

CONNECTICUT LAW REVIEW

VOLUME 49

NOVEMBER 2016

NUMBER 1

Article

Empirical Analyses of Judicial Opinions: Methodology, Metrics, and the Federal Circuit

JASON RANTANEN

Despite the popularity of empirical studies of the Federal Circuit's patent law decisions, a comprehensive picture of those decisions has only recently begun to emerge. Historically, the literature has largely consisted of individual studies that provide just a narrow slice of quantitative data relating to a specific patent law doctrine. Even studies that take a more holistic approach to the Federal Circuit's jurisprudence primarily focus on their own results and address the findings of other studies only briefly. While recent developments in the field hold great promise, one important but yet unexplored dimension is the use of multiple studies to form a complete and rigorously supported understanding of particular attributes of the court's decisions.

ARTICLE CONTENTS

INTRODUCTION	229
I. A FRACTURED, MULTIDIMENSIONAL JIGSAW PUZZLE: REVERSAL-RATE STUDIES OF THE FEDERAL CIRCUIT	238
A. OVERALL REVERSAL-RATE STUDIES	238
B. CLAIM CONSTRUCTION REVERSAL RATES	240
II. THE CAUSES OF INTER-STUDY VARIABILITY	242
A. SELECTION EFFECTS AND STUDIES OF APPELLATE DECISIONS	242
B. CHOICES IN STUDY DESIGN	244
III. HOW STUDY DESIGN CHOICES CAN PRODUCE VARIATION IN DATA	259
A. STUDY DESIGN AND OVERALL REVERSAL RATE STUDIES	259
B. CLAIM CONSTRUCTION STUDIES	265
IV. POTENTIAL IMPACTS ON OTHER STUDIES	276
A. FIELD AND JUDICIAL HYPERACTIVITY	276
B. HENRY AND TURNER AND THE FEDERAL CIRCUIT’S IMPACT ON PATENT LITIGATION	279
CONCLUSION: TOWARD MORE MEANINGFUL EMPIRICAL STUDIES OF JUDICIAL OPINIONS	281
APPENDIX A	283
APPENDIX B.1	288
APPENDIX B.2	290
APPENDIX C	291



Empirical Analyses of Judicial Opinions: Methodology, Metrics, and the Federal Circuit

JASON RANTANEN *

INTRODUCTION

The United States Court of Appeals for the Federal Circuit is quite possibly the most scrutinized court of all time. Nearly every written decision the Federal Circuit issues involving patents is pounced on, dissected, and criticized within hours of release.¹ Academic centers and legal journals have dedicated entire conferences to this court.² While the term “legion” is often

* Professor of Law, University of Iowa College of Law. Thanks to the many folks who provided data for and commented on this project including, but not limited to, Jonas Anderson, Chris Cotropia, Peter Menell, Lee Petherbridge, David Schwartz, Chris Seaman, Ted Sichelman, Shawn Miller, John L. Turner, Matthew D. Henry, and the participants in the 3rd Annual Empirical Patent Law Conference at Northwestern School of Law and The Patent Conference 5 at the University of Kansas Law School. This paper benefited greatly from the work of my research assistants Brianna Chamberlin, Peter Kline, Alex Lodge, Raj Patel, and Andrew Stanley.

¹ See, e.g., Gene Quinn, *Federal Circuit Ignores Jury Finding of Non-Obviousness*, IPWATCHDOG (Aug. 21, 2014, 10:17 AM), <http://www.ipwatchdog.com/2014/08/21/federal-circuit-ignores-jury-finding-of-non-obviousness/id=50839> [<https://perma.cc/76W6-F7QP>] (commenting on the Federal Circuit’s “infinite wisdom” in “substituting its own decision for that of the decision maker at the district court level”); Jason Rantanen, *Federal Circuit Reminds Courts of Discretion on Fee Awards*, PATENTLY-O, (Aug. 26, 2014), <http://www.patentlyo.com/patent/2014/08/federal-reminds-discretion.html> [<https://perma.cc/HQ2H-TB8A2HRC-AJHU>] (discussing a Federal Circuit decision to remand to the district court following a Supreme Court reversal); Donald Zuhn, *AbbVie Inc. v. Mathilda & Terence Kennedy Institute of Rheumatology Trust* (*Fed. Cir. 2014*), PATENT DOCS (Sept. 4, 2014, 11:59 PM), <http://www.patentdocs.org/2014/09/abbvie-inc-v-mathilda-terence-kennedy--institute--of-rheumatology-trust-fed-cir-2014.html> [<https://perma.cc/64EM-YL5S3PUA-CVPB>] (opining on a Federal Circuit decision about AbbVie’s blockbuster drug, Humira, which upended more than a century’s worth of obviousness-type double-patenting precedent).

² See, e.g., FED. CIR. B. ASS’N, <http://www.fedcirbar.org> [<https://perma.cc/DWY2-B64438NK-XJHT>] (last visited Feb. 26, 2016) (describing purpose of the Federal Circuit Bar Association); FED. CIR. B.J., <https://www.law.gwu.edu/federal-circuit-bar-journal> [<https://perma.cc/PYD7-Q68M>] (last visited Oct. 11, 2016) (explaining academic and practical goals of the Federal Circuit Bar Journal); see also S.J. Plager, *The Federal Circuit as an Institution: On Uncertainty and Policy Levers*, 43 LOY. L.A. L. REV. 749, 750 (2009) (“[W]hen the conversation turns to the Federal Circuit, it turns to patent law, at least within this subset of the legal and academic community.”); MO. L. REV., *Evolving the Court of Appeals for the Federal Circuit and its Patent Law Jurisprudence* (Feb. 25, 2011), <http://law.missouri.edu/faculty/event/evolving-the-court-of-appeals-for-the-federal-circuit-and-its-patent-law-jurisprudence/> [<https://perma.cc/PY6S-FPE5S9KZ-X49N>] (describing a symposium on the Federal Circuit’s patent-related decisions from previous year); AM. U. L. REV., *Federal Circuit Symposium* (Mar. 24, 2016), <https://www.wcl.american.edu/secle/founders/2016/20160324a.cfm> [[https://perma.cc/VUC3-GUHR\]3T3R-YCSX](https://perma.cc/VUC3-GUHR]3T3R-YCSX)] (describing a panel on “the reviewability of the Patent Trial and Appeal Board’s decision to institute post-grant proceedings”).

used only as hyperbole, it is surprisingly accurate to use it to refer to articles discussing the Federal Circuit.³ One Federal Circuit judge remarked on the level of attention the court's opinions receive: "As a district court judge, your judgments are scrutinized. As a court of appeals judge, it's not just your judgments, but how you phrase your judgments that get scrutinized."⁴

Our collective fascination with the Federal Circuit parallels a general rise in interest in empirical studies of courts and their opinions. Scholars have honed in on the Federal Circuit, using it as the subject of a host of empirical studies covering an array of topics from jurisprudential content, judicial behavior, and institutional structure. These studies measure and probe virtually every aspect of the court's opinions, from outcomes to textual content to subtextual considerations. Over the past several years, the production of empirical studies of the Federal Circuit's decisions has simply exploded.⁵ Outside of the Supreme Court, the Court of Appeals for the Federal Circuit is probably the most empirically analyzed court in history, with nearly every aspect of its decisions measured and reported.

The increasing use of quantitative empirical methods to analyze judicial precedent is a healthy evolution of conventional legal scholarship techniques.⁶ Empirical legal scholarship typically involves reading judicial opinions in a systematic manner, carefully recording information about those opinions, and reporting and analyzing the results. In other words,

³ A Roman "legion" was a unit of 3,000–6,000 soldiers. *Legion*, NEW OXFORD AMERICAN DICTIONARY 997 (3d ed. 2010). A search of WestlawNext performed on September 1, 2014, using the search phrase "adv: ATLEAST2 ("Court #of Appeals #for the Federal Circuit")"—in other words, searching Westlaw's database for results with a least two uses of the phrase—resulted in 2,988 hits in the "Law Reviews & Journals" category. Although highly superficial, this search result still adds some flavor to the basic proposition offered here: There are many, many articles about the Federal Circuit.

⁴ Doug Sherwin, *Federal Circuit Judge Speaks at PatCon4 on Rise in IP Issues*, DAILY TRANSCRIPT (Apr. 8, 2014), http://www.sddt.com/News/article.cfm?SourceCode=20140404czj&_t=Federal+Circuit+judge+speaks+at+PatCon4+on+rise+in+IP+issues#.V_1Cwzub-f4 [<https://perma.cc/KMK5-4Q4T>].

⁵ This would normally be the place for a lengthy footnote listing all such studies. But that footnote would span multiple pages. Instead, to give a sense of the number, and publishing dates, of quantitative empirical studies of the Federal Circuit, I have listed over eighty such studies found in preparing this article in Appendix A. See Ryan G. Vacca, *The Federal Circuit as an Institution*, in RESEARCH HANDBOOK ON THE ECONOMICS OF INTELLECTUAL PROPERTY LAW – VOL. II: ANALYTICAL METHODS (Peter S. Menell, Ben Depoorter & David L. Schwartz, eds., forthcoming 2016) [hereinafter RESEARCH HANDBOOK].

⁶ See, e.g., Kevin M. Clermont & Theodore Eisenberg, *CAFA Judicata: A Tale of Waste and Politics*, 156 U. PA. L. REV. 1553, 1554 (2008) (analyzing cases decided under the Class Action Fairness Act of 2005 to determine whether "courts match[ed] the congressional ardor for class action reform"); Mark A. Hall & Ronald F. Wright, *Systematic Content Analysis of Judicial Opinions*, 96 CALIF. L. REV. 63, 64 (2008) ("We propose in this Article that one standard social science technique—content analysis—could form the basis for a uniquely legal empirical methodology."). But see Harry T. Edwards & Michael A. Livermore, *Pitfalls of Empirical Studies that Attempt to Understand the Factors Affecting Appellate Decisionmaking*, 58 DUKE L.J. 1895, 1903–05 (2009) (describing problems with empirical studies of judicial decisions).

precisely what lawyers do every day but on a more systematic basis.⁷ The ultimate goal of such scholarship is to add another perspective; one that looks less at individual trees and more at the forest of judicial opinions.⁸ While the value of this enterprise is not without criticism,⁹ and particular forms of the methodology may be more or less like traditional legal research,¹⁰ it can nevertheless be a very useful tool, especially when combined with other forms of legal research and analysis.

Given the number of empirical studies, one would reasonably expect that we should have a pretty good sense of the Federal Circuit from a quantitative standpoint. After all, if every inch of the court's decisions has been prodded, examined, and measured, the result should be a clear picture of the court's jurisprudence; or at the very least, that portion of it which can be quantified. And, perhaps by using this data, we could finally answer some of the most pressing questions about the court: Is patent law jurisprudence more uniform today than when Congress created the Federal Circuit? Do specialist courts work better than generalist courts? Is the claim construction reversal rate too high? And so on. We should, one would think, be able to take all of these studies together to build a complete picture of the Federal Circuit; or, at least, to put together enough of the jigsaw that we can spot what information is missing.

But it turns out that we still are not so certain about the answers to these questions. And even many of the answers that seem to be relatively certain really are not. Rather, until only very recently, each new article mainly sought to add its own empirical data to the heap, to report its individual study

⁷ See R. Polk Wagner & Lee Petherbridge, *Is the Federal Circuit Succeeding? An Empirical Assessment of Judicial Performance*, 152 U. PA. L. REV. 1105, 1126–27 (2004) (“[M]ining a large number of judicial opinions to collect relevant information is not so different from what lawyers do every day. The difference here is that the data was collected more systematically and measured by more rigid criteria.”).

⁸ Hall & Wright, *supra* note 6, at 64 (explaining that content analysis is “more than a better way to read cases. It brings the rigor of social science to our understanding of case law, creating a distinctively legal form of empiricism.”).

⁹ See Edwards & Livermore, *supra* note 6, at 1903 (“In recent years, there has been a greater effort on the part of legal scholars to apply various empirical methodologies to the study of judicial decisionmaking.”); see also Dennis Crouch, *An Empirical Study of the Role of the Written Description Requirement in Patent Examination*, 104 NW. U. L. REV. 1665, 1667 (2010) (publishing the results of a “retrospective empirical study of the role of the written description requirement in patent office examination practice”); Jay P. Kesan & Gwendolyn G. Ball, *Judicial Experience and the Efficiency and Accuracy of Patent Adjudication: An Empirical Analysis of the Case for a Specialized Patent Trial Court*, 24 HARV. J.L. & TECH. 393, 396 (2011) (examining “whether there is empirical evidence supporting the assumption that increasing judicial patent specialization will result in ‘better’ patent adjudication”).

¹⁰ For example, content analysis studies that look closely at the substantive content and analysis of the court's opinions, such as that of Lee Petherbridge and R. Polk Wagner, tend to be more closely akin to traditional legal research than studies that measure and report only outcomes or case counts based on a keyword search. See Lee Petherbridge & R. Polk Wagner, *The Federal Circuit and Patentability: An Empirical Assessment of the Law of Obviousness*, 85 TEX. L. REV. 2051, 2055–56 (2007) (reporting that courts' obviousness analyses “appear[] relatively stable and increasingly flexible”).

results, and—to the extent it addressed other studies at all—to show why *that* study's approach was superior to the other studies.¹¹ In a way, this makes perfect sense: each new data point should add to and advance our collective understanding of patent law. The risk, though, is that the conversation turns into a cacophony rather than a symphony, especially when those data points are inconsistent with one another.

Recently, a group of leading empirical researchers came together to produce a literature review of *all* empirical studies of intellectual-property law.¹² Much of that literature focuses on patent law, with chapters on claim construction,¹³ nonobviousness,¹⁴ inequitable conduct, and patent misuse¹⁵—even on the Federal Circuit as an institution.¹⁶ Thus, one can no longer say that the field lacks any attempts to consider all of the empirical studies of the Federal Circuit's jurisprudence in a comprehensive manner. Rather, there is a thorough literature review of those studies—an extremely important development in the field.

Equally important, however, is an understanding of the degree to which the studies described in the literature actually provide meaningful data. And patent law offers a unique opportunity to examine this issue due to the existence of multiple studies purportedly all measuring the same thing. Looking at these studies comprehensively should give a sense of the degree to which researchers can accurately measure given attributes. That comparison also tells us something about the extent to which those measurements can be relied upon for purposes of assessing theoretical frameworks and supporting normative arguments.¹⁷

¹¹ But see Ted Sichelman, *Myths of (Un)certainty at the Federal Circuit*, 43 LOY. L.A. L. REV. 1161, 1163 (2010) (situating claim construction reversal rates into the broader set of reversal rates generally).

¹² This project, initiated and developed by Peter Menell, Dave Schwartz, and Ben DePoorter, is manifest in the RESEARCH HANDBOOK, *supra* note 5.

¹³ J. Jonas Anderson & Peter Menell, *Empirical Studies of Claim Construction*, in RESEARCH HANDBOOK, *supra* note 5.

¹⁴ Christopher Cotropia & Ronald Mann, *Empirical Analysis of Patent Validity*, in RESEARCH HANDBOOK, *supra* note 5.

¹⁵ Jason Rantanen & Lee Petherbridge, *Inequitable Conduct and Misuse*, in RESEARCH HANDBOOK, *supra* note 5.

¹⁶ Ryan Vacca, *The Federal Circuit as an Institution*, in RESEARCH HANDBOOK, *supra* note 5.

¹⁷ In other fields, an extensive literature employing meta-analysis techniques has emerged. A meta-analysis, sometimes called a “metastudy,” is a form of statistical research that involves analyzing the results of multiple prior studies in order to obtain a clear picture of the actual characteristics of a given population. Generally, this involves averaging the quantitative outcomes of those prior studies. See, e.g., J.E. HUNTER ET AL., *METHODS OF META-ANALYSIS: CORRECTING ERROR AND BIAS IN RESEARCH FINDINGS* xxvii (2d ed. 2004). The literature of empirical legal studies of patent law is not nearly so well developed, nor are standard meta-analysis techniques necessarily appropriate, as its studies generally involve entire populations rather than samples of a population. The only meta-analysis of judicial opinions of which I am aware is Daniel R. Pinello, *Linking Party to Judicial Ideology in American Courts: A Meta-Analysis*, 20 JUST. SYS. J. 219 (1999). Furthermore, meta-analyses typically involve multiple samples of a population all being combined to more accurately measure a characteristic of that population. In contrast, empirical studies of judicial opinions in patent law typically examine the entire

This Article seeks to bring a measure of structure to this discussion by examining the mostly heavily studied (from an empirical perspective) aspects of Federal Circuit decision-making: the rates at which the court reverses lower tribunals in patent infringement suits. To be sure, while reversal-rate studies are certainly not unique to the Federal Circuit,¹⁸ studies of that court dominate the contemporary empirical literature examining reversal rates.¹⁹

Central to some of the most intense debates about the patent system is quantitative data about two subjects: (1) the rates at which the Federal Circuit reverses the courts it reviews and (2) the outcomes of Federal Circuit decisions on particular doctrinal issues. These discussions question and scrutinize the court's review of claim construction decisions, its nonobviousness and inequitable conduct jurisprudence and even the very idea of having a Federal Circuit (to name just a few topics).²⁰ Inevitably, these debates draw upon statistical information about the Federal Circuit's decisions—since those decisions are, after all, the very fuel that feeds the fire. The following quotation from a treatise on patent litigation is typical:

The Federal Circuit reverses the decisions of district courts at a relatively high rate. Some 40% of the appeals from district court patent cases were reversed in whole or in part in the 12-month period ending in September 2009. The reversal rate, in whole or in part, of appeals from district courts to the Federal Circuit was as high as 53% in 1997. On claim construction issues alone, where the Federal Circuit gives de novo review, the reversal rate has been reported as high as 40%.²¹

population. In other words, these studies are all—on the surface at least—measuring the same attribute of the same set of things.

¹⁸ See, e.g., Robert Steinbuch, *An Empirical Analysis of Reversal Rates in the Eighth Circuit During 2008*, 43 LOY. L.A. L. REV. 51, 52 (2009) (studying “likelihood of reversal by the U.S. Court of Appeals for the Eighth Circuit”); Corey Rayburn Yung, *Flexing Judicial Muscle: An Empirical Study of Judicial Activism in the Federal Courts*, 105 NW. U. L. REV. 1, 5 (2011) (analyzing “whether activism was correlated with the political party of the President, the identity of the appointing President, the ideology of the judge, the political composition of the Senate, and the scenario of the President and the Senate majority being of the same party at the time of judicial appointment”).

¹⁹ Of the top fifty hits for “empirical /s ‘reversal rate’” in the “Law Review & Journals” directory on WestlawNext, by my count only six focused on courts other than the Federal Circuit. WESTLAW, www.westlaw.com (search conducted Jan. 20, 2016) (results available upon request).

²⁰ See Craig Allen Nard & John F. Duffy, *Rethinking Patent Law's Uniformity Principles*, 101 NW. U. L. REV. 1619, 1623–24 (2007) (arguing for a “decentralized decisionmaking” approach to appellate review of patent cases); Diane P. Wood, *Keynote Address: Is It Time to Abolish the Federal Circuit's Exclusive Jurisdiction in Patent Cases?*, 13 CHI.-KENT J. INTELL. PROP. 1, 2–3 (2013) (analyzing the historical justifications behind the Federal Circuit's exclusive jurisdiction over patent appeals).

²¹ DON W. MARTENS & JOHN B. SGANGA JR., PRE-LITIGATION PATENT ENFORCEMENT § 10:50 (2013–2014 ed.) (footnotes omitted); see also ERIC M. DOBRUSIN & KATHERINE E. WHITE, INTELLECTUAL PROPERTY LITIGATION: PRETRIAL PRACTICE § 4.03[A] (3d ed. 2014); 7 HON. TIMOTHY

Scholars and other commentators frequently draw upon reversal rate statistics for a variety of points, including both intra-circuit analyses and comparisons with other circuits.²² Even judges draw upon and reference the rate at which the Federal Circuit reverses district courts. In his dissent in *Cybor Corp. v. FAS Technologies, Inc.*,²³ for example, Judge Rader observed that “[t]he Federal Circuit, according to its own official 1997 statistics, reversed in whole or in part 53% of the cases from district courts (27% fully reversed; 26% reversed-in-part),”²⁴ as support for his broader concerns about the Federal Circuit’s standard of review for claim construction.

Claim construction—the process of interpreting the words of a patent’s claims—is an especially strong attractor of statistics on reversals.²⁵ In the last two years alone, the standard of review for claim construction has been the subject of both an en banc Federal Circuit opinion²⁶ and a Supreme Court decision.²⁷ Much of the debate surrounding claim construction has focused on the perception that the Federal Circuit reverses district courts’ claim constructions at a rate that is too high.²⁸ Unsurprisingly, statistics about that rate are frequently deployed.²⁹ Nor is citation of these figures limited to

B. DYK & SAMUEL F. ERNST, BUSINESS AND COMMERCIAL LITIGATION IN THE FEDERAL COURTS, § 86:1 (3d ed. 2011) (“The Federal Circuit has heard an average of over 400 patent infringement appeals a year over the last 10 years. This is probably due in large part to the high reversal rate of district court judgments by the Federal Circuit, where at least 40% of patent cases have been reversed, at least in part, for each of the last 10 years.”) (footnotes omitted).

²² See, e.g., Jeanne C. Fromer, *Patentography*, 85 N.Y.U. L. REV. 1444, 1446 (2010) (“The Federal Circuit reverses district courts twenty-two percent of the time”) (citations omitted); Jeffrey A. Lefstin, *The Measure of the Doubt: Dissent, Indeterminacy, and Interpretation at the Federal Circuit*, 58 HASTINGS L.J. 1025, 1064 (2006) (comparing reversal rates between tribunals).

²³ 138 F.3d 1448 (Fed. Cir. 1998) (en banc), *abrogated by* 35 U.S.C. § 112 (2006) *as recognized in* *Teva Pharm. USA, Inc. v. Sandoz, Inc.*, 789 F.3d 1335, 1340 (Fed. Cir. 2015).

²⁴ *Id.* at 1476.

²⁵ See, e.g., J. Jonas Anderson & Peter S. Menell, *Informal Deference: A Historical, Empirical, and Normative Analysis of Patent Claim Construction*, 108 NW. U. L. REV. 1, 4 (2013) (“Notwithstanding the critical importance of claim boundaries to both patentees and competitors, the processes and doctrines governing the construction of patent claims are notoriously amorphous and uncertain.”).

²⁶ *Lighting Ballast Control L.L.C. v. Philips Elecs. N. Am. Corp.*, 744 F.3d 1272, 1276–77 (Fed. Cir. 2014), *vacated sub nom.* *Lighting Ballast Control L.L.C. v. Universal Lighting Techs., Inc.*, 135 S. Ct. 1173 (2015) (confirming the *Cybor* standard of de novo review for claim construction).

²⁷ *Teva Pharm. USA, Inc. v. Sandoz, Inc.*, 135 S. Ct. 831, 844 (2015) (vacating and remanding after finding that the Federal Circuit incorrectly applied the clearly erroneous standard to a review of a claim construction ruling).

²⁸ Indeed, any discussion of claim construction reversal rates would be incomplete without noting the set of articles in the *Loyola Los Angeles Law Review* discussing how these rates should be interpreted. See, e.g., Richard S. Gruner, *How High Is Too High?: Reflections on the Sources and Meaning of Claim Construction Reversal Rates at the Federal Circuit*, 43 LOY. L.A. L. REV. 981, 985 (2010) (arguing that reversal-rate statistics are “outliers” and “selectively filtered”); David L. Schwartz, *Pre-Markman Reversal Rates*, 43 LOY. L.A. L. REV. 1073, 1075–76 (2010) (reporting reversal rates before and after *Markman* and *Cybor*); Sichelman, *supra* note 11, at 1163 (likening the Federal Circuit’s claim construction decisions to “a coin flip” or “roll [of] the dice”) (internal citations omitted).

²⁹ See, e.g., Sichelman, *supra* note 11, at 1169 (advocating for the Federal Circuit’s retaining exclusivity over patent disputes due to interjurisdictional differences in deciding key issues); see also Lee Petherbridge, *The Claim Construction Effect*, 15 MICH. TELECOMM. & TECH. L. REV. 215, 216

academic commentary. Litigants cite to them.³⁰ So do judges. For instance, Judge Rader continued on in *Cybor* to observe that:

[O]ne study shows that the plenary standard of review has produced reversal, in whole or in part, of almost 40% of all claim constructions since *Markman I*. A reversal rate in this range reverses more than the work of numerous trial courts; it also reverses the benefits of *Markman I*. In fact, this reversal rate, hovering near 50%, is the worst possible. Even a rate that was much higher would provide greater certainty.³¹

Similarly in *Lighting Ballast*,³² Judge Newman drew upon Jonas Anderson and Peter Menell's empirical study of claim construction reversal rates for the proposition that "[n]ow the reversal rate for claim construction appeals is much closer to that of other patent-related issues."³³

Outcomes, too, are a frequent and important subject of empirical examination. These types of studies help establish a broader understanding of what the Federal Circuit is actually doing. For example, the outcome of appeals involving claims of inequitable conduct are a frequent subject of interest, as are outcomes of nonobviousness,³⁴ infringement,³⁵ and disclosure issues.³⁶

(2008) (citations omitted) ("The premise of the Federal Circuit *is* widespread dissatisfaction with the confusion and uncertainty that followed from regional circuit involvement [in substantive patent disputes].").

³⁰ *Raylon, LLC v. Complus Data Innovations, Inc.*, 700 F.3d 1361 (Fed. Cir. 2012).

³¹ *Cybor Corp. v. FAS Techs., Inc.*, 138 F.3d 1448, 1476 (Fed. Cir. 1998) (en banc) (Rader, J., dissenting), *abrogated by* 35 U.S.C. § 112 (2006) *as recognized in* *Teva Pharm. USA, Inc. v. Sandoz, Inc.*, 789 F.3d 1335, 1340 (Fed. Cir. 2015).

³² 744 F.3d 1272 (Fed. Cir. 2014), *vacated sub nom.* *Lighting Ballast Control LLC v. Universal Lighting Techs., Inc.*, 135 S. Ct. 1173 (2015).

³³ *Id.* at 1290 (quoting Anderson & Menell, *supra* note 25, at 57).

³⁴ See, e.g., Christopher A. Cotropia, *Nonobviousness and the Federal Circuit: An Empirical Analysis of Recent Case Law*, 82 NOTRE DAME L. REV. 911, 913–14 (2007) (examining, over a four-year period, Federal Circuit cases considering obviousness of a patent claim); Petherbrige & Wagner, *supra* note 10, at 2056 ("[T]he jurisprudence of obviousness, as developed by the Federal Circuit, appears relatively stable and increasingly flexible."); Jason Rantanen, *The Federal Circuit's New Obviousness Jurisprudence: An Empirical Study*, 16 STAN. TECH. L. REV. 709, 710–11 (2013) (testing hypothesis about the impact of *KSR* on obviousness determinations by examining an expansive data set of pre- and post-*KSR* obviousness decisions).

³⁵ See, e.g., Matthew D. Henry & John L. Turner, *The Court of Appeals for the Federal Circuit's Impact on Patent Litigation*, 35 J. LEGAL STUD. 85, 88–89 (2006) (analyzing "courts' tendencies with respect to the validity and infringement inquiries" over a nearly sixty-year period to determine whether cases turn more frequently on validity or infringement); Glynn S. Lunney, Jr., *Patent Law, the Federal Circuit, and the Supreme Court: A Quiet Revolution*, 11 SUP. CT. ECON. REV. 1, 2–3 (2004) (highlighting the Supreme Court's general reluctance towards altering the Federal Circuit's holdings regarding the doctrine of equivalents).

³⁶ See, e.g., John R. Allison & Lisa Larrimore Ouellette, *How Courts Adjudicate Patent Definiteness and Disclosure*, 65 DUKE L.J. 609, 611 (2016) (offering novel empirical data about how the Patent Act's disclosure requirement is applied in practice); Timothy R. Holbrook, *Possession in Patent Law*, 59 SMU L. REV. 123, 146 (2006) ("[W]hile stating that patents *never* serve a teaching function is too strong of a

Reversal rates and outcome measures are also used in assessing legal realism-based claims about appellate decision making. Using reversal rate measures, for example, empiricists have tested claims that the Federal Circuit exhibits “judicial hyperactivity.”³⁷ Additionally, by studying nonobviousness outcomes, empiricists have examined whether judges’ political ideology is liberal or conservative, or whether their “patent expertise” affects their decisions.³⁸ They have also considered whether “sitting by designation” affects the rate at which the Federal Circuit reverses district court judges on claim construction issues.³⁹

Litigants may also consider these measures when deciding whether to appeal or what strategy to follow. For example, it probably does not make sense, in most instances, to appeal a factfinder’s determination of no inequitable conduct; historically, the Federal Circuit has almost never reversed such a determination.⁴⁰ While statistics alone do not compel this conclusion, they help provide context that can be useful in developing appellate strategy.

All that said, it is important to keep firmly in mind that reversal rate and outcome studies come with substantial limitations.⁴¹ Population bias, in particular, is an important consideration to take into account, as are the limitations of predicting future behavior based on historical data. This

statement, the courts have grossly overstated the true extent of this function. Given the teaching function’s inconsistency and irrelevance to patent theory, basing patent policy and doctrine of the quid pro quo, as the courts do, is a flawed approach.”); Mark A. Lemley & Ragesh K. Tangri, *Ending Patent Law’s Willfulness Game*, 18 BERKELEY TECH. L.J. 1085, 1087 (2003) (questioning whether pre-litigation infringement and validity opinions constitute privileged communication when used as a defense to willful infringement).

³⁷ See Ted L. Field, “Judicial Hyperactivity” in the Federal Circuit: An Empirical Study, 46 U.S.F. L. REV. 721, 723 (2012) (finding that “the Federal Circuit’s reversal rates are . . . greater than those of the other circuits” to suggest that “the Federal Circuit in patent cases is more judicially hyperactive than other circuits”).

³⁸ Banks Miller & Brett Curry, *Expertise, Experience, and Ideology on Specialized Courts: The Case of the Court of Appeals for the Federal Circuit*, 43 LAW & SOC’Y REV. 839, 840–41 (2009) [hereinafter Miller & Curry, *Expertise, Experience*] (associating Federal Circuit judges’ technical training with their substantive decision-making tendencies); see also Banks Miller & Brett Curry, *Experts Judging Experts: The Role of Expertise in Reviewing Agency Decision Making*, 38 LAW & SOC. INQUIRY 55, 56 (2013) [hereinafter Miller & Curry, *Experts Judging Experts*] (determining “whether expert and nonexpert judges review” decisions in different ways in view of Federal Circuit decisions).

³⁹ Mark A. Lemley & Shawn P. Miller, *If You Can’t Beat ‘Em, Join ‘Em? How Sitting by Designation Affects Judicial Behavior* (Stan. Pub. Law, Working Paper No. 2449349, 2014), <http://ssrn.com/abstract=2449349> [https://perma.cc/886V-YWXS63D-CQC2].

⁴⁰ Lee Petherbridge et. al., *The Federal Circuit and Inequitable Conduct: An Empirical Assessment*, 84 S. CAL. L. REV. 1293, 1318–19 (2011).

⁴¹ See, e.g., Wagner & Petherbridge, *supra* note 7, at 1128–29 (discussing limitations of looking at claim construction reversal rates); see also Edward L. Rubin, *The Concept of Law and the New Public Law Scholarship*, 89 MICH. L. REV. 792, 801 (1991) (“[Judges] justify [their] conclusion[s] by showing that [they] proceed[] from accepted sources by legitimate, properly argued steps.”); William M. Sage, *Judicial Opinions Involving Health Insurance Coverage: Trompe L’Oeil or Window on the World?*, 31 IND. L. REV. 49, 61–65 (1998) (identifying small sample size, long time lags, and selection biases in an empirical study sample).

Article does not, however, explore these limitations in depth. Rather, it takes as its starting point the proposition that these studies *are* important and that meaningful inferences can be drawn from the underlying attributes they measure.

Indeed, despite their limitations, objective measures of the rate at which the Federal Circuit reverses the tribunals it reviews, both generally and on a doctrine-specific basis, have been a frequent subject of measurement and quantification, a trend that shows no signs of waning. Improvements in technology, coding methodologies, and expertise will make these studies more and more feasible, and the corpus of judicial opinions continues to beckon—particularly as the Supreme Court’s high degree of activity in patent law over the past several years inspires hypotheses about the impact of its rulings.

This Article proceeds as follows. Part I presents aggregate data from studies and sources reporting reversal rates, both generally and for claim construction. When viewed together, surprisingly wide variations in some of the reported data are revealed, even when the same legal issue and time period are examined. But not all metrics reveal extreme inter-study variability; others show much less variation among studies.

Part II examines reasons why empirical data on the Federal Circuit’s decisions can differ from source to source even when examining what nominally seems to be same thing. Central to this variability are choices made during study design, both those whose structure directionally affect results and those that inherently produce variation in results. Many of these issues of study design reflect choices made during study development; some are controllable, others less so. Understanding their role is critical for any multi-study comparison of results. Part III applies these issues in study design to the data on reversal rates and doctrinal outcomes described in Part I, examining how and to what extent particular study-design choices affect the reported data, and Part IV extends these insights to other studies, particularly those that use decisional outcomes to assess judicial behavior and characteristics.

Ultimately, as I discuss in the conclusion, addressing study design issues in a consistent way is critical for developing meaningful studies, interpreting findings, and using results. In addition, such studies are likely to gain even more attention as scholars explore the Supreme Court’s recent opinions affecting both core patent law doctrines and appellate standards of review. Indeed, if history is any guide, claim construction will inevitably be the subject of multiple future studies, especially given the Supreme Court’s recent decision in *Teva v. Sandoz*⁴² that altered the standard of review used by the Federal Circuit when reviewing district courts’ claim constructions.⁴³

⁴² 135 S. Ct. 831 (2015).

⁴³ *Id.* at 835.

I. A FRACTURED, MULTIDIMENSIONAL JIGSAW PUZZLE: REVERSAL-RATE STUDIES OF THE FEDERAL CIRCUIT

A. Overall Reversal-Rate Studies

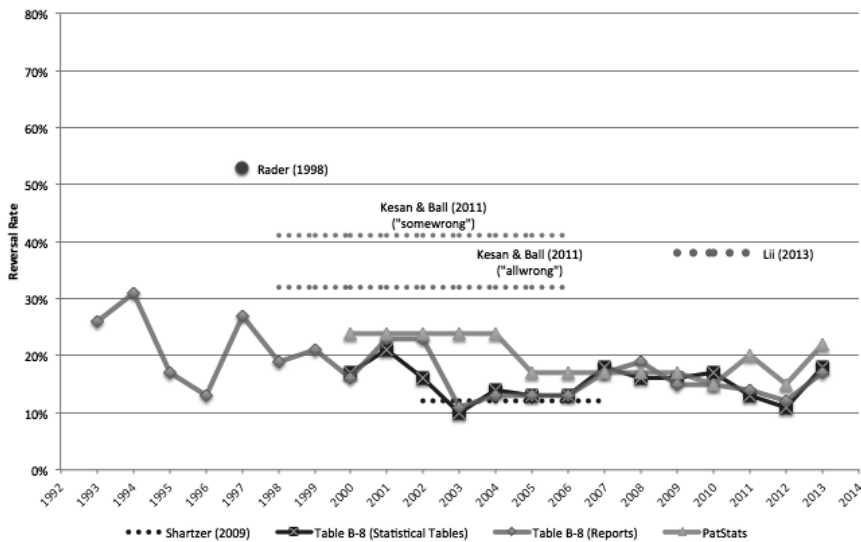
At the heart of quantitative measures of the Federal Circuit lies a deceptively simple metric: the rate at which the appellate court reverses the courts and other tribunals it reviews—particularly district courts, since that is where most patent infringement litigation takes place. While there are many articles that report reversal rates, the data actually comes from relatively few original sources.⁴⁴ Figure 1 presents those sources for appeals arising from the district courts.⁴⁵ All data is on a calendar year basis, with the exception of Table B-8 (FY Reports). Sources that report reversal rates on an aggregate basis—that is, a reversal rate for the entire period of the study—are depicted below as a dotted line, while sources that report reversal rates on an annual basis are presented as solid lines.⁴⁶

⁴⁴ Indeed, I was surprised how few original sources there actually are for this data.

⁴⁵ Tables containing the underlying data used for all figures are provided in Appendices B1 and B.2. See Appendices B.1 & B.2.

⁴⁶ Data from Figure 1 comes from the following sources: *Cybor Corp. v. FAS Techs., Inc.*, 138 F.3d 1448, 1476 (Fed. Cir. 1998) (Rader, J., dissenting); Kesan & Ball, *supra* note 9; Teresa Lii, *Shopping for Reversals: How Accuracy Differs Across Patent Litigation Forums*, 12 CHI.-KENT. J. INTELL. PROP. 31 (2013); Adam Shartzter, *Patent Litigation 101: Empirical Support for the Patent Pilot Program's Solution to Increase Judicial Experience in Patent Law*, 18 FED. CIR. B.J. 191 (2009); U.S. COURTS, STATISTICS & REPORTS: DATA TABLES, B-8, <http://www.uscourts.gov/data-table-numbers/b-8> [<https://perma.cc/6SPK-VJQU9DFF-VRKN>]; U.S. PATENT LITIGATION STATISTICS PROJECT, UNIV. OF HOUSTON LAW CTR., <http://www.Patstats.org/Patstats2.html> [<https://perma.cc/UY7Q-LH84G5VC-YYM3>] (last visited Aug. 8, 2016) [hereinafter PATSTATS]. These sources are discussed in more detail *infra* at Part III.

Figure 1: Reported Rates at Which the Federal Circuit Reverses District Courts



A glance at Figure 1 reveals substantial variation in the reported rates at which the Federal Circuit reverses district courts. On the low end are the reversal rates provided by the Federal Judicial Center’s *Statistical Tables* and *Reports*, which range from 11% to 27%, collected from Patstats (15%–24%),⁴⁷ and provided by Shartzter⁴⁸ for the high-frequency patent case districts (12%). On the high end, by contrast, are the rates provided by Lii⁴⁹ for the “busiest districts for patent litigation” (38%), Kesan and Ball⁵⁰ at the per-issue level (41% using the “somewrong” standard; 32% using the “allwrong” standard),⁵¹ and Rader (54%).

What might be the sources of this variation in reversal rates? To preview Parts II and III, some of the answer turns on the answer to the threshold question of *what* is a reversal. Is it reversal of any aspect of the lower tribunal’s handling of the case? Or just a reversal-in-all? Reversals in written decisions only? Are reversals counted on a per-case or per-issue level? What about appeals that settle? Do those count as reversals, affirmances, or not at all? How are remands treated? Decisions that reverse-in-part and affirm-in-part? The same as an outright reversal? What about Rule 36 affirmances? Are these counted in the dataset or not? Differences in the answers to these

⁴⁷ PATSTATS, *supra* note 46.

⁴⁸ Shartzter, *supra* note 46, at 233.

⁴⁹ Lii, *supra* note 46, at 43, 45.

⁵⁰ Kesan & Ball, *supra* note 9, at 434.

⁵¹ The rate cited by Judge Rader in his dissent in *Cybor* is the highest reported rate; Judge Rader offered no explanation as to source other than it was from the court’s “own official 1997 statistics” *Cybor*, 138 F.3d. at 1476 (Rader, J., dissenting). He does add, however, that this consists of “27% fully reversed[,] 26% reversed-in-part.” *Id.*

questions can affect the reported reversal rates—sometimes in a substantial way. Worse, a failure to understand their causes renders Figure 1 (let alone reliance on any individual statistic) essentially meaningless.

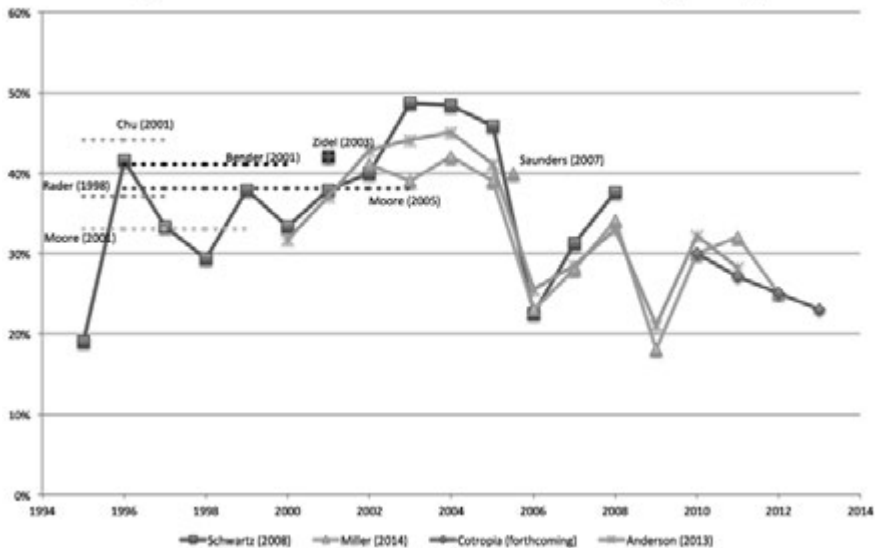
B. *Claim Construction Reversal Rates*

The claim construction reversal rate holds a special place in patent law lore. The subject of numerous studies—one by a professor who became a Federal Circuit judge—it appears in some form in almost every discussion of the Federal Circuit's review of district courts' claim constructions.

In examining reversal rates on claim construction, contemporary researchers have been careful to define the precise unit of analysis studied. Thus, the claim construction metric has come to take three distinct forms: (1) reversals on individual *claim terms*; (2) reversals of any claim term *in the opinion*; and (3) reversals that *impact the outcome*.⁵² Of these statistics, the most data is available on the second. That data is provided in Figure 2.⁵³

⁵² Reversals that impact the outcome result in either a decision to reverse the case, or remand it to the district court. Anderson & Menell, *supra* note 25, at 40 n.244.

⁵³ The table containing the underlying data for these figures is provided in Appendices B.1 & B.2. Data for Figure 2 comes from *Cybor Corp.*, 138 F.3d. at 1476 n.4 (Rader, J., dissenting); Anderson & Menell, *supra* note 25; Gretchen Ann Bender, *Uncertainty and Unpredictability In Patent Litigation: The Time Is Ripe for a Consistent Claim Construction Methodology*, 8 J. INTELL. PROP. L. 175, 207–08 (2001); Christian A. Chu, *Empirical Analysis of the Federal Circuit's Claim Construction Trends*, 16 BERKELEY TECH. L.J. 1075, 1125 n.180 (2001); Cotropia & Mann, *supra* note 14; Shawn P. Miller, "Fuzzy" Software Patent Boundaries and High Claim Construction Reversal Rates, 17 STAN. TECH. L. REV. 809, 811–12 (2014); Kimberly A. Moore, Markman *Eight Years Later: Is Claim Construction More Predictable?*, 9 LEWIS & CLARK L. REV. 231 (2005); Michael Saunders, *A Survey of Post-Phillips Claim Construction Cases*, 22 BERKELEY TECH. L.J. 215, 232–35 (2007) (studying claim construction reversal rates for a one year period during 2005–06); David L. Schwartz, *Practice Makes Perfect? An Empirical Study of Claim Construction Reversal Rates in Patent Cases*, 107 MICH. L. REV. 223 (2008); Andrew T. Zidel, *Patent Claim Construction in the Trial Courts: A Study Showing the Need for Clear Guidance From the Federal Circuit*, 33 SETON HALL L. REV. 711, 737–54 (2003). Thanks to David Schwartz, Shawn Miller, Chris Cotropia, Jonas Anderson, and Peter Menell for providing underlying data from their recent published and unpublished work. Values from Kimberly A. Moore, Markman *Eight Years Later: Is Claim Construction More Predictable?*, *supra*, were estimated from a figure on page 246 of that article and are best estimates to within a percentage point or two.

Fig. 2: Federal Circuit Claim Construction Reversal Rates (per-case)

Although there remains some variation among reported reversal rates, especially for the earlier period, many of the studies report relatively consistent results. In particular, as has been observed elsewhere,⁵⁴ all of the post-*Phillips* studies indicate a general decline in the rate at which the Federal Circuit is reversing district courts on claim construction. And yet, some variation in results among studies remains even when nominally the same thing—reversal rates on a per-case level—is measured.

The data presented above is subject to many important limitations. First, the figures contain information only on overall reversal or outcome data. They are thus subject to all of the limitations that apply when interpreting this type of data, limitations that are discussed extensively in some of the studies that the data come from.⁵⁵ Thus, apparent trends should be approached with caution.

But the purpose of this Article is not to draw inferences from these studies; that is a separate project.⁵⁶ Rather, the purpose is to examine the differences among these studies.⁵⁷ After all, they are all purportedly measuring the same underlying data. Why, then, do they report different results? Much of the answer lies in the issue of study design.

⁵⁴ Cotropia & Mann, *supra* note 14, at 1099; *see also* Anderson & Menell, *supra* note 25, at 6 (showing the falling reversal rates since the *Phillips* decision).

⁵⁵ *See, e.g.*, Schwartz, *supra* note 28, at 1098–1101.

⁵⁶ *See, e.g.*, Jason Rantanen & Lee Petherbridge, *Inequitable Conduct and Misuse*, in RESEARCH HANDBOOK, *supra* note 5.

⁵⁷ For a discussion of issues that commonly arise in cross-study comparisons, *see* HUNTER ET. AL., *supra* note 17, at 34–71.

II. THE CAUSES OF INTER-STUDY VARIABILITY

A. *Selection Effects and Studies of Appellate Decisions*

When studying judicial opinions, particularly those of an appellate court such as the Federal Circuit, the standard approach is to collect data on what is typically described as the “entire population.” By this the authors usually mean all of the court’s decisions or written opinions.⁵⁸ There are several reasons for this approach. First, and perhaps most importantly, it is a task that is feasible. It is not a Sisyphean task to collect the population of all relevant decisions. This is not to say that collecting such a population does not require a substantial amount of work; unquestionably it does. But because it is possible to collect the entire population, scholars of the Federal Circuit typically do so. Second, because scholars can collect the population, use of this approach offers greater authority to information reported in the study.⁵⁹ The dataset *is* the population, and there is no need to rely on sampling techniques and the additional limitations of those techniques.⁶⁰

Importantly, however, collection of the “population” does *not* mean that selection effects can be disregarded in empirical studies of judicial opinions. First, and most relevant to this Article, the data collected and reported on by a study can be affected by choices in study design. A classic example of study design selection comes from the claim construction reversal rate literature cited above. Early post-*Markman* studies focused only on the court’s written opinions: that is, opinions that resulted in a substantive discussion of the outcome.⁶¹ The primary reason for this was that such opinions were relatively clear as to whether they addressed claim construction as an issue. As discussed in more depth below, however, this study design choice had the effect of excluding a set of decisions in which the Federal Circuit affirmed lower courts: its set of summary affirmances under Federal Circuit Rule 36. Rule 36 allows the court to “enter a judgment of affirmance without opinion . . . when it determines that any of [several] conditions exist and an opinion would have no precedential value.”⁶² By

⁵⁸ See, e.g., John R. Allison & Mark A. Lemley, *Empirical Evidence on the Validity of Litigated Patents*, 26 AIPLA Q.J. 185, 235 (1998).

⁵⁹ *Id.*

⁶⁰ Although, for a variety of reasons, scholars will often treat the population as a sample of a super-population for purposes of inferential statistics. See John R. Allison et al., *Our Divided Patent System*, 82 U. CHI. L. REV. 1073, 1092 n.57 (2015).

⁶¹ See, e.g., John R. Allison & Mark A. Lemley, *How Federal Circuit Judges Vote in Patent Cases*, 27 FLA. ST. U. L. REV. 745, 746 (2000).

⁶² FED. CIR. R. 36. The conditions are: (a) the judgment, decision, or order of the trial court appealed from is based on findings that are not clearly erroneous; (b) the evidence supporting the jury’s verdict is sufficient; (c) the record supports summary judgment, directed verdict, or judgment on the pleadings; (d) the decision of an administrative agency warrants affirmance under the standard of review in the statute authorizing the petition for review; or (e) a judgment or decision has been entered without an error of law. *Id.*

systematically excluding these affirmances through the way the studies were designed (many of which involved claim construction disputes),⁶³ the result was a reversal rate⁶⁴ that was higher than the rate in studies that *did* account for Rule 36 affirmances, which are generally understood to better reflect the “true” reversal rate.⁶⁵

A related, but distinct, selection issue is that of population biases: that is, even if a study is careful to capture an entire population and to limit its conclusions to just that population, it is still possible that underlying structural elements of the population could affect information observed about that population.⁶⁶ When I say population biases, I’m referring to the idea that even when the population is viewed as a single whole (i.e., not as a sample of a larger population), it is affected by hidden characteristics that can affect its composition and impact the results of any study that reports statistical data about the population. In other words, even if the population is treated *as a population*, it is still necessary to account for how that population came to be in the first place when interpreting study results. Because population biases affect the underlying data source itself, however, they do not inherently produce differences in study results. Consequently, they are beyond the scope of this Article.

One could argue that the distinction between study-design selection issues and population biases is not meaningful because a population of appellate opinions is not really a population; it is actually a sample from a larger set of legal disputes (all appealable decisions of the lower tribunal) that itself is a sample from a yet larger set of legal disputes (all litigated disputes). Thus, the very act of conducting a study on appellate opinions is, effectively, a study design decision that results in sampling—and nonrandom sampling at that. The issue of Rule 36 summary affirmances illustrates this point: one might define the population as being only written opinions or one might define the population as being all decisions of the court. If the population is defined as being all written opinions, exclusion of Rule 36 affirmances could be described as a population bias. On the other hand, if the population is defined as being all decisions of the Federal

⁶³ Moore, *supra* note 53, at 234–36.

⁶⁴ *Id.* at 236 (observing a difference between a reversal rate of 34.5% and 40.8%, depending on whether Rule 36 summary affirmances were included or excluded).

⁶⁵ *Id.*; see Schwartz, *supra* note 53, at 238; see also Miller, *supra* note 53, at 816 (stating that Moore (2001, 2005) and Schwartz (2008, 2010) provide the most accurate measures of claim construction reversal rates because their studies only analyze Rule 36 summary affirmances).

⁶⁶ An example of such a study is the Priest-Klein hypothesis and other more modern examinations of population selection in judicial opinions. See, e.g., Daniel M. Klerman & Yoon-Ho Alex Lee, *Inferences from Litigated Cases*, 43 J. LEGAL STUD. 209 (2014) (explaining the Priest-Klein hypothesis, which argues that the end result of cases is equal for either side, regardless of the legal standard favoring one side); George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1, 5 (1984) (explaining the Priest-Klein hypothesis). For a discussion of population biases, see Schwartz, *supra* note 53, at 241–45.

Circuit, the exclusion of Rule 36 affirmances would likely be articulated as a study-design selection effect.

Nevertheless, it is conceptually useful to draw a distinction between study-design selection effects and population biases. First, when studying the entire set of appellate opinions, we are not talking about a dataset that can be corrected through better sampling practices: what you have is what there is. In other words, there are immutable characteristics of the population of appellate opinions that we cannot change through methodology alone (although perhaps we can account for them through multivariate regression techniques). Furthermore, although the set of appellate decisions can be viewed as a sample of a larger population of all litigated disputes, that is not usually how a study treats it; rather, empirical studies of Federal Circuit decisions tend to draw inferences within the context of how the Federal Circuit resolves this particular issue, *not* in terms of how all litigated disputes play out. Thus, the “population” really is being treated like a “population” and not as a sample of the larger set of all litigated disputes. Third, population biases do not produce inter-study variability as long as the same characteristic is being measured; study-design selection does. This distinction matters for purposes of interpretation because it means that even when two studies follow the same methodology and report the same results, those results could be affected by population bias effects.

Finally, researchers will occasionally note the potential for selection-effect biases within their methodologies but then conclude that these biases effectively cancel each other out.⁶⁷ Without knowing the magnitude of the biases, however, such an assumption is unfounded. Worse, biases that might on their surface seem to cancel each other out can sometimes operate in a synergistic way, resulting in more bias rather than less.⁶⁸ Consequently, researchers should be careful not to assume that two biases that seem to operate in opposite directions will “cancel” each other out.

B. *Choices in Study Design*

Below, I explore common potential sources of intra-study variation introduced by choices in study design. Although I discuss them in the context of the Federal Circuit, most of these choices affect studies of appellate courts generally, and may also impact studies of trial courts. This section begins at the level of document identification and collection before

⁶⁷ See, e.g., Christian E. Mammen, *Controlling the “Plague”: Reforming the Doctrine of Inequitable Conduct*, 24 BERKELEY TECH. L.J. 1329, 1348–50 (2009) (explaining the methodology of the study and the problems that the methodology faced in the collection of data).

⁶⁸ Jason Rantanen, *Recalibrating Our Empirical Understanding of Inequitable Conduct*, 3 IP THEORY 98, 104–05 (2013). Another example is the issue of record unit choices. Some record unit choices carry the potential to give both too much and too little weight to individual determinations by the court. That some empiricists give greater weight to some determinates and less weight to others does not mean the biasing disappears.

exploring the decisions researchers make when translating those documents into data that can be analyzed through empirical techniques.

1. *Data Source*

The main sources for the Federal Circuit decisions used in most studies are the Westlaw and Lexis databases.⁶⁹ These databases are vast and are generally available to most legal scholars. Generally, these databases are highly complete in terms of appellate opinions,⁷⁰ containing every precedential, and many non-precedential, written opinions. There are some limitations with respect to non-precedential opinions, however. Prior to 2001, non-precedential opinions were not published in an official reporter and many were not released for electronic distribution.⁷¹ While West Publishing began publishing tables of cases “disposed of by unpublished order or memorandum” in the 1970s, those tables contain very little information about the dispositions.⁷² The incompleteness of non-precedential opinions applies to the Federal Circuit: while many non-precedential opinions are available from the pre-2001 period, through Westlaw for example,⁷³ some are reported only in minimal form as table dispositions.⁷⁴ Thus, while Westlaw and Lexis may be the most complete electronic sources of Federal Circuit decisions, there are limits in a researcher’s ability to obtain pre-2001 non-precedential decisions from those sources.⁷⁵

⁶⁹ See *supra* note 6 (citing studies).

⁷⁰ But see Schwartz, *supra* note 53, at 269 n.221 (observing minor instabilities in a Lexis database).

⁷¹ See Michael Hannon, *A Closer Look at Unpublished Opinions in the United States Courts of Appeals*, 3 J. APP. PRAC. & PROCESS 199, 212 (2001); Sarah E. Ricks, *The Perils of Unpublished Non-Precedential Federal Appellate Opinions: A Case Study of the Substantive Due Process State-Created Danger Doctrine in One Circuit*, 81 WASH. L. REV. 217, 221 (2006); Thomas Reuters, *Federal Appendix (National Reporter System)*, <http://legalsolutions.thomsonreuters.com/law-products/Reporters/Federal-Appendix-National-Reporter-System/p/100000796> [https://perma.cc/59US-2AXM] (explaining that the Federal Appendix was first published in 2001).

⁷² Hannon, *supra* note 71, at 210. The extent to which these tables include rulings on non-dispositive motions is also unclear.

⁷³ *In re Chavez*, 899 F.2d 1229 (Fed. Cir. 1990). The opinion is also available as 1990 WL 33122. The text, however, is not available on Lexis. See *In re Chavez*, No. 89-1746, 1990 U.S. App. LEXIS 4368 (Fed. Cir. Mar. 27, 1990). Other unpublished opinions are available, though. See, e.g., *Am. Mobilphone, Inc. v. Am. Paging, Inc.*, No. 90-1205, 1990 U.S. App. LEXIS 19966 (Fed. Cir. Nov. 14, 1990); *Am. Mobilphone, Inc. v. Am. Paging, Inc.*, No. 90-1205, 1990 WL 177458 (Fed. Cir. Nov. 14, 1990).

⁷⁴ See, e.g., *In re Beaudet*, 795 F.2d 1013 (Fed. Cir. Apr. 16, 1986) (Table). While sometimes the basis for affirmances is indicated as Federal Circuit Rule 36, in others (such as *Beaudet*), no basis is available. This may be because what is available in Westlaw is all that the court issued, or it may be because Westlaw does not possess the additional documents. However, given that the “Opinion” field in *Beaudet* says “Remanded for Clarification,” a reasonable inference is that there was a more explicit opinion that explained what clarification was being sought. *Id.*

⁷⁵ This Article is not intended to be a comprehensive examination of the completeness of Westlaw or Lexis as data sources for non-precedential Federal Circuit opinions. However, based on exploratory

Sources besides Westlaw and Lexis are sometimes used. I have collected decisions from the Federal Circuit's website directly.⁷⁶ One drawback of opinions collected directly from the Federal Circuit's website is that these opinions are not yet published and could potentially be changed by the court prior to official publication.⁷⁷ A benefit of collecting from the Federal Circuit directly is that it provides a public domain version of each opinion,⁷⁸ which can then be archived and searched. However, the Federal Circuit's own website is temporally limited.⁷⁹ Written opinions on the Federal Circuit's website begin in 2004,⁸⁰ and the first Rule 36 summary affirmance appears in 2007.⁸¹

Another possible data source is Public Access to Court Electronic Records (PACER).⁸² While PACER is the most complete source possible for recent time periods (since it is linked to the electronic docketing system used by courts), using PACER directly can be extraordinarily expensive and time-consuming due to its per-page cost and essentially nonexistent search capabilities.⁸³ While the Administrative Office of the United States Courts may give researchers no-cost access to PACER, that access typically comes with substantial restrictions on the researcher's use of the database.⁸⁴ Furthermore, as discussed below, PACER is limited in terms of historical records and contains virtually no pre-21st century records. Instead of PACER, contemporary empirical studies of patent litigation may use an intermediary, such as LexMachina or Docket Navigator.⁸⁵ These sources, however, do not presently contain Federal Circuit decisions.

Researchers who may lack access to Westlaw or Lexis or who wish to examine a relatively long time period may utilize the United States Patent

searching I conducted while researching this article, it appears that Westlaw, at least, contains a substantial collection of the Federal Circuit's unpublished opinions dating back to 1987 or earlier.

⁷⁶ Jason Rantanen & Lee Petherbridge, *Disuniformity*, 66 FLA. L. REV. 2007, 2017–18 (2014).

⁷⁷ See Richard J. Lazarus, *The (Non)Finality of Supreme Court Opinions*, 128 HARV. L. REV. 540, 547–55 (2014) (discussing the history of revisions occurring even in precedential (and important) Supreme Court opinions).

⁷⁸ See 17 U.S.C. § 105 (2012) (“Copyright protection under this title is not available for any work of the United States Government.”).

⁷⁹ Rantanen & Petherbridge, *supra* note 76.

⁸⁰ *Vaizburd v. United States*, 384 F.3d 1278 (Fed. Cir. 2004).

⁸¹ *Oilfield Equip. Mktg., Inc. v. New Tech. Sys.*, 227 Fed. App'x 925 (Fed. Cir. 2007).

⁸² *Homepage*, PACER, <https://www.pacer.gov> [<https://perma.cc/B3QA-KWZFT2CJ-6683>] (last visited Sept. 1, 2016).

⁸³ See Steve Schultze, *Making Excuses for Fees on Electronic Public Records*, FREEDOM TO TINKER (Feb. 7, 2013), <https://freedom-to-tinker.com/blog/sjs/making-excuses-for-fees-on-electronic-public-records> [<https://perma.cc/76TF-GF8G>] (reviewing the most popular reasons made by the Administrative Office for why PACER implements fees).

⁸⁴ See, e.g., *Electronic Public Access Fee Schedule*, PACER, https://www.pacer.gov/documents/epa_feesched.pdf [<https://perma.cc/YB9Y-NUQDJS2VA-DZ45>] (describing limitations on exemptions from PACER fees).

⁸⁵ See, e.g., John R. Allison et al., *Understanding the Realities of Modern Patent Litigation*, 92 TEX. L. REV. 1769, 1772–73 (2014) (listing the benefits of using LexMachina as a primary data source).

Quarterly (U.S.P.Q.). The U.S.P.Q. is a publication that contains intellectual property cases, including patent opinions issuing from the district and appellate courts. Its primary advantage is that it includes opinions beginning in 2004.⁸⁶ Several studies use the patent cases cited in the U.S.P.Q. as their population.⁸⁷

The U.S.P.Q., however, comes with its own limitations that can potentially affect both the interpretation of studies relying on its opinions and cross-study comparisons. First, the U.S.P.Q. does not contain every patent-related opinion of the Federal Circuit or district courts. For example, an examination of Federal Circuit decisions issued in January 2000 and available in Westlaw reveals three substantive written opinions in appeals arising from infringement actions in the district courts that were not reported in the U.S.P.Q.⁸⁸ Nor does the U.S.P.Q. include summary affirmances, a study-design issue discussed further below. In addition, the cases included in the U.S.P.Q. likely exhibit the selection preferences of the reporter's editors. Henry and Turner note that the U.S.P.Q. contains cases "deemed, by legal editors, to be potentially precedential or to include noteworthy fact patterns,"⁸⁹ criteria not denied by the current editors.⁹⁰ This may mean, for

⁸⁶ See, e.g., *Vaizburd v. United States*, 384 F.3d 1278 (Fed. Cir. 2004).

⁸⁷ Lii, *supra* note 46, at 41–42; Miller & Curry, *Experts Judging Experts*, *supra* note 38, at 62; Matthew D. Henry et al., *Dynamics of Patent Precedent and Enforcement: An Introduction to The UGA Patent Litigation Datafile 7* (Nat'l Sci. Found., Working Paper, 2013); see also Jay P. Kesan & Gwendolyn G. Ball, *How Are Patent Cases Resolved? An Empirical Examination of the Adjudication and Settlement of Patent Disputes*, 84 WASH. U. L. REV. 237, 250 (2006) (stating that the U.S.P.Q. is a main source of data used for empirical research on patent litigation).

⁸⁸ *NFA Corp. v. Asheboro Elastics Corp.*, 230 F.3d 1377 (Fed. Cir. 2000) (unpublished); *Rivera-Davila v. Asset Conservation, Inc.*, 230 F.3d 1378 (Fed. Cir. 2000) (unpublished); *Cellnet Data Sys., Inc. v. Itron, Inc.*, 230 F.3d 1379 (Fed. Cir. 2000) (unpublished). During that month, the Federal Circuit also issued three decisions in infringement actions that were published in the U.S.P.Q. A full comparison of the U.S.P.Q. to Westlaw or Lexis or other sources is beyond the scope of this Article; the point is that the U.S.P.Q. is not a comprehensive collection of appellate decisions on patent law. A better characterization would be that it is a collection of appellate decisions that the editors deem to be important to the development of patent law.

⁸⁹ Henry et al., *supra* note 87, at 7. Henry and Turner observe that "[v]irtually every appellate court decision is recorded in the USPQ, as well as a large sample of district court decisions," but do not cite any testing or source for this observation. *Id.* Based on the comparison described above, however, as well as an examination of the U.S.P.Q. itself, I am skeptical of the completeness of the U.S.P.Q. for use as a data source for measuring outcomes.

⁹⁰ My U.S.P.Q. inquiry about methodology received the following response:

As a general principle, we strive to publish decisions in the U.S. Patents Quarterly that advance or clarify the law of intellectual property in some way. Such cases come from three main sources: the federal courts of appeal (including the Supreme Court), the federal district courts, and the U.S. Patent and Trademark Office.

We publish all precedential opinions from the federal appeals courts and the PTO. The balance of opinions published in PQ are federal district court cases, and, less frequently, decisions from state courts (mostly trade secrets and right-of-publicity cases), that are viewed as newsworthy, especially those that have been reported in BBNA's news publications such as the Patent, Trademark, and Copyright Journal. In

example, that certain subject-matter areas are more or less represented, that there is a weighting towards Federal Circuit decisions (since these decisions carry more judicial weight) as opposed to district court decisions, and that because the most “potentially precedential” cases are selected for inclusion, the outcomes may be biased in the sense that the U.S.P.Q. exclude decisions that are mundane and not precedent-setting in the editors’ opinions. Despite (and keeping) these limitations in mind, however, the U.S.P.Q. may be the best dataset available for studying longitudinal shifts in patent decisions.

A common issue that arises with most sources of judicial decisions is that older data tend to be less complete. This is less of a problem when studying only published appellate opinions, as these opinions are by their nature published and widely available. Difficulties arise, however, with unpublished opinions, or worse, summary dispositions—i.e., those that are not accompanied by a written opinion at all. These summary dispositions, such as those made under the Federal Circuit’s Rule 36 summary affirmance mechanism, may be underreported for this reason. No researcher to date has conducted a formal comparison of the contents of Westlaw or Lexis to the Federal Circuit’s own records to determine the completeness of those data sources with respect to Rule 36 summary affirmances. Based on my own searches,⁹¹ though, I can confirm that Westlaw contains Rule 36 summary affirmances dating back to 1989,⁹² the year in which the rule was adopted.⁹³

Beyond the difficulty in finding Rule 36 summary affirmances is the problem that they contain much less information than written opinions—for example, there is no identification of what issues are being decided or the procedural posture of the appeal.⁹⁴ Most researchers who include Rule 36 affirmances address this concern by reviewing the parties’ briefs to

publishing our weekly issues/updates, we also try whenever possible to balance the number of cases from the main intellectual property subject areas (patents, copyrights, and trademarks and unfair competition).

Decisions that are not generally published include nonprecedential opinions from the federal district courts, appellate courts, and the PTO, as well as opinions that resolve routine matters of procedure, jurisdiction, or discovery unrelated to intellectual property issues.

Email from Jackie McCloud, Law School Relationship Manager, Bloomberg BNA, to Theodore A. Potter, Reference Librarian, University of Iowa Law Library (July 9, 2015, 4:36 PM) (quoting Bill McKey, Managing Editor, U.S.P.Q. & MEDIA L. REP.) (on file with author).

⁹¹ Using “Fed.Cir. R.36”, “Fed.Cir. Rule 36”, and “Fed.Cir.R. 36” as search terms.

⁹² Specifically, the oldest affirmance under Fed. Cir. R. 36 that I have observed is *Sublet v. U.S. Postal Serv.*, 883 F.2d 1026 (Fed. Cir. 1989) (unpublished).

⁹³ Donald R. Dunner et al., *A Statistical Look at the Federal Circuit’s Patent Decisions: 1982-1994*, 5 FED. CIR. B.J. 151, 153 n.16 (1995).

⁹⁴ See, e.g., *Rates Tech., Inc. v. Mediatrix Telecom, Inc.*, 688 F.3d 742, 750–51 (2012); 251 F.3d 170 (table of Rule 36 affirmances) (Dec. 8, 2000). This problem also arises with unpublished dispositions for which no written opinion is available in Westlaw or Lexis. Based on my exploratory searches in Westlaw, this affects all such dispositions prior to April 13, 1987.

determine which issues were the subject of the appeal—a strategy that works as long as the briefs are actually available.⁹⁵ Researchers may look up the district court opinions involved in the appeal, or contact the parties directly.⁹⁶ At least one study author has attempted to estimate the extent to which Rule 36 affirmances apply,⁹⁷ but there is reason to be cautious about the empirical foundation for such an estimate.⁹⁸

There are also potential artifacts resulting from the digitization of federal litigation practice over the first decade of the twenty-first century. District court records, in particular, can present this issue because different districts implemented electronic filing at different points in time. The result is that electronic records for some districts are available from earlier periods than other districts.⁹⁹

When comparing study results drawn from sources such as Westlaw and Lexis, it is also important to keep in mind that because these databases are designed for practitioners and judges, rather than assembled for the use of academic researchers, the contents of those databases can change over time. Generally, this is because additional older records are added, resulting in a dataset that is more comprehensive. While a more comprehensive database is generally a good thing, it can result in variation among studies conducted at different times, even when they otherwise employ identical methodologies.¹⁰⁰

A recent problem is that of obsolescence. As an extreme example of this concern, in the fall of 2014, the Administrative Office of the United States Courts removed several circuits from the PACER database.¹⁰¹ Among the

⁹⁵ Moore, *supra* note 53.

⁹⁶ Schwartz, *supra* note 53, at 238–39.

⁹⁷ Field, *supra* note 37, at 736–37.

⁹⁸ Field, *supra* note 37, uses an assumption of even distribution, but this assumption is questionable given that Moore, *supra* note 53, observed that Rule 36 affirmances were less likely in claim construction appeals than on other issues.

⁹⁹ For example, the Western District of Missouri was one of a handful of pilot districts to implement electronic case filing as early as 1999. Tom O'Connor, *E-filing Projects Around the Nation*, 1 NO. 1 E-FILING REP. 9 (2000). As of September 2002, CM/ECF systems were in use in ten federal district courts. William A. Fenwick & Robert D. Brownstone, *Electronic Filing: What Is It? What Are Its Implications?*, 19 SANTA CLARA COMPUTER & HIGH TECH. L.J. 181, 194 (2002). Arizona implemented its ECF system in August 2005. Mohyeddin K. Abdulaziz, *Slowly But Surely?*, 7 NO. 8 E-FILING REP. 1 (2006). It appears that the Western District of Wisconsin was one of the last districts to implement the CM/ECF system, beginning their implementation of the system in May of 2006. *Case Management/Electronic Case Files (CM/ECF) Implementation Schedule*, ADMIN. OFFICE OF THE U.S. COURTS (Aug. 1, 2006), https://www.pacer.gov/documents/cmecf_impl.pdf [<https://perma.cc/M3B5-ZZVY>] (last visited Oct. 13, 2016).

¹⁰⁰ These changes can operate at both a macro level as well as a micro level. See Schwartz, *supra* note 53, at 269 n.221 (noting minor instability issues when searches of LexisNexis were run just a week apart).

¹⁰¹ See *Changes to Information Available On Pacer*, ADMIN. OFFICE OF THE U.S. COURTS, <https://www.pacer.gov/announcements/general/webpacer.html> [<https://perma.cc/BB7D-L7MA>] [copy archived as of Sept. 3, 2014] (requiring a PACER login to view restored access for several circuits); Joe

data removed were all cases filed prior to March 1, 2012 at the Federal Circuit.¹⁰² While the documents supposedly remained available at the individual courts in hard copy form,¹⁰³ the review and access to such records is much more limited than the records in electronic form. In addition, it is obviously not possible to conduct electronic searches on physical copies of documents. After substantial public outcry, electronic access to the documents was restored.¹⁰⁴ Given the uncertainty inherent in future access to documents, however, researchers may need to assemble their own archives of the documents to the extent permissible while respecting the intellectual property rights of relevant stakeholders.¹⁰⁵

2. Time Period

A related concern is the choice of time period. As the data in Part I indicate, substantial temporal variation can exist for a given metric.¹⁰⁶ All studies of the Federal Circuit report their temporal components; thus, generally speaking, the most important thing to do when interpreting a study is to do so within the context of its time period. But there is one temporal concern to pay particular attention to: short study periods. Because there is so much temporal variability in these results, perhaps simply as a result of high amounts of random chance, empiricists should avoid study designs that use only a very narrow slice of time due to the risk of capturing an

Mullin, *US Courts Trash a Decade's Worth of Online Documents, Shrug It Off*, ARS TECHNICA (Aug. 26, 2014, 5:45 PM), <http://arstechnica.com/tech-policy/2014/08/us-courts-trash-a-decades-worth-of-documents-shrug-it-off> [<https://perma.cc/79A4-EEA6>] (noting that certain federal courts were removed from the PACER database due to incompatibility with the new PACER system); see also Andrea Peterson, *Online Court Archive PACER Says It Will Restore Access to Missing Records*, WASH. POST (Sept. 19, 2014), <https://www.washingtonpost.com/news/the-switch/wp/2014/09/19/online-court-archive-pacer-says-it-will-restore-access-to-missing-records/> [<https://perma.cc/W2GU-RKKK>] (reporting that the Administrative Office will restore online access to documents removed from PACER).

¹⁰² Mullin, *supra* note 101.

¹⁰³ See Vera Eidelman & Amul Kalia, *Right to Know: The PACER Mess and How to Clean It*, ELEC. FRONTIER FOUND. (Sept. 2, 2014), <https://www.eff.org/deeplinks/2014/09/right-know-pacer-mess-and-how-clean-it> [<https://perma.cc/9FDT-ZPDZ>] (stating that PACER claims hard copies of deleted records can be obtained directly from the court); Andrea Peterson, *Why PACER Removed Access to Case Archives of Five Courts*, WASH. POST (Aug. 26, 2014), <https://www.washingtonpost.com/news/the-switch/wp/2014/08/26/why-pacer-removed-access-to-case-archives-of-five-courts/> [<https://perma.cc/P8BN-VYZZ>]

¹⁰⁴ See, e.g., Mullin, *supra* note 101; Andrea Peterson, *Senate Judiciary Committee Chairman Urges PACER To Restore Access To Removed Case Archives*, WASH. POST (Sept. 12, 2014), <https://www.washingtonpost.com/blogs/the-switch/wp/2014/09/12/senate-judiciary-chairman-urges-pacer-to-restore-access-to-removed-case-archives> [<https://perma.cc/AZZ4-BLP3>] (referencing an Administrative Office spokesman's implication that online access will be restored once changes to PACER are fully implemented).

¹⁰⁵ For a discussion of issues involved in assembling databases of legal documents, see *White v. West Publ'g Corp.*, No. 12 Civ. 1340 (JSR), 2012 U.S. Dist. LEXIS 69463 (S.D.N.Y. Mar. 28, 2012) and *White v. West Publ'g Corp.*, 29 F. Supp. 3d 396 (S.D.N.Y. 2014).

¹⁰⁶ See *supra* Figures 1, 2 (showing substantial variation in reversal rates).

atypical time period through random chance alone.¹⁰⁷

3. Collection and Filtering Methodology

After a data source is identified, relevant records must be collected. Simply collecting all records from a source such as, say, Westlaw is impossible. There are just too many. Even a topically assembled data source, such as the set of intellectual property-related opinions in the U.S.P.Q., must be winnowed down to only the relevant records for analysis. Although data collection and filtering for some historical studies, such as P.J. Federico's 1956 study of adjudicated patents,¹⁰⁸ was done by hand, modern studies typically employ a combination of electronic and manual methodologies.

Due to the size and complexity of databases such as Westlaw, studies that draw upon this database necessarily use some form of keyword or algorithm-based search. While algorithm searches of these databases offer seemingly easy replicability, this advantage is limited because both Westlaw and Lexis change their interface periodically, such as in 2010 for Westlaw¹⁰⁹ and in 2011 for Lexis.¹¹⁰ Changes to the interfaces can make reliance on methodologies that depend on those interfaces problematic.¹¹¹ Recent changes to the interfaces of Westlaw and Lexis that resulted in the elimination of some functionality and the inability to use certain forms of Boolean searches that were previously available is one example of a problematic change. For example, Westlaw "Classic" was organized by database;¹¹² WestlawNext is organized by content categories.¹¹³ Search algorithms in Westlaw Classic could include different databases as

¹⁰⁷ See *supra* text accompanying note 31 (discussing empirical findings which show the Federal Circuit's higher reversal rate than other circuit courts). Other examples of short study time periods include: W. Thad Adams, III & J. Derel Monteith, Jr., *The Continuing Saga of Federal Circuit Patent Claim Construction Jurisprudence: Extrinsic Evidence and Other Stories*, 8 FED. CIR. B.J. 83 (1999) (studying claim construction reversal rates for 1998 and a portion of 1999); Michael Saunders, *A Survey of Post-Phillips Claim Construction Cases*, 22 BERKELEY TECH. L.J. 215, 232–35 (2007) (examining studies focused on claim construction reversal rates over a multi-year period); Zidel, *supra* note 53, at 747 (reporting the reversal rate on claim construction for 2001).

¹⁰⁸ *Adjudicated Patents, 1948–54*, 38 J. PAT. OFF. SOC'Y 233, 234 (1956).

¹⁰⁹ Paula J. Hane, *Thomas Reuters Launches WestlawNext—The Next Chapter in Legal Research*, INFO. TODAY (Feb. 4, 2010), <http://newsbreaks.infotoday.com/NewsBreaks/Thomson-Reuters-Launches-WestlawNext-The-Next-Chapter-in-Legal-Research-60975.asp> [<https://perma.cc/85KR-X5N9>].

¹¹⁰ Press Release, LexisNexis Launches New Release of Lexis Advance, LEXISNEXIS (Dec. 5, 2011), <http://www.lexisnexis.com/en-us/about-us/media/press-release.page?id=1323111249773407> [<https://perma.cc/TT7A-5F2U>].

¹¹¹ Another example is the change in how LexisNexis categorized opinions of the Federal Circuit observed in the studies reviewed. In 2001, Christian Chu used the "US Court of Appeals Cases – Federal Circuit" database in Lexis. Chu, *supra* note 53, at 1091 n.77. Seven years later, David Schwartz used the LexisNexis "CAFC" database. Schwartz, *supra* note 53, at 269 n.221.

¹¹² WESTLAWNEXT QUICK REFERENCE GUIDE: COMPARING WESTLAW CLASSIC AND WESTLAWNEXT 2 (2013) [hereinafter WESTLAWNEXT QUICK REFERENCE GUIDE].

¹¹³ WESTLAWNEXT QUICK REFERENCE GUIDE (on file with author).

modifiers;¹¹⁴ search procedures in WestlawNext use a different procedure that may or may not produce identical results. Studies that relied upon those search terms in order to provide reproducibility of their data collections are thus no longer directly replicable.

One way to address this problem is to ensure that a complete study population is collected at the time of the study, rather than relying on the continuing viability of a particular interface. Thus, preferably, all cases in the study population would be downloaded into an archive so that the entire archive could be reviewed by future researchers interested in replicating the study. More preferably, a centralized archive would be developed to minimize the duplication of opinion downloading. Contractual limitations imposed by services such as Westlaw and Lexis, however, limit extensive collection of decisions and specifically prohibit copying these records for archival purposes.¹¹⁵

The precise choice of keyword searches when filtering can also affect the composition of the dataset. For example, one might begin searching for all opinions involving the doctrine of nonobviousness by searching for the term “obvious” with a root expander. The searcher would quickly realize, however, that this search is not practical; the word “obvious” appears in many opinions—even many patent opinions—that have nothing to do with the legal doctrine of nonobviousness. So instead, one might search for just the term “nonobvious!” But this search might run into the opposite problem: although many cases involving the legal doctrine of nonobviousness might use that term, it is possible that some author—for whatever reason—chose not to use that particular term.¹¹⁶ Ultimately, empiricists should pay careful attention to this issue, recording the collection and filtering methodologies they employ with the recognition that they may be sources of inter-study variability.

¹¹⁴ Lefstin, *supra* note 22, at 1053 n.123.

¹¹⁵ *Westlaw Subscriber Agreement (2016)*, WESTLAW (last visited Sept. 24, 2016), <https://lawschool.westlaw.com/marketing/display/mi/75> [https://perma.cc/FC73-EMMP] (“Downloaded Data shall not be stored or used in an archival database or other searchable database except as expressly permitted by this Agreement or as quoted in Subscriber’s work product.”); *see also Terms & Conditions for Use of the LexisNexis Services*, LEXISNEXIS (Dec. 9, 2011), <http://www.lexisnexis.com/terms/supplemental.aspx> [https://perma.cc/X6KZ-DQ5Q] (stating that information obtained on LexisNexis cannot be published, broadcasted, or sold for commercial purposes). It is possible for legal academics to discuss alternatives and exceptions to these general policies with the respective services, although a common restriction is that the large-scale data collection must be limited to academic uses. For a discussion of this issue, see Todd Vare & Michael Mattoli, *Big Business, Big Government and Big Legal Questions*, MANAGING INTELL. PROP. (Sept. 26, 2014), http://image.exct.net/lib/fefa1276756603/m/1/Vare_ManagingIP_October.pdf [https://perma.cc/3W6A-CZ4M].

¹¹⁶ *See, e.g., Siemens AG v. Seagate Tech.*, 369 Fed. App’x 118, 119 (Fed. Cir. 2010) (unpublished decision) (supporting the jury’s obviousness determination in favor of Seagate based on substantial evidence).

4. *Record Unit*

This study design choice, and those that follow, relate to the translation of documents into analyzable data.

A key issue in many empirical studies of the Federal Circuit is how the record unit is defined: that is, the level at which data will be recorded. A common initial approach is to decide to treat individual opinions as record units. The basic idea behind this approach is that each opinion constitutes a discrete, easily identifiable record. Information about that opinion—the record—is then recorded. An opinion, however, may not contain actual substantive analysis. When summary affirmances are involved, for example, an opinion contains nothing more than a one-line affirmance.

Due to the fact that, as standard parlance, lawyers often use the word “opinion” to refer to the court’s discussion of the substantive reasons for its decision, some ambiguity surrounds the terminology for the approach described in the preceding paragraph. Sometimes it is referred to as “per case,”¹¹⁷ which is confusing because a single lawsuit (colloquially, a “case”) might result in multiple appellate opinions.¹¹⁸ Other times, the terminology “per decision” is used. That, too, can be confusing because it could also refer to an individual decision *within* an opinion. To avoid confusion, this Article uses the terminology “per opinion,” recognizing that an “opinion” may be nothing more than a summary affirmance or table disposition.

While a per-opinion approach is generally adequate when performing outcome counts at the per-opinion level, it can present problems when the subject to be studied manifests itself at a more granular level. In patent cases, this is a regular issue, as many patent appeals involve multiple discrete issues that the court may analyze and resolve independently, and sometimes differently.¹¹⁹ As a result, some studies use individually identifiable analyses of, or decisions on, an issue as the record unit.¹²⁰ This approach, however, comes with its own challenges. In particular, deciding what constitutes an individual analysis can be difficult and, to some extent, subjective. The problem of figuring out just how thinly to slice the unit of analysis may be particularly likely to occur in studies that involve issues similar to those of nonobviousness and anticipation or novelty issues, as these doctrines tend to invite multiple grounds of attack (and thus analyses that must be addressed

¹¹⁷ See, e.g., Chu, *supra* note 53, at 18.

¹¹⁸ Schwartz, *supra* note 53, at 238.

¹¹⁹ Lemley, *The Fractioning of Patent Law*, in *INTELLECTUAL PROPERTY AND THE COMMON LAW* (Shyamkrishna Balganesh ed., 2013).

¹²⁰ See, e.g., Petherbridge & Wagner, *supra* note 10, at 2072 (“Because the core unit of measurement here was the way the Federal Circuit analyzes obviousness, each observable analysis of obviousness was a record in the database.”); Petherbridge et. al, *supra* note 40, at 1305 (“In most instances, a case is equivalent to a record entry.”); Rantanen, *supra* note 34, at 728 (“In many instances, a case is equivalent to a record entry.”).

by the court). This Article uses the terminology “per analysis” to refer to this type of record unit.

Some studies attempt to deal with the subjectivity inherent in coding at a per analysis level by treating only differing outcomes as individual units of analysis. For example, one study of nonobviousness classified all patent claims or patents within a judicial opinion that had the same outcome (obvious, nonobvious, or undecided) as a single decision.¹²¹ A drawback to this approach is that it effectively treats the court’s individual analyses in different and potentially biasing ways. For example, if there were five proposed theories of obviousness that the court rejected, and a sixth that the court accepted, then the result would be coded as one count of no obviousness and one count of obviousness, even though most of the analyses resulted in a conclusion of no obviousness. There is no perfect solution here. A researcher can reduce the amount of subjectivity involved in slicing the level at which the unit of analysis is to be measured, but that approach runs the risk of underweighting some sets of the court’s analyses.

A third way of defining the record unit is to operate at the level of a “patent case”—that is, the resolution of an individual patent in an appeal.¹²² Since patents may be analyzed and resolved differently by the court, this approach helps mitigate some of the problems with analyzing on a per-case level while avoiding the challenges of defining an individual analysis. A limitation of this approach is that, on the one hand, there may be multiple distinct issues analyzed by the court for a given patent (such as distinct obviousness arguments), while on the other, there may be multiple patents analyzed as a unit. The consequence is that the “patent case” approach carries the potential to overvalue or undervalue individual judicial analyses.

A fourth way of defining the record unit is specific to claim construction studies; it treats each claim term as the record unit.¹²³ This is a useful approach, as it is usually fairly clear what constitutes a particular claim term and because data, once collected, can usually be aggregated, thus allowing for per-case results to be reported as well as per-term results. Therefore, the “per-term” approach tends not to involve quite the same subjectivity involved in the “per-analysis” slicing discussed above. Several of the analyses of claim construction reversal rates described in Part I follow this approach.

¹²¹ See Jennifer Nock & Sreekar Gadde, *Raising the Bar for Nonobviousness: An Empirical Study of Federal Circuit Case Law Following KSR*, 20 FED. CIR. B.J. 369, 387 (2010) (“For the purposes of this study, a decision includes all patent claims within a case that result in the same outcome (obvious, nonobvious, or undecided) [If] there were multiple claims or patents at issue within a certain case, but all were found obvious, this was classified as a single decision [If], however, the court found some claims obvious and other claims nonobvious, each of the outcomes is treated as a separate decision, and analyzed separately.”).

¹²² Studies using the “patent case” approach include Cotropia, *supra* note 34, at 914.

¹²³ Examples of the “per-term” approach include Anderson & Menell, *supra* note 25, at 35–36, and Schwartz, *supra* note 53, at 238.

5. Parameter Definitions

Perhaps one of the most difficult study-design tasks is the creation of parameter definitions. Yet, these definitions are critical to many important aspects of a study, including its accuracy, reliability, replicability, and even its meaning.

For example, it turns out to matter a great deal how “reversal” is defined in reversal rate studies. One possible definition of “reversal” is that a decision is counted as such when the appellate court reverses and affirms the reviewed tribunal in part. Another possibility for the definition of “reversal” is that a decision only counts as a reversal when the appellate court reverses the lower tribunal in its entirety. Kesan and Ball’s work demonstrates how big a difference this can make: when defining reversals as the former, they found the Federal Circuit to reverse the district courts 41% of the time; when defining a reversal as the latter, they found the Federal Circuit to reverse the district courts 32% of the time.¹²⁴ The only difference between the two figures was the definition of reversal; the data itself was the same.

When resolving an appeal, an appellate court has several choices of outcome. The court can agree that the lower tribunal got the decision right—or, at least, not wrong, depending on the standard of review—and affirm the decision.¹²⁵ The court can reverse the lower tribunal, saying that its decision should have come out the other way.¹²⁶ It can also vacate the reviewed decision. Adding complexity, a court can do a combination: it can reverse and vacate a decision, or it can affirm in part and vacate in part.

Vacated decisions obviously pose a particular challenge for researchers interested in studying reversal rates, but they also complicate deeper analyses of legal doctrines. Sometimes, a “vacate” means that the decision

¹²⁴ Kesan & Ball, *supra* note 9. Kesan and Ball’s results are even more impressive when one takes into consideration that their data is reported on a per-subject matter basis, as opposed to the opinion as a whole. *Id.* at 434. When the opinion as a whole—and the possibility that it involves multiple subject matters—is taken into account, the difference is even greater, as discussed *supra* in Part III.

¹²⁵ On some issues, typically questions of law, the appellate court reviews the lower tribunal without deference. *See, e.g.,* *DePuy Spine, Inc. v. Medtronic Sofamor Danek, Inc.*, 567 F.3d 1314, 1323 (Fed. Cir. 2009) (stating that the ensnarement is decided as a question of law and reviewed *de novo*). On other issues, usually involving issues of fact or a judge’s exercise of her discretion, the appellate tribunal reviews with deference to the judge or other factfinder. *See id.* at 1324 (“[W]e review the district court’s resolution of those factual issues for clear error.”) (citation omitted). When the appellate court is reviewing without deference, an affirmance means that the appellate court itself concludes that the district court’s outcome was correct; when reviewing with deference, however, the appellate court affirm even if it would not have reached that result had it been operating on a clean slate.

¹²⁶ ROBERT E. LARSEN, NAVIGATING THE FEDERAL TRIAL § 14:3 (2015 ed.) (“If an appellate court wants to reverse a lower tribunal, it characterizes the issue as a mixed issue of law and fact, thereby allowing *de novo* review.”) (quoting Kelly Kunsch, *Standards of Review (State and Federal): A Primer*, 18 SEATTLE U. L. REV. 11, 12 (1994)).

of the district court disappears, and the court is told to try again.¹²⁷ Other times, a “vacate” occurs when the appellate court rejects early dismissal of the case and instructs the district court to keep going.¹²⁸ In both situations, however, the appellate decision is merely a temporary resting point, with the case heading back to the lower tribunal for additional proceedings.

Most patent law scholars who study substantive doctrinal issues tend to parse out appellate decisions to vacate and instead focus on the final decisions of the court: those decisions that effectively represent the end of the line for the case (subject to rehearing by the court en banc or further appeal to the Supreme Court).¹²⁹ This is because the final decisions are the ones where the court is making a decision one way or the other on a given patent or determinative issue. Regardless of which approach is taken, however, the effect of this study design choice should be kept in mind when interpreting a given study or comparing results across studies. As a general proposition, the effect of including non-final decisions produces a lower rate of a particular outcome, since the effect is to lower both the proportion of affirmances and the proportion of vacated decisions.¹³⁰ Or, if researchers are studying outcomes (such as decisions that a patent claim is or is not obvious), the effect of including non-final decisions will be to lower the apparent rate at which the court reaches a conclusion of *both* obvious and nonobvious.

One way to deal with the complexity of outcome categories is to record the procedural posture of the appeal. Procedural posture often matters on appeal—and so it can be important to develop parameters that capture that posture. In patent infringement appeals before the Federal Circuit, several procedural postures are common: jury trial, bench trial, grant of judgment as a matter of law (JMOL) following a jury trial, grant of summary judgment, and, increasingly, grant of a motion to dismiss pursuant to Fed. R. Civ. P. 12(b)(6). Different procedural postures can apply to different issues in a case, a problem that reinforces the benefit of defining record units at a level more granular than the appeal itself.

In addition to the procedural posture itself, it can be useful to record outcomes at the trial court (for example, whether a given set of patent claims is obvious or nonobvious). When combined with the procedural posture and the procedural outcome of the case, this can help researchers better

¹²⁷ See, e.g., *United States v. Microsoft Corp.*, 253 F.3d 34, 46–47 (D.C. Cir. 2001) (vacating and remanding the District Court’s remedial order in part because the District Court failed to hold an evidentiary hearing).

¹²⁸ See, e.g., *Phillips v. AWH Corp.*, 415 F.3d 1303, 1337 (Fed. Cir. 2005) (reversing the summary judgment of noninfringement based on disagreement with the district court’s claim construction, and remanding for further proceedings).

¹²⁹ See, e.g., Rantanen, *supra* note 34, at 723 (noting how content analysis differs from practices where scholars only focus on a small set of opinions that are deemed “important”).

¹³⁰ Sichelman, *supra* note 11.

understand what is actually going on in the data. A high reversal rate for grants of summary judgment of inequitable conduct, for example, might suggest that Federal Circuit is relatively hostile to assertions of this doctrine.¹³¹ I have called the combination of the procedural posture of the appeal and the decision made by the lower tribunal the “substrate” of the appeal, as it is the matter on which the appellate court acts.¹³²

6. Coding Issues

As a general practice of good study execution, coding should be performed according to an established methodology. Nonetheless, some errors are inevitable, particularly when human beings are involved.¹³³ Since there are works that discuss this issue in depth,¹³⁴ and coder errors are hardly unique to empirical studies of judicial opinions, this section provides only a general discussion of the issue and identifies specific types of coding errors that can arise when working with the content of judicial opinions.

Errors by coders typically take one of a few forms. Coders may have difficulty in ascertaining the “correct” answer to a particular piece of information. This difficulty sometimes is due to the coder’s lack of familiarity with the subject matter. For example, a coder without substantial knowledge of patent law who is tasked with recording whether cases involve the doctrine of nonobviousness might include Section 102 (anticipation/novelty) cases as nonobviousness cases simply due to a lack of familiarity with the legal doctrines.¹³⁵ Good coding instructions can help mitigate this issue to some extent, but instructions can only go so far.¹³⁶ Even

¹³¹ See Petherbridge et al., *supra* note 40, at 1315 (noting that the Federal Circuit holds a stricter standard for inequitable conduct than some lower courts, and that when reviewing summary judgment cases, the Federal Circuit does not find inequitable conduct 92.86% of the time).

¹³² See Rantanen, *supra* note 34, at 740 (describing the obviousness posture of the patents being reviewed as the substrate from which the Federal Circuit works).

¹³³ See, e.g., Richard O. Lempert, *The Significance of Statistical Significance: Two Authors Restate an Incontrovertible Caution. Why a Book?*, 34 LAW & SOC. INQUIRY 225, 227–28 (discussing common coding errors from economic studies in recent decades).

¹³⁴ See Hall & Wright, *supra* note 6, at 109–11 (discussing the challenges and errors inherent in the act of coding, as well as process of selecting and training coders).

¹³⁵ See Edwards & Livermore, *supra* note 6, at 1922–23, 1926 (noting coding mistakes based on the misclassification and erroneous classification of court decisions).

¹³⁶ Timothy Hagle observes this issue in characterizing litigants or votes as liberal or conservative in empirical studies of the Supreme Court. See Timothy M. Hagle, Response, *Two Worlds, Neither Perfect: A Comment on the Tension Between Legal and Empirical Studies*, 63 BUFF. L. REV. 379, 382 (2015) (“Researchers can take steps to avoid overly subjective coding or bias on the part of the coders, but this does not guarantee that all such problems are eliminated At the very least, we may still be left with differences of opinion as to whether the coding rules are appropriate or whether they have been correctly followed in particular instances.”); see also Todd E. Pettys, *Free Expression, In-Group Bias, and the Court’s Conservatives: A Critique of the Epstein-Parker-Segal Study*, 63 BUFF. L. REV. 1, 5 (2015) (discussing some of the issues and errors with the study by LEE EPSTEIN ET. AL., DO JUSTICES DEFEND THE SPEECH THEY HATE? IN-GROUP BIAS, OPPORTUNISM, AND THE FIRST AMENDMENT (2015), <http://epstein.wustl.edu/research/InGroupBias.pdf> [<https://perma.cc/LZU5-2LP7>]).

a coder with a relatively high degree of sophistication and a set of good instructions might find it difficult to maintain consistency across hundreds of judicial opinions when recording information about the court's legal reasoning in an opinion addressing nonobviousness. One solution is to identify particular linguistic features of the legal doctrine that can serve as a touchstone. In nonobviousness, for example, the "[teaching], suggestion or motivation" requirement was one such feature of the Federal Circuit's nonobviousness jurisprudence during the first half of the 2000s.¹³⁷ This touchstone can help orient empirical researchers when examining the content of particular legal doctrine.

This leads to a related problem: the more sophisticated or complex the information that the coder is asked to record, the more likely there is to be variability in the study results and the more difficult it is to replicate study results. Particularly sophisticated studies, such as Wagner and Petherbridge's study of claim construction methodologies, are likely to be challenging to replicate without the coder developing a deep understanding of the methodologies the authors hypothesized were involved in Federal Circuit claim construction analyses.¹³⁸ That it would be possible at all is because they included an extremely detailed explanation of the holistic and procedural approaches.¹³⁹

Humans are not perfect, and neither is their ability to record data. Humans can make mistakes even on relatively straightforward, simple issues. For example, a coder who is recording the names of the panel judges involved in a decision might make an error while coding, substituting one judge's name for another. No amount of coding instructions is likely to prevent these types of errors. Fortunately, particularly with simple parameters, researchers can catch and address these types of errors, and reliability statistics can help determine whether these errors are a significant problem.¹⁴⁰ Generally speaking, however, minor variations in study coding are to be expected as long as human beings are involved.

Human-coder errors can also be reduced by computer-based data collection. Many forms of basic case information, such as case name and lower tribunal, can be automatically collected. Some researchers have developed sophisticated algorithms to collect more complex information contained within opinions such as citations to scholarship,¹⁴¹ and the art has

¹³⁷ Cotropia, *supra* note 34, at 917; Rantanen, *supra* note 34; Wagner & Petherbridge, *supra* note 7.

¹³⁸ See Wagner & Petherbridge, *supra* note 7.

¹³⁹ *Id.*

¹⁴⁰ See, e.g., Lee Petherbridge & David L. Schwartz, *An Empirical Assessment of the Supreme Court's Use of Legal Scholarship*, 106 NW. U. L. REV. 995, 1003 n.25 (2012) (discussing coding reliability statistics using the Cohen's kappa (*k*) statistical argument).

¹⁴¹ See, e.g., David L. Schwartz & Lee Petherbridge, *Legal Scholarship and the United States Court of Appeals for the Federal Circuit: An Empirical Study of a National Circuit*, 26 BERKELEY TECH. L.J.

advanced to the point where it is possible to conduct more sophisticated analyses of the text of judicial opinions.¹⁴²

Because individual coder errors are generally not detectable when reviewing study methodologies alone, the discussions below do not directly examine their potential impact. Scholars in other areas, however, do review coding decisions and sometimes conclude that they may affect the study results in a substantial way.¹⁴³

III. HOW STUDY DESIGN CHOICES CAN PRODUCE VARIATION IN DATA

After reading about the array of study design choices that legal empiricists must make and their potential impact on results, one might be a bit less taken aback by the degree of inter-study variation in Figure 1, the overall reversal rate chart. But merely knowing that results can vary is not particularly helpful or interesting. What really matters when comparing or interpreting studies is understanding: (1) whether these variations *matter*—that is, if they are likely to produce meaningful differences in results; and (2) whether study design choices produce a substantial directional effect.

A. Study Design and Overall Reversal Rate Studies

First, one word of caution when reviewing the figures below: because different studies can contain multiple differences in methodology, each figure should not be considered in isolation. Rather, it is possible that results for a given study reflect multiple design-driven causes. Below, I focus on those sources that I consider to be the most likely drivers of the greatest inter-study variability.

1. Data Source

The earliest data in Figure 1 comes from Table B-8 of the *Judicial Business of the United States Courts Annual Report of the Director* (the “Reports”).¹⁴⁴ The Reports are produced annually and, along with a summary discussion of the highlights of that fiscal year’s statistical data, contain an extensive set of tables reporting statistical information about the

1561, 1576–77 (2011) (discussing study’s dataset relying on a computer-made program designed to detect citations to legal scholarship).

¹⁴² See, e.g., Hall & Wright, *supra* note 6, at 117 (describing studies that conduct a computer analysis of judicial texts and look for revealing patterns in syntax or semantics or both).

¹⁴³ See Pettys, *supra* note 136, at 5 (stating that coding errors adversely influenced the accuracy of Epstein’s, Segal’s, and Parker’s findings concerning Supreme Court justices’ voting patterns).

¹⁴⁴ UNITED STATES COURTS: JUDICIAL BUSINESS OF THE UNITED STATES COURTS, <http://www.uscourts.gov/statistics-reports/analysis-reports/judicial-business-united-states-courts> [<https://perma.cc/T3AF-G56D>] [hereinafter REPORTS].

United States court system,¹⁴⁵ including (in Table B-8) the Federal Circuit.¹⁴⁶ Table B-8 of the *Reports* is also provided by the Federal Circuit on its website,¹⁴⁷ and its reversal rate data is frequently cited by commentators.¹⁴⁸ The version of Table B-8 published in the *Reports* and released on the Federal Circuit's website cover the court's fiscal year (e.g., October 1, 2002 to September 30, 2003 constitutes the "2003" fiscal year).¹⁴⁹ In addition, Table B-8 is available in calendar year form in the *Statistical Tables for the Federal Judiciary* publication.¹⁵⁰ For purposes of clarity, the fiscal year data is referred to as Table B-8 (FY Reports), while the calendar year data is referred to as Table B-8 (Statistical Tables).¹⁵¹ Table B-8 (FY Reports) tends to be more widely cited by commentators, possibly because it is more easily found.¹⁵²

Another official source is referenced by Judge Rader in his dissent in *Cybor Corporation v. FAS Technologies, Inc.*¹⁵³ Judge Rader wrote that "[t]he Federal Circuit, according to its own official 1997 statistics, reversed

¹⁴⁵ Among these are the B-series tables, which report information about the Federal Appellate courts. Table B-5 contains data on the regional circuit courts of appeal but not the Federal Circuit Court of Appeals. *Id.*

¹⁴⁶ Unlike the data for the regional circuits, which is collected by the Administrative Office of the U.S. Courts, Table B-8 is provided directly by the Federal Circuit. Telephone Interview with Administrative Office, U.S. Courts (June 25, 2015).

¹⁴⁷ *Appeals Filed, Terminated, and Pending*, U.S. COURT OF APPEALS FOR THE FEDERAL CIRCUIT, <http://www.cafc.uscourts.gov/the-court/statistics> [<https://perma.cc/8HCZ-XM24>] [hereinafter STATISTICS].

¹⁴⁸ Sometimes expressly so, sometimes not. Jeffrey Lefstin, for example, uses it as the basis for comparing reversal rates by tribunal. Lefstin, *supra* note 22, at 1063 n.140; *see also* Chu, *supra* note 53, at 1125 n.180 (discussing the Federal Circuit's reversal rate in 1999, as reported in Table B-8); Teresa Lii, Note, *Shopping For Reversals: How Accuracy Differs Across Patent Litigation Forums*, 12 CHL-KENT J. INTEL. PROP. 31, 40 n.66 (2013); Sichelman, *supra* note 11, at 1171–72. In Moore, *supra* note 53, n.60, Moore cites to the Reports for reversal rates per year. *But see* Kimberly A. Moore, *Judges, Juries, and Patent Cases: An Empirical Peek Inside the Black Box*, 99 MICH. L. REV. 365, 396 (2000) (showing no citations for the same data).

¹⁴⁹ The Federal Courts system produces four versions of Table B-8 each year, ending on March 31, June 30, September 30, and December 31. Because the September and December reports are the most cited, I focus on those.

¹⁵⁰ Tables back to 2001 can be found at UNITED STATES COURTS: STATISTICAL TABLES FOR THE FEDERAL JUDICIARY, U.S. COURTS, <http://www.uscourts.gov/statistics-reports/analysis-reports/statistical-tables-federal-judiciary> [<https://perma.cc/XJC7-V3ZJ>].

¹⁵¹ In general, these two data are roughly the same. This is unsurprising given that their time periods largely overlap and come from the same source.

¹⁵² *See* Michael Risch, *The Availability Heuristic and IP*, WRITTEN DESCRIPTION BLOG (Sept. 18, 2015), <http://writtendescription.blogspot.com/2015/09/the-availability-heuristic-and-ip.html> [<https://perma.cc/FE8C-6YVT>] (stating how perceptions of the importance of *Lenz v. Universal Music* is very much dependent on availability, rather than data). The fiscal year version of the data is easily found, as it is published on the Federal Circuit's own website. Uncovering the fiscal year data required substantially more digging.

¹⁵³ 138 F.3d 1448, 1476 n.16 (Fed. Cir. 1998) (en banc) (Rader, J., dissenting), *abrogated by* 35 U.S.C. § 112 (2006) *as recognized in* *Teva Pharm. USA, Inc. v. Sandoz, Inc.*, 789 F.3d 1335, 1340 (Fed. Cir. 2015).

in whole or in part 53% of the cases from district courts (27% fully reversed; 26% reversed-in-part).¹⁵⁴ The 27% matches up with the reversal rate reported in Table B-8 (FY Reports),¹⁵⁵ but that table does not contain information on reversals-in-part.¹⁵⁶ While the Federal Circuit no longer publishes data on reversals-in-part, at one point it did.¹⁵⁷

Several sources report only an aggregated reversal rate; that is, a reversal rate for the entire period of the study. Studies reporting data in this form include Jay Kesan and Gwendolyn Ball's extensive study of district court decisions (mid-2000s),¹⁵⁸ Adam Shartzer's study of the top fifteen patent districts in terms of patent case filings for 2007 (2002 to 2007),¹⁵⁹ and Teresa Lii's study of "the busiest districts for patent litigation" (January 2009 to March 2012).¹⁶⁰ Sources that report reversal rates on an aggregate basis are depicted below as a dotted line, while sources that report reversal rates on an annual basis are presented as solid lines. Kesan and Ball drew from PACER,¹⁶¹ Lii drew from the WestlawNext,¹⁶² and Shartzer drew from an unstated source, possibly Lexis.¹⁶³

A final source is the Patstats dataset maintained by the University of Houston Law Center.¹⁶⁴ Patstats is a web service that provides quarterly and annual statistics for rulings on various issues in patent cases.¹⁶⁵ This data includes Federal Circuit affirmances and reversals in patent cases on an issue-by-issue basis. Although Patstats does not itself report reversal rates,

¹⁵⁴ *Id.*

¹⁵⁵ See Table B-8: *U.S. Court of Appeals for the Federal Circuit – Appeals Filed, Terminated, and Pending During the 12-Month Period Ending March 31, 2001*, U.S. COURTS (Aug. 5, 2016), <http://www.uscourts.gov/statistics/table/b-8/federal-judicial-caseload-statistics/2001/03/31> [<https://perma.cc/JK5Q-3XNP>] (listing 22% of cases appealed from U.S. District Courts as reversed by the Federal Circuit between April 1, 2000 and March 31, 2001).

¹⁵⁶ It appears that at one point the Federal Circuit released data on reversals-in-part. The only such reports I have been able to find are for 2006–2009.

¹⁵⁷ Though the Federal Circuit recently updated their statistics page, removing much of that data, traces of the data remain. See *United States Court of Appeals for the Federal Circuit: Affirmance and Reversal Rates for District Court Patent Infringement Appeals*, U.S. COURTS, <http://www.cafc.uscourts.gov/sites/default/files/announcements/2009/March5thInnpresentation.pdf> [<https://perma.cc/U8AL-VZTM>].

¹⁵⁸ Kesan & Ball, *supra* note 9, at 457 tbl. X. Due to the methodology of their study, Kesan and Ball do not report reversals on an annual basis. Rather, they report reversals based on the year that the appeal was filed and only for appeals arising from district court patent cases filed between 1998 and 2003. In addition, as discussed further below, Kesan and Ball report data on a per-issue basis.

¹⁵⁹ Shartzer, *supra* note 46, at 233.

¹⁶⁰ Lii, *supra* note 46, at 43–44. In her Note, Lii identified the twenty districts with the most cases dealing with patents and that decided the merits of the patent-in-suit contained in volumes 93 through 96 of the United States Patent Quarterly.

¹⁶¹ Kesan & Ball, *supra* note 9, at 430.

¹⁶² Lii, *supra* note 46, at 42.

¹⁶³ Shartzer, *supra* note 46, at 227–28.

¹⁶⁴ PATSTATS, *supra* note 46. Patstats's collection and reporting methodology is available at <http://www.Patstats.org/Patstats3.html> [<https://perma.cc/QB2N-DEM6>] (copy on file with author).

¹⁶⁵ *Id.*

its data on affirmances and reversals is sometimes drawn upon by other scholars.¹⁶⁶ The Patstats reversal rate data presented in Figure 1 is based on a compilation of this issue-by-issue coding, following a methodology similar to that used by Ted Sichelman.¹⁶⁷ Note that the first two periods of Patstats data (2000 to 2004 and 2005 to 2009) consist of aggregate data, while the latter two periods consist of annual data.

At this point, it is difficult to say whether the specifics of the data source have a meaningful effect on the reversal rate metrics. Certainly they carry some potential for variation. Based on a review of the published methodologies, all of the sources depicted in Figure 1 included summary affirmances, at least to the extent that information was available about the issues decided by the court.¹⁶⁸ Inclusion or exclusion of Rule 36 summary affirmances does not appear to play a substantial role in the inter-study variability depicted in Figure 1, but there are potential differences among these sources. In addition to the issue of record unit discussed below, the sources may vary in terms of the completeness of their data.

2. *Collection and Filtering Methodologies*

Discussion of potential differences among collection and filtering methodologies for the data provided in Figure 1 is confounded by the lack of detail provided by many of those sources. Rader, for example, simply references the data with only a loose citation to its source. Shartzer, similarly, provides only a small amount of information on data collection. And while the Federal Courts undoubtedly have a detailed data collection methodology, only limited methodological data is provided with the data tables themselves. Other studies—notably, the Kesan and Ball study—contain extensive discussion of methodology. Unfortunately, it is difficult to compare something to nothing, no matter how detailed that something is.

3. *Record Unit*

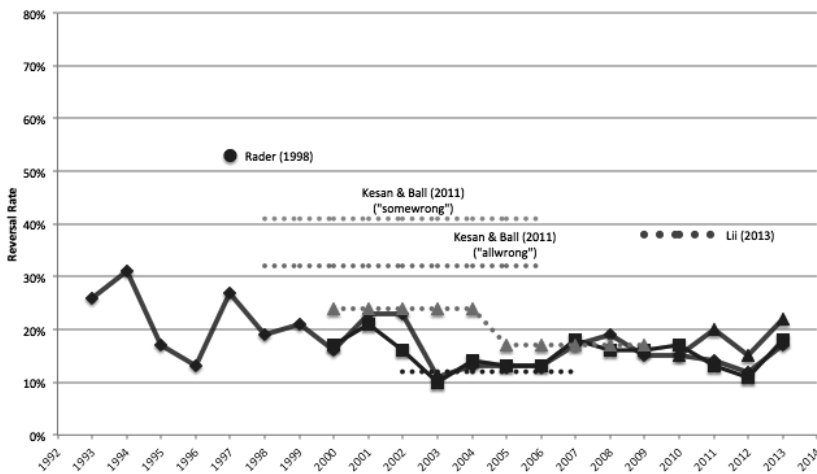
Another difference in study methodology is between those studies that examine reversal rates on a per-opinion basis and those that examine reversal rates on a per-issue basis, an approach that is roughly commensurate to the per-analysis approach discussed above. Figure 3.a distinguishes between these approaches. Reversal rates based on a per-case or appeal unit are shown in black; those on a per-issue unit are depicted as gray.

¹⁶⁶ See e.g., Sichelman, *supra* note 11, at 1174–75 n.58 (noting that the data was gathered from Patstats).

¹⁶⁷ *Id.* at 1175. The coded data is available at <https://www.empirical.law.uiowa.edu>.

¹⁶⁸ PATSTATS, *supra* note 46; Lii, *supra* note 46, at 42; Shartzer, *supra* note 46, at 232–33. Kesan and Ball included Rule 36 summary affirmances as long as they could be classified as to issue by reading the briefs. Kesan & Ball, *supra* note 9, at 431 n.212, 432 n.217, 457 n.263.

Figure 3.a: Reported Rates at which the Federal Circuit Reverses District Courts (per-issue vs. per case)



The Patstats and Kesan and Ball data are reversal rates per issue, based on their descriptions of the methodologies, while Table B-8 (Reports/Statistical Tables), Lii, Shartzer, and Rader numbers are per-case/appeal reversal rates.¹⁶⁹ Due to the limited methodological information available for Table B-8, it is not clear exactly what the denominator is, but it appears to be appeals terminated by a merits panel.¹⁷⁰ If this is so, it adds another wrinkle to the reversal rate data because multiple docketed appeals may be decided in a single appellate opinion, particularly when the appellee files a cross-appeal. Note that some descriptions of methodologies are much clearer and more complete than others, so the quality of these inferences varies.

By itself, Figure 3.a does not really tell us much about whether the record unit has a substantial effect on the direction of the reversal rate metric. The effect of the record unit is explored in more depth below in Part III.B.

4. Parameter Definitions

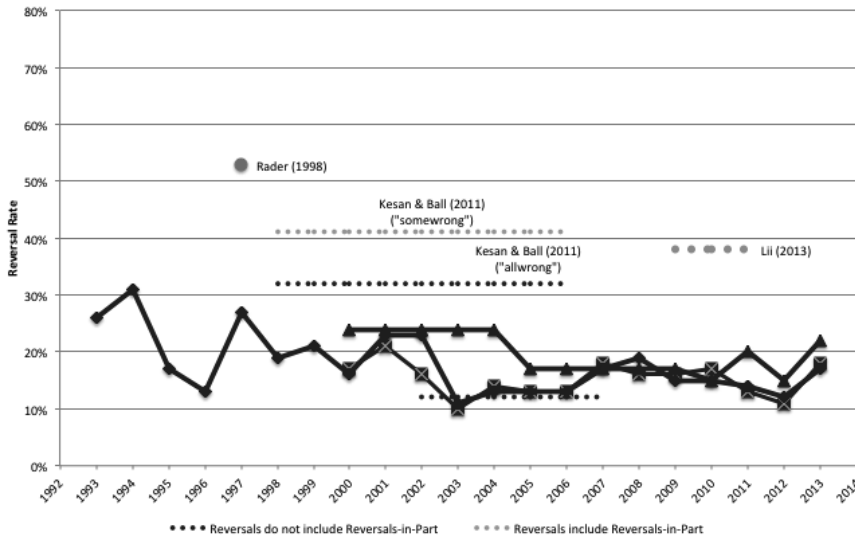
When comparing reversal rates across studies, or interpreting their results, a key study-design component is the way in which a given study defines a reversal. As described above, a study might count a decision that affirms-in-part and reverses-in-part as a reversal, or it might count it as an affirmance. Figure 3.b illustrates how this study-design difference can produce systematic variation in reported reversal rates. Data that treats

¹⁶⁹ *Cybor Corp. v. FAS Techs., Inc.*, 138 F.3d 1448, 1476 n.4 (Fed. Cir. 1998) (en banc) (Rader, J., dissenting), *abrogated by* 35 U.S.C. § 112 (2006) *as recognized in* *Teva Pharm. USA, Inc. v. Sandoz, Inc.*, 789 F.3d 1335, 1340 (Fed. Cir. 2015); Kesan & Ball, *supra* note 9, at 440; Lii, *supra* note 46, at 44 n.86; Shartzer, *supra* note 46; PATSTATS, *supra* note 46; REPORTS, *supra* note 144.

¹⁷⁰ See REPORTS, *supra* note 144.

affirmances-in-part and reversals-in-part as affirmances is depicted in black; those that treats these decisions as reversals are depicted in gray.

Figure 3.b: Reported Rates at Which the Federal Circuit Reverses District Courts (Reverse-in-Part vs. Reverse-in-All)



All of the studies in Figure 3.b that treat affirm-in-part decisions as reversals report much higher reversal rates than all of the sources treating affirm-in-part decisions as affirmances.¹⁷¹ Kesan and Ball, in particular, illustrate how applying a different definition of “reversal” can dramatically affect results. In all other respects their methodology for the two reported sets of data were identical. The Rader data also, which combines a 27% reverse-in-all component with a 26% reverse-in-part component, illustrates the powerful effect parameter definition can produce.¹⁷²

The choice of what constitutes a reversal is particularly important when comparing results against other circuits’ behavior—particularly when considering the nature of patent litigation, which involves multiple issues that may be more likely than other types of cases to produce a mixed appellate decision. Thus, for example, comparing studies treating affirmances-in-part as reversals to the reversal rates of other circuits reported by the administrative office of the U.S. Court system (which treats

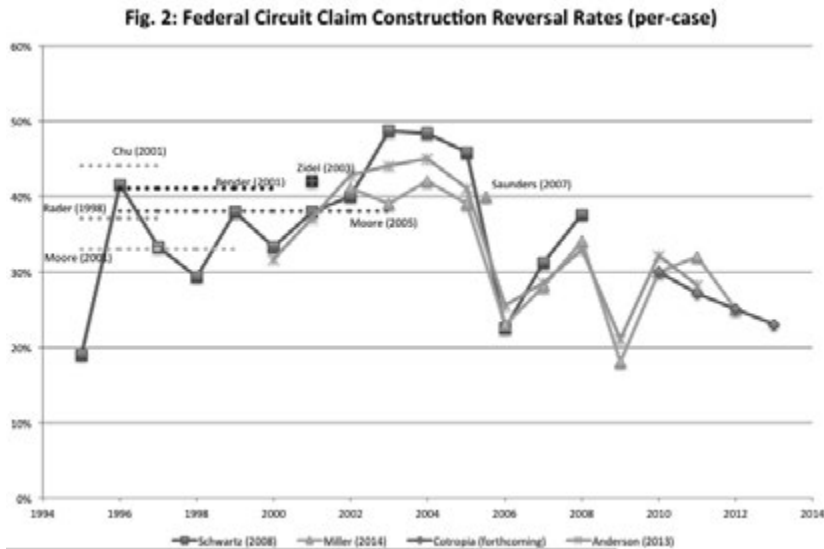
¹⁷¹ A result that, obviously, follows as a matter of logic. See, e.g., Sichelman, *supra* note 11, at 1185 (explaining how, when partial affirmance rates are excluded, the reported reversal rate at the Federal Circuit mirrors the reported reversal rate for private civil cases in other circuits); Ted Sichelman, *Are Appeals at the Federal Circuit a “Coin Flip”?*, PATENTLY-O (Apr. 9, 2010), www.patentlyo.com/patent/2010/04/are-appeals-at-the-federal-circuit-a-coin-flip.html. [https://perma.cc/4EBQ-R7ZU].

¹⁷² *Cybor*, 138 F.3d at 1476 (Rader, J., dissenting).

affirmances-in-part as affirmances) will result in erroneous conclusions.¹⁷³

B. Claim Construction Studies

In contrast with the overall reversal rate data discussed above, studies of claim construction contain substantially less inter-study variability. Yet differences remain. For reference, Figure 2, originally provided at Section I.B., is reproduced below. In general, these studies provide much more detail on their methodologies, and thus offer more opportunity for comparison and analysis. This is a very good thing, as it allows for a greater understanding and contextualization of this data.



1. Data Source

As with reversal rates generally, a possible source of inter-study variation might be the researchers' different choices about the sources they mine. These sources carry potential to either expressly or inherently cabin the claim construction decisions that form the study population.

Express choices about data sources are fairly obvious. If the studies were only to draw from the U.S.P.Q., for example, they might not be collecting the entirety of claim construction decisions—particularly those seen as relatively mundane. But all of the studies depicted in Figure 2 draw from the sources that are generally acknowledged to be the most complete repository of appellate decisions: the online Westlaw and Lexis databases.

Less obvious are the inherent limitations of the data sources. As discussed above, not all data sources contain records of the summary

¹⁷³ See, e.g., Lii, *supra* note 46, at 43 (describing the methodology of the study, which recorded cases that were reversed – both in part and in whole – as “reversed”).

affirmances, let alone the materials necessary to identify whether a Rule 36 summary affirmation involves a claim construction dispute. Some design choices to not include Rule 36 summary affirmances were made consciously¹⁷⁴ (perhaps due to resource constraints), but the lack of summary affirmances in other studies may be due to simple oversight or a failure to appreciate their potential impact.¹⁷⁵

Of course, the least useful studies are those that include little or no methodological data, thus making it difficult or impossible to tell if they include summary affirmances. Judge Rader's "study" of claim construction from his dissent in *Cybor* is illustrative. In that study Judge Rader surveyed

every patent decision rendered by the Court of Appeals for the Federal Circuit between 5 April 1995 (the date *Markman I* was decided) and 24 November 1997. A total of 246 patent cases, originating in the Board of Patent Appeals and Interferences (BPAI), the district courts, and the Court of Federal Claims were evaluated. Of the 246 cases, 141 cases expressly reviewed claim construction issues.¹⁷⁶

No source is provided, and aside from the quotation above, no description of methodology is offered; the study was presumably conducted by Judge Rader himself or under his direction. Read that passage as you will.

In contrast, Christian Chu's 2001 study provides an extensive discussion of methodology, explaining that he drew from the Lexis database and that he initially collected but did not include Rule 36 summary affirmances in his claim construction analysis—likely because of the difficulty of obtaining the underlying briefs at the time.¹⁷⁷ As Kimberly Moore observed a few years later, information about issues on appeal "cannot be obtained from a quick search on Westlaw or Lexis, but instead requires resort to the briefs filed with the Federal Circuit."¹⁷⁸

Other contemporaneous studies followed a similar approach. Gretchen Bender's 2001 study of claim construction reversals did not provide a source, but did list all decisions counted for purposes of calculating the reversal rate.¹⁷⁹ That list contains Westlaw citations for non-precedential opinions, suggesting that Bender drew from Westlaw. But like Chu, Bender

¹⁷⁴ Chu, *supra* note 53, at 1097 n.12.

¹⁷⁵ Moore, *supra* note 53, at 234.

¹⁷⁶ *Cybor*, 138 F.3d at 1476 n.4 (Rader, J., dissenting).

¹⁷⁷ At the time Chu (then a student at the University of California at Berkeley School of Law) performed his study, the only way to obtain the briefs was to go to the Federal Courts Building and review the paper dockets—a time-consuming and expensive undertaking later done by Kimberly Moore (at the time, a Professor at George Mason Law School). Moore, *supra* note 53, at 236.

¹⁷⁸ *Id.*

¹⁷⁹ Bender, *supra* note 53, at 203–07 n.215.

did not include Rule 36 summary affirmances.¹⁸⁰ Andrew Zidel's 2003 study of claim construction decisions also does not specify a source, but the list of cases in his Appendix A contains Lexis citations for some non-precedential decisions, suggesting the obvious inference. Zidel's data also does not include Rule 36 summary affirmances.¹⁸¹ Michael Saunders's 2007 study also collected opinions, in his case for the period July 13, 2005 through September 13, 2006, but he did not include summary affirmances either.¹⁸²

A major advance made by Kimberly Moore's 2001 and 2005 studies was the identification of a data source that would allow Rule 36 summary affirmances to be analyzed: the parties' briefs themselves. The Rule 36 affirmances themselves were not difficult to collect, as they contain nothing more than a one-line affirmation of the lower tribunal. Moore also undertook painstaking collection of the paper briefs, reviewing them all to determine whether the appeal involved a claim construction dispute.¹⁸³ Consequently, Moore's studies of claim construction reversals are the first to include summary affirmances.¹⁸⁴

More recent researchers have used sources similar to Moore's, with the advantage of being able to access electronic copies of briefs for at least a portion of the periods studied. David Schwartz's 2008 study used a combination of LexisNexis and Westlaw databases, collecting both opinions (from the LexisNexis "CAFC" database) and Rule 36 summary affirmances (from Westlaw).¹⁸⁵ As with Moore, Schwartz obtained the appellate briefing for the Rule 36 summary affirmances to use in determining whether a summary affirmation involved a claim construction issue.¹⁸⁶ When the appellate briefing was not available, Schwartz reviewed additional sources.¹⁸⁷ Shawn Miller's 2014 study¹⁸⁸ "replicated the methodology of Schwartz (2008, 2010)"¹⁸⁹ for the period January 1, 2002 to December 31,

¹⁸⁰ Moore, *supra* note 53, at 235.

¹⁸¹ Zidel, *supra* note 53, at 755–60.

¹⁸² Saunders, *supra* note 53, at 235. Saunders indicates that he used the same approach as a previous study that drew from Westlaw.

¹⁸³ In the 2001 article, Moore simply explains that she collected the briefs, without providing further details. Kimberly A. Moore, *Are District Court Judges Equipped to Resolve Patent Cases?*, 15 HARV. J.L. & TECH. 1, 8 (2001). In the 2005 article, Moore indicates that the briefs came from the Federal Circuit itself. *See* Moore, *supra* note 53, at 236 ("Obtaining the actual briefs is both time consuming and expensive. This study did just that . . .").

¹⁸⁴ J. Jonas Anderson & Peter Menell, *Empirical Studies of Claim Construction*, in RESEARCH HANDBOOK 7, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2661005 [<https://perma.cc/M86W-RR46>].

¹⁸⁵ David L. Schwartz's data was used in a series of articles addressing different aspects of claim construction. *See, e.g.*, Schwartz, *supra* note 53, 238–39. The article provides the methodology used in the study and is used for identification purposes.

¹⁸⁶ Schwartz, *supra* note 53, at 269.

¹⁸⁷ *Id.* at 270.

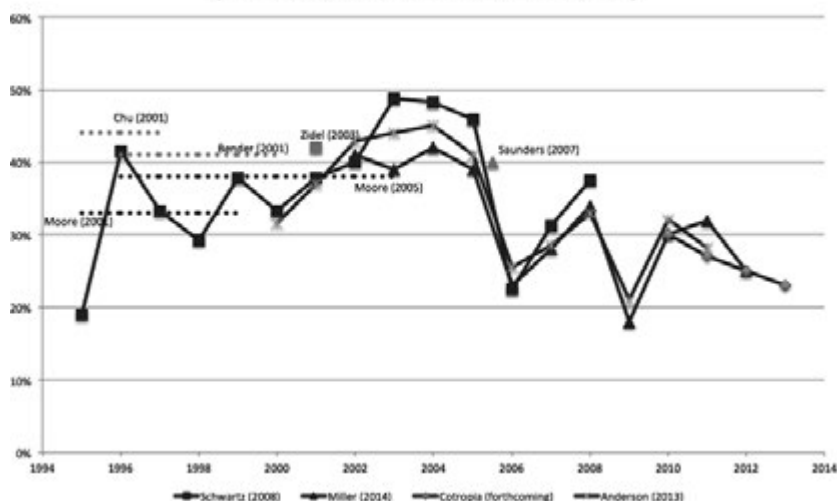
¹⁸⁸ Miller, *supra* note 53.

¹⁸⁹ *Id.* at 822 n.80.

2012,¹⁹⁰ thus including summary affirmances. Similarly, Jonas Anderson and Peter Menell's 2014 study used LexisNexis for both opinions and Rule 36 summary affirmances,¹⁹¹ and the briefs for Rule 36 summary affirmances obtained from an unstated source (possibly LexisNexis itself) were reviewed. And Christopher Cotropia's 2015¹⁹² study states that "[a]ll Federal Circuit decisions involving utility patents from January 1, 2010, to December 31, 2013, were collected."¹⁹³ Cotropia does not otherwise provide a source for these decisions. Cotropia did, however, collect Rule 36 summary affirmances.¹⁹⁴

How much does the inclusion or exclusion of Rule 36 summary affirmances actually matter in the claim construction context? Unlike in overall reversal rate studies, there is a substantial cost to including these decisions. Figure 4.a provides a comparison between those studies that included Rule 36 affirmances and those that did not. Studies shown in black included Rule 36 affirmances; those shown in gray did not.

**Fig. 4.a: Federal Circuit Claim Construction Reversal Rates Per-case
(Inclusion vs. Exclusion of Rule 36 affirmances)**



¹⁹⁰ Miller, working with Mark Lemley, later expanded the data through June 2014. Lemley & Miller, *supra* note 39.

¹⁹¹ J. Jonas Anderson & Peter Menell, *Informal Deference: A Historical, Empirical, and Normative Analysis of Patent Claim Construction*, 108 NW. L. REV. 1, 35 (2014).

¹⁹² Christopher Cotropia, *Patent Claim Interpretation Review: Deference or Correction Driven?*, 2014 BYU L. REV. 1095 (2015).

¹⁹³ *Id.* at 1100.

¹⁹⁴ *Id.*

The conclusion is that it probably matters a fair amount, perhaps more so today given the Federal Circuit's increasing use of the Rule 36 vehicle.¹⁹⁵ With the exception of Saunders (whose data encompass part of 2005 and part of 2006, a period when the claim construction reversal rate was falling dramatically), the studies that do not include Rule 36 affirmances report a reversal rate that is consistently higher than that of the studies that did include Rule 36 affirmances. Post-Moore, study authors have consistently expended an effort to include Rule 36 affirmances.¹⁹⁶ In her 2005 study, Moore compared her results when including Rule 36 summary affirmances with those when they were excluded.¹⁹⁷ With the summary affirmances, the reversal rate on a per-term basis was 34.5%; without them it was 40.8%.¹⁹⁸ Anderson and Menell observed that the court "issued 18.7% of claim construction cases before *Phillips* under Rule 36" and "30.2% of cases since that time."¹⁹⁹ Similarly, Cotropia found that of the "314 cases involving claim interpretation reviews by the Federal Circuit, . . . 118 were reviews done via Rule 36 determinations."²⁰⁰ Excluding those 118 decisions—all affirmances of the lower court—transforms Cotropia's observed 26% reversal rate into a 42% per-case reversal rate.

2. Collection and Filtering Methodology

While overall reversal rate studies can simply collect all decisions of the court for a given time period, studies of claim construction decisions (and studies that focus on discrete legal doctrines) pose a greater collection challenge: how to identify the decisions that involve claim construction disputes.

As with overall reversal rate studies, claim construction studies must first identify the originating court. The studies depicted in Figure 2 in this Article primarily originate from the district courts, but some studies also include appeals arising from the Court of Federal Claims (CFC), the International Trade Commission (ITC), and sometimes even the United States Patent and Trademark Office (USPTO), through the Board of Patent Appeals and Interferences (BPAI) (now the Patent Trial and Appeal Board). Although CFC and ITC appeals are unlikely to affect the results substantially, as patent appeals from these courts make up only a tiny portion of the claim construction appeals considered by the Federal Circuit,²⁰¹

¹⁹⁵ See Anderson & Menell, *supra* note 25, at 61–62 (noting the increased use of summary affirmance).

¹⁹⁶ *Id.* at 35; Cotropia, *supra* note 192, at 1100; Moore, *supra* note 53, at 236.

¹⁹⁷ Moore, *supra* note 53, at 236.

¹⁹⁸ *Id.*

¹⁹⁹ Anderson & Menell, *supra* note 25, at 55.

²⁰⁰ Cotropia, *supra* note 192, at 1101.

²⁰¹ See, e.g., Chu, *supra* note 53, at 1122 n.175 ("[T]he Federal Circuit heard fewer than 10 cases from either the CFC (9 cases) or the ITC (7 cases).").

appeals from the BPAI could potentially impact the results—particularly because the patent office has long used an arguably broader scope for claim construction than the other three courts.²⁰²

More challenging is the question of how to collect decisions relating to claim construction. As with reversal-rate studies generally, studies of claim construction typically use some form of electronic algorithm followed by human coding according to a particular criteria. But often these algorithms incorporate some form of content searching to make at least a rough pass at identifying the relevant decisions. The approaches taken by the studies in Figure 2 are described below.

Rader, Bender, and Zidel provide relatively little information on the collection methodology they employed. Bender simply indicates that the cases arose from the district courts,²⁰³ while Zidel indicates that the data included appeals from the district courts, BPAI, CFC, and ITC.²⁰⁴

Each of the remaining studies depicted in Figure 2 used some form of search algorithm to collect claim construction decisions. Chu collected “all patent decisions” from the “US Court of Appeals Cases – Federal Circuit” dataset within the Lexis database.²⁰⁵ To collect these decisions, Chu electronically searched for all Federal Circuit opinions filed within the defined time period.²⁰⁶ Chu then reviewed each of the opinions to screen patent cases from non-patent cases. The cases thus included appeals from district courts, the BPAI, and ITC. For purposes of identifying relevant opinions, Chu defined “claim construction review” as “any instance in which the Federal Circuit *expressly* reviewed the lower tribunal’s claim interpretation.”²⁰⁷ This definition excluded Rule 36 summary affirmances.²⁰⁸

In her 2001 article, Moore conducted an electronic search on Westlaw using a search algorithm, followed by a manual examination to determine “whether the district court judge’s claim construction was being appealed to the Federal Circuit.”²⁰⁹ This search focused on the period from April 23, 1996 to December 31, 2000. Moore’s dataset was limited to appeals from district court judges. Moore also manually reviewed the appellate briefs for the Rule 36 summary affirmances issued by the court during this period to determine which summary affirmances challenged claim construction determinations.²¹⁰ Moore’s subsequent study, published in 2005, expanded

²⁰² Cotropia, *supra* note 192, at 1100 n.18.

²⁰³ Bender, *supra* note 53, at 203.

²⁰⁴ Zidel, *supra* note 53, at 746.

²⁰⁵ Chu, *supra* note 53, at 1092.

²⁰⁶ *Id.*

²⁰⁷ *Id.* at 1094.

²⁰⁸ *Id.*

²⁰⁹ See Moore, *supra* note 183, at 8 n.36 (“I conducted a search on Westlaw using the query: ‘patent & claim /s interp! or constr!’”).

²¹⁰ *Id.*

the dataset through 2003.²¹¹ Saunders also used an electronic keyword algorithm provided in the paper to identify and collect district court cases from Westlaw, but did not capture Rule 36 summary affirmances.²¹²

Schwartz, Miller, Anderson and Menell, and Cotropia all followed a similar methodology to that employed by Moore, with some differences in the set of courts collected and the relevancy criteria. Following use of a search algorithm, Schwartz employed manual review to determine “whether the case contained a resolution by the Federal Circuit of an issue of patent claim construction that a district court had previously decided.”²¹³ Schwartz’s study focused on district courts.²¹⁴ Miller “replicated the methodology of Schwartz (2008, 2010).”²¹⁵ As with Schwartz, Miller collected decisions arising from the district courts.²¹⁶ Anderson and Menell used electronic keyword algorithms to collect opinions “dealing with claim construction,”²¹⁷ followed by human review for relevancy.²¹⁸ Rule 36 decisions were also collected through electronic searches, followed by human review of the briefs to determine whether the Rule 36 summary affirmances involved a claim construction issue.²¹⁹ Anderson and Menell excluded appeals arising from the BPAI from their data set; they included appeals from the International Trade Commission, the district courts, and the Court of Federal Claims.²²⁰ The period studied ran from 2000 to 2011. Cotropia collected claim construction appeals arising from all lower tribunals except the BPAI.²²¹ With Rule 36 summary affirmances, Cotropia examined the briefs of both parties to determine “if either party argued in their brief that the appeal’s resolution required review of a claim interpretation by the lower court”²²²

3. *Record Unit*

Three different types of record units dominate the claim construction

²¹¹ Moore, *supra* note 53, at 239.

²¹² Saunders, *supra* note 53, at 235.

²¹³ Schwartz, *supra* note 53, at 269. “The exact search query was “claim w/10 (constru! or interp!) and date(geq (04/24/1996) and leq (6/30/2007).” *Id.* at 269 n.220.

²¹⁴ *Id.* at 269.

²¹⁵ Miller, *supra* note 53, at 822 n.80.

²¹⁶ *Id.* at 822–23.

²¹⁷ Anderson & Menell, *supra* note 25, at 78. No search algorithms are provided in the paper itself.

²¹⁸ *See id.* at 78 n.346 (defining relevancy as “any case in which the meaning of a claim term was challenged on appeal”). Earlier in the paper, the authors offer a slight variation on this relevancy category. *Id.* at 35 (“A case was deemed relevant if the Federal Circuit reviewed the district court’s construction of a claim term.”).

²¹⁹ *See id.* at 35–36 (“A Rule 36 case was deemed relevant for the study if the briefs challenged the district court’s construction of a claim term.”).

²²⁰ *See id.* at 38 n.236, 78–79 (using Appendix A to show which cases were excluded in the final results).

²²¹ E-mail from Christopher Cotropia, Professor, U. Richmond School of Law, to Jason Rantanen, Assoc. Professor, Univ. Iowa Coll. Law (June 25, 2015) (on file with author).

²²² Cotropia, *supra* note 192, at 1100–01.

studies. Some claim construction reversal rate studies record data at the level of the opinion, while others provide more granular data by collecting information on each individual claim term or element in dispute (the “per-term” level). The studies that operate at this level of granularity often report on a different metric: whether the appellate court’s change altered the outcome of the appeal.

Some studies report data only at the per-opinion level. Based on their methodological discussions, Rader, Bender, Zidel, Chu, and Cotropia all appear to have used a per-opinion record unit in their studies.²²³ Miller also reports data only at the per-opinion level, although information was collected at the per-patent level as well.²²⁴ The remaining studies recorded data at multiple levels. Moore, Saunders, Schwartz and Anderson report results on both a per-term basis and a per-opinion basis, with Moore and Anderson also reporting on whether the change in claim construction also changed the outcome of the case.

Although there is some difference depending on whether the claim construction reversal data is considered at the per-term or per-decision level, that difference is both relatively small and largely due to the standard practice in claim construction studies of treating the decision as a “reversal” if the Federal Circuit reverses on *any* of the terms. This is similar to the “reversal-in-part counting as a reversal” approach discussed above in the overall reversal rates. By treating *any* reversal on a term in a decision as a reversal for that decision, the consequence is that reversal numbers in data at the per-decision level will necessarily be higher than that at the per-term level (since some claim construction decisions involve more than one term at issue).

When distinguishing between the per-opinion and per-case approaches, Moore observed a 3% difference for the period from 1996 to 2003,²²⁵ and Saunders observed approximately a 6% difference for the approximately

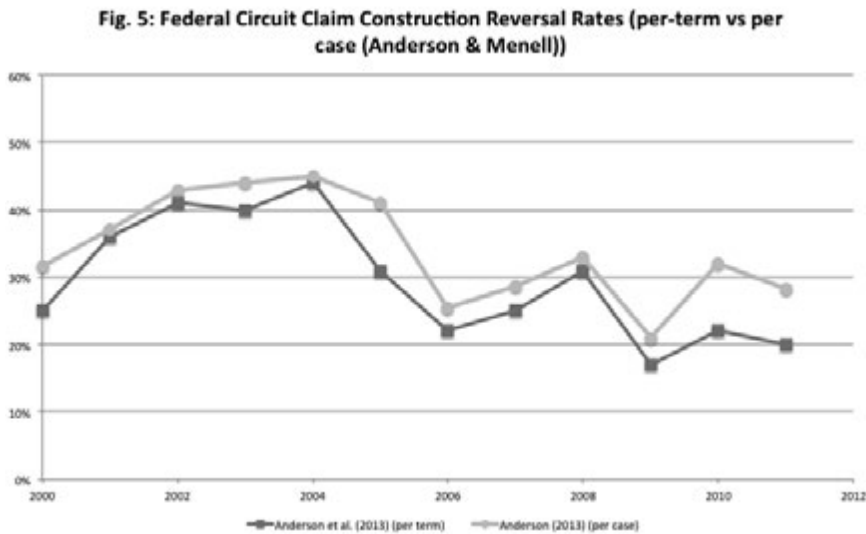
²²³ *Cybor Corp. v. FAS Techs., Inc.*, 138 F.3d 1448, 1476 n.4 (Fed. Cir. 1998) (en banc) (Rader, J., dissenting), *abrogated by* 35 U.S.C. § 112 (2006) *as recognized in* *Teva Pharm. USA, Inc. v. Sandoz, Inc.*, 789 F.3d 1335, 1340 (Fed. Cir. 2015) (“Among these 141 decisions, this court reversed, in whole or in part, 54 or 38.3% of all claim constructions.”). This implies that “decision” and “claim constructions” are operating at the opinion level. *See* Bender, *supra* note 53, at 206–07 (“Of these 160 decisions, the Federal Circuit has changed the trial court’s claim interpretation in nearly 65 cases (although the final judgment itself may have been affirmed.)”; Chu, *supra* note 53, at 1092 (implying that “decisions” are synonymous with written opinions and summary affirmances). Chu also determined the rate at which the Federal Circuit “changed” the construction of individual claim construction “elements” (which seems to be synonymous with “terms”), but reports only “normalized” monthly percentiles. Chu, *supra* note 53, at 1105; *see also* Cotropia, *supra* note 192, at 1095, 1101 (explaining that any claim term under review that was reversed was marked as such). Zidel does not specify whether the reversal rate is on a per-opinion or per-term basis, but the context suggests that the reversal rate is on a per-opinion basis. Zidel, *supra* note 53, at 739–40.

²²⁴ Miller, *supra* note 53, at 823.

²²⁵ Moore, *supra* note 53, at 239.

year period from mid-2005 to mid-2006.²²⁶ Schwartz observed a 6% difference for the period 1996–2007.²²⁷ Anderson and Menell observed a 3% difference (pre-*Phillips*) and a 5.5% difference (post-*Phillips*).²²⁸

The following figure, showing Anderson and Menell’s data for the period of their study, illustrates the two different approaches, one showing reversal rates at the per-term level and the other showing reversal rates at the per-decision level. It illustrates the direction of the difference, but suggests that it has not led to great variation in results.²²⁹



4. Parameter Definitions

Much of the differences in parameter definitions among claim construction studies depend on how the record unit is defined. When it is defined at a per-term level, a reversal is typically just a changed claim construction for that term. When it is defined at the per-opinion level, the standard approach that has emerged is to identify *any* change in claim construction as a reversal, with some variation in language.²³⁰ Only one

²²⁶ Saunders, *supra* note 53.

²²⁷ Schwartz, *supra* note 53, at 249.

²²⁸ Anderson & Menell, *supra* note 25, at 6.

²²⁹ Thanks to Jonas Anderson and Peter Menell for the data for this figure. Actual percentages are provided in Appendix B.2.

²³⁰ Bender reports the rate at which the Federal Circuit “changes” the trial court’s claim construction. Bender, *supra* note 53, at 207. Zidel (2003) reports the percentage of “decisions” in which the district court reversed either in whole or in part. Zidel, *supra* note 53, at 745–46. Saunders (2007) reports the rate at which the Federal Circuit “changed” the district courts’ claim constructions, which Saunders uses synonymously with “reversed.” Saunders, *supra* note 53, at 236. Chu reports the rate at which the CAFC “changed” at least one claim interpretation in a case (in other words, akin to treating

researcher, Cotropia, has distinguished between reversals-in-part and reversals on all terms by including a “mixed” category.²³¹ Some studies also examine the link between claim construction and overall outcome by creating a separate parameter to capture information on whether the reversal was outcome-determinative: in other words, whether the “claim construction error actually resulted in reversal of the appealed judgment.”²³²

5. *Coding and Other Sources of Inter-Study Variation*

Because Miller (2014) replicated Schwartz (2008)’s methodology for an overlapping time period, it is possible to examine how other sources of inter-study variation—particularly, perhaps, coding differences—can affect reversal rates. The figure below compares Miller’s observed reversal rates with those of Schwartz for the overlapping period.²³³

reversals-in-part as reversals). Chu, *supra* note 53, at 1104. Moore (2001, 2005) reports the rate at which district court claim constructions were “wrong” according to the Federal Circuit. Moore, *supra* note 53, at 243. When reporting per-decision data, Moore reports the rate at which the district court “held at least one term was wrongly construed.” *Id.* at 239. Schwartz provides data on the per-term level (“wrongly construed”), the per-case level (“cases with at least one wrongly construed term”) and case-determinative level (“cases reversed, vacated, and/or remanded because of erroneous claim construction.”). Schwartz, *supra* note 53, at 249. Anderson and Menell provide data on the per-term level (affirmed or reversed), the per-case level (“cases with at least one reversed claim term”), and the case-determinative level (“cases resulting in remand, reversal, or vacation due to claim construction error”). Anderson & Menell, *supra* note 25, at 40. Miller (2014) provides data on the per-case level (“[w]hether the Federal Circuit explicitly rejected at least part of the district court’s claim construction”) and case-determinative level (“whether claim construction error required the Federal Circuit to reverse, vacate or remand the district court judgment that led to the appeal”). Miller, *supra* note 53, at 823. Cotropia (2015) examined:

[W]hether the Federal Circuit affirmed or reversed a claim interpretation determination by the lower court. This included the interpretation of at least one claim term’s meaning, and also included determinations of whether or not a claim term was indefinite or whether a claim was governed by § 112(f) and written in means plus-function or step-plus-function language.

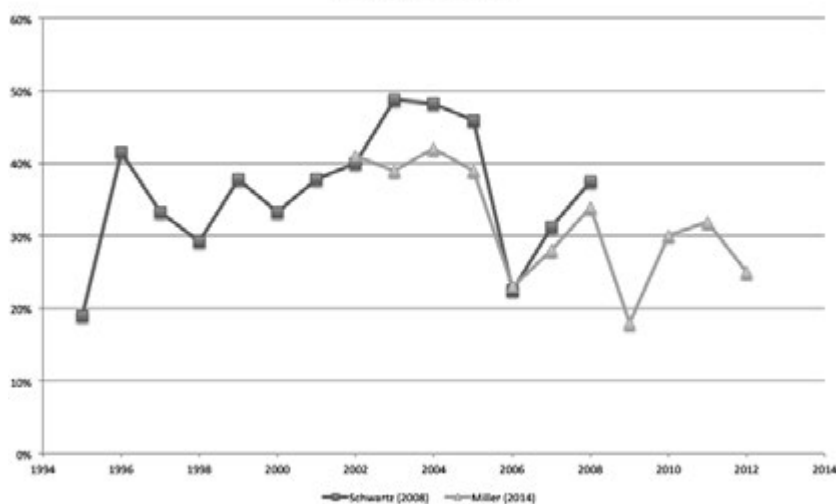
Cotropia, *supra* note 192, at 1100.

²³¹ Cotropia, *supra* note 192, at 1101. When multiple claim terms were involved, Cotropia coded the result as “affirmed” only if all claim interpretations were affirmed, “reversed” only if all the claim interpretations were reversed, and “mixed” when some claim interpretations were affirmed and others reversed. *Id.* Out of the 314 claim construction decisions recorded, Cotropia observed only sixteen “mixed” decisions, confirming that this group of decisions is relatively small. Cotropia, *supra* note 192, at 1103.

²³² Moore linked the claim construction errors to the rate at which the Federal Circuit reversed on a substantive issue, observing a 27% (2001) and 29.7% (2005) reversal/vacate rate directly attributable to errors in district court claim construction. Moore, *supra* note 53, at 238–39.

²³³ Thanks to Dave Schwartz and Shawn Miller for providing the data underlying this figure. Actual percentages are available in Appendix B.2.

Figure 5: Comparison between Schwartz (2008) and Miller (2014)'s Per-case reversal rates



Some points are relatively consistent across both studies; others vary by as much as 10%. Given Miller's attempt to replicate Schwartz's methodology, the implication is that coding variation can play a relatively noticeable role in the results observed by these studies. Perhaps the difference comes from the limitations of what electronic briefing was available in the PACER database from the 2000s. Or perhaps it comes from the authors' differences in identifying whether a claim term was in dispute in a Rule 36 briefing (and thus whether the Rule 36 opinion would be included as an affirmance in the data set). Yet another possibility is that the authors—both experienced attorneys and sophisticated empirical researchers—had different interpretations of whether a claim term was changed by the court. Data points for these studies were, as a general matter, relatively substantial.²³⁴ Curiously, using a slightly different data set that included decisions from the ITC and CFC, Anderson and Menell's data produced a reversal rate falling between those provided by Schwartz and Miller.²³⁵

Setting aside the potential effect of coding differences, the main drivers of variation among these studies appears to be the inclusion or exclusion of Rule 36 summary affirmances and the treatment of reversals-in-part as reversals.

²³⁴ For 2003, for example, Schwartz had a total of 113 discrete record units and Miller had a total of 118 record units. (Data provided by David Schwartz and Shawn Miller).

²³⁵ See Anderson & Menell, *supra* note 25, at 38–39 (comparing their results for the period 2001–2003 with that of Moore for that period and observing a 0.3% difference in the calculated per-term reversal rate). But also note the minor differences in collection and filtering methodology discussed above and descriptions of methodologies that are incomplete. *Id.*

IV. POTENTIAL IMPACTS ON OTHER STUDIES

Reversal rates and outcome measures are sometimes used in assessing broader institutional claims about the Federal Circuit and judicial decision-making generally. While outcome-focused studies represent a valuable step forward in empirically testing these hypotheses, their conclusions are not without caveat. Just as choices in study design can produce differences in studies that purport to measure the same characteristics of judicial decisions, so too do these choices carry the potential to affect the results of studies assessing institutional or realism-based claims. For reasons of space and interest, the following discussions focus specifically on those studies that employ reversal rates as a tool for testing particular hypotheses about the Federal Circuit.

A. *Field and Judicial Hyperactivity*

In a 2012 study, Ted Field conducted an empirical examination of whether the Federal Circuit judges were “hyperactive” as compared to the judges of other circuits.²³⁶ Field used the courts’ reversal rates to assess this claim. For the period studied, January 2010 to June 2010, Field reported an overall reversal rate of 27.7% when Rule 36 summary affirmances were not included and 24.0% when they were accounted for.²³⁷

A few observations on these numbers are necessary. First, Field’s study examined *all* decisions by the Federal Circuit during the relevant period, not just appeals arising from the district courts. Appeals from the district courts are economically significant and are the single largest category of appeals that the Federal Circuit historically hears (they amounted to 32% of the Federal Circuit’s judicial determinations in FY 2010).²³⁸ But they do not make up the majority of the court’s docket. The Federal Circuit also hears appeals from the Merit Systems Protection Board (23%), CFC (16%), the Court of Veterans Claims (10%), USPTO (6%), and a handful of other tribunals collectively comprising about 8% of judicial terminations for Financial Year 2010.²³⁹ Thus, Field’s reversal rate cannot be directly compared to the reversal rates discussed in Part I.

Fortunately, in addition to the data on reversal rates in appeals arising from the District Courts, the Federal Circuit also reports reversal rates for the court overall. For Financial Year 2010 (i.e., October 1, 2009 to September 30, 2010), the Federal Circuit’s own statistic is a reversal rate of 12%—substantially below both of Field’s reported reversal rates.

²³⁶ Field, *supra* note 37, at 738–40.

²³⁷ *Id.* Field does not provide the search algorithms employed, except in collecting the Rule 36 summary affirmances. *Id.* at 744–45.

²³⁸ *U.S. Court of Appeals for the Federal Circuit Statistical Tables for the Federal Judiciary*, U.S. COURTS (Dec. 31, 2010), <http://www.uscourts.gov/statistics-reports/statistical-tables-federal-judiciary-december-2010> [https://perma.cc/WRX7-C7SE].

²³⁹ *Id.*

Setting aside the ultimate question of whether the Federal Circuit judges are hyperactive, the difference between these two values raises the question of whether methodological differences might be involved. And if so, then what does that mean for the reversal rate that Field observed (or that the Federal Circuit reports)? And—given that Field’s use of the metric is to compare the Federal Circuit to other circuits—do those choices in study design matter?

The first difference between Field and the Federal Circuit’s data may be that they are not drawing from exactly the same data source. The Federal Circuit is using whatever internal data it collects on its decisions; Field is collecting from Lexis or Westlaw.²⁴⁰ This difference may matter if the Federal Circuit’s source consists of each individual docket (with a single “appeal” potentially consisting of multiple dockets)²⁴¹ as opposed to the court’s decisions themselves (with each decision constituting the relevant document). It is not clear whether this difference would produce a systematic bias, however, unless there is a relationship between reversals or affirmances and multiple dockets being decided in a single appeal.

A related issue is the question of time. Field collected the first 299 Federal Circuit decisions of 2010, a period running from January 2010 to June 2010.²⁴² In contrast, the Federal Circuit’s data period runs for an entire year. So the time periods, although overlapping, are not fully coextensive.

A bigger issue with time is the relatively short period of both of these sources. As illustrated in Figure 1, reversal rates exhibit temporal variation, even when viewed on an annual basis. Monthly or bi-annual reversal rates could exhibit even more variation. The consequence is that the reversal rate of the Federal Circuit over any given six-month period may not be representative of the court’s reversal rate over a longer time period.

The second potential difference between Field and the Federal Circuit’s data is in the different collection and filtering methodologies. As discussed above, the Federal Courts system uses its own internal procedures to assemble this data. Field collected decisions through a combination of electronic searches and manual review.²⁴³ As part of the manual review, Field excluded “issues for which the court did not articulate a standard of review.”²⁴⁴ The Article does not indicate how many Federal Circuit issues or decisions were discarded on the basis that the court did not articulate a standard of review. But if articulation of a standard of review is somehow

²⁴⁰ Field does not specify which cases came from which source. *See* Field, *supra* note 37, at 738–39 n.109 (“These cases were retrieved using either Westlaw or Lexis.”).

²⁴¹ *See, e.g.,* *Koopmann v. United States*, 454 F. App’x 890, 891 (Fed. Cir. 2010) (dismissing twenty-two separately docketed appeals).

²⁴² Field, *supra* note 37, at 738–40.

²⁴³ *Id.*

²⁴⁴ *Id.* at 739.

tied to reversals, the result could be a systematic exclusion of affirmances.

The third potential difference between the two methodologies is in the unit of record. The Federal Circuit data is on the level of the appeal. Field records data at a mixture of the per-issue and per-case level. He first reports reversals at the per-issue level.²⁴⁵ In order to account for Rule 36 summary affirmances, he then adds the summary affirmances, recorded at the per-case level.²⁴⁶ But just as appeals leading to opinions can involve multiple opinions, so too can appeals leading to summary affirmances. The consequence is that Field's methodology systematically undervalues the summary affirmances, producing an inflated reversal rate relative to that in Table B-8. Each issue that could be either affirmed or reversed in a written opinion would be counted, while multiple issues in a summary affirmance would be treated collectively as a single unit.

Aside from the Rule 36 affirmances, however, the difference in record unit helps to explain some of the difference between Table B-8 and Field. Table B-8 reports a result as a reversal *only* if the case is reversed in its entirety. In other words, an affirm-in-part or reverse-in-part result does not show up in the "Percent Reversed" column. The consequence, as discussed *supra* Part II.B, is to produce a lower apparent reversal rate.

In contrast, because Field records data on the per-issue level, a reversal on any issue will show up as a "reversal." This is the case even if the Federal Circuit affirmed on other issues in the opinion. Consequently, Field's methodology will inherently produce a higher reversal rate than that of the Federal Circuit.²⁴⁷

Of these sources of variation, however, the only ones that actually matter for purposes of the *overall* reversal rate²⁴⁸ in Field's study concern temporal variability, the unknown filtering methodologies Field employed, and the treatment of Rule 36 summary affirmances. This is because these study-design choices are ones that are likely to produce differences that limit the quality of the comparison between the Federal Circuit and the other circuits Field studied. Due to the short period studied (two months in the case of the other circuits),²⁴⁹ it is difficult to say with confidence that any of the circuits' reversal rates were representative. Indeed, it is entirely possible that the Federal Circuit's reversal rate over the sample period was merely the product of random chance. Second, if the filtering methodologies employed by Field systematically eliminated more Federal

²⁴⁵ *Id.* at 738–40.

²⁴⁶ *Id.* at 740 n.115, 744–45.

²⁴⁷ See *supra* note 230. This is because the Federal Circuit counts an affirmance-in-part, reversal-in-part as an affirmance whereas Field's methodology counts it as a reversal.

²⁴⁸ As discussed *infra* at Part II, Field's distributional allocation of the Rule 36 summary affirmances is questionable, thus creating issues for the more granular reversal rate numbers.

²⁴⁹ See Field, *supra* note 37, at 738–39 n.109 (explaining that the time period ran from January 2010 to February 2010).

Circuit decisions that affirmed then those of other circuits, the result would be to inflate the apparent reversal rate. And, as discussed, the effect of the combination of the per-case Rule 36 affirmances with the per-issue affirmances and reversals could be to inflate the apparent reversal rate, particularly if patent cases are more likely to involve multiple issues than the cases reviewed by other circuits. Instead, the data should be treated consistently (probably requiring that it all be treated on a per-case level).²⁵⁰

B. *Henry and Turner and the Federal Circuit's Impact on Patent Litigation*

A second example of how factors leading to inter-study variability can affect the interpretation of study results is that of Henry and Turner's 2006 *The Court of Appeals for the Federal Circuit's Impact on Patent Litigation*.²⁵¹ Henry and Turner assessed the Federal Circuit's impact on patent litigation by studying, among other characteristics of patent decisions, the rate at which the Federal Circuit affirmed district court decisions of invalidity and infringement.

While Henry and Turner do not report an overall affirmance rate for the periods studied in the article itself, the above framework can be used, in connection with their underlying data set, to assess the study-design-related impacts on their results.

Data Source: The researchers drew on the United States Patent Quarterly (U.S.P.Q.) for their set of district and appellate decisions, with the additional modification of tracking the case history of the reported district court opinions through Westlaw (to the extent possible).

Collection and Filtering: All records were collected by hand and subjected to various filtering processes to identify opinions that involve the issues of patent validity or infringement.²⁵² As a result, some appeals from the district courts are not included.

Record Unit: A "patent case," or, a particular patent in a particular case. The main effect of this design choice is in conjunction with the definition of affirmance, discussed below. Because the data is recorded on a unit that is smaller

²⁵⁰ This comparison does not express any opinion on the quality of the data or coding itself or the meaningfulness of comparison between the Federal Circuit and that of other circuits given differences in the types of appeals heard by those courts. Ted Sichelman, for example, argued that such comparisons are problematic unless the point of comparison is to "other complex cases in other circuits." See Sichelman, *supra* note 11, at 1189.

²⁵¹ See Henry & Turner, *supra* note 35, at 96, 111–14 (discussing results of their study and the reliance on the CAFC).

²⁵² *Id.* at 96.

than the level of the case itself, the effect of the definition of affirmance will be smaller than it would be if the unit had been the level of the case.

Parameter Definition: Henry and Turner define an affirmance as “an endorsement of the district court decision and a clear defeat for the appellant, while any other decision represents some unwillingness of the circuit court to endorse the district court decisions.”²⁵³ Thus, vacate and reversal-in-part decisions are not treated as affirmances.²⁵⁴

Of these study-design characteristics, the one with the greatest potential to impact the observed results is the use of the U.S.P.Q. as a data source. As discussed in Part II, the U.S.P.Q. is not a complete set of all patent appellate decisions as it does not include at least some non-precedential opinions or any summary affirmances under Federal Circuit Rule 36. This carries the potential for systematic exclusion of Rule 36 summary affirmances, an exclusion that will bias the apparent reversal rate upwards.²⁵⁵ When comparing affirmances with pre-Federal Circuit courts, the consequence of this choice is that the data may be missing a substantial category of affirmances from the Federal Circuit era—particularly if the Federal Circuit used these affirmance mechanisms more often than the pre-1982 circuit courts. These affirmances could change the shape of the distribution observed by Henry and Turner.²⁵⁶

An examination of the researchers’ dataset confirms that these study-design decisions probably do impact the reported reversal rates. The figure below incorporates into Figure 1 the annual reversal rates for appeals arising from the district courts based on the dataset that they subsequently released for general academic use (the “UGA Dataset”).²⁵⁷

²⁵³ *Id.* at 99.

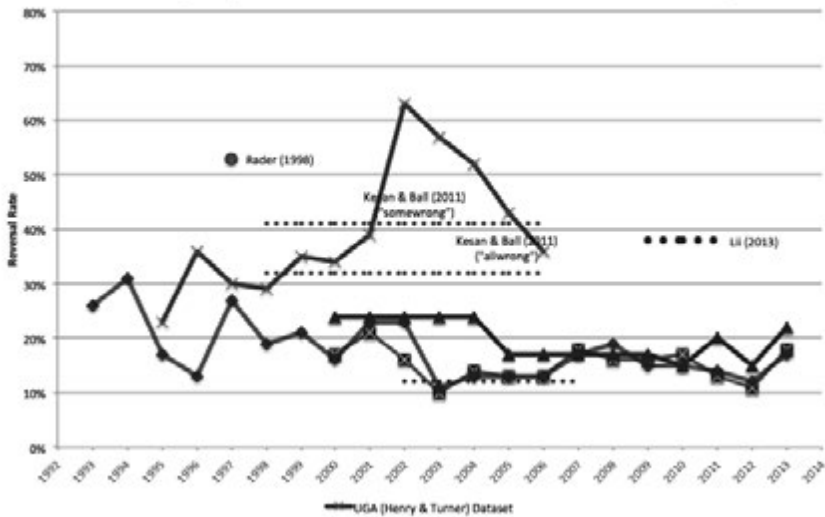
²⁵⁴ JOHN L. TURNER, THE UGA PATENT LITIGATION DATAFILE: CODEBOOK FOR MASTER LITIGATION DATA 14 (2013), <http://people.terry.uga.edu/jlturner/patentlitigationdata/PatentDataCodebook1March2013> [https://perma.cc/W4M4-MHM8] (separating affirmances from all other results).

²⁵⁵ See, e.g., Moore, *supra* note 53, at 239 (altering the study to account for Rule 36 affirmances, post-*Markman*).

²⁵⁶ As discussed above, Henry and Turner also worked from the district court dataset (obtained from the U.S.P.Q.) to identify the outcomes of those decisions that did not have a U.S.P.Q. citation for inclusion within the appellate dataset. However, such a process inherently selects for district court opinions that are more significant and thus potentially more likely to be appealed. In addition, since these decisions were selected for inclusion in the U.S.P.Q., they may be more likely to involve issues of law that the court is considering for the first time, and thus may be more likely to be reversed than the typical appealed case.

²⁵⁷ TURNER, *supra* note 254. The affirmance rates per year are provided in Appendix C. Categories used in this calculation were the “ACDate” and “Affirmed” fields.

Figure 6: Reported Rates at Which the Federal Circuit Reverses District Courts (Incorporates Reversal Rate Calculated from UGA Dataset)



As is apparent from the figure, the reversal rate calculated from the UGA dataset exceeds, and in some cases, vastly exceeds, the reversal rates from other reported sources. In addition to the systematic exclusion of summary affirmances and non-precedential decisions, some of this difference may be due to the record unit that the study used in conjunction with the definition of affirmance (a definition that is similar to the “reversal-in-part is a reversal” definition discussed *supra* Part II.B). Some of it may also be due to the criteria for inclusion, which could select for issues that are more conducive to reversal than other issues, such as “housekeeping” decisions by the district court.²⁵⁸

CONCLUSION: TOWARD MORE MEANINGFUL EMPIRICAL STUDIES OF JUDICIAL OPINIONS

Here, I offer a few closing thoughts on empirical studies of judicial opinions generally.

First, variation in study results due to study design is not necessarily undesirable as long as the study methodology is clear and replicable. Different study designs allow for different perspectives on the underlying data. For example, the three definitions of “reversal” that scholars have developed in the context of claim construction allow for a deeper understanding of what is going on at the Federal Circuit. Methodological diversity can thus be a positive characteristic in the literature.

Yet, while methodological diversity can provide value, it is *critical* that key methodological decisions be explained and identified. Transparent

²⁵⁸ Henry & Turner, *supra* note 35, at 96.

methodologies are necessary for understanding sources of inter-study variation, for replicating studies, and for interpreting their results. Nor is it clear that all methodological differences are beneficial. In some cases, there are clear best practices, such as the inclusion of Rule 36 summary affirmances, the recognition of temporal variation, and the use of procedures to minimize coder subjectivity.

One possibility for obtaining methodological clarity and transparency is for researchers to develop and make available a detailed coding manual that explains how data was collected and coding options. Providing such a manual allows authors to not get bogged down in methodological detail in articles reporting on results, but to still ensure that future scholars can understand the methodology employed. In addition, a coding manual can be written in a way that would be inappropriate for an article but that provides for clearer explanation in a format more conducive for the information it provides.

Along with greater transparency of methodology should be greater transparency of study data. Ideally, the dataset analyzed in the paper itself should be archived and made available for future researchers—as Henry and Turner did. Not only does this allow for a higher level of scrutiny of empirical study results, but it prevents valuable, labor-intensive datasets from being lost through obsolescence of one’s individual data stores and accidental deletion. Indeed, this is a major limitation of “data on request” policies (i.e., when a researcher indicates that they will provide their data upon request by another researcher): such a policy does nothing to guard against accidental deletion of the data.

Empirical studies in patent law have made great leaps forward over the past two decades. The field is at a pivotal moment, with the recognition of problems that only become apparent with maturity. For meaningful growth to continue, legal empiricists must take the next step towards adding quality data to our understanding of patent law opinions by ensuring the methodological clarity and reproducibility of results that characterize high quality studies in well-developed empirical fields.

APPENDIX A

Title	Author	Year	Citation
<i>What the Federal Circuit Has Done and How Often: Statistical Study of the CAFC Patent Decisions — 1982 to 1988</i>	Coolley	1989	71 J. PAT. & TRADEMARK OFF. SOC'Y 385
<i>The Impact of the Creation of the Court of Appeals for the Federal Circuit on the Availability of Preliminary Injunctive Relief Against Patent Infringement</i>	Morrison	1990	23 IND. L. REV. 169
<i>A Statistical Look at the Federal Circuit's Patent Decisions: 1982-1994</i>	Dunner, Jakes & Karceski	1995	5 FED. CIR. B.J. 151
<i>Shedding Light on Chevron: An Empirical Study of The Chevron Doctrine in the U.S. Courts of Appeals</i>	Kerr	1998	15 YALE J. ON REG. 1
[Judge Rader's Dissent in <i>Cybor</i>]	Rader	1998	138 F.3d 1448
<i>How Federal Circuit Judges Vote in Patent Validity Cases</i>	Allison & Lemley	2000	27 FLA. ST. U. L. REV. 744
<i>Judges, Juries, and Patent Cases — An Empirical Peek Inside the Black Box</i>	Moore	2000	99 MICH. L. REV. 365
<i>Empirical Analysis of the Federal Circuit's Claim Construction Trends</i>	Chu	2001	16 BERKELEY TECH. L.J. 1075
<i>E-Obviousness</i>	Lunney, Jr.	2001	7 MICH. TELECOMM. & TECH. L. REV. 363
<i>Are District Court Judges Equipped to Resolve Patent Cases?</i>	Moore	2001	15 HARV. J.L. & TECH. 1
<i>Toward a Cautious Approach to Obeisance: The Role of Scholarship in Federal Circuit Patent Law Jurisprudence</i>	Nard	2002	39 HOUS. L. REV. 667
<i>Patent Law, the Federal Circuit, and the Supreme Court: A Quiet Revolution</i>	Lunney, Jr.	2003	11 SUP. CT. ECON. REV. 1
<i>The Judicial Doctrine of Equivalents</i>	Paul	2003	17 HARV. J.L. & TECH. 247
<i>Patent Claim Construction in the Trial Courts: A Study Showing the Need for Clear Guidance from the Federal Circuit</i>	Zidel	2003	33 SETON HALL L. REV. 711
<i>An Empirical Analysis of the Patent Court</i>	Landes & Posner	2004	71 U. CHI L. REV. 111
<i>The Selection Effects (and Lack Thereof) in Patent Litigation: Evidence from Trials</i>	Marco	2004	4 TOPICS IN ECON. ANALYSIS & POL'Y 1
<i>Please Ignore This Case: An Empirical Study of Nonprecedential Opinions in the Federal Circuit</i>	Shaw	2004	12 GEO. MASON L. REV. 1013
<i>Is the Federal Circuit Succeeding? An Empirical Assessment of Judicial Performance</i>	Wagner & Petherbridge	2004	152 U. PA. L. REV. 1105

<i>The Court of Appeals for the Federal Circuit's Impact on Patent Litigation</i>	Henry & Turner	2005	35 J. LEGAL STUD. 85
<i>Who Wins Patent Infringement Cases?</i>	Janicke & Ren	2005	34 AIPLA Q.J. 1
<i>The Proven Key: Roles and Rules for Dictionaries at the Patent Office and the Courts</i>	Miler & Hilsenteger	2005	54 AM. U. L. REV. 829
<i>Empirical Statistics on Willful Patent Infringement</i>	Moore	2005	14 FED. CIRCUIT B.J. 227
<i>Markman Eight Years Later: Is Claim Construction More Predictable?</i>	Moore	2005	9 LEWIS & CLARK L. REV. 231
<i>The Measure of the Doubt: Dissent, Indeterminacy, and Interpretation at the Federal Circuit</i>	Lefstin	2006	58 HASTINGS L.J. 1025
<i>The (Unnoticed) Demise of the Doctrine Of Equivalents</i>	Allison & Lemley	2007	59 STAN. L. REV. 955
<i>Nonobviousness and the Federal Circuit: An Empirical Analysis of Recent Case Law</i>	Cotropia	2007	82 NOTRE DAME L. REV. 911
<i>The Federal Circuit and Patentability: An Empirical Assessment of the Law of Obviousness</i>	Petherbridge & Wagner	2007	85 TEX. L. REV. 2051
<i>A Survey of Post-Phillips Claim Construction Cases</i>	Saunders	2007	22 BERKELEY TECH. L.J. 215
<i>The Economics of a Centralized Judiciary: Uniformity, Forum Shopping, and the Federal Circuit</i>	Atkinson, Marco, & Turner	2008	52 J.L. & ECON. 411
<i>Assessing Bias in Patent Infringement Cases: A Review of International Trade Commission Decisions</i>	Hahn & Singer	2008	21 HARV. J.L. & TECH. 457
<i>Does Practice Make Perfect? An Examination of Congress's Proposed District Court Patent Pilot Program</i>	Olson	2008	55 UCLA L. REV. 745
<i>The Claim Construction Effect</i>	Petherbridge	2008	15 MICH. TELECOMM. & TECH. L. REV. 215
<i>Practice Makes Perfect? An Empirical Study of Claim Construction Reversal Rates in Patent Cases</i>	Schwartz	2008	107 MICH. L. REV. 223
<i>The Well-Pleaded Complaint Rule and Jurisdiction over Patent Law Counterclaims: An Empirical Assessment of Holmes Group And Proposals For Improvement</i>	Chen	2009	8 NW. J. TECH. & INTELL. PROP. 94
<i>Should the United States Designate Specialist Patent Trial Judges? An Empirical Analysis of H.R. 628 in Light of the English Experience and the Work of Professor Moore</i>	Gitter	2009	10 COLUM. SCI. & TECH. L. REV. 169

<i>Effects of the Federal Circuit Judges on Hatch-Waxman Litigation</i>	Masar, Martin S. III	2009	19 DEPAUL J. ART, TECH. & INTELL. PROP. L. 315
<i>Expertise, Experience, and Ideology on Specialized Courts: The Case of the Court of Appeals for the Federal Circuit</i>	Miller & Curry	2009	43 LAW & SOC'Y REV. 839
<i>Patent Law Uniformity?</i>	Petherbridge	2009	22 HARV. J.L. & TECH. 421
<i>Courting Specialization: An Empirical Study of Claim Construction Comparing Patent Litigation Before Federal District Courts and the International Trade Commission</i>	Schwartz	2009	50 WM. & MARY L. REV. 1699
<i>Patent Litigation 101: Empirical Support for the Patent Pilot Program's Solution to Increase Judicial Experience in Patent Law</i>	Shartzter	2009	18 FED. CIR. B.J. 191
<i>The TSM Test Is Dead! Long Live The TSM Test! The Aftermath of KSR, What Was All the Fuss About?</i>	Simic	2009	37 AIPLA Q.J. 227
<i>Determining Uniformity Within the Federal Circuit by Measuring Dissent and En Banc Review</i>	Cotropia	2010	43 LOY. L.A. L. REV. 801
<i>Unrepresentative Randomization: An Empirical Study of Judging Panels of USPTO Appeals to the CAFC</i>	Froats	2010	19 TEX. INTELL. PROP. L.J. 79
<i>Not So Technical: An Analysis of Federal Circuit Patent Decisions Appealed from the ITC</i>	Lance	2010	17 MICH. TELECOMM. & TECH. L. REV. 243
<i>A Definite Claim on Claim Indefiniteness: An Empirical Study of Definiteness Cases of the Past Decade with a Focus on the Federal Circuit and the Insolubly Ambiguous Standard</i>	Laser	2010	10 CHI.-KENT J. INTELL. PROP. 25
<i>Controlling the "Plague": Reforming the Doctrine of Inequitable Conduct</i>	Mammen	2010	24 BERKELEY TECH. L.J. 1329
<i>A New Look at Patent Quality: Relating Patent Prosecution to Validity</i>	Mann	2010	Working paper
<i>An Empirical Study of The Effect of KSR v. Teleflex on the Federal Circuit's Patent Validity Jurisprudence</i>	Mojibi	2010	20 ALB. L.J. SCI. & TECH. 559
<i>On the Decline of the Doctrine of Equivalents</i>	Petherbridge	2010	31 CARDOZO L. REV. 1371
<i>Pre-Markman Reversal Rates</i>	Schwartz	2010	43 LOY. L.A. L. REV. 1073
<i>Myths of (Un)Certainty at the Federal Circuit</i>	Sichelman	2010	43 LOY. L.A. L. REV. 1161

<i>Toward Certainty and Uniformity in Patent Infringement Cases After Festo and Markman: A Proposal for a Specialized Patent Trial Court with a Rule of Greater Deference</i>	Wallace	2010	77 S. CAL. L. REV. 1383
<i>Investigating Patent Law's Presumption Of Validity, PAn Empirical Analysis of How Unconsidered Evidence and Evidentiary Standards Affect Jury Verdicts</i>	Chatlynne, Kenny & Watkins	2011	2011 CARDOZO L. REV. DE NOVO 46
<i>Judicial Experience and the Efficiency and Accuracy of Patent Adjudication: An Empirical Analysis of the Case for a Specialized Patent Trial Court</i>	Kesan & Ball	2011	24 HARV. J.L. & TECH. 393
<i>Raising the Bar For Nonobviousness: An Empirical Study of Federal Circuit Case Law Following KSR</i>	Nock & Gadde	2011	20 FED. CIR. B.J. 369
<i>What Are the Sources of Patent Inflation? An Analysis of Federal Circuit Patentability Rulings</i>	Ouellette	2011	121 YALE L.J. ONLINE 347
<i>The Federal Circuit and Inequitable Conduct: An Empirical Assessment</i>	Petherbridge, Rantanen & Mojibi	2011	84 S. CAL. L. REV. 1293
<i>The Impact of Dickinson v. Zurko on Federal Circuit Review of USPTO Board Decisions: An Analytic and Empirical Analysis</i>	Samuels & Samuels	2011	20 FED. CIRCUIT B.J. 665
<i>Claim Construction and Technical Training: An Empirical Study of the Reversal Rates of Technically Trained Judges in Patent Claim Construction Cases</i>	Schuster	2011	29 QUINNIPIAC L. REV. 887
<i>Explaining the Demise of the Doctrine of Equivalents</i>	Schwartz	2011	26 BERKELEY TECH. L.J. 1157
<i>Legal Scholarship and the United States Court of Appeals for the Federal Circuit: An Empirical Study of a National Circuit</i>	Schwartz & Petherbridge	2011	26 BERKELEY TECH. L.J. 1561
<i>The Use of Legal Scholarship by the Federal Courts of Appeals: An Empirical Study</i>	Schwartz & Petherbridge	2011	96 CORNELL L. REV. 1345
<i>Acting Like an Administrative Agency: The Federal Circuit En Banc</i>	Vacca	2011	76 MO. L. REV. 733
<i>Secondary Considerations in Nonobviousness Analysis: The Use of Objective Indicia Following KSR v. Teleflex</i>	Thomas	2011	86 N.Y.U. L. REV. 2070
<i>"Judicial Hyperactivity" in the Federal Circuit: An Empirical Study</i>	Field	2012	46 U.S.F. L. REV. 721
<i>Injunctions as More (or Less) than "Off Switches": Patent-Infringement Injunctions' Scope</i>	Golden	2012	90 TEX. L. REV. 1399

<i>Beyond Patents: The Supreme Court's Evolving Relationship with the Federal Circuit</i>	Kazhdan	2012	94 J. PAT. & TRADEMARK OFF. SOC'Y 275
<i>Not So Obvious After All: Patent's Nonobviousness Requirement, KSR, and the Fear of Hindsight Bias</i>	Lunney, Jr. & Johnson	2012	47 GA. L. REV. 41
<i>Informal Deference: A Historical, Empirical, and Normative Analysis of Patent Claim Construction</i>	Anderson & Menell	2013	108 NW. U. L. REV. 1
<i>Technically Speaking, Does It Matter? An Empirical Study Linking the Federal Circuit Judges' Technical Backgrounds to How They Analyze the Section 112 Enablement and Written Description Requirements</i>	Barnes	2013	88 CHI.-KENT L. REV. 971
<i>What Close Cases and Reversals Reveal About Claim Construction at the Federal Circuit</i>	Krause & Auyang	2013	12 J. MARSHALL REV. INTELL. PROP. L. 583
<i>Shopping for Reversals: How Accuracy Differs Across Patent Litigation Forums</i>	Lii	2013	12 CHI.-KENT J. INTELL. PROP. 31
<i>Experts Judging Experts: The Role of Expertise in Reviewing Agency Decision Making</i>	Miller & Curry	2013	38 LAW & SOC. INQUIRY 55
<i>The Federal Circuit's New Obviousness Jurisprudence: An Empirical Study</i>	Rantanen	2013	16 STAN. TECH. L. REV. 709
<i>Ten Years of Inter Partes Patent Reexamination Appeals: An Empirical View</i>	Rogers	2013	29 SANTA CLARA COMPUTER & HIGH TECH. L.J. 305
<i>Is Patent Claim Interpretation Review Deference or Correction Driven?</i>	Cotropia	n/a	B.Y.U. L. R. (forthcoming)
<i>If You Can't Beat 'Em, Join 'Em? How Sitting by Designation Affects Judicial Behavior</i>	Lemley & Miller	2014	Stanford Public Law Working Paper No. 2449349
<i>"Fuzzy" Software Patent Boundaries and High Claim Construction Reversal Rates</i>	Miller	2014	17 STAN. TECH. L. REV. 809
<i>Disuniformity</i>	Rantanen & Petherbridge	2014	66 FLA. L. REV. 2007
<i>Understanding The Realities of Modern Patent Litigation</i>	Allison, Lemley, & Schwartz	2014	92 TEX. L. REV. 1769
<i>Fee-Shifting and Octane Fitness: An Empirical Approach to Understanding "Exceptional"</i>	Jiam	2015	30 BERKELEY TECH. L.J. 611

APPENDIX B.1

Appendix B.1: Overall Reversal Rates of the Federal Circuit in Appeals Arising from the District Courts													
	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
<i>UGA (Henry & Turner) Dataset</i>			23%	36%	30%	29%	35%	34%	39%	63%	57%	52%	43%
<i>Table B-8 (Statistical Tables)</i>								17%	21%	16%	10%	14%	13%
<i>Table B-8 (Reports)</i>	26%	31%	17%	13%	27%	19%	21%	16%	23%	23%	11%	13%	13%
<i>PatStats</i>								24%	24%	24%	24%	24%	17%
<i>Lii (2013)</i>													
<i>Field (2012)</i>													
<i>Kesan & Ball (2011)</i> <i>("somewrong")</i>						41%	41%	41%	41%	41%	41%	41%	41%
<i>Kesan & Ball (2011)</i> <i>("alhwrong")</i>						32%	32%	32%	32%	32%	32%	32%	32%
<i>Shartzer (2009)</i>										12%	12%	12%	12%
<i>Rader (1998)</i>					53%								
	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005

2006	2007	2008	2009	2010	2011	2012	2013
36%							
13%	18%	16%	16%	17%	13%	11%	18%
13%	17%	19%	15%	15%	14%	12%	17%
17%	17%	17%	17%	15%	20%	15%	22%
			38%	38%	38%		
				38%			
41%							
32%							
12%	12%						
2006	2007	2008	2009	2010	2011	2012	2013

APPENDIX B.2

	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2005.5	2006	2007	2008	2009	2010	2011	2012	2013
<i>Corrobia (2015)</i>																	30%	27%	25%	23%
<i>Miller (2014)</i>								41%	39%	42%	39%		23%	28%	34%	18%	30%	32%	25%	
<i>Anderson et al. (2013) (per term)</i>						25%	36%	41%	40%	44%	31%		22%	25%	31%	17%	22%	20%		
<i>Anderson (2013)</i>						32%	37%	43%	44%	45%	41%		26%	29%	33%	21%	32%	28%		
<i>Schwartz (2008)</i>	19%	41%	33%	29%	38%	33%	38%	40%	49%	48%	46%		22%	31%	38%					
<i>Saunders (2007)</i>												40%								
<i>Moore (2005)</i>		38%	38%	38%	38%	38%	38%	38%	38%											
<i>Zidel (2003)</i>							42%													
<i>Bender (2001)</i>		41%	41%	41%	41%	41%														
<i>Chu (2001)</i>				44%	44%	44%														
<i>Moore (2001)</i>		33%	33%	33%	33%	33%														
<i>Rader (1998)</i>	37%	37%	37%																	
	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2005.5	2006	2007	2008	2009	2010	2011	2012	2013

APPENDIX C

Reversal Rates Calculated Based on UGA Dataset				
	Affirmed	Not Affirmed	Total	% Not Affirmed
1995	49	15	64	23%
1996	45	25	70	36%
1997	65	28	93	30%
1998	55	22	77	29%
1999	55	30	85	35%
2000	50	26	76	34%
2001	91	58	149	39%
2002	36	60	96	63%
2003	49	64	113	57%
2004	54	58	112	52%
2005	67	51	118	43%
2006	43	24	67	36%

