:** Among lawsuits filed in 2014 that did not end in consolidation or transfer, 81% of PAE and 62% of non-PAE suits had settled by the beginning of 2016.
- **Merit Wins and Losses:** Across our entire 20% random sample, Category 1 patent asserters won a mere 13% of judgments on the merits, compared to 35% for practicing entities.

As our results—including first-ever trends in the share of litigation attributable to practicing entities versus various types of NPEs over

21. See *infra* Section II.1 for the full description of our coding methodology.

time—suggest, the Dataset can dramatically improve understanding of the role that different types of patent owners play in the patent system. Accordingly, we hope that researchers will utilize the Dataset in independent analyses to determine if reform aimed at decreasing PAE patent enforcement is necessary and, if so, what policy action would be both proportionate and effective. We similarly encourage Congress, the PTO, FTC, GAO, and other decision-makers to utilize the Dataset to enhance patent policy. In that spirit, the Stanford PAE Symposium (May 10-11, 2017) brought together private-sector stakeholders, policy makers, judges, lawyers, scholars, and researchers to engage with the data, share findings, and think dynamically about the future of patent policy. Additionally, at the Symposium, sixteen teams of academic researchers presented their initial findings using our Dataset. The links for their work and future projects by other researchers will be on our project website after they are finalized and published.²²

II. CREATING THE STANFORD NPE LITIGATION DATASET

In partnership with Lex Machina and prominent scholars who have conducted research on NPE litigation,²³ our project objectives are: (1) to review every patent infringement lawsuit, including declaratory judgments, filed in U.S. district court since 2000 and categorize the party (or parties) asserting the patent(s) in each case (hereinafter called “patent asserters”) as a practicing entity or as one of 11 types of NPEs (see Table 1); and (2) to conduct a preliminary analysis of the data to determine whether litigation trends differ by patentasserter type and whether there is variation in the characteristics of litigation across patentasserter type. Our categories are adapted from the taxonomy of Allison, Lemley, and Walker (2009)²⁴ with the names of the categories listed in Table 1 below.

Part II.A discusses the patentasserter taxonomy used in the Stanford NPE Litigation Patent Dataset. Part II.B describes the methodology we use to categorize patent asserters.

22. See Shawn Miller, *Stanford NPE Litigation Dataset*, STANFORD LAW SCH. (Dec. 16, 2017), <https://law.stanford.edu/projects/stanford-npe-litigation-dataset/#slnav-brief-dataset-methodology> [<https://perma.cc/RT8Y-7SEQ>].

23. Scholars who shared how they categorized entities in prior work include: Colleen Chien, Associate Professor of Law, Santa Clara Law School; Chris Cotropia, Professor of Law, University of Richmond School of Law; Jay Kesan, Professor of Law, University of Illinois College of Law; Mark Lemley, William H. Neukom Professor of Law, Stanford Law School; Shawn Miller, Lecturer in Law, Stanford Law School; Michael Risch, Professor of Law, Charles Widger School of Law at Villanova University; and David Schwartz, Professor of Law, Pritzker School of Law at Northwestern University.

24. John R. Allison, Mark A. Lemley & Joshua Walker, *Extreme Value or Trolls on Top? The Characteristics of the Most-Litigated Patents*, 158 U. PA. L. REV. 1, 10 (2009).

A. Taxonomy

TABLE 1. PATENT ASSERTER CATEGORIES IN STANFORD NPE LITIGATION DATASET

Category	Description
1	Acquired patents
2	University heritage or tie
3	Failed startup
4	Corporate heritage
5	Individual-inventor-started company
6	University/Government/Non-profit
7	Startup, pre-product
8	Product company
9	Individual
10	Undetermined
11	Industry consortium
12	IP subsidiary of product company
13	Corporate-inventor-started company

Of the thirteen categories, only Category 8 (product company) consists of practicing entities, i.e. firms that actually make products or offer services for sale. The remaining categories are different types of NPEs.²⁵ We believe those who use the term PAE or “patent troll” are generally referring to entities that fall within Category 1 (acquired patents), Category 4 (corporate heritage), or Category 5 (individual-inventor started company).

Category 1 includes any NPE primarily in the business of asserting patents it has acquired from other entities. We include in this category large patent aggregators such as Acacia and Intellectual Ventures.²⁶ By contrast, Category 5 is comprised of firms primarily in the business of asserting patents, where the original inventor of the patents is the founder and/or owner of the NPE. Most Category 5 entities are limited liability companies owned by the inventor(s) of the asserted patents and exist solely to hold and enforce those patents. For example, we coded Ronald A. Katz Technology Licensing, L.L.C., as a Category 5 entity.

25. We reserve Undetermined (Category 10) for the very few cases without sufficient evidence to determine patent asserter type.

26. Note that Cotropia, Kesan & Schwartz, *Unpacking Patent Assertion Entities (PAEs)*, 99 MINN. L. REV. 649 (2014), and others use a taxonomy with a separate category for large patent aggregators. Most entities that Cotropia, Kesan, and Schwartz categorize as “patent holding companies” will also fall in our Category 1. They describe patent holding companies as follows: “[U]sually limited liability companies, that appear to have been formed solely to hold and enforce a patent or small portfolio of patents. As far as we can tell, these companies are not owned by the original inventor. Frequently, these companies were formed shortly before litigation commenced.” *Id.* at 670.

Category 6 includes any patent asserter that is a university, government entity, or a not-for-profit institution. Relatedly, Category 2 entities are not themselves universities but are rather “IP subsidiaries” of universities or separate licensing firms known to primarily assert patent rights on behalf of universities. Examples include the Wisconsin Alumni Research Foundation, the University of Colorado Foundation, and Competitive Technologies, Inc.

Category 8 and Category 12 have a similar relationship as Category 6 and Category 2. Category 8 product companies manufacture products, sell products, or deliver services (unrelated to patent enforcement). We do not analyze whether a particular company is making use of the patent it is asserting in its products or services, but simply whether the company sells goods or services generally. Category 12 consists of the IP enforcement subsidiaries of practicing entities. For example, we coded AT&T Intellectual Property I, L.P., as a Category 12 entity.

Category 9 includes any individual litigant asserting patents. Usually these are the original inventors suing in their own name rather than through an L.L.C. or other company they own. We also include inventors’ family trusts in Category 9.

Category 13 “corporate-inventor-started” companies generally consist of what Cotropia, Kesan, and Schwartz (2014) describe as technology development companies.²⁷ These entities engage in substantial research and development in-house and obtain patents from that research. They are generally much more likely to negotiate exclusive licenses with producers and have recurring relationships with the same producers. Valinge is an example of a corporate-inventor-started company that develops technology through R&D and licenses its IP to producing companies.²⁸

Failed startups in Category 3 are entities that originally intended to commercialize their inventions but were unable to bring their product to market before ceasing operations. In defining Category 3, we included entities that achieved only *de minimis* sales or production of prototypes before shutting down. One example is CardSoft.²⁹ Category 7 “pre-product” startups are companies that were still in operation at the time of the lawsuit but did not yet possess any relevant products or make significant sales. One example is T5 Labs.³⁰ For a firm to be a

27. *Id.* at 670.

28. For a description of Valinge and its R&D, see *About Us*, VÄLINGE, <https://www.valinge.se/about> [<https://perma.cc/96MG-QJ4Z>].

29. See Aaron S. Lukas, *Will the Inventor of the Handheld Cash Register Ever Cash-In?*, INVENTOR’S DIG. (Dec. 2014), https://www.cozen.com/Templates/media/files/Inventors%20Digest%20-%20December_.pdf [<https://perma.cc/TN7P-ZP79>].

30. See Dean Takahashi, *Who Invented Cloud Gaming? T5 Labs Tangles with*

“startup” and fit into either Category 3 or Category 7, the entity had to be described as a startup, possess venture funding, or possess other indicators of being a startup as conventionally understood.

The final two categories are Category 4 and Category 11, and occur rarely. Category 4 “corporate heritage” entities are firms that were successful producers for a sustained period of years but then transitioned to a PAE business model. Examples include Encyclopedia Britannica, Inc., IMX, Inc., and PDL Biopharma, Inc. Finally, Category 11 “industry consortiums” are industry groups that sue on behalf of their members. One example is MPEG L.A.

By the end of 2018, we anticipate completing our coding of the nearly 63,000 patent lawsuits filed between 2000 and 2017. Thereafter we will continue to release subsequent versions of the Dataset that include more recently filed lawsuits and corrections to previous coding as revealed to us through public usage and comment. To date, we have completed our review of over 50,000 lawsuits, including a 20% random sample of lawsuits filed between 2000 and 2015. We made this completed random sample publicly available in May 2017. As we describe in the preliminary results section below, the random sample is sufficient to report observations on the extent of patent litigation trends attributable to each type of patent owner, as well as significant differences in the characteristics of litigation involving these different types of patent asserters. Before delving into those results, however, we detail the methodology we employ to categorize patent infringement lawsuits.

B. Coding Methodology

Every patentasserter involved in a lawsuit corresponds to one of the thirteen categories (see Table 1). Each data point in the Dataset is a lawsuit; some lawsuits contain multiple patent asserters and some of these multiple-asserter cases contain patent asserters that fit into different categories.³¹ Those cases are assigned to multiple patentasserter categories. For example, if an infringement action is filed by an individual inventor and a product company that has an exclusive license to use the technology, then that lawsuit fits in both Category 9 and Category 8 and is categorized both as an “8” and a “9.” Accordingly, as we review each lawsuit, we assign each entity in the case that asserted a patent a category and code it into our in-house user interface.³²

OnLive (Exclusive), VENTUREBEAT (Feb. 15, 2011), <http://venturebeat.com/2011/02/15/t5labs-patent-onlive/> [<https://perma.cc/EK8H-4QVN>].

31. As we will report in our results section below, about six percent of lawsuits reviewed so far include multiple patent asserters that also fit into multiple entity categories.

32. Note that the Dataset is now publicly available. *See supra* note 18.

To determine how to categorize each entity, student researchers consult several different sources. First, they review the docket for each lawsuit and read the initial complaint and other pleadings. The complaint in particular often contains useful identifying information, such as a company touting the success of their products in the marketplace (suggesting a Category 8), or the prowess of the individual inventor in starting the company (suggesting a Category 5). In addition, reviewing the complaint also enables students to distinguish declaratory judgments, where the coded patentasserter is the defendant in the action.

Second, student researchers utilize Lex Machina's links to the patents asserted in a lawsuit to gather relevant information. For example, Lex Machina indicates whether the same patent was asserted in other cases, enabling students to review the pleadings in those cases to learn additional facts about the entity. Additionally, the Lex Machina interface allows students to determine whether the patent was asserted by the inventor, or if the patent has been assigned and asserted by non-inventors. Lex Machina's patent page also contains links to Google Patents and the PTO's online patent databases; student researchers use these links to gather further information about patent assignments and the geographic location of the inventors and assignees.

Third, student researchers conduct a comprehensive web search of each unknown patentasserter. These searches usually reveal useful sources like an entity's website, which may offer background or applicable product information. Additionally, web searches may reveal media and industry news articles discussing the entity or its litigation, public corporate information such as SEC filings identifying the entity owners and sources of income, or other legal websites (like RPX) that identify certain cases as those involving producers or non-practicing entities. Table 2 summarizes examples of the evidence used to identify each entity category.

TABLE 2. STANFORD NPE LITIGATION DATASET PATENT ASSERTER TAXONOMY

Category	Evidence
1. Acquired patents	<ul style="list-style-type: none"> • The entity states in the pleadings or on its website that its purpose is to license patents or generate revenue from licensing patents • There is no evidence in the pleadings or on the entity's website that it made, sold, or offered a product or service • The entity has been identified in news reports or legal websites as an NPE

	<ul style="list-style-type: none"> • The entity has never been sued for patent infringement • The entity has filed many lawsuits and/or has sued many different alleged infringers • The entity has filed suit in the Eastern District of Texas or the District of Delaware • The entity's address was within the Eastern District of Texas • The entity has a small number of employees • None of the entity's employees or founders match the names of the inventors on its asserted patents • A PTO patent assignment search shows that the original patent owner (e.g., individual or producing company) transferred the patent to the entity that appears unrelated • The patent had been assigned to the entity shortly before a large increase in litigation • The entity generates largely variant amounts of income
<p>2. University Heritage or Tie</p>	<ul style="list-style-type: none"> • The pleadings, company website, news reports, or other filings indicate that the company has ties to a university and/or exists to license university intellectual property • There is no evidence suggesting the entity is actually the university itself or an alias for the university
<p>3. Failed startup</p>	<ul style="list-style-type: none"> • The pleadings, company website, news reports or other filings indicate the entity's former intention to make products or offer services, such as the creation of prototypes or attempts at marketing

	<ul style="list-style-type: none"> • The complaint, company website, news reports, or other filings indicate that the entity had received venture funding • The complaint, company website, news reports, or other filings indicate that the entity is a “startup” • There is evidence from the complaint, company website, news report or other sources that the entity failed or was failing before litigation
4. Corporate Heritage	<ul style="list-style-type: none"> • Websites, news, or other reports indicate that the entity formerly made products or offered services • Websites, news, or other reports indicate that the entity was commercially viable for more than a <i>de minimis</i> period of time • Websites, news, or other reports indicate that the entity shifted its business to IP monetization
5. Individual-Inventor-Started Company	<ul style="list-style-type: none"> • The pleadings, company website, news reports, or other filings indicate that the entity is owned, operated, or founded by the inventor of the held patent(s) • There is no evidence available to suggest that the entity itself offers products or services for sale • A PTO PAIR search shows the inventor(s) resided in the same geographic location as the entity • Public company information indicates that the inventor(s) listed on the asserted patent owns, operates, founded or is closely related to the entity (e.g., familial relative to the company owner)
6. University/Government/NGO	<ul style="list-style-type: none"> • The pleadings or website of the entity states that it is an institution for higher education, a government body, or a non-profit organization

7. Startup, pre-product	<ul style="list-style-type: none"> • The pleadings, company website, news reports, or other filings indicate that the entity is a “startup” but has not yet developed a product or service • The complaint, company website, news reports, or other filings indicate that the entity has received venture funding • There is no evidence available to suggest that the entity actually offered products or services at the time of litigation
8. Product Company	<ul style="list-style-type: none"> • The pleadings, company website, court filings, or SEC filings indicate that the plaintiff made a product, sold a product, or offered a service at the time the lawsuit was filed • The entity has many employees • The entity operates brick and mortar retail establishments or a functioning website offering actual products or services for sale • News or other reports describe the entities products or services
9. Individual	<ul style="list-style-type: none"> • The entity identifies itself in the complaint as an individual or trust • The entity has the name of a natural person • The entity is a family trust (typically for the inventor(s)’ family) • Usually, the name of the entity matches the name of the inventor on the patent
10. Undetermined	<ul style="list-style-type: none"> • There is no evidence available to conclusively classify the entity asserting the patent
11. Industry Consortium	<ul style="list-style-type: none"> • The pleadings, entity website, news reports, or other filings indicate the entity is a standards setting or other type of industry

	<p>organization operating on behalf of independent companies</p> <ul style="list-style-type: none"> • The evidence suggests that the entity does not itself produce products or services beyond IP holding or assertion on behalf of its members
12. IP Subsidiary of Product Company	<ul style="list-style-type: none"> • The pleadings, company website, news reports, or other filings indicate that the entity is a holding company for another company's intellectual property • There is no evidence available to suggest that the entity itself offers products or services
13. Corporate-Inventor-Started Company	<ul style="list-style-type: none"> • The pleadings, company website, news reports, or other filings indicate that the entity performs in-house research and development • The pleadings, company website, news reports, or other filings indicate that the entity has made significant expenditures on research and development • The pleadings, company website, news reports, or other filings indicate that the entity exclusively licenses its technology to manufacturers for commercialization • The entity has a larger number of employees than needed merely to enforce IP and these employees include scientists and engineers

After determining the proper categories for the patent asserters in a lawsuit, student researchers assign the numbers for those categories to the patent asserters in our in-house database user interface. Before moving to the next unknown entity, the coders review every other lawsuit involving the entities and assigns them to their appropriate category in every lawsuit where they appear. Occasionally, a student finds that another coder has previously categorized an entity differently in other lawsuits involving the same patentasserter. When this occurs, the students resolve the inconsistency by reviewing the evidence and reach the final categorization as a group.

When a student is unable to categorize an entity after reviewing all the sources of evidence, the student notes the entity and enlists the aid of the group through quality control measures described below.

Quality Control

Our project employs quality control (“QC”) procedures with several levels of review to ensure reliable coding. First, researchers are instructed and encouraged to bring ambiguous or challenging classifications to the entire research group for discussion at weekly meetings. For example, if a researcher suspects—but is not certain—that an entity is a licensing firm that acquires patents (Category 1) as opposed to an individual-inventor-started licensing firm (Category 5), that researcher emails the entire research group the evidence he/she collected and asks for their thoughts and suggestions. Only after the group achieves consensus, with input from project lead Shawn Miller, does the researcher code the entity in the Dataset.

Additionally, the project lead regularly reviews random samples of each researcher’s work product and, if necessary, directs that student researcher to correct errors and revisit any categories that were consistently miscoded. Since the second year of the research project, more experienced law students have also performed QC for those newer to the project. The senior member reviews roughly twenty lawsuits out of an assigned batch of 200 (10% random sample). Inter-coder reliability between researchers almost always ranges from 85 to 100% (i.e., researchers disagree on zero to three of the twenty cases they reviewed). When agreement is less than 90%, or it appears that mistakes are due to a misunderstanding of our categorization taxonomy, the project lead and/or senior team members instruct the researcher to review their batch again and correct mistakes. Further, if the reviewing senior members are unsure of an entity coded by a junior member, the members discuss the issue with the entire research team. Importantly, researchers are not rewarded for the number of cases they ultimately code. Instead, our project emphasizes the primacy of accuracy over speed.

We believe that the methodology and quality control measures we employ have resulted in a reliable dataset. However, the best mechanism for ensuring reliability is public accessibility and review. We encourage all who use the Dataset to contact us with any categorization disagreements. As policy, we always review the evidence in such flagged cases and change entity categorization if the review supports such a change.

III. DESCRIPTION OF THE DATA: DESCRIPTIVE STATISTICS, TIME TRENDS, AND PRELIMINARY FINDINGS

Now that we have described the Stanford NPE Litigation Dataset and our methodology for creating it, we report in Part III.A descriptive

statistics and time trends revealed solely by information in the Dataset. To demonstrate the power of the Dataset when combined with other data sources, in Part III.B we report several findings related to differences in technology asserted, lawsuit duration, and lawsuit outcome across cases involving different types of patent asserters. Our description of the Dataset and findings utilize our complete 20% random sample of lawsuits filed from 2000 through 2015—10,812 total lawsuits. This sample is large enough that our reported statistics and differences across groups of patent asserters reflect the population of all lawsuits during this sixteen-year period within a 3% margin of error.

A. *Stanford NPE Litigation Dataset Descriptive Statistics and Time Trends*

We begin this section reporting the number of lawsuits that include each of our thirteen categories of patent asserters. We then report the total number of distinct defendant-lawsuit pairs attributable to each asserter category in order to account for the America Invents Act's change in joinder rule. In the second half of this section, we reveal changes in the share of lawsuits and defendant-lawsuit pairs attributable to the most common asserter types over time.

1. *Counts of Lawsuits by Patent Asserter Type*

Table 3 reports the number of lawsuits and percent of all lawsuits in our 20% random sample that include a patent asserter in each of our thirteen categories. The first thirteen rows of Table 3 add up to more than the total—10,812 cases—because there are 750 lawsuits (6.9%) where patents were asserted by more than one of the thirteen types of patent asserter. There are dozens of combinations of entity types in these multiple category cases. However, about six combinations account for the vast majority (84%). These include:

- A Category 8 entity and a Category 9 entity (297 lawsuits);
- A Category 8 entity and a Category 12 entity (134 lawsuits);
- A Category 5 entity and a Category 9 entity (73 lawsuits);
- A Category 8 entity and a Category 2 or 6 entity (70 lawsuits); and
- A Category 8 entity and a Category 5 entity (56 lawsuits).

Four out of five of these most common multiple-category disputes are a combination of an NPE and a Category 8 product company, and in almost all of these cases the NPE is the owner of the asserted patent while the product company is the exclusive licensee. Over one third of

these multiple-category disputes include a Category 8 product company and a Category 9 individual who is almost always the inventor-owner of the asserted patents.

Turning again to Table 3, we see that a product company was the sole patentasserter—or multiple product companies were the sole asserters—in 58% of lawsuits filed over this sixteen-year period. This might suggest that talk of the “rise of the patent troll” is overblown, but the time trends revealed by the Dataset will reveal a different story. Seven of the categories appear as the soleasserter type in less than 1% of lawsuits. These include our Category 10 “undetermined” category as well as Category 2 university heritage or tie, Category 3 failed startup, Category 4 corporate heritage, Category 7 startup pre-product, Category 11 industry consortium and Category 13 corporate inventor-started company. Category 6 universities, governments or non-profit entities appear in 1.4% of lawsuits and Category 12 IP subsidiaries of product companies in just under 2% of cases. With the last category, however, we note from our discussion of multiple-category cases above that 134 of the 205 cases (65%) involving a Category 12 entity also involved a product company and overwhelmingly the Category 8 product company owns the IP subsidiary party. Thus, most litigation involving Category 12 entities can be thought of as essentially product-company litigation.

TABLE 3. NUMBER OF LAWSUITS BY CATEGORY (2000-2015)

Entity Category	# of Lawsuits	% of Lawsuits
1. Acquired patents	2,327	21.5
2. University heritage or tie	22	0.2
3. Failed startup	63	0.6
4. Corporate heritage	62	0.6
5. Individual-inventor-started company	1,376	12.7
6. University/Government/Non-profit	153	1.4
7. Startup, pre-product	19	0.2
8. Product company	6,295	58.2
9. Individual	1,014	9.4
10. Undetermined	4	0.0
11. Industry consortium	6	0.0
12. IP subsidiary of product company	205	1.9
13. Corporate-inventor-started company	38	0.4
Total	10,812	100%

NOTE: 10,812 total lawsuits with 750 lawsuits possessing more than one category of patentasserter. 3 percent margin of error in percentages as compared with population of lawsuits filed 2000-2015.

This leaves us with three types of NPEs that frequently assert patents in litigation—Category 1 acquired patents (21.5%), Category 5 inventor-started (12.7%) and Category 9 individual (9.4%). In most of the remainder of our analysis we focus on these three types of NPEs and Category 8 product companies because at least one of these four categories is present in all but 278 lawsuits in our random sample (i.e., present in 97.4% of lawsuits). Further, as Category 1, 4 and 5 entities track closest to our definition of PAEs, we combine lawsuits involving only these three categories in reporting statistics on PAE litigation.

Because of important changes in the joinder rules effective after enactment of the America Invents Act (“AIA”),³³ it is necessary to account for the AIA in time trend analysis. We do so by obtaining, from Lex Machina, the number of alleged infringers in each of the lawsuits in our random sample and calculating the number of defendant-lawsuit pairs involving each of the thirteen types of patentasserter in the Dataset.³⁴ We calculate a total of 25,507 defendant-lawsuit pairs in the suits in our random sample.

In addition to controlling for changes in the joinder rule, counting defendant-lawsuit pairs, rather than lawsuits, provides a better tally of distinct patent disputes in some situations—including pre-AIA U.S. patent litigation where plaintiffs were permitted to join defendants with little or nothing in common other than allegedly infringing the same patent. Thus, most, but not all, pre-AIA lawsuits against multiple unaffiliated defendants are better considered as comprised of distinct patent disputes.

In Table 4, we report the number (and percentage) of the 25,507 defendant-lawsuit pairs in our random sample that included each of our thirteen patentasserter types. We also include the average number of defendants per lawsuit in the last column and note that in our 10,812 lawsuits there are, on average, 2.4 defendants per lawsuit over the sixteen years—a time period that straddles the enactment of the AIA. Further, in comparison with Table 3, we note that both of our main PAE

33. The Leahy-Smith America Invents Act provides that joinder of accused infringers is permitted only where the claims against the defendants arise out of “the same transaction, occurrence, or series of transactions, or occurrences relating to the making, using, importing into the United States, offering for sale, or selling of the *same accused product or process*” and “questions of fact common to all defendants or counterclaim defendants will arise in the action.” 35 U.S.C. § 299(a)(1)-(2) (emphasis added). Accused infringers may *not* be joined “based solely on allegations that they each have infringed the patent or patents in suit.” 35 U.S.C. § 299(a)(2).

34. Determining total defendant-lawsuit pairs for each category of cases involves only simple math: First, identify all lawsuits of a given type; second, weigh each suit by the number of alleged infringers; and third, sum the weighted totals for each cases. For example, if we have ten lawsuits involving product companies with five involving only one defendant and five involving three defendants we would have a total of twenty defendant-lawsuit pairs for those ten cases.

categories, Category 1 and Category 5, make up a larger fraction of pairs than lawsuits (23.1% versus 21.5%, and 13.9% versus 12.7%, respectively). This shows that pre-AIA, PAEs tended to join more defendants in a single lawsuit than the average patentasserter. The same is true of our third most common NPE type—Category 9 individuals (11.3% versus 9.4%). The reverse is true of product companies who are the soleasserter type in 53% of pairs versus 58% of lawsuits.

The average defendants per lawsuit for some of our less common categories are even more extreme. Category 3 failed startups sued about one more defendant per suit than average, perhaps because they assert while winding down their business and pre-AIA sued all possible infringers in a few cases in this relatively short time frame. Perhaps surprisingly, Category 6 universities, government and non-profits include more than two additional defendants per suit than average. Category 13 and Category 11 entities also have higher than average defendants per suit. We suspect these types of entities are working in tech spaces that involve collaboration between multiple entities to commercialize and that these patent asserters are tending to sue all the entities working together at once. Deeper analysis of what drives certain categories of patent asserters to select a certain number of accused infringers to sue in a single action is beyond the scope of this paper, and we further note that outliers may explain some of these high averages.³⁵

35. E.g., for Category 6 universities, government, and non-profits, three of 153 lawsuits account for half of the total defendant-lawsuit pairs: 1) *Mass. Inst. of Tech. v. Abacus Software*, No. 5:01-CV-00344, 2004 WL 5268128 (E.D. Tex. Aug. 24, 2004) (221 defendants); 2) *Thermolife Int'l v. Myogenix Corp.*, No. 3:13-CV-00651 JLS (MDD), 2014 WL 12160740 (S.D. Cal. Dec. 2, 2014) (83 defendants); and 3) *Bd. of Regents of Univ. of Tex. Sys. v. Benq America Corp.*, 1:05-CV-00181, 2006 WL 6112210 (W.D. Tex. July 14, 2006) (33 defendants).

TABLE 4. NUMBER OF DEFENDANT-LAWSUIT PAIRS BY CATEGORY (2000-2015)

Entity Category	# of Pairs	% of Pairs	Defendants / Lawsuit
1. Acquired patents	5,881	23.1	2.5
2. University heritage or tie	53	0.2	2.4
3. Failed startup	204	0.8	3.2
4. Corporate heritage	149	0.6	2.4
5. Individual-inventor-started company	3,536	13.9	2.6
6. University/Government/Non-profit	717	2.8	4.7
7. Startup, pre-product	48	0.2	2.5
8. Product company	13,576	53.2	2.2
9. Individual	2,882	11.3	2.8
10. Undetermined	12	0.0	3.0
11. Industry consortium	33	0.1	5.5
12. IP subsidiary of product company	562	2.2	2.7
13. Corporate-inventor-started company	147	0.6	3.9
Total	25,507	100%	2.4

NOTE: 2,248 defendant-lawsuit pairs possess more than one category of patent assertor with those multiple-asserter type lawsuits against 3.0 defendants per lawsuit on average. 3% margin of error in percentages as compared with population of lawsuits filed 2000-2015.

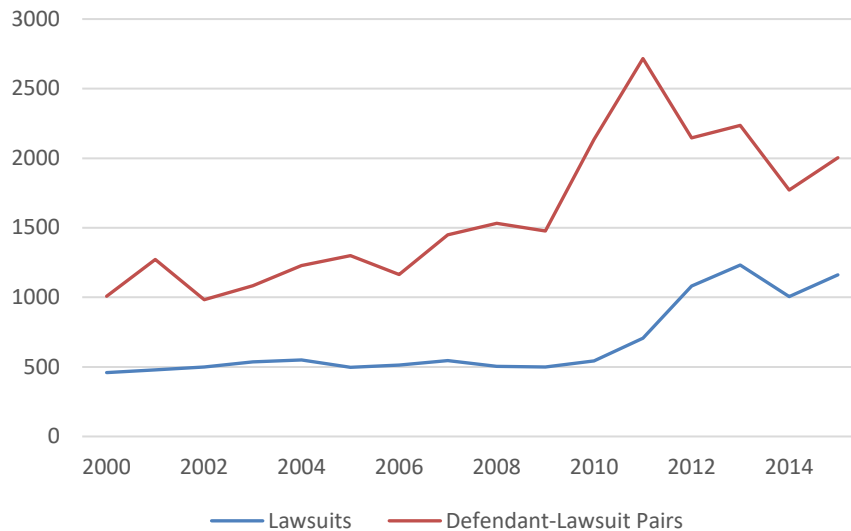
2. Trends in Lawsuits by Patent Asserter Type

How has litigation distribution by patent asserter type changed over time? Armed with both counts of lawsuits and defendant-lawsuit pairs we now answer that question by reporting time trends. First, in Figure 1 we chart the change in total lawsuits and defendant-lawsuit pairs over sixteen years, 2000 through 2015. Concerning lawsuits, we see that the number of lawsuits per year remained relatively constant until the AIA was passed in 2011. After an initial fluctuation post-AIA, the number of lawsuits has somewhat stabilized at a higher level of total lawsuits. Given that we are analyzing a 20% sample of all lawsuits, scaling up the totals in Figure 1 by five to 100% of the population there were about 2,500 patent lawsuits filed each year pre-AIA and over 5,000 per year post-AIA.

The trend in total annual defendant-lawsuit pairs revealed in Figure 1 is similar to that of lawsuits but not the same. There is in fact a mild upward trend in the number of pairs even before passage of the AIA. We believe growth in patent disputes as measured by pairs actually began to take off between 2006 and 2007 or perhaps even as early as 2002. Thereafter, we see a large spike in pairs around passage of the AIA and

as others have reported, there was a rush to court before passage, particularly by NPEs.³⁶ After 2011 the total number of pairs stabilized and appears to trend with the number of lawsuits.

FIGURE 1. NUMBER OF LAWSUITS VERSUS DEFENDANT-LAWSUIT PAIRS OVER TIME



NOTE: 20% random sample of 10,812 lawsuits filed from 2000 through 2015.

Given our discussion of the spike in pairs and increase in the level of lawsuits due to the AIA joinder rule, we next investigate whether the joinder rule had disparate effect on different types of patent plaintiffs. In Table 5 we report the number of pairs, lawsuits, and average defendants per lawsuit for key categories of patent plaintiffs in the two-year period before and the two-year period after passage of the AIA on September 16, 2011. The results are dramatic.

In Table 5, we see that across all lawsuits, the average number of defendants per suit was cut in half—from 4.1 to 2.0. The most dramatic decrease in the number of defendants per suit is for our Category 1 firms, which are classic patent licensing firms that acquire their patents

36. See Tracie L. Bryant, *The America Invents Act: Slaying Trolls, Limiting Joinder*, 25 HARV. J.L. & TECH. 673, 674 (2012) (citing Dennis Crouch & Jason Rantanen, *Rush to Judgment: New Dis-Joinder Rules and Non-Practicing Entities*, PATENTLY-O (Sept. 20, 2011), <https://patentlyo.com/patent/2011/09/rush-to-judgment-new-dis-joinder-rules-and-non-practicing-entities.html> [https://perma.cc/WG8W-9VXC]) (describing how the day before the AIA went into effect, on September 15, 2011, over fifty patent litigation suits accusing an average of sixteen entities were filed, creating the all-time high in the number of entities sued for patent infringement in a single day at over 800 entities).

from third parties. Pre-AIA these firms joined an average of 9.9 different defendants per suit and in the two years after the AIA that average dropped to 1.8. There is a similar but milder decline with Category 5 licensing firms started or owned by inventors (dropping from 4.6 to 1.6 defendants per suit). Because Category 1 and Category 5 firms comprise the bulk of our PAE lawsuits, the change for PAEs as a whole is intermediate to those two numbers—dropping from 7.6 to 1.7 defendants per suit.

For universities, the average number of defendants per suit did decline by one (from 4.7 to 3.8), but the fact that the average is twice that of all lawsuits post-AIA suggests our intuition above was correct—university lawsuits tend to be against multiple defendants allegedly infringing in concert. The average for product companies declined from 2.7 pre-AIA to 2.1 post-AIA, suggesting that they too joined independent infringers pre-AIA. Finally, and perhaps surprisingly, the AIA had no effect on the average number of defendants per suit for Category 9 individual. The differential impact of the AIA should be kept in mind as we now turn to time trends for different types of patent asserters.

TABLE 5. DEFENDANT-LAWSUIT PAIRS, LAWSUITS, AND AVERAGE DEFENDANTS PER SUIT TWO YEARS BEFORE AND AFTER PASSAGE OF AIA, BY SELECT PATENT ASSERTER CATEGORY

Patent Asserter Type	Before AIA			After AIA		
	Pairs	Suits	De- fend ants / Suit	Pairs	Suits	De- fend ants / Suit
Cat. 1: Acquired	1606	163	9.9	1360	765	1.8
Cat. 5: Inventor- started	639	139	4.6	585	367	1.6
Cat. 6: Univ./Gov't/Non- profit	90	19	4.7	158	42	3.8
Cat. 8: Product com- pany	2024	751	2.7	1966	920	2.1
Cat. 9: Individual	303	109	2.8	326	115	2.8
PAE	2189	289	7.6	1891	1100	1.7
All lawsuits	4676	1148	4.1	4249	2175	2.0

NOTE: 3% margin of error in percentages as compared with population of lawsuits filed 2000-2015.

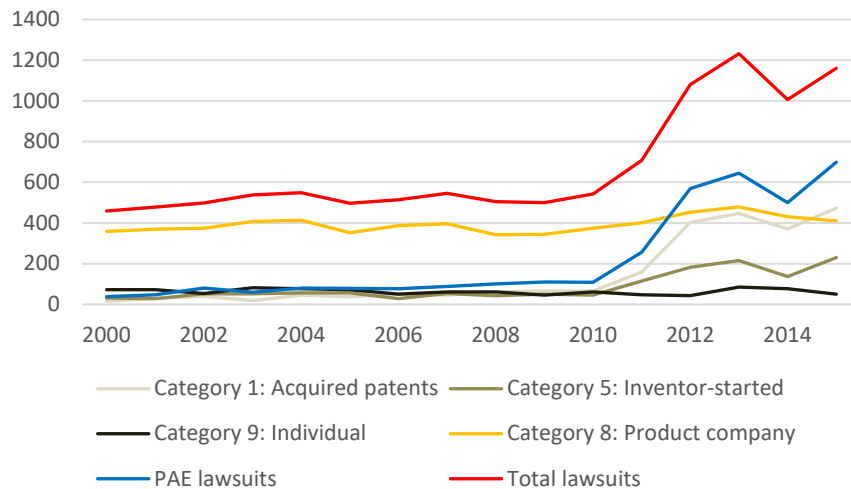
In Figures 2 and 3, we plot the annual number of lawsuits and defendant-lawsuit pairs, respectively, for the most common categories of patent owners. In Figure 2 we see the sharp effect of the AIA on Category 1 and 5 and PAE lawsuits and, in turn, their effect on total lawsuits.

Interestingly, the annual number of lawsuits involving Category 8 product companies remained flat—or at most has seen a very small upward trend—during the entire sixteen-year period. Finally, the number of lawsuits brought by individuals has also remained quite constant.

In Figure 3, we see a clear spike in total, PAE, Category 1, and to a lesser extent Category 5 and Category 8 defendant-lawsuit pairs around passage of the AIA. Again, we believe counting by defendant-lawsuit pairs better captures the total number of distinct litigated patent disputes. For each of these categories, the spike is brief with the annual number of pairs trending back down.

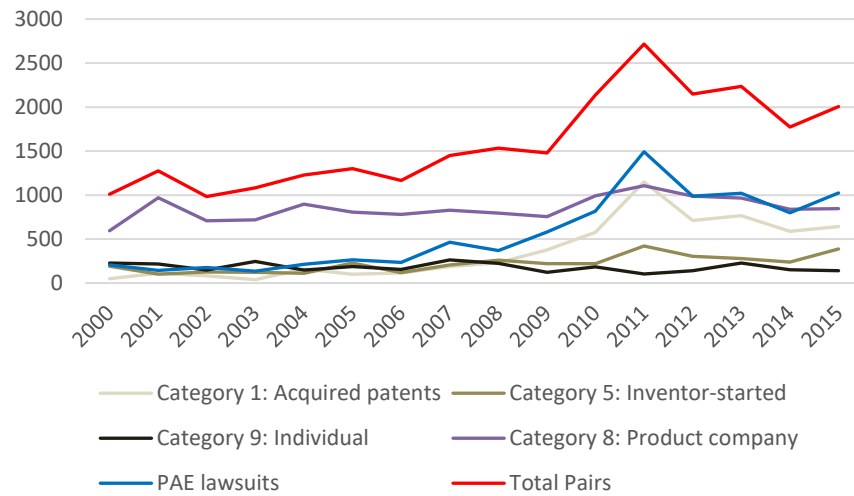
The trend for annual product company pairs appears constant such that the entire increase in pairs that we believe began around 2006 or 2007 is attributable to PAEs. In other words, there was a “rise of the patent troll” as a prevalent litigant type that began around the mid-2000s and that may or not be continuing to the present. Interestingly, this increase in PAE litigation is mostly explained by our Category 1 licensing firms who acquire the patents they assert from third parties. From the pair totals in Figure 3, we know that in 2000 20% of pairs were attributable to PAEs only, 5% of pairs involved a Category 1 firm, and 19% involved a Category 5 PAE. From 2012 through 2015, these percentages were 47% of PAE only pairs, 33% of Category 1 pairs and 15% of Category 5 pairs. Thus, while the percentage of distinct disputes attributable to Category 5 entities remained relatively constant, that of Category 1 patent licensing firms skyrocketed.

FIGURE 2. NUMBER OF LAWSUITS OVER TIME BY KEY CATEGORY



NOTE: 20% random sample of 10,812 lawsuits filed from 2000 through 2015. The “PAE lawsuits” trend comprises cases including only Category 1, 4 and 5 patent asserters. Category 1, 5, 9 and 8 trends include all cases with at least one patent asserter of this type.

FIGURE 3. NUMBER OF DEFENDANT-LAWSUIT PAIRS OVER TIME BY KEY CATEGORY



NOTE: 20% random sample of 25,507 distinct defendant-lawsuit pairs in 10,812 lawsuits filed from 2000 through 2015. The “PAE lawsuits” trend comprises cases including only Category 1, 4 and 5 patent asserters. Category 1, 5, 9 and 8 trends include all cases with at least one patentasserter of this type.

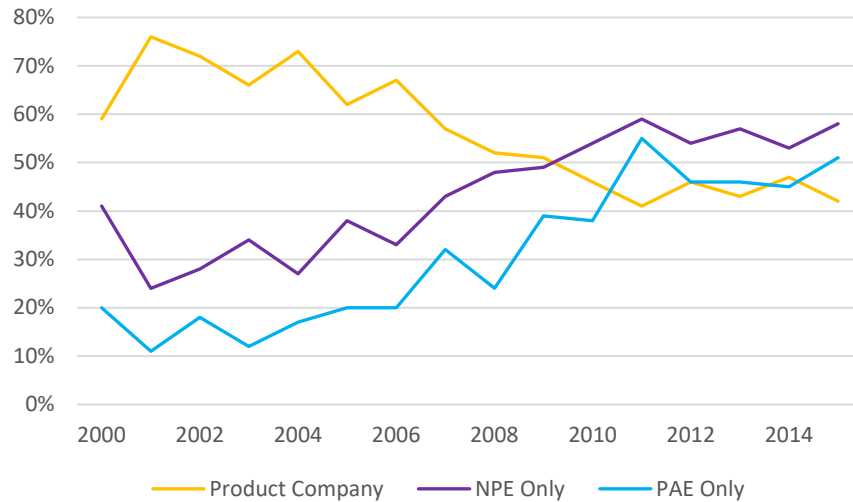
What has been the effect of this increase in PAE (and by extension NPE) litigation on the ecology of U.S. patent litigation since 2000? In Figure 4, we see that the share of distinct litigated patent disputes—as measured by pairs—attributable to product companies and PAEs has each been about 45% since 2010 or 2011. In other words, simply measured by the frequency of suit, over the last six years alleged infringers have been sued by PAEs as frequently as they have been sued by product companies.

While the relative share between PAEs and product companies appears flat since 2010 or 2011, by comparison, the relative share of these two groups was very different at the beginning of the 2000s when the share of disputes attributable to product companies averaged around 70% and that attributable to PAEs averaged around 15%. Something or some things clearly happened during the mid-2000s to dramatically change the ecology of U.S. patent litigation. One side effect is that PAEs are likely consuming a larger percentage of patent-related judicial resources than ever before, at least at the district court level.

Finally, we include a trend line in Figure 4 for NPE-only lawsuits. Recall that each of our categories of asserters except for Category 8 are NPEs. Thus, our NPE trend line includes all cases without a Category 8 asserter. Since PAE litigation is a subset of NPE litigation, the “PAE-Only” trend line is always below the “NPE-Only” trend line. We further note that the share of all litigation attributable to NPEs only overtook product company litigation after 2009 and has hovered between 55 and

60 percent of all disputes since. Before 2009 most patent disputes were product company disputes and since most have been NPE disputes.

FIGURE 4. PERCENT OF DEFENDANT-LAWSUIT PAIRS OVER TIME



NOTE: Percent shares derived from 20% random sample of 25,507 distinct defendant-lawsuit pairs in 10,812 lawsuits filed from 2000 through 2015. PAE Only includes cases including only Category 1, 4 and 5 patent asserters. Product Company includes all cases with at least one Category 8 patentasserter and NPE Only includes all cases without a Category 8 patentasserter. 3% margin of error in percentages as compared with population of lawsuits filed 2000-2015.

We believe the descriptive statistics and first-ever time trends by assenter type in this section provide important evidence on the functioning and ecology of the U.S. patent system. However, the true power of the Dataset in investigating our patent system will be revealed by future researchers as they merge our patent assenter categorization of lawsuits with other datasets and resources. We provide a glimpse of this potential in the next section.

B. Findings Combining Stanford NPE Litigation Dataset with other Data Sources

In addition to the number of defendants in each of the lawsuits in our 20% random sample, Lex Machina also provided us with the identifying numbers of the patents asserted in each lawsuit as well as each case's date of termination and outcome. From this information we can provide a set of pertinent comparative statistics, including technology and case outcome differences across patent assenter type. We first investigate technology differences.

1. Technology

With the patent numbers provided by Lex Machina, we utilize the PTO's online patent number search database³⁷ to determine the primary 3-digit technology class for each patent asserted in each case in our 20% random sample that was filed in 2014. With this technology class information, we categorize each of our lawsuits as asserting a patent or patents in a variety of technology areas.

First, following Hall et al.³⁸, we categorize each lawsuit as fitting into one or more of the following six categories: 1) Chemical; 2) Computers & Communications; 3) Drugs & Medical; 4) Electrical & Electronic; 5) Mechanical; and 6) Others. The number of lawsuits with at least one patent fitting into each of these categories is included in Table 6.

Because of their particular relevance in policy debate, we further identify lawsuits containing at least one drug patent or at least one software patent. We define pharmaceutical patents as any patent assigned either U.S. technology class 514 or 424.³⁹ For software, we use Bessen's definition of a software patent as one assigned a class that either includes data processing in the class title or is "reliant on software and in which software companies obtain patents."⁴⁰ Finally, we determine whether or not a design patent was asserted in each of the lawsuits.

In Table 6 we see that there is significant technological variation across lawsuits involving different types of patent asserters. Across the six NBER categories, we see that Category 8 product company litigation includes the most technological variety with percentages ranging from 8.7% of electrical patents to 31.9% of drugs and medical patents. Only 18.5% of product company lawsuits assert a computer and communications patent. In contrast, about three quarters of PAE (and separately Category 1 and Category 5 patent asserter) lawsuits assert a computer and communications patent, and these categories assert far fewer chemical, drugs and medical and mechanical patents than product companies. One interesting caveat is that Category 5 inventor-started patent licensing firms have the second highest rate of asserting drugs and medical patents. We believe many of these cases are brought by medical device

37. This database is publicly available. *USPTO Patent Full-Text and Image Database*, USPTO, <http://patft.uspto.gov/netahtml/PTO/srchnum.htm>.

38. Bronwyn H. Hall, Adam B. Jaffe & Manuel Trajtenberg, *The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools* 41 (Ctr. for Econ. Policy Research, Discussion Paper No. 3094, 2001).

39. Both PTO classes 514 and 424 are named "Drugs, Bio-Affecting and Body Treating Compositions". See *Select US Classes by Number with Title*, USPTO, <https://www.uspto.gov/web/patents/classification/selectnumwithtitle.htm>.

40. James Bessen, *A Generation of Software Patents* 14 (Boston Univ. Sch. of Law, Law and Economics Research Paper No. 11-31, 2011) (listing PTO classes 341, 345, 370, 375, 380, 381, 382, 700-707, 715-717, 726, and 902).

patent owning inventors sophisticated enough to create an LLC to hold their IP rights.

Lawsuits with Category 9 individual asserters are interestingly closer in patent technology distribution to product companies than the other NPE categories. However, a majority of individual lawsuits include a computer and communications patent—just like the other NPE categories. Further, of all the categories, individuals least frequently assert drugs and medical patents (just 1.4% of 2014 Category 9 lawsuits).

From the NBER rates in Table 6, it appears that litigation of drugs and medical patents, and separately chemical patents, is the domain of product companies while computer and communications patent litigation is dominated by NPEs and PAEs. Undoubtedly, much of this is due to the relatively higher costs of innovation in the technology areas where product companies prevalently assert.⁴¹ Further, these most disparate technology categories (across assenter type) track with the two technologies of patents that receive the most policy attention—pharmaceuticals and software.⁴² The first is considered the technology for which patents “work” the best,⁴³ while the second is the area that has received the most criticism.⁴⁴

Starting with pharmaceuticals, 116 of 1006 cases filed in 2014 in our random sample asserted at least one drug patent (11.5%). Referencing Table 6 and Figure 5, we see that 111—a whopping 95.7% of cases with a pharmaceutical patent—were product company lawsuits. Less than one percent of pharmaceutical lawsuits involved a Category 1 or Category 9 patent assenter. Further, only 2.6% of drug cases involved a Category 5 assenter, and only 1.7% of drug lawsuits did not include a Category 1, 5, 8, or 9 assenter but another type of NPE. Thus, pharmaceutical patent litigation is almost exclusively the domain of product companies.

In contrast, there is much more diversity in the type of patent assenter in software patent litigation. 296 of 1006 cases filed in 2014 involved software (29%). Nearly 58% of software suits included a Category 1 entity and nearly 20% included a Category 5 entity. Thus, about 80% of software patent litigation involves PAEs. Product companies are involved in over 15% of software suits—an economically significant cut but a relatively small share compared to NPEs.

41. Dan L. Burk & Mark A. Lemley, *Is Patent Law Technology-Specific*, 17 BERKELEY TECH. L.J. 1155, 1156 (2002); Dan L. Burk & Mark A. Lemley, *Policy Levers in Patent Law*, 89 VA. L. REV. 1575, 1581-83 (2003).

42. *Policy Levers in Patent Law*, *supra* note 41; see also Brian Kahin, *Patents and Diversity in Innovation*, 13 MICH. TELECOMM. TECH. L. REV. 389 (2007).

43. Kahin, *supra* note 42, at 389-90 (describing how the pharmaceutical industry fosters a “discrete-product environment” with a “close relationship between a high-value product and the protection of a high-value patent”).

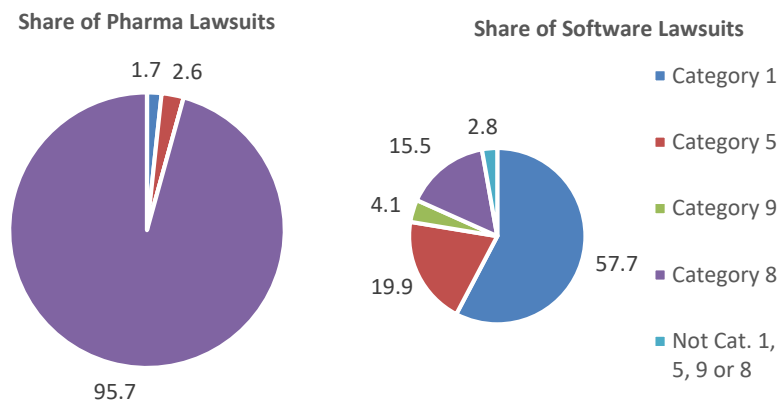
44. *Id.* (describing the environment of the information technology industry as a “extreme complex-product environment,” and how the information technology industry strives for “freedom of action,” but is prone to inadvertent infringement).

TABLE 6. LITIGATED PATENT TECHNOLOGY BY TYPE (2014)

Technology Type	N	Cat. 1	Cat. 5	Cat. 9	PAE only	Cat. 8
Chemical	49	0.8% (3)	0.0% (0)	1.4% (1)	0.6% (3)	10.6% (45)
Computer & Comm.	518	78.1% (289)	74.5% (102)	56.8% (42)	78.0% (390)	18.5% (79)
Drugs & Medical	154	2.2% (8)	6.6% (9)	1.4% (1)	3.2% (16)	31.9% (136)
Electrical	87	5.9% (22)	9.5% (13)	20.3% (15)	7.0% (35)	8.7% (37)
Mechanical	63	1.9% (7)	5.1% (7)	12.2% (9)	2.2% (11)	11.0% (47)
Other	184	17.8% (66)	5.8% (8)	24.3% (18)	14.2% (71)	23.7% (101)
Software	296	46.2% (171)	43.1% (59)	16.2% (12)	46.4% (232)	10.8% (46)
Pharmaceuticals	116	0.5% (2)	2.2% (3)	0.0% (0)	1.0% (5)	26.1% (111)
Design	48	0.0% (0)	0.7% (1)	6.8% (5)	0.2% (1)	10.1% (43)

NOTE: 20% random sample of 1006 lawsuits filed in 2014. 370 lawsuits included at least 1 Category 1 patent assertor and 636 did not. 137 lawsuits included at least 1 Category 5 assertor and 869 did not. 77 lawsuits included at least 1 Category 9 patent assertor and 929 did not. 497 lawsuits included only PAE assertors (Categories 1, 4 and 5) and 509 did not. 426 lawsuits included at least 1 Category 8 assertor and 580 did not. A lawsuit fits into a technology category if at least one of the patents asserted in the lawsuit possesses one of the primary 3-digit PTO technology classes that make up that NBER category of patent. Thus, many lawsuits fit into multiple technology categories. All reported statistics using our 20% random sample possess a 3% margin of error in comparison with the unreported population.

FIGURE 5. PHARMACEUTICAL AND SOFTWARE PATENT LITIGATION BY ASSERTER TYPE (2014)



NOTE: 3% margin of error in percentages as compared with population of lawsuits filed 2000-2015.

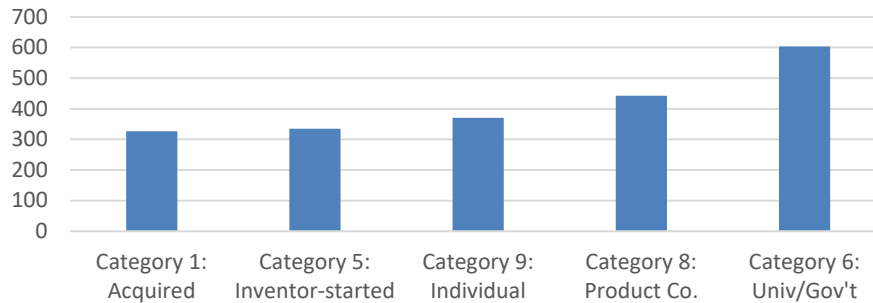
2. Lawsuit Duration and Outcome

Beyond patent numbers, Lex Machina also provided us data for each of the cases in our random sample that enable us to determine differences in lawsuit duration, settlement and outcome. We begin with duration as measured by time to termination from filing date.

Time to Termination

Time to termination is an important proxy for lawsuit cost. Via selection effects it may also capture case complexity and private value to the parties.⁴⁵ Using the lawsuit filing and termination dates available from each of Lex Machina's case-specific pages, we calculate the total days to termination for all lawsuits in our random sample that had in fact terminated. In Figure 6, we see that cases with only Category 6 patent asserters (universities/government/NGOs) possess the longest duration (604 days). The category with the next longest time to termination is lawsuits with only product company patent asserters (443 days). Cases with only Category 1 or Category 5 patent asserters have the shortest duration (327 and 335 days respectively) and Category 9 lawsuits brought by individuals possess average durations in between product companies and our two PAE categories (443 days).

FIGURE 6. LAWSUIT TIME TO TERMINATION BY PATENT ASSERTER CATEGORY (2000-2015)



NOTE: Average number of days to termination for five categories of lawsuits using 20% random sample of cases. Data excludes lawsuits with multiple categories of plaintiffs. 3% margin of error in averages as compared with population of lawsuits filed 2000-2015.

45. See also Alan C. Marco, Shawn P. Miller & Ted M. Sichelman, *Do Economic Downturns Dampen Patent Litigation?*, 12 J. EMPIRICAL LEGAL STUD. 481, 485 (2015); Christopher A. Cotropia, Jay P. Kesan, & David L. Schwartz, *Heterogeneity Among Patent Owners in Litigation: An Empirical Analysis of Settlement, Case Progression, and Adjudication* 22 (Hoover Inst. Working Grp. on Intellectual Prop., Innovation & Prosperity, Working Paper No. 16008, 2016) (finding “individual inventors who litigated in their personal capacity survived shorter,” meaning their “cases were resolved faster”).

Settlement

Among lawsuits filed in 2014, those involving PAEs overwhelmingly settled by February 2016 while a smaller majority of non-PAE lawsuits settled by February 2016.⁴⁶ This could be explained by PAEs having a weaker hand, having less interest in “vindicating” their rights, assessing the risk/reward of litigation differently than practicing entities, or other factors. We believe each of these explanations plays a role, but that much of the higher settlement rate is due to the common—but not the only—PAE business strategy of pursuing nuisance value settlements from their targets, for whom settlement is cheaper than litigation through adjudication.⁴⁷

TABLE 7. NUMBER OF SETTLEMENTS AND SETTLEMENT RATE BY TYPE (2014)

Category	Lawsuits settled	Lawsuits not settled	Settlement rate
1: Acquired patents	311	60	84%
5: Individual-inventor-started co.	97	40	71%
8: Product company	272	159	63%
9: Individual	42	35	55%
PAE lawsuits	403	95	81%
Non-PAE lawsuits	317	191	62%
Total	720	286	72%

NOTE: 20% random sample of 1006 lawsuits filed in 2014. Settlement status determined via Lex Machina on February 19, 2016. All reported statistics using our 20% random sample possess a 3% margin of error in comparison with the unreported population.

Beyond PAEs, we see in Table 7 that Category 9 individuals possessed the lowest settlement rate and it is important to remember that the vast majority of these cases are brought by the inventors of the patents asserted. Our result is consistent with Allison et al. (2017).⁴⁸ Ad-

46. On February 19, 2016, we collected from Lex Machina settlement data for all lawsuits in our 20% random that were filed in 2014.

47. Lemley, *supra* note 20, at 2126 (describing one type of business model used by PAEs where the PAE relies on the “high cost of patent litigation . . . to induce the parties they sue to settle for small amounts of money rather than pay millions to their lawyers.” The authors suggest that most PAEs practice this model.)

48. See also John R. Allison, Mark A. Lemley & David L. Schwartz, *How Often*

ditional research is needed to fully understand the differences in behavior and motivation of inventor patent plaintiffs but we believe vindication of property rights and the desire for recognition of personal contribution to innovation play a role.⁴⁹ In contrast with individuals, the remaining categories of patent asserters are firms that we would expect to take a more dispassionate, economically rational approach to the choice of additional litigation or settlement.

Merits Outcomes

Using Lex Machina's case outcome tags, we also determine for our entire 20% random sample how many lawsuits different categories of patent asserters won and lost on the merits. We count a lawsuit as a "win" where the court found at least one of the asserted patents infringed and not invalid, and count a lawsuit as a "loss" where all asserted patents were found not infringed and/or invalid.

Our results are summarized in Table 8, where we see that of the 429 lawsuits with a Category 8 patent asserter that ended in a merits win or loss, the product company won 35% of the time. In contrast, lawsuits with a Category 1 or Category 9 entity only result in a patent asserter win 13% of the time. Category 5 patent asserters have greater success than our other two NPE categories, winning 20% of lawsuits that ended in merit decisions.

Clearly, product firms are much more successful in cases that go to trial or summary judgment than NPEs and PAEs, and this in turn may partially explain why PAEs settle a much higher percentage of the time than practicing entities, as previously discussed. Scholars have theorized that PAEs frequently assert patents that possess arguably broad scope—or cover widely adopted technology.⁵⁰ The downside of asserting such patents is they might also be more likely to be found anticipated

Do Non-Practicing Entities Win Patent Suits?, 32 BERKELEY TECH. L.J. 237; Christopher A. Cotropia, Jay P. Kesan, & David L. Schwartz, *Heterogeneity Among Patent Owners in Litigation: An Empirical Analysis of Settlement, Case Progression, and Adjudication* 22 (Hoover Inst. Working Grp. on Intellectual Prop., Innovation & Prosperity, Working Paper No. 16008, 2016) (finding "individual inventors who litigated in their personal capacity survived shorter," meaning their "cases were resolved faster").

49. Colleen Chien, *Of Trolls, Davids, Goliaths, and Kings: Narratives and Evidence in the Litigation of High-Tech Patents*, 87 N.C. L. REV. 1571, 1587 (2008) ("[S]ome independent inventors are perceived as seeking not only money, the main objective of licensing shops, but also justice or vindication by a court.").

50. See James Bessen, Jennifer Ford & Michael J. Meurer, *The Private and Social Costs of Patent Trolls*, 34 REG., Winter 2011-2012, at 26, 26 ("The critics call NPEs 'patent trolls,' claiming that they buy up vaguely worded patents that can be construed to cover established technologies and use them opportunistically to extract

or obvious.⁵¹ It may be that the small minority of cases where PAEs allow their patents to be litigated through summary judgment or trial tend to possess two characteristics: First, their targets refuse to settle early because they do possess significant validity issues.⁵² Second, the patents arguably cover valuable technology so that the PAE expects a large money judgment in the minority of cases where their patents are found valid.⁵³ Under these conditions, it can be worthwhile to litigate through judgment despite the high chance of losing.

TABLE 8. NUMBER OF MERIT WINS, LOSSES AND WIN % BY PATENT ASSERTER CATEGORY

Category	Wins	Losses	Win Rate
1: Acquired patents	8	54	13%
5: Individual-inventor-started co.	11	43	20%
8: Product company	149	280	35%
9: Individual	6	40	13%
Total	174	417	29%

NOTE: 20% random sample of all lawsuits filed 2000-2015. Excludes lawsuits with multiple categories of patent asserters. 3 percent margin of error in percentages as compared with population of lawsuits filed 2000-2015.

licensing fees from the real innovators. Indeed, there has been a general and dramatic rise in patent litigation that some analysts attribute to rapid growth in the number of patents with unclear or unpredictable boundaries.”); Anthony Biller, *A Cure for the Common Troll*, OPENSOURCE.COM (Feb. 20, 2012), <http://opensource.com/law/12/2/cure-common-troll> [https://perma.cc/B6HR-HNPT] (“Their club of choice is the broad, complex, and vague patent claim.”). And, Acacia directs its subsidiaries to “identify core, patented technologies that have been or are anticipated to be widely adopted by third-parties in connection with the manufacture, sale or use of products and services.” *Acacia Research Corporation, Annual Report Pursuant to Section 13 or 15(d) of the Securities Exchange Act of 1934 for the Fiscal Year Ended December 31, 2015*, 3 (Dec. 21, 2015), <https://www.sec.gov/Archives/edgar/data/934549/000093454916000100/actg2015123110-k.htm> [https://perma.cc/694E-DHUB].

51. See Donald S. Chisum, *Anticipation, Enablement and Obviousness: An Eternal Golden Braid*, 15 AIPLA Q.J. 57, 58 (1987) (demonstrating the interrelated nature of patent scope with the obviousness and enablement requirements).

52. As an example, in 2013, HTC refused to settle with patent-licensing firm Wi-Lan (a “Category 1” entity in our Dataset) and prevailed at trial, invalidating three of Wi-Lan’s patents. See Jeffrey Sistrunk, *HTC, Sony, Others Beat Wi-LAN’s Wireless Patent Claims*, LAW360.COM (July 15, 2013), <https://www.law360.com/articles/457454/htc-sony-others-beat-wi-lan-s-wireless-patent-claims> [https://perma.cc/2QMV-89PU].

53. Lemley, *supra* note 20, at 2126.

IV. IMPACT ON AND SUGGESTIONS FOR FUTURE RESEARCH

We see the immediate value of the Dataset to patent researchers as twofold: First, given the increase in patent litigation—and the change in mix of patent plaintiffs—we reveal occurred over the last two decades sophisticated time series analysis is sorely lacking in the patent literature. Such analysis is impossible without a dataset like ours that is comprehensive and spans many years. Second, while there are many empirical studies utilizing multivariate regression analysis, scholars continue to dispute the role that NPEs and PAEs play in the patent system. We believe some of the dispute will be resolved if prior studies are redone—and future studies are completed—using more than small litigation samples. In this Part, we explain the value of the Dataset in these two types of studies through several examples.

Time Series Analysis

Time series analysis includes empirical techniques that allow researchers to identify changes in trends as well as isolate the causes of those changes. The 2000s and 2010s have been an era of heightened change in the patent policy space with passage of the America Invents Act, key Supreme Court patent cases including *eBay*, *Alice* and *TC Heartland*, and various executive orders focused on increasing transparency and reducing costs in patent suits. In-depth multivariate time series analysis has the ability to measure the impact of such changes over time. Further, where we see clear changes in trend—such as the share of patent litigation attributable to NPEs and PAEs we identify in this paper—time series analysis beyond the scope of our project has the potential to determine *why* the trend changed.

Perhaps the increase in NPE and PAE litigation we identify as accelerating during the mid-2000s was caused by specific legal changes like *Phillips* or *eBay*. Alternatively, perhaps some business change such as the clearing of the market for patents after the dotcom bubble burst led to more of the “right type” of patent being available for use by PAEs during this period. The empirical tools for such investigation of causes and effects exist, and with this Dataset we now have the data to make such studies possible.

The importance of time series analysis is revealed in one recent study that does not use it. In *Unpacking Patent Assertion Entities (PAEs)*,⁵⁴ Christopher Cotropia, Jay Kesan, and David Schwartz show that accounting for the AIA change in joinder rule there was no explosion in NPE litigation between 2010 and 2012 as others had reported.⁵⁵ The researchers hand-coded every patent-owning litigant as an operating

54. Christopher A. Cotropia, Jay P. Kesan, & David L. Schwartz, *Unpacking Patent Assertion Entities (PAEs)*, 99 MINN. L. REV. 649 (2014).

55. *Id.* at 655.

company, university, individual inventor, patent aggregator, technology development company, failed start-up, IP holding subsidiary of an operating company, or patent holding company. Their data includes 2,520 lawsuits from 2010 and 5,185 suits from 2012. They show that if all NPEs are combined into one category, then NPEs sued the majority of accused patent infringers in 2012.⁵⁶ This result is consistent with our findings from our Dataset.

The authors attribute most of the rise in patent litigation to the AIA joinder rules,⁵⁷ and we suspect that their prior intuition regarding the effect of the AIA on lawsuit counts drove their decision to categorize patent owners in 2010 and 2012 only (given how excruciatingly slow it can be to categorize patent plaintiffs). This method gave them a snapshot before and after a legal change, but these snapshots are inadequate to determine either the long-term change in trends caused by the AIA or short-term spikes. For example, we see that the number of distinct disputes, as measured by lawsuit pairs, actually spiked in the year the AIA was enacted.⁵⁸ Cotropia et al.'s data misses this spike entirely because they did not gather data for 2011. Furthermore, they would have no way of knowing, *ex ante*, whether the two snapshots they took would be most relevant to their investigation of a purported NPE litigation explosion. For example, if the AIA had set off an increase in the growth rate of NPE litigation rather than—as appears from our data—a step up in level of NPE lawsuits, 2012 would be too soon to measure increases in NPE litigation in subsequent years. A continuous dataset over time, like ours, removes these limitations.

Other Regression Analysis

Turning from time series analysis, the Dataset provides immediate value to scholars who wish to replicate past empirical studies using the full population of lawsuits over a significant span of years. Scholars who have previously conducted studies on PAE litigation can confirm past research conclusions by utilizing our comprehensive Dataset. However, the data might also reveal that some findings derived from small, non-random samples are not supported by analysis of the full population or are not representative of other subsets. With this possibility, it is especially important to repeat, with more data, past studies that have influenced vital policy decisions. We explain these issues through discussion of three prior studies that we believe are excellent in their design and the questions they ask, but that would be improved with more data.

Our first example shows how the Dataset will allow scholars to study PAEs of varying sizes, rather than just the largest and most easily

56. *Id.* at 655-656.

57. *Id.* at 655.

58. *See supra* Figure 1.

recognizable PAEs. In *Patent Troll Myths* (2011), Michael Risch offers a rich empirical study of a group of PAEs to explain how patent trolls operate.⁵⁹ Risch's sample, however, is restricted to lawsuits by the ten most litigious PAEs at the time—around 1,000 cases that encompassed 350 unique patents over a 23-year period (from 1986 to 2009).⁶⁰ We agree with Risch that this is an economically important group of litigants worthy of individualized study.⁶¹ However, we also agree with him that there are reasons to doubt that his important findings apply to all PAEs or NPEs more generally.⁶²

In categorizing the initial assignees of the most litigious PAE patents,⁶³ Risch offers a valuable addition to prior work that sheds light on the origins of litigated patents. Risch finds that a large percentage of patents asserted by the most litigious PAEs (28%) were originally owned by their inventors.⁶⁴ Thus, he concluded that PAEs might serve the important role of helping inventors enforce and monetize their patent rights.⁶⁵ He further found that only 21% of the original assignees of PAE-litigated patents could be described as NPEs other than individual inventors.⁶⁶ The remainder—just over half—of original assignees were product companies.⁶⁷

But does this mix of patent origin hold for the far more numerous group of less litigious PAEs? We do not know, but the answer to this question may have important implications for policy makers seeking to design the optimal patent system. For example, if a large percentage of less-litigious PAE patents were also purchased from inventors, this may support the argument that PAEs play a beneficial role by helping inventors recoup the costs of their innovations. Risch's study—particularly its methodology and creative use of explanatory variables—is valuable. However, his limited sample restricts broad inferences about the patent system as a whole.

Our second example of prior regression analysis that we would like to see repeated using the Dataset is *Extreme Value or Trolls on Top? The Characteristics of the Most-Litigated Patents*⁶⁸ (2009). In this work, John Allison, Mark Lemley, and Joshua Walker examine the “most-litigated

59. Michael Risch, *Patent Troll Myths*, 42 SETON HALL L. REV. 457 (2012).

60. *Id.* at 469-472.

61. *Id.* at 467.

62. *Id.* at 468.

63. *Id.* at 473.

64. *Id.* at 497.

65. *Id.*

66. *Id.* at 486.

67. *Id.*

68. See John R. Allison, Mark A. Lemley, & Joshua Walker, *Extreme Value or Trolls on Top? The Characteristics of the Most-Litigated Patents*, 158 U. PA. L. REV. 1 (2009).

patents”—106 patents asserted in eight or more separate lawsuits between 2000 and 2007—and compare their characteristics with those patents asserted in a single lawsuit during the same time period.⁶⁹ They conclude that the most-litigated patents are more likely to be software and telecommunications and are disproportionately owned by NPEs.⁷⁰ Specifically, the authors find 72% of the most-litigated patents cover software while the same is true of only 27% of once-litigated patents.⁷¹ Further, more than 80% of the once-litigated patents were owned by product companies while NPEs owned nearly 55% of the most-litigated patents.⁷² The authors also reveal that over 75% of most-litigated patents owned by NPEs were owned by inventor-led PAEs.⁷³

Like Risch, Allison, Lemley, and Walker study an economically important but small and unrepresentative group of patents and litigants. Neither study reveals much about the long tail of litigation by PAEs and other patent litigants who assert their patents in fewer lawsuits. We hope our comprehensive dataset will enable studies like this to be replicated using all PAE lawsuits over a longer period of time spanning the early 2000s through the present.

Our third example of an insightful empirical study we believe can reveal even more about the patent system if replicated using the Dataset is *How Often Do Non-Practicing Entities Win Patent Suits?*⁷⁴(2016) In it, John Allison, Mark Lemley, and David Schwartz evaluate a comprehensive dataset of every patent suit filed in 2008 and 2009 and report many interesting results from these two years of data. For example, they find that operating companies fare better in litigation than NPEs overall⁷⁵; that whereas 22.8% of practicing entities cases litigated to judgment involved software patents, 65.9% of NPE suits did⁷⁶, and that once pharmaceutical cases were removed from their calculations, no patent plaintiff fared very well.⁷⁷

Yet despite its comprehensiveness over two years and interesting findings, we have to ask if the same results hold for lawsuits filed in other years. 2008 and 2009 were the peak years of the Great Recession, and prior work shows that patent lawsuit filing behavior is different during economic downturns.⁷⁸ Perhaps litigating firms, and investors

69. *Id.* at 5.

70. *Id.* at 3.

71. *Id.* at 18.

72. *Id.* at 24.

73. *Id.* at 23-24 (with Table 7 categorizing 43 of the 103 classified most-litigated patents as asserted by individual-inventor-started NPEs).

74. John R. Allison, Mark A. Lemley, & David L. Schwartz, *How Often Do Non-Practicing Entities Win Patent Suits?*, 32 BERKELEY TECH. L.J. 235 (2016).

75. *Id.* at 269.

76. *Id.* at 263.

77. *Id.* at 275.

78. Alan C. Marco, Shawn P. Miller, & Ted M. Sichelman, *Do Economic Downturns Dampen Patent Litigation?*, 12 J. EMPIRICAL LEGAL STUD. 481, 485 (2015).

less desperate to monetize their IP assets, fare better during economic expansions. Maybe, as our results in Table 8 suggest, non-pharmaceutical patent plaintiffs fared poorly in all recent years. If Allison, Lemley and Schwartz' analysis was replicated with data covering a longer period of time, we could confirm that their results hold beyond a narrow window impacted by particular macroeconomic and legal conditions.

Overall, we can infer much less about the patent system than is possible from current empirical studies because they analyze relatively small and often unrepresentative subsets of patent litigation. We believe that a comprehensive analysis of litigation outcomes and statistics can further inform the policy debate about PAEs. Our Dataset, when complete, will categorize every patentasserter in every lawsuit filed in U.S. district courts since 2000, and we are sharing the Dataset publicly in the hopes that researchers will utilize it to complete more exhaustive empirical studies in the future.

V. CONCLUSION

Motivated by the intense debate surrounding the impact of PAEs on the patent system, innovation, and the macro economy, we began our work creating the Stanford NPE Litigation Dataset five years ago. Since then, we have carefully categorized patent asserters in over 50,000 cases including a random sample of 20 percent of all lawsuits filed between 2000 and 2015 (10,812 cases). Our preliminary data release consists of these cases,⁷⁹ and during Summer 2018 we are launching a new website that will enable researchers to utilize the entire Dataset.⁸⁰

Preliminary findings from our Dataset offer surprises. For one, licensing firms that acquired their patents from third parties are the type most responsible for the doubling of total defendant-lawsuit pairs since 2000. In 2000, these licensing firms asserted their patents in less than 5% of distinct defendant-plaintiff disputes, but this number has risen to encompass more than 30% of all disputes. Other findings reinforce what other researchers have already revealed. For example, PAE suits tend to involve software patents, terminate more quickly, and end in settlement more frequently than practicing entity cases. These facts demonstrate that PAEs have been responsible for a growing share of patent enforcement, and also lend support to the view that PAEs frequently assert patents of questionable validity (or more uncertain scope) and thus pursue a strategy of seeking quick nuisance value settlements.

The Dataset provides immediate value to scholars who wish to replicate past empirical studies using the full population of lawsuits over a

79. Available for download at <https://law.stanford.edu/projects/stanford-npe-litigation-dataset/#slnav-brief-dataset-methodology> [<https://perma.cc/96Y4-8TP9>].

80. The full Dataset will be accessible at <https://npe.law.stanford.edu>.

significant span of years. Scholars who have previously conducted studies on PAE litigation can confirm past research conclusions utilizing our comprehensive Dataset. The data might also reveal that some findings derived from small non-random samples do not extend to the population or other subsets. For example, scholars can now study how PAEs of varying sizes operate, rather than just how the largest and most easily recognizable PAEs operate.

Dataset research will allow scholars to develop a deeper understanding of the realities of litigation in the patent system. Scholars can examine the near doubling of distinct defendant-lawsuit pairs revealed over the 16-year period of the random-sample analysis. A more in-depth analysis of the entire Dataset might reveal the catalyst of increased PAE litigation. Scholars may also be able to draw more robust conclusions about PAE settlement rates and the amount of time respective entities take to litigate.

By providing scholars with party names along with patent asserter types, the Dataset also offers scholars the ability to focus on industry-specific players. For example, at least half of all PAE lawsuits filed in 2014 asserted a software patent. In contrast, only a small minority (11%) of Category 8 product company lawsuits asserted a software patent in 2014. Researchers can hone in on particular industries and their impact on the overall patent system. For example, who asserts pharmaceutical patents and how are these lawsuits resolved?

These preliminary findings and suggestions for further research underscore the potential value of our Dataset. We encourage scholars, policymakers, and other decision-makers to utilize our Dataset to test new theories, evaluate policy, and engage the PAE debate from a data-driven perspective. Our shared goal is to bring transparency to patent litigation by reshaping that landscape with new tools and understandings that help to promote the public interest through innovation, entrepreneurship, and competition across sectors and industries.