 Printed in U.S.A.
Vol. 119, No. 1

SETTLEMENT AS CONSTRUCT: DEFINING AND COUNTING PARTY RESOLUTION IN FEDERAL DISTRICT COURT

Charlotte S. Alexander, Nathan Dahlberg & Anne M. Tucker

ABSTRACT—Most civil cases settle. Yet generating a definitive settlement rate presents complex definitional and empirical problems, both in what should count as a settlement and how to count it. This Essay makes three contributions to better understanding and defining settlement. First, we propose a flexible, empirically informed, operationalizable definition of settlement as party resolution. Second, we exploit a new federal litigation data source to count party resolutions using machine learning models trained on 11 million docket sheet entries. Third, we offer new findings on party resolution frequency and distribution in the federal courts. Settlement is more widely and differently deployed than previously understood. We uncovered almost 40,000 additional party resolutions that were missing from the main existing source of administrative data on federal litigation. We also observe party resolution appearing alongside other dispositive events, functioning as a trimming device to drop parties, claims, or both as a lawsuit proceeds. We conclude by mapping directions for future work, integrating our settlement findings into a larger discussion of litigation and the courts' role in achieving resolution, whether by a judge, jury, or the parties themselves.

AUTHORS—Charlotte S. Alexander: Professor of Law and Ethics, Scheller College of Business, Georgia Institute of Technology; Co-Principal Investigator, SCALES–OKN Project. Nathan Dahlberg: Founder, Docket Analyzer; former Data Scientist and Deep Learning Engineer, SCALES–OKN Project. Anne M. Tucker: Professor of Law, Georgia State University College of Law; Research Team Member, SCALES–OKN Project.

The authors contributed equally to this work. An early version of this research benefited from feedback from the Smart Law Hub faculty at HEC Paris and at the Conference on Settlement at the University of Illinois College of Law's Kimball R. and Karen Gatsis Anderson Center for Advocacy and Professionalism. We thank Jennifer Robbennolt and Verity Winship for organizing the conference and others who have provided valuable comments: Christina Boyd, Dan Klerman, Alexandra Lahav, Margo

Schlanger, and Nancy Welsh. The Civil Litigation Events Ontology Working Group of the larger SCALES–OKN project was instrumental in developing the litigation events ontology and classification process described here, including Chris Cotropia, Dave Schwartz, and Adam Pah. We also thank Kristin Garri of the Federal Judicial Center for her assistance with our IDB questions. Finally, we thank Sahil Dharamrup, Reid Hanson, Jim Hicks, and Kellyann Malone for their able research support, along with Pam Brannon in the Georgia State University College of Law Library. This work was supported by funding from the National Science Foundation Convergence Accelerator grant program, A7123: Transform the Transparency and Accessibility of Court Records.

INTRODUCTION

How many civil cases settle? Many accounts put the rate as high as 95%.[1] But this is wrong. It is litigation lore that originated in a set of studies

[1] In a Forbes personal injury suit guide, Forbes cites a U.S. Department of Justice study claiming that only 4% to 5% of tort cases go all the way to trial. This study is cited widely across online legal memoranda. Jeffrey Johnson & Adam Ramirez, *Personal Injury Settlement Amounts Examples (2024 Guide)*, FORBES ADVISOR (Sept. 22, 2022, 12:54 PM) (citing THOMAS H. COHEN, U.S. DEP'T OF JUST., BUREAU OF JUST. STATS., NCJ 228129, TORT BENCH AND JURY TRIALS IN STATE COURTS 2005 (Nov. 2009), https://bjs.ojp.gov/content/pub/pdf/tbjtsc05.pdf. [https://perma.cc/R929-

that subtracted the number of trials from the total number of cases filed; they called the remainder "settlement."[2] Scholars have since made substantial progress in defining and counting settlement more precisely, whittling away at and complicating the mythical 95% figure.[3]

Yet tallying settlement outcomes continues to present complex definitional and empirical problems regarding what should count as a settlement and how to count it.[4] Faced with this complexity, some scholars

M7YG]), https://www.forbes.com/advisor/legal/personal-injury/personal-injury-settlement-amounts [https://perma.cc/EH9X-UFPV]; *see also Litigation Risk Assessment*, PERKINS COIE LLP, https://www.perkinscoie.com/en/litigation-risk-assessment-1.html [https://perma.cc/XU8P-JJ2B] ("Well over 95% of civil litigation cases settle."); *Settlements*, N.Y. STATE UNIFIED CT. SYS., https://nycourts.gov/Courthelp/goingtocourt/settlements.shtml [https://perma.cc/Y28P-X6ZG] ("Most court cases are settled."); *How Courts Work*, AM. BAR ASS'N (Sept. 9, 2019), https://www.americanbar.org/groups/public_education/resources/law_related_education_network/how_courts_work/cases_settling/ [https://perma.cc/PJQ6-FJMH] ("Relatively few lawsuits ever go through the full range of procedures and all the way to trial. Most civil cases are settled by mutual agreement between the parties."); *see also What Percentage of Lawsuits Settle Before Trial? What Are Some Statistics on Personal Injury Settlements?*, LAW DICTIONARY, https://thelawdictionary.org/article/what-percentage-of-lawsuits-settle-before-trial-what-are-some-statistics-on-personal-injury-settlements/ [https://perma.cc/Z7Q2-MPZ6] ("According to the most recently-available statistics, about 95 percent of pending lawsuits end in a pre-trial settlement.").

[2] John Barkai, Elizabeth Kent & Pamela Martin, *A Profile of Settlement*, 42 CT. REV. 34, 34–35 (2006).

[3] Past empirical studies of settlement have examined termination rates in patent cases, employment cases, or within specific jurisdictions, finding settlement rates that range from 20% to 84% depending on the case type and geography. *Id*. at 35 (documenting settlement rates for cases in Hawaii). Another study focused on settlement rates in corporate veil piercing cases and found that 77% were settled (452 out of 585 cases). Christina L. Boyd & David A. Hoffman, *Litigating Toward Settlement*, 29 J.L., ECON. & ORG. 898, 909 (2012). In another study focused on employment arbitration, 59.1% of cases settled. Alexander J.S. Colvin, *An Empirical Study of Employment Arbitration: Case Outcomes and Processes*, 8 J. EMPIRICAL LEGAL STUD. 1, 6 (2011).

[4] Settlement scholarship has noted continued problems in defining and counting settlements. Theodore Eisenberg & Charlotte Lanvers, *What Is the Settlement Rate and Why Should We Care?*, 6 J. EMPIRICAL LEGAL STUD. 111, 114 (2009) (noting the role of difficult judgment calls in defining and counting settlement); Gillian K. Hadfield, *Where Have All the Trials Gone? Settlements, Nontrial Adjudications, and Statistical Artifacts in the Changing Disposition of Federal Civil Cases*, 1 J. EMPIRICAL LEGAL STUD. 705, 716–17 (2004) (noting the murkiness of settlement and how hard it is to observe reliably); Tom Vacek & Frank Schilder, *A Sequence Approach to Case Outcome Detection*, *in* SIXTEENTH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW: PROCEEDINGS OF THE CONFERENCE 209, 210–11 (2017) (noting that the multiplicity of claims and potential outcomes in a single case complicate settlement research); J.J. Prescott & Kathryn E. Spier, *A Comprehensive Theory of Civil Settlement*, 91 N.Y.U. L. REV. 59, 64–65, 67 (2016) (noting that settlements can be full or partial); Boyd & Hoffman, *supra* note 3, at 901–02 (noting the overlap between settlement agreements and consent decrees in recording party resolutions); Christopher A. Cotropia, Jay P. Kesan & David L. Schwartz, *Heterogeneity Among Patent Plaintiffs: An Empirical Analysis of Patent Case Progression, Settlement, and Adjudication*, 15 J. EMPIRICAL LEGAL STUD. 80, 97 (2018) (noting the interchangeability of settlement agreements and voluntary dismissal in recording party resolutions).

have called off the "futile" search for a "single settlement rate, or even a single reasonable definition of settlement."[5]

We offer both an operationalizable definition of settlement as party resolution and a reliable settlement rate for federal civil lawsuits. This definition was generated by using machine learning models to analyze millions of federal court docket sheet entries made newly accessible by the Systematic Content Analysis of Litigation EventS Open Knowledge Network (SCALES or SCALES–OKN).[6] Deploying our definition, we find that party resolution occurred in approximately 57% of the cases in our dataset, resolving the case in full or in part.[7]

Although not as common as the 95% figure above, our work suggests that party resolution is utilized more widely and differently than previously understood. Specifically, our models uncovered almost 40,000 party resolutions that are missing from the main existing source of administrative data on federal litigation, the Federal Judicial Center's Integrated Database (IDB).[8] Moreover, we find that party resolution often appears alongside other dispositive events in the life of a case and is used as a trimming device to drop parties or claims as litigation proceeds.[9] Many of these partial party resolutions are lost when researchers code only one dispositive event per case or rely on the IDB's single-disposition data structure. Our approach thus enables the study of the hydraulic relationship between party resolution and various court-driven adjudications.

---

[5] Eisenberg & Lanvers, *supra* note 4, at 146.

[6] This Essay is limited to federal civil litigation due to the data available. As explained further below, the federal courts have a single electronic court records system, enabling aggregation of court records and data across the federal system. In contrast, each state and county maintains its own electronic court records system, with varying degrees of standardization and public access. This fragmentation makes the assembly of a single state court litigation dataset across jurisdictions extremely costly and difficult.

[7] *Infra* Section IV.A.

[8] *Integrated Database (IDB)*, FED. JUD. CTR., https://www.fjc.gov/research/idb [https://perma.cc/97PH-L9FK]; *see also Facts About FJC's Integrated Database*, FREE L. PROJECT, https://free.law/idb-facts [https://perma.cc/7DHU-SEJE].

[9] The Institute for the Advancement of the American Legal System found in its 2009 work that many cases settle shortly after denial of a motion to dismiss or a motion for summary judgment, suggesting that the parties look to the court for answers that affect settlement questions and that denying motions to dismiss and for summary judgment provides valuable information to the parties about the strength of their respective claims and defenses. INST. FOR THE ADVANCEMENT OF THE AM. LEGAL SYS., CIVIL CASE PROCESSING IN THE FEDERAL DISTRICT COURTS 7 (2009); *see also* Cotropia et al., *supra* note 4, at 85 (studying litigation outcomes in patent cases); Eugene F. Lynch & Lawrence C. Levine, *The Settlement of Federal District Court Cases: A Judicial Perspective*, 67 OR. L. REV. 239, 243–46 (1988) (describing a judge's role in settlement); Marc S. Galanter, *The Federal Rules and the Quality of Settlements: A Comment on Rosenberg's, the Federal Rules of Civil Procedure in Action*, 137 U. PA. L. REV. 2231, 2232 (1989) (describing settlement as "intimately and inseparably entwined" with adjudication).

This sort of empirical analysis of civil settlements and other litigation events matters beyond the exercise of counting jellybeans in a jar. Settlement is a common litigation outcome and is heavily favored by the Federal Rules of Civil Procedure (FRCP) through features such as the expansive role of discovery.[10] As legal historian John H. Langbein recounts, when the FRCP were revised in 1983, "'facilitating settlement' of the case [became] an express objective of pretrial conferences. Managerial judging thus expanded 'from a set of techniques for narrowing issues to a set of techniques for settling cases.'"[11] For example, the FRCP require pretrial conferences to promote settlement in Rule 16, facilitate class action settlements in Rule 23(e), provide for voluntary dismissals in Rule 41(a), and enable offers of judgment in Rule 68. Federal Rule of Evidence 408(a) similarly excludes compromises or settlement offers as evidence of an admission of fault.[12] Our system of federal civil procedure has thus been designed to achieve settlement wherever possible, but we know very little about settlement's frequency, distribution, and characteristics.

Further, settlement is often value-enhancing for the parties.[13] As Professors Theodore Eisenberg and Charlotte Lanvers concluded, "[s]ettlement is not only the modal litigation outcome, it is also the most common *successful* outcome for plaintiffs."[14] They argue that settlement no longer happens as an alternative disposition in the shadow of trial. Settlement is not just the most frequent outcome, it is the preferred outcome.[15] This sentiment is bolstered by recent empirical work on lay persons' perceptions of settlement, where respondents preferred settlement to trial because it is "cheaper and quicker," avoiding the costs, time delays, and uncertainties of going to court.[16]

---

[10] John H. Langbein, *The Disappearance of Civil Trial in the United States*, 122 YALE L.J. 522, 548 (2012) (describing the views of 1,938 Federal Advisory Committee on Civil Rules members favoring discovery practices over pleadings because the quantum of proof, or lack thereof, drives settlement).

[11] *Id.* at 559 (citations omitted).

[12] FED. R. CIV. P. 16, 23(e), 41(a), 68; FED. R. EVID. 408(a).

[13] Prescott & Spier, *supra* note 4, at 78 (arguing that settlements add value by "reducing adjudication costs, mitigating losses due to risk, and maximizing ex ante expected returns"). *But cf.* Richard A. Posner, *An Economic Approach to Legal Procedure and Judicial Administration*, 2 J. LEGAL STUD. 399, 429 (1973) (arguing that settlement introduces error costs because, without trials, the relative efficiency of different substantive legal rules cannot be evaluated where settlements deprive future litigants of the basis to estimate likelihoods of success and damages ranges).

[14] Eisenberg & Lanvers, *supra* note 4, at 112 (emphasis added); *see also* Katheryn E. Spier, *Litigation*, *in* 1 HANDBOOK OF LAW AND ECONOMICS 281 (A. Mitchell Polinsky & Steven Shavell eds., 2007) (describing the plaintiff as "better off through settlement than she would be going to trial" because if not, the plaintiff would "simply refuse to settle and go to trial instead").

[15] Eisenberg & Lanvers, *supra* note 4, at 112–13.

[16] Jennifer K. Robbennolt, Jessica Bregant & Verity Winship, *Settlement Schemas: How Laypeople Understand Civil Settlement*, 20 J. EMPIRICAL LEGAL STUD. 488, 513–14 (2023).

The parties' and courts' preference for compromise over litigation drives civil cases towards settlement, but much remains murky about cases' pathways and their final destinations. Resolving questions such as what counts as settlement, how often it occurs, and which antecedent events promote or forestall settlement would inform a litigant's journey.[17] Settlement rates also inform our understanding of how well the FRCP function to accomplish their various goals and how the procedural hand shapes the substantive outcomes of cases.[18] Last but not least, settlement data provide the building blocks for answering equity-based questions about who wins, who loses, and how often in federal civil litigation.

Our work builds on and extends the empirical literature developed around these questions.[19] We offer our empirically informed,

[17] Settlement data may also add insight into the selection effects in litigation studies.

[18] For articles discussing predictors of settlement, see, for example, Kevin M. Clermont & Theodore Eisenberg, *Litigation Realities*, 88 CORNELL L. REV. 119, 138 (2002), which describes selection effects in settlement outcomes whereby disputes and cases favoring one party are more likely to settle, but hard cases "falling close to the applicable decisional criterion" do not settle because parties disagree about the predicted outcomes. *See also* Christina L. Boyd, *She'll Settle It?*, 1 J.L. & CTS. 193, 213 (2013) (finding that a judge's gender matters because "female judges are more successful at fostering settlements than men"); Theodore Eisenberg & Henry S. Farber, *The Litigious Plaintiff Hypothesis: Case Selection and Resolution*, 28 RAND J. ECON. S92, S110–11 (1997) (finding that individual defendants are more likely to settle than organization defendants); Orley Ashenfelter, Theodore Eisenberg & Stewart J. Schwab, *Politics and the Judiciary: The Influence of Judicial Background on Case Outcomes*, 24 J. LEGAL STUD. 257, 281 (1995) (finding that neither judge characteristic nor political party was predictive of trial or settlement outcomes); *cf.* Barkai et al., *supra* note 2, at 36 (finding no clear predictors of settlement versus nonsettlement from the facts of a case).

[19] Empirical accounts of settlement first appeared in the mid-2000s, documenting settlement rates and situating settlements in the context of changing federal civil procedure and litigation landscapes. *See, e.g.*, Marc S. Galanter, *The Vanishing Trial: An Examination of Trials and Related Matters in Federal and State Courts*, 1 J. EMPIRICAL LEGAL STUD. 459, 460–61 (2004) (tracing the decline in the portion of cases terminated by trial and the absolute number of trials in various American judicial fora); Hadfield, *supra* note 4, at 711–12 (investigating whether the trend of decreasing trials in federal courts is a reflection of an increase in private settlements or an increase in public non-trial adjudication); Barkai et al., *supra* note 2, at 35–36 (examining settlement rates using data from Hawaii); Clermont & Eisenberg, *supra* note 18, at 120–21 (presenting new observations on each of the six phases of a civil lawsuit and drawing a series of "lessons for understanding and using empirical methods in the study of the legal system's operation"); Eisenberg & Lanvers, *supra* note 4, at 114–15 (creating different measures of settlement to examine "(1) settlement as a proxy for plaintiffs' litigation success, or (2) settlement as a measure of litigated disputes resolved without final adjudication"); Langbein, *supra* note 10, at 524–26 (situating the declining trial in the history of U.S. litigation and the Federal Rules of Civil Procedure). Other scholarship built on this empirical foundation with new data-informed inquiries into what drives settlement rates, including case type, geography, motions practice, and judge characteristics. *See, e.g.*, Gillian K. Hadfield, *Exploring Economic and Democratic Theories of Civil Litigation: Differences Between Individual and Organizational Litigants in the Disposition of Federal Civil Cases*, 57 STAN. L. REV. 1275, 1280–81 (2005) (examining differences between individual and organizational litigants in the disposition of federal civil cases, including settlement); Prescott & Spier, *supra* note 4, at 67–68 (introducing a comprehensive view of settlement including partial settlements); Boyd & Hoffman, *supra* note 3, at 908–10 (examining

operationalizable definition of settlement as party resolution, our novel methodological approach to counting settlement, and new findings on settlement frequency and distribution in the federal courts.[20] In addition, throughout the Essay, we compare and crosswalk our results to the IDB.[21] Despite some well-documented problems,[22] the IDB has long served as the primary source for settlement scholarship.[23] We offer our data, process, and results as a complement to the IDB and hope that other scholars of litigation, civil procedure, and federal courts will continue to mine the SCALES–OKN data to study not only settlement but also the myriad other litigation pathways and outcomes.[24]

The Essay proceeds as follows. In Part I, we survey existing literature to understand how other scholars have defined "settlement" and explain why such a simple-seeming task has frustrated many researchers. In Part II, we present our own definition. This definition—which focuses on non-adjudicated, full, and partial party resolutions—is built from extensive docket sheet reviews and takes into account settlement's many different

federal trial court litigation from case filing through settlement by looking at the time from motion filing to termination or time of settlement); Cotropia et al., *supra* note 4, at 85 (examining litigation behavior of patent assertion entities, "focusing on cases that resulted in a settlement or other voluntary disposition"); Boyd, *supra* note 18, at 204 (finding female district court judges more likely than male district court judges to foster intracourt case settlements successfully).

[20] The definition of settlement as party resolution offered in this Essay and the methodologies outlined could be extended to state court civil settlement, but it is outside this project's scope.

[21] *Integrated Database (IDB)*, *supra* note 8; *see also Facts About FJC's Integrated Database*, *supra* note 8. For an explanation of the data crosswalk, see *infra* note 114 and accompanying text, and *infra* Section IV.B.1.

[22] *See, e.g.*, Hadfield, *supra* note 4, at 709–10 (noting that there is little auditing of the IDB data or process and documenting widespread errors in the reported data).

[23] *See generally* Theodore Eisenberg & Margo Schlanger, *The Reliability of the Administrative Office of the U.S. Courts Database: An Initial Empirical Analysis*, 78 NOTRE DAME L. REV. 1455 (citing many studies relying on data from the federal Administrative Office of the U.S. Courts (AO or AOUSC)). For additional examples of empirical work focused on pathways to trial that are not extensively cited throughout this Essay, see generally Jason Scott Johnston & Joel Waldfogel, *Does Repeat Play Elicit Cooperation? Evidence from Federal Civil Litigation*, 31 J. LEGAL STUD. 39 (2002), which studies repeated attorney pairs in federal civil litigation outcomes. *See also* Peter Siegelman & Joel Waldfogel, *Toward a Taxonomy of Disputes: New Evidence Through the Prism of the Priest/Klein Model*, 28 J. LEGAL STUD. 101 (1999) (modeling trial and plaintiff win rates); Joel Waldfogel, *Reconciling Asymmetric Information and Divergent Expectations Theories of Litigation*, 41 J.L. & ECON. 451 (1998) (studying drivers of trial and win rates while accounting for pretrial adjudication and settlement); Joel Waldfogel, *The Selection Hypothesis and the Relationship Between Trial and Plaintiff Victory*, 103 J. POL. ECON. 229 (1995) (modeling trial and win rates with judge data on contracts, property rights, and torts cases).

[24] *See, e.g.*, Frank Fagan, *Natural Language Processing for Lawyers and Judges*, 119 MICH. L. REV. 1399, 1399–1400 (2021) (discussing developments in law and data that "can improve life and legal practice"); *see also* Keren Weinshall & Lee Epstein, *Developing High-Quality Data Infrastructure for Legal Analytics: Introducing the Israeli Supreme Court Database*, 17 J. EMPIRICAL LEGAL STUD. 416, 417–18 (2020) (arguing that weak data infrastructures hamper the ability to study litigation outcomes and proposing a framework for litigation data infrastructure).

procedural manifestations. In Part III, we turn from what counts as a settlement to how to count it. We describe the process of generating settlement data from the raw text of the docket sheets contained within the SCALES–OKN corpus and the limitations of our methods. Part IV presents our findings and compares them to published IDB settlement figures, and Part V discusses directions for future work.

I. EXISTING SETTLEMENT DEFINITIONS

"Settlement" is a construct, an idea comprising various conceptual elements.[25] Scholars, practicing lawyers, and the public may have differing views of settlement's essential features, including absence of adjudication, some degree of moral and legal compromise, and finality.[26] But there is a need for clarity to create a workable definition of settlement that can be deployed for empirical purposes. This Part considers three definitions of settlement suggested by the literature: settlement as IDB coding, non-adjudication, and compromise.[27]

---

[25] Jessica Bregant, Jennifer K. Robbennolt & Verity Winship, *Perceptions of Settlement*, 27 HARV. NEGOT. L. REV. 93, 97 (2021) ("At its simplest, settlement is merely an agreement to resolve a legal dispute. But in the law and in public consciousness, the way settlement is conceptualized is often more complicated."); *see generally* Robbennolt et al., *supra* note 16 (discussing the wide array of settlement outcomes and the public's consistent perception of settlement as an agreed outcome, but for which there is much variation in meaning and consequent).

[26] In a study of 1,000 U.S. adults, many respondents described settlement as when parties "come to an agreement" or "agree on a solution." Robbennolt et al., *supra* note 16, at 490. These views reflect that parties must agree to the outcome—the agreement is a *moral* compromise. The authors of the same study describe settlement as when "the parties to a dispute reach an agreement that ends the lawsuit." *Id.* An agreement in principle to end a suit does not, by itself, end the litigation. A second step is required to convert the *moral* compromise to a legal one. We return to this distinction several times throughout the Essay.

Some personal injury attorneys described settlement in the following terms: "In some cases you might not even have to go into court or could be eligible to resolve your lawsuit outside of court before a judge or jury makes a final determination. This is known as a settlement and while it does not happen in every case, can [*sic*] be a beneficial option to pursue or at least explore when you are in the midst of a suit." *Settlements: Is a Settlement Considered a Win?*, MORGAN & MORGAN, https://www.forthepeople.com/practice-areas/personal-injury-lawsuits/is-a-settlement-considered-a-win [https://perma.cc/XP2T-JSJQ].

[27] We also acknowledge the large theoretical and experimental literature about the preconditions and triggers of settlement, mostly from the law-and-economics tradition. *See, e.g.*, Paul Pecorino & Mark Van Boening, *An Empirical Analysis of the Signaling and Screening Models of Litigation*, 20 AM. L. & ECON. REV. 214 (2018) (studying the effect of asymmetric information on various litigation models); Sean P. Sullivan, *Why Wait to Settle? An Experimental Test of the Asymmetric-Information Hypothesis*, 59 J.L. & ECON. 497 (2016) (testing whether asymmetric information contributes to settlement delays). Those analyses are irrelevant here; we focus on the conceptual and empirical questions of reliably defining and identifying settlements. This work can then feed into future investigations of the topics of interest in the broader theoretical literature, for example, testing theories of when and under what circumstances settlement might be expected to occur.

A. Settlement as IDB Coding

The IDB is the main administrative data set produced by the federal courts.[28] It is a "tantalizingly complete long-term data set"[29] and often serves as the primary source cited in settlement scholarship.[30] The database contains a variety of fields for every civil case filed in federal district court, including a code for a case's disposition or termination, one value of which is "settled."[31]

A brief detour into the mechanics of the IDB's assembly is necessary before considering the definitional potential of the IDB's settlement disposition code. The Federal Judicial Center (FJC), the research and education arm of the federal courts, produces the IDB through a relationship with the Administrative Office of the U.S. Courts (AO or AOUSC), the entity that gathers and maintains court-related statistics and oversees the

---

[28] Eisenberg & Schlanger, *supra* note 23, at 1496 ("The AO database is likely to remain one of the major sources for civil justice research."). An additional data source on civil litigation in state courts was the Civil Justice Survey of State Courts (CJSSC) (inactive since 2005), which collected data in individual state courts through a joint project of the National Center for State Courts and the Bureau of Justice Statistics, surveying tort, real property, and contract cases in a sample of 45 of the 75 most populous counties. *See Civil Justice Survey of State Courts (CJSSC)*, BUREAU OF JUST. STATS., https://bjs.ojp.gov/data-collection/civil-justice-survey-state-courts-cjssc [https://perma.cc/M2DU-PRWE].

[29] Hadfield, *supra* note 19, at 1282.

[30] *See, e.g.*, Eisenberg & Schlanger, *supra* note 23, at 1458 (calling the IDB "'by far the most prominent' database used by legal researchers for statistical analysis of case outcomes" (quoting Frank B. Cross, *Comparative Judicial Databases*, 83 JUDICATURE 248, 248 (2000))); Clermont & Eisenberg, *supra* note 18, at 136 (analyzing 259,637 federal civil cases from the AO database that were terminated during fiscal year 2000); Kevin M. Clermont, *Litigation Realities Redux*, 84 NOTRE DAME L. REV. 1919, 1955 (2009) (analyzing 271,753 federal civil cases terminated in all federal districts during fiscal year 2005); Hadfield, *supra* note 19, at 1282–83 (using the AO database for all federal cases recorded in 2000); Gary M. Fournier & Thomas W. Zuehlke, *Litigation and Settlement: An Empirical Approach*, 71 REV. ECON. & STAT. 189, 191 (1989) [hereinafter Fournier & Zuehlke, *Litigation and Settlement*] (using data taken from AO survey); Gary M. Fournier & Thomas W. Zuehlke, *The Timing of Out-of-Court Settlements*, 27 RAND J. ECON. 310, 314 (1996) [hereinafter Fournier & Zuehlke, *Out-of-Court Settlements*] (using data prepared by the AO).

Before 2000, the data presently available in the IDB were only accessible by special permission by the Inter-University Consortium for Political and Social Research (ICPSR), a data repository housed at the University of Michigan. Early studies may refer to ICPSR rather than the IDB. *See Federal Court Cases: Integrated Data Base, 1970–2000*, FED. JUD. CTR., https://www.icpsr.umich.edu/web/NACJD/studies/8429 [https://perma.cc/YS79-6S6S]. Settlement scholarship also draws insights and conclusions from smaller, hand-collected data sets. *See, e.g.*, Ashenfelter et al., *supra* note 18, at 265 (analyzing 2,258 cases total—including nearly every federal civil rights and prison case filed in the Central District of California, Eastern District of Pennsylvania, and Northern District of Georgia from fiscal year 1981); Cotropia et al., *supra* note 4, at 108 (studying settlement in patent cases); Boyd & Hoffman, *supra* note 3, at 900 (studying settlement rates in 585 piercing the corporate veil cases); Colvin, *supra* note 3, at 7, 8, 16–17 (observing a settlement rate of 59% in the dataset, a lower mean time to resolution among the settled cases, and the role of counsel in settled cases).

[31] *IDB Civil 1988–Present*, FED. JUD. CTR., https://www.fjc.gov/research/idb/interactive/21/IDB-civil-since-1988 [https://perma.cc/4PKX-QNNL].

courts' electronic case filing and public access to court records systems.[32] As the IDB's website explains:

> The IDB contains data on civil case and criminal defendant filings and terminations in the district courts, along with bankruptcy court and appellate court case information from 1970 to the present. The FJC receives regular updates of the case-related data that are routinely reported by the courts to the AOUSC. The FJC then post-processes the data, consistent with the policies of the Judicial Conference of the United States governing access to these data, into a unified longitudinal database, the IDB.[33]

The provenance of the IDB data on civil litigation thus begins at each of the ninety-four U.S. district courts, where parties submit lawsuit filings through each court's Case Management/Electronic Case Filing (CM/ECF) system. Judges and clerks also enter opinions, memoranda, orders, and notices on the docket sheets. Court personnel convert those real-time records of litigation activity into structured data that are reported to the centralized AOUSC, which then forwards litigation data snapshots four times per year to the FJC for post-processing and publication in the IDB.[34] The fields contained in the IDB include "[d]ates of filing and termination (if applicable)," "the type of termination," and "case-level information relevant to each area of litigation. . . . [This includes] nature of suit, jurisdiction, origin codes, the names of plaintiffs and defendants, class action allegation, the procedural progress of the case at termination, and the nature and amounts of judgment."[35]

Turning from mechanics to definitions, settlement might simply be defined as occurring whenever the IDB records the termination or disposition

---

[32] *Integrated Database (IDB)*, *supra* note 8.

[33] *Id.* Post-processing is described as follows:

> First, data values that are out of range for the variable are recoded as missing. Second, some information is redacted—for example, the names of criminal defendants in criminal and appellate files. Relatedly, information on the judge or judges presiding over the case is redacted pursuant to Judicial Conference policy. Third, the IDB integrates three types of case records: filings, pendings, or terminations. New cases are filings records. Cases that were filed previously but not yet terminated are pending records. Cases that were previously filed, or filed and terminated in the same quarter, are terminations records. Each quarterly update reflects the current status of the case records, including new case-related information (such as conversion of a bankruptcy from a Chapter 11 to a Chapter 7).

FED. JUD. CTR., THE INTEGRATED DATABASE: A RESEARCH GUIDE 1, https://www.fjc.gov/sites/default/files/IDB-Research-Guide.pdf [https://perma.cc/Y35M-VHMC].

[34] *Id.*; E-mails from Kristin Garri, Senior Rsch. Assoc., Fed. Jud. Ctr., to Charlotte S. Alexander, Professor of L. & Ethics, Ga. Inst. of Tech. Scheller Coll. of Bus. (Jun. 9, 2023–Sept. 15, 2023) (on file with authors).

[35] FED. JUD. CTR., *supra* note 33, at 2.

code as "settled." Indeed, some settlement scholarship tends to take the IDB's settlement coding as self-defining.[36]

There are at least two problems with this approach. First, although the IDB provides a count of cases with the termination or disposition code "settled," it does not *define* what counts. As Table 1 shows, the IDB code book lists twenty-two disposition options grouped into three larger categories.[37] The rows shaded in gray in Table 1 could all potentially indicate settlements, but no official explanation is provided for the differences among these dispositions or why a case might receive one code versus another.[38]

TABLE 1: IDB CODES FOR DISPOSITION ("DISP") FIELD, POSSIBLE SETTLEMENT CODES SHADED

Cases Transferred or Remanded	
0	transfer to another district
1	remanded to state court
10	multi-district litigation transfer
11	remanded to U.S. Agency
Dismissals	
2	want of prosecution
3	lack of jurisdiction
12	voluntarily
13	settled

[36] *See, e.g.*, Eisenberg & Farber, *supra* note 18 at S99, S107–09 (drawing settlement coding from the civil case termination codes assigned by the AO, which populate the IDB).

[37] *Federal Judicial Center Integrated Data Base Civil Documentation*, FED. JUD. CTR., https://www.fjc.gov/research/idb/civil-cases-filed-terminated-and-pending-sy-1988-present (click "Civil Codebook 1988 Forward 10252023.pdf") [https://perma.cc/T7AS-9GL9].

[38] The IDB's civil documentation reads, "See Appendix A: CIVIL CODE SHEETS, under disposition, for explanation of the three manners." *Id.* The referenced civil code sheets are not publicly available, but in response to an inquiry, FJC personnel provided a version of the sheets from 2007, which states in relevant part:

> (12) Voluntarily. Plaintiff voluntarily withdrew the action from judicial review in accordance with Rule 41(a), F.R.Cv.P. . . . (13) Settled. The action was disposed of after settlement between parties out of court . . . (05) Consent. The action was disposed of by an order of judgment agreed to by all parties and signed by the judge or magistrate, which grants some form of affirmative relief to one of the parties. This category should be indicated even though the agreement was entered into after a trial began.

E-mails from Kristin Garri, Senior Rsch. Assoc., Fed. Jud. Ctr., to Charlotte S. Alexander, Professor of L. & Ethics, Ga. Inst. of Tech. Scheller. of Coll. Bus. (Aug. 2, 2023–Sept. 15, 2023) (on file with authors). Although these guidelines create some distinctions between the categories, they also create definitional confusion. What happens, for example, when the parties settle but give effect to their settlement via a motion for voluntary dismissal under Rule 41 of the Federal Rules of Civil Procedure or a consent judgment? We return to this problem below.

14	other
Judgment On	
4	default
5	consent
6	motion before trial
7	jury verdict
8	directed verdict
9	court trial
15	award of arbitrator
16	stayed pending bankruptcy
17	other
18	statistical closing
19	appeal affirmed (magistrate judge)
20	appeal denied (magistrate judge)
-8	missing

It may be possible to reverse engineer the IDB by matching disposition codes to the events and entries present on the underlying docket sheets that record litigation activity in each case. We attempt this in Part III, but the crosswalking exercise is complex.[39] We suspect that the IDB, in fact, reflects a variety of approaches to—and local clerks' office norms around—defining settlement by the court-level personnel who generate case-level data.[40] This suspicion is borne out by previous research into IDB disposition codes. Professor Gillian Hadfield, for example, compared the events on docket sheets with those cases' IDB disposition codes and found large numbers of settlements in cases labeled "14: Dismissals, Other" and "6: Judgment on

[39] *See infra* Part III.

[40] In research that involves "hand coding"—applying a set of codes or labels to a particular data type—assigning the same coding or labeling tasks to multiple coders is standard practice. This ensures that any individual coder does not impose their idiosyncratic interpretation or process on the data collection process. When all coders agree, this is known as "intercoder reliability," and researchers can have confidence in the validity of their coding methodology. Where coders disagree, a reconciliation process is used, such as an additional coder acting as a tiebreaker. *See, e.g.*, LEE EPSTEIN & ANDREW D. MARTIN, AN INTRODUCTION TO EMPIRICAL LEGAL RESEARCH 50, 110 (2014) (describing and applying fundamental principles of social science methodology to legal research); Cliodhna O'Connor & Helene Joffe, *Intercoder Reliability in Qualitative Research: Debates and Practical Guidelines*, 19 INT'L J. QUALITATIVE METHODS 1, 5 (2020) (describing hand coding best practices and methods for assessing intercoder reliability). The IDB results from a hand coding enterprise, where clerks take in the raw material of litigation events and choose the appropriate code or codes to apply. However, unlike the best practices described above, we are unaware of efforts to ensure intercoder reliability among clerks, introducing the potential for errors. *See, e.g.*, FED. JUD. CTR., *supra* note 33, at 4 ("A process of error correction and detection does not cover all the variables in the IDB . . . .").

Motion Before Trial," in addition to the shaded labels in Table 1.[41] Thus, scholars who rely on the IDB for a measure of settlement frequency in federal civil litigation skip the crucial first step of defining what they mean by "settlement," as there is no clear line-drawing around the IDB "settled" code or its relationship to other dispositions.

The second problem with defining settlement according to IDB coding concerns partial resolutions.[42] A lawsuit may completely settle, ending all claims against all parties and therefore concluding the entire case. Alternatively, specific parties may settle, for example, an insurance company and the tort victim, but the claims between the insurance company and the insured or tortfeasor may persist.[43] Similarly, parties to a business dispute may settle the contract claims in a case but leave the tort claims open for subsequent motions practice and possibly trial. These resolutions are undoubtedly of interest to scholars of settlement, but they are obfuscated by the IDB's data collection on a case-level basis, with only one disposition code per case.[44]

Thus, the IDB is a poor source for a definition of settlement, which complicates its empirical reliability—or at least raises caution flags. We explore this further in connection with our findings in Part IV below. We now turn to two alternative definitional possibilities derived from the civil litigation literature: settlement as non-adjudication and settlement as compromise.

B. Settlement as Non-Adjudication

As noted in the Introduction, some of the first settlement studies viewed litigation outcomes as either trial or non-trial, and then labeled all non-trial outcomes as "settlement."[45] Although this definition is erroneously overinclusive,[46] the absence of adjudication by a judge or jury is a key component of many settlement definitions offered by leading legal scholars[47]

---

[41] Hadfield, *supra* note 4, at 719–20.

[42] For a discussion of full versus partial settlements, see Prescott & Spier, *supra* note 4, at 64.

[43] Vacek & Schilder, *supra* note 4, at 210 (noting the complications of settlement research where there are different outcomes for different parties and claims); Prescott & Spier, *supra* note 4, at 64 (discussing the "virtually infinite menu of potential [settlement] arrangements," including partial settlements).

[44] *See Integrated Data Base Civil Documentation*, *supra* note 37 (explaining the disposition code as the singular "manner in which the case was disposed of"); E-mails from Kristin Garri to Charlotte S. Alexander, *supra* note 38.

[45] *See supra* note 3 and accompanying text.

[46] Eisenberg & Lanvers, *supra* note 4, at 146 ("[D]iscussions of settlement rates tend to be overly simplistic. . . . [And the rates] find[] little support in actual practice.").

[47] *Id.* at 115 (describing, as one possible framework for considering settlement, "settlement as a measure of litigated disputes resolved without final adjudication").

and practicing attorneys alike.[48] Adjudication and settlement "run in opposite directions along the dispute resolution continuum: at one extreme, a case is fully settled, with nothing left to adjudicate; at the other end of the continuum," the case proceeds toward trial.[49] Between these two litigation poles lie other forms of adjudication: successful motions to dismiss for failure to state a claim, granted motions for summary judgment, and default judgments, for example.

One might, therefore, tally up all case resolutions from adjudicated outcomes and subtract that number from the number of cases filed.[50] *Et voilà*: a settlement rate and a definition at the same time—settlement as non-adjudication. But complexities remain. What counts as an "adjudicated outcome?" For example, court-connected mediation programs, wherein parties attempt settlement through a court-mandated neutral program, became "widespread" after the federal Alternative Dispute Resolution Act passed in 1998.[51] Judges themselves may also weigh in on their view of the strength of each party's case and likely dispositions in court. A judge's adjudicatory power often lurks behind outcomes where parties memorialize their compromised resolution in a settlement agreement.[52]

Further, some settlements resolve only a subset of claims or parties while leaving others to proceed. It is unclear whether these partially settled cases would be considered "settled," "adjudicated," or some combination of both under a settlement-as-non-adjudication approach. Where researchers use the IDB as the source for the non-adjudication definitional approach, its data do not reliably track partial dispositions, as discussed in the previous Section.[53] Given the difficulties of defining settlement in the negative, as non-adjudication, some scholars have attempted a more affirmative definition: settlement as compromise.[54]

---

[48] *Settlements: Is a Settlement Considered a Win?*, *supra* note 26 ("[R]esolve your lawsuit outside of court before a judge or jury makes final determination. This is known as a settlement . . . .").

[49] Prescott & Spier, *supra* note 4, at 66.

[50] This could be accomplished by the IDB's disposition codes, but the IDB's reliability problems would undermine such an exercise.

[51] Langbein, *supra* note 10, at 561 (citing Alternative Dispute Resolution Act of 1998, Pub. L. No. 105-315, § 3, 112 Stat. 2993, 2993 (codified as amended at 28 U.S.C. § 651(b))).

[52] *See* INST. FOR THE ADVANCEMENT OF THE AM. LEGAL SYS., *supra* note 9, at 65 (noting that settlement is facilitated by court-sponsored or court-ordered ADR, scheduling conferences, and setting early trial dates); *see also* Judith Resnik, *A2J/A2K: Access to Justice, Access to Knowledge, and Economic Inequalities in Open Courts and Arbitrations*, 96 N.C. L. REV. 605, 626 n.86 (2018) (describing how settlements can conceal relevant information from the public).

[53] *Supra* Section I.A.

[54] *See, e.g.*, Lynch & Levine, *supra* note 9, at 244 (footnote omitted) (describing settlement as "a voluntary, consensual resolution by counsel and clients based upon an enlightened assessment of risks");

C. Settlement as Compromise

Compromise and mutual agreement offer a third possible definition of settlement.[55] Here, the focus is the parties' intent behind the resolution rather than the absence of a trial or other adjudicatory event. Settlement as moral compromise has traction with practitioners and legal scholars who define settlement by focusing on the agreement or concession of the parties.[56]

Yet compromise as a defining feature of settlement has at least two limitations. First, scholars have raised doubts about whether settlement outcomes reliably reflect intent, agreement, and uncoerced concessions.[57] Parties may decide to end a case for various reasons, including full agreement on the underlying facts and law or a recognition that settlement can reduce risks, costs, uncertainty, and reputational harm.[58] They may also simply run out of resources—time, energy, information, or clout—to continue litigation or to negotiate for the outcome they really want in a settlement.[59] Analogies have been made between civil settlement and criminal plea bargaining, with scholars calling court-sanctioned alternative dispute resolution programs in civil litigation "Cajolery Conferences."[60]

Marc Galanter & Mia Cahill, *"Most Cases Settle": Judicial Promotion and Regulation of Settlements*, 46 STAN. L. REV. 1339, 1350 (1994) (noting the role of compromise in settlement); Boyd & Hoffman, *supra* note 3, at 925 ("Cases settle because the parties choose to compromise rather than contest.").

[55] Prescott & Spier, *supra* note 4, at 78 ("[S]ettlements are simply agreements between parties to a dispute that offer value to both on one or more of the following dimensions: reducing adjudication costs, mitigating losses due to risk, and maximizing ex ante expected returns.").

[56] *Settlements*, N.Y. STATE UNIFIED CT. SYS. (May 4, 2021), https://www.nycourts.gov/courthelp/goingtocourt/settlements.shtml [https://perma.cc/9VKG-GNG6] ("In a settlement both sides agree to the outcome of the case and there is no trial before a judge or a jury. The settlement can be made with or without the help of the court."); *see also* Eisenberg & Lanvers, *supra* note 4, at 114 (dismissing the trial versus settlement trope in favor of a richer view of party resolution); Clermont, *supra* note 30, at 1955 n.180 (including "compromise by private negotiations or through ADR" in the definition of settlements).

[57] *See, e.g.*, Owen M. Fiss, *Against Settlement*, 93 YALE L.J. 1073, 1076–78, 1085 (1984) (observing that settlement may be unfair when the parties have unequal resources to conduct litigation and that it prevents the generation of precedent); *see also* Jay Tidmarsh, *Auctioning Class Settlements*, 56 WM. & MARY L. REV. 227, 241 (2014) (describing the problem of "reverse auction[s]" in class actions as a "race to the bottom" where parties underbid settlement claims in the hopes of settling first and earning fees).

[58] Prescott & Spier, *supra* note 4, at 62, 78.

[59] Jennifer K. Robbennolt, *Litigation and Settlement*, *in* THE OXFORD HANDBOOK OF BEHAVIORAL ECONOMICS AND THE LAW 624 (2014) (describing the standard "economic model" of litigation as an inquiry into the costs of litigation and the likely outcome driving settlement when the costs—in time and money—become too great).

[60] Langbein, *supra* note 10, at 562 (quoting Albert W. Alschuler, *Mediation with a Mugger: The Shortage of Adjudicative Services and the Need for a Two-Tier Trial System in Civil Cases*, 99 HARV. L. REV. 1808, 1828 (1986)).

One critique levied against plea bargains is that the criminal adjudication process is a party to the negotiation.[61] Bargaining from jail and in the shadow of a criminal trial shapes the bargain struck.[62] Likewise, negotiations in civil cases are intentionally shaped by formal litigation processes.[63] The FRCP discovery process was designed with an eye towards facilitating settlement.

> [P]art of the rationale that motivated the drafters of the Federal Rules to prefer discovery over pleading was the expectation that the change would promote settlement. . . . [They] wrote that "one of the greatest uses of judicial procedure is to bring parties to a point where they will seriously discuss settlement." . . . ["]It is not what a party asserts, but what he can establish by proof" that shapes settlement.[64]

Further, the 1983 revision to the Federal Rules made "facilitating settlement" an express objective of pretrial conferences enhanced by other features of civil adjudication such as case management, scheduling orders, and conferences that expose the remaining sticking points and reveal paths to resolution.[65] "[B]y acquainting a neutral observer with the evidence, judicial case management can function as 'mediation by another name.'"[66]

Even if we were to set aside questions about the reason or process for the concession and simply focus on the fact of compromise, a second problem would remain: observability.[67] Because the moral compromise

---

[61] *See* Russell D. Covey, *Toward a More Comprehensive Plea Bargaining Regulatory Regime*, 101 OR. L. REV. 257, 267–68 (2023).

[62] *Id.* at 271 ("Since plea bargaining outcomes are typically justified as rational because they are negotiated in the 'shadow of trials,' plea bargaining without trials is a negotiation process in which bargaining outcomes are negotiated in a vacuum.").

[63] Robbennolt, *supra* note 59, at 627 (describing litigation bargaining as happening in the "shadow of the law").

[64] Langbein, *supra* note 10, at 548 (citations omitted); *see* FED. R. CIV. P. 26. Other scholars see motion practice as influencing settlement. Boyd and Hoffman found that substantive, non-discovery motions speed up settlement in a case, especially when granted. Boyd & Hoffman, *supra* note 3, at 915.

[65] Langbein, *supra* note 10, at 559.

[66] *Id.* (citing E-mail from Geoffrey C. Hazard Jr., Thomas H. Miller Distinguished Professor of L., Univ. of Cal. Hastings Coll. of L., to John H. Langbein, Sterling Professor of L. & Legal Hist., Yale Univ. (Nov. 22, 2011) (on file with John H. Langbein)).

[67] Christopher A. Cotropia, Jay P. Kesan, Kyle Rozema & David L. Schwartz, *Progression and Workload in Civil Litigation: An Empirical Analysis of Patent Disputes* 7 (Univ. Ill. Coll. of L. & Legal Stud., Research Paper No. 17-37, 2017), https://papers.ssrn.com/a=3021903 [https://perma.cc/6XF7-TQFY] ("[I]n the vast majority of cases, it is not possible to capture from the public record why the parties settled and the terms of the settlement."); *see* Clermont & Eisenberg, *supra* note 18, at 137–38 (describing the tendency of empirical research to focus on the "readily observable" and the failure of this approach to capture underlying factors affecting the outcome of a case accurately).

occurs outside the formal litigation mechanisms, it is like a litigation black hole: "fundamentally unseeable."[68]

To transform a moral compromise into a legal one, the parties must invoke the power of the court. They must inform the court of their agreement via a notice, motion, or other filing, and the court must take action to terminate the relevant party, claim, or entire case. Researchers can observe party filings and court actions, just as scientists can observe the effect of black holes on surrounding matter.[69] But measurement by proxy can be imperfect. Consider a plaintiff's voluntary motion to dismiss filed under FRCP 41, with no indication that settlement is the real reason for dismissal. A plaintiff may have simply lost interest in litigating, or a defendant may wish to avoid the word "settlement" on the record to avoid the hint of liability (or even worse, a suggestion of deep pockets).[70] Is this bare-bones voluntary dismissal a reliable signal of moral compromise? Should it, and the court's order granting dismissal, be considered a compromise and, therefore, settlement?[71] We return to these questions, implicating both the definition and counting problems in Parts II and III.

Next, we introduce our own definition of settlement, which draws from the approaches outlined above but is also informed by an extensive on-the-ground review of docket sheets. Importantly, our definition is designed to be operationalizable—deployable in an empirical sense—as we move from considering what qualifies as a settlement to counting it.

II. DEFINING SETTLEMENT AS PARTY RESOLUTION

Building on the conceptualizations of settlement described above, we define settlement as party resolution, or an agreement between parties to resolve some or all claims held by a party or parties. This definition does not differentiate between agreements that occur entirely "out of court" and those reached as part of some court-annexed process. In other words, though we define the act of settlement as non-adjudicated by a judge or jury, our concept

---

[68] *What Do Black Holes Look Like?*, HARV. & SMITHSONIAN CTR. FOR ASTROPHYSICS, https://www.cfa.harvard.edu/big-questions/what-do-black-holes-look [https://perma.cc/3BQM-8GNB].

[69] *Black Hole Basics*, NASA, https://science.nasa.gov/astrophysics/focus-areas/black-holes [https://perma.cc/25N6-AFTC].

[70] *See, e.g.*, Bregant et al., *supra* note 25, at 96 (describing the signals sent by settlement, including guilt, wrongdoing, and largess for lay people and judges); *see also* Robbennolt, *supra* note 59, at 624 (describing nonmonetary influences for settlement, including reputation).

[71] Court personnel probably come to different answers to questions such as these, which introduces variation and unreliability in assigning IDB disposition codes. Some clerks may mark anything labeled "voluntary dismissal" on the docket sheet as disposition code "12-voluntary dismissal"; others may look elsewhere on the docket for indications of settlement and mark the subsequent voluntary dismissal as "13-settlement."

of party resolution is agnostic as to the acceptable level of judicial influence.[72] This flexible definition also encompasses both full and partial resolutions of parties' claims. Nor are we purists regarding the parties' reasons or motivations for settlement.

We exclude from this definition party agreements around procedural matters, such as limitations on discovery, an agreed remand to state court, or a joint decision to submit claims to a magistrate or arbitrator. These would fall into what Professors J.J. Prescott and Kathryn Spier have labeled "procedure-modification agreements."[73] These types of agreements are directed more toward the rules of litigation than the substantive claims that triggered litigation in the first place, making them different in kind from the party resolutions that are the topic of this Essay.[74]

Despite—or maybe because of—its relatively big-tent approach, our definition is not immediately operationalizable toward any empirical end. As discussed above, the agreement between the parties typically occurs outside official court proceedings. Most settlements manifest procedurally only when the parties convert their moral compromise to a legal one by invoking the court's power to dismiss a party, claim, or case. That invocation can take a bewildering variety of forms, an incomplete list of which includes:

- One party filing a settlement agreement with a joint motion for approval by the court;
- The parties notifying the court that a settlement has been reached and jointly moving to dismiss;
- The plaintiff filing a motion for voluntary dismissal under Rule 41(a) (sometimes bare-bones, other times agreed, joint, consent, or stipulated);
- The parties approaching the bench during a hearing or trial and having the settlement entered from the bench;

---

[72] With the benefit of the full set of litigation events gleaned from docket sheets as part of the SCALES–OKN project, future work can explore the presence of many different indicators of judicial influence on docket sheets before settlement. These might include scheduling orders, status conferences, settlement conferences, mediations, and partial rulings on dispositive motions that might signal the judge's view of the strengths and weaknesses of the parties' cases.

[73] Prescott & Spier, *supra* note 4, at 114.

[74] We recognize that this substance–procedure distinction may be flimsy when poked. Prescott and Spier write that "each category of partial settlement agreements (as well as full settlement agreements) are just different instances of the same thing, with the precise mix being determined by the particulars of the parties, the nature of the dispute, and real-world frictions, such as negotiation costs and cognitive and behavioral biases." *Id.* at 83. However, Prescott and Spier also recognize that procedural modification agreements are a "hodgepodge." *Id.* We found the same and encountered enormous difficulty in identifying and counting these types of agreements on the docket sheets, which provided another reason for excluding them here, as the language used by the parties and the court to describe these sorts of events is even more varied than the settlement language that we targeted in our project.

- The parties recording resolution through a consent judgment, agreed judgment, or consent decree; and
- The parties terminating proceedings via an accepted offer of judgment under Rule 68.[75]

This list captures only a fraction of all possible procedural configurations, which may involve motions, notices, or other filings by the parties and notices, orders, and judgments by the court; may or may not use the word "settlement"; may or may not cite or follow the procedures for dismissal or final judgment established by the FRCP; may be with or without prejudice; and many other variations on the theme.[76]

Our method accounts for the procedural and lexical differences in how settlement manifests on docket sheets to reduce our concept of settlement to an observable and measurable phenomenon. Part III further describes our data and methodology.

III. COUNTING PARTY RESOLUTION

We now move from what to count as a settlement to how to count it. Above, we presented our broad conceptual definition. Here, we describe our data and methods and their limitations. Part IV then presents our findings and squares them with IDB disposition codes.

A. Data

The data source for this project is the SCALES–OKN, a multi-university, multidisciplinary collaboration that is working to "develop a suite of tools to enable access to court records and analytics."[77] With funding from

[75] The authors gleaned this list of party resolution examples via our process of constructing litigation event labels, described in Section III.B. This process required extensive manual review of thousands of docket entries to construct a set of labels that our classification models would apply, to validate the models' output, and to troubleshoot as necessary.

[76] Other scholars have recognized settlement's many complexities. Professors Prescott and Spier, for example, describe settlement as a "virtually infinite menu of potential arrangements against the default litigation background." Prescott & Spier, *supra* note 4, at 64. They advocate for a rich, nonmonolithic conceptualization of settlement along a continuum of outcomes. *See id.* This view is common in the more recent settlement literature, including a patent litigation study by Professor Cotropia and his coauthors, where they broadly define settlement as a "resolution" to reflect the many procedural paths to that outcome. *See* Cotropia et al., *supra* note 67, at 7.

[77] *Transforming the Accessibility and Transparency of Federal Courts*, SCALES, https://scales-okn.org/ [https://perma.cc/BB6M-9GPN]. *See generally* Adam R. Pah, David L. Schwartz, Sarah Sanga, Zachary D. Clopton, Peter DiCola, Rachel Davis Mersey, Charlotte S. Alexander, Kristian J. Hammond & Luís A. Nunes Amaral, *How to Build a More Open Justice System*, 369 SCIENCE 134 (2020) (describing the values and principles behind the SCALES project). All three of the present authors have been involved in the larger SCALES–OKN collaboration. Alexander is a co-principal investigator and lead of the Civil

the National Science Foundation and private foundations, the SCALES project has amassed approximately 1.3 million docket sheets and court documents from criminal and civil cases filed in U.S. district courts.[78] The present study used a subset of docket sheets from the full SCALES corpus. Docket sheets, our source material, contain a real-time chronological record of litigation as it unfolded in each case and enable the study of settlement in all its forms and procedural guises.[79] While access to the underlying party and court-filed documents would have added detail and richness to the analysis, the SCALES project funding only enables the assembly of small court document sets, in contrast to the less expensive docket sheets.[80]

We began with all cases of all types filed in all ninety-four U.S. district courts in 2016 and 2017, downloaded by SCALES from the federal courts' Public Access to Court Electronic Records (PACER) system as of 2020 and 2021, or at least four years after each case was initiated. This 2016–2017 set, consisting of 676,815 cases, was the most comprehensive tranche of SCALES docket sheets, as it included a complete set of all cases filed nationwide.[81] Restricting the set to civil cases brought the total to 550,439; restricting further to cases that were closed at the time of download reduced the number to 513,064. For reasons explained further in Part IV, the "closed case" limitation enables us to crosswalk our docket sheet-based data most effectively to the IDB and other settlement scholarship.

Ontology Working Group. Dahlberg worked as a data scientist and deep-learning engineer across the SCALES project. Tucker helped lead the portion of the research focusing especially on settlement and led the early stages of the civil litigation ontology work.

[78] *SCALES Awarded NSF Grant to Build the Integrated Justice Platform Proto-OKN*, SCALES (Oct. 12, 2023), https://scales-okn.org/2023/10/12/scales-awarded-nsf-grant-to-build-the-integrated-justice-platform-proto-okn/ [https://perma.cc/VY9Q-HU4G]; Scott Daniel, *Notes on Our Internal Data Pipeline*, SCALES (Mar. 26, 2024), https://docs.scales-okn.org/guide/pipeline/#federal-courts-pacer [https://perma.cc/7Q67-ANKH] (listing "1,278,268 total PACER cases in [the SCALES] dataset").

[79] Our approach follows the work of other noted settlement scholars such as Christina L. Boyd and David A. Hoffman. *See* Boyd & Hoffman, *supra* note 3, at 900 (examining federal trial court litigation, from case filing through settlement, by looking at the time from motion filing to termination and time of settlement); *see also* Cotropia et al., *supra* note 67, at 5–7 (assessing patent infringement lawsuit settlement data); Barkai et al., *supra* note 2, at 35–36 (using 3,000 docket sheets as one of two data sources in the study); Hadfield, *supra* note 4, at 710 (examining case-level data on federal terminations between 1970 and 2001).

[80] The federal courts' Public Access to Court Electronic Records (PACER) system charges $0.10 per page with a cap at $3 per document, an "antiquated pricing structure with origins in the era of photocopied court documents." *What Does a Year of PACER Actually Cost?*, SCALES (Dec. 13, 2020), https://scales-okn.org/2020/12/13/what-does-a-year-of-pacer-actually-cost/ [https://perma.cc/JN42-L268]. SCALES estimated the cost of "query[ing] and download[ing] every publicly accessible docket sheet and attached document pertaining to civil and criminal cases" in 2016 to be "somewhere between $5.5 million and $5.75 million." *Id.*

[81] SCALES holds additional docket sheets and court documents that span longer timeframes, but this coverage is isolated to certain courts or "Nature of Suit" codes, or its use is otherwise restricted under the terms-of-data-access agreements with courts.

Next, we dropped two subsets of civil cases governed by procedural rules that differ from the FRCP in ways that complicate settlement counting. First, we excluded cases with any of the "Nature of Suit" (NOS) codes listed in Online Appendix A. When plaintiffs file suit in federal court, they must choose an NOS code from a designated list to characterize the claims in the case.[82] The NOS codes represent cases that come to U.S. district courts in unusual ways—for example, bankruptcy and Social Security, where the district court acts as an appellate body. It also includes cases where the parties' relationship with one another diverges from the typical adversarial civil litigants' relationship, such as federal foreclosure or habeas corpus cases. Therefore, the concept of "settlement" that we operationalize in this work is inapt in those cases. After excluding cases with these NOS codes, 393,035 cases remained.

Finally, we dropped cases that we identified as part of multi-district litigation (MDL) because of difficulties in correctly associating the docket sheet for each case coming from a transferor court with the docket sheet for the consolidated case in the transferee court.[83] In these cases, settlement might happen in either the transferor or transferee court and could apply to one or many cases, so we chose to drop MDL cases entirely rather than under- or over-count settlement across a set of MDL-related dockets. This last step produced a final corpus of 346,916 docket sheets comprising 12,108,851 individual docket entries.

B. Methods

We now describe the methods that the larger SCALES research team employed to develop and apply litigation event labels to the raw text of docket sheets, as well as the particular methods we used to generate the set of party resolution labels relevant to the present settlement project.

1. SCALES Civil Litigation Events Ontology

One goal of the larger SCALES collaboration is to create a civil litigation events ontology or a unitary conceptual understanding of the major mileposts in a federal civil lawsuit.[84] To this end, the SCALES team developed a set of deep-learning classifiers built on an existing general-

[82] *See Nature of Suit*, PACER, https://pacer.uscourts.gov/help/faqs/what-nature-suit-code (click "list of nature of suit codes" PDF hyperlink under "What is nature of suit code?" heading) [https://perma.cc/MK8N-U2XE] (listing nature of suit codes).

[83] For an explanation of the MDL procedure, see Clay D. Land, *Multidistrict Litigation After 50 Years: A Minority Perspective from the Trenches*, 53 GA. L. REV. 1237, 1238–40 (2019).

[84] Adam R. Pah, David L. Schwartz, Sarath Sanga, Charlotte S. Alexander, Kristian J. Hammond, Luís A.N. Amaral & SCALES OKN Consortium, *The Promise of AI in an Open Justice System*, 43 AI MAG. 71 (2022) (describing civil litigation ontology).

purpose English language model and trained further on the text of millions of docket entries to apply ontological labels to docket entries.[85]

We focus here on the portion of the SCALES ontology that covers dispositive events or those that terminate a party, claim, or entire case. The SCALES team developed nine labels, listed in Table 2, along with a brief definition of each. The labels relevant to the present settlement project are shaded in gray; we explain them further below.

TABLE 2: SCALES DISPOSITIVE EVENT LABELS,
POSSIBLE PARTY RESOLUTION CODES SHADED

Label	Definition
Administrative Closing	A procedural action taken by the court to temporarily remove a case from its active docket, usually pending the resolution of a related matter or awaiting further developments.
Default Judgment	A ruling in favor of the plaintiff when the defendant fails to respond or appear in court, resulting in an automatic decision without a full trial.
Outbound Transfer or Remand	An event where a case is transferred from one court to another, typically due to a change in jurisdiction or venue, and is removed from the transferring court's docket. A court order sending a case back to a lower court for further action or reconsideration, often due to procedural errors or new evidence.
Rule 12(b) Dismissal	A motion to dismiss a case for specific reasons outlined in the FRCP, such as lack of jurisdiction, improper venue, or failure to state a claim upon which relief can be granted. Actions on motions to dismiss that partially dismiss the case, including some of the claims or parties, are included.
Summary Judgment	A court ruling that decides a case without a full trial when there are no genuine disputes over material facts and the moving party is entitled to judgment as a matter of law. Actions on motions for summary judgment that partially dismiss the case, including some of the claims or parties, are included.

[85] For a general explanation of classification tasks such as this, see FRANCOIS CHOLLET, DEEP LEARNING WITH PYTHON 95–152, 309–63 (Jennifer Stout, Frances Buontempo, Aleksandar Dragosavljevic, Keri Hales & Andy Carroll eds., 2d ed. 2021), which explains classification from pages 95 to 152 and covers applications to text from pages 309 to 363. For a description of the training process for language models more generally, see JACOB DEVLIN, MING-WEI CHANG, KENTON LEE & KRISTINA TOUTANOVA, BERT: PRE-TRAINING OF DEEP BIDIRECTIONAL TRANSFORMERS FOR LANGUAGE UNDERSTANDING (May 24, 2019), https://arxiv.org/pdf/1810.04805v2 [https://perma.cc/EDM3-22CSV]. The full set of SCALES litigation event labels is available on the SCALES–OKN documentation site. *See Litigation Ontology*, SCALES, https://docs.scales-okn.org/guide/ontology/ [https://perma.cc/2895-8MVH].

Label	Definition
Trial	A formal legal proceeding where parties present evidence and arguments to a judge or jury to determine the outcome of a case, either in a criminal prosecution or a civil lawsuit.
Settlement	Settlement includes explicit indicators that the case has been resolved via settlement. It also includes "strong bilateral" dismissals of a case, which includes cases that are dismissed bilaterally using the terms "agreed," "joint," or "consent."
Voluntary Dismissal	Voluntary dismissal includes any of the following: notices of dismissal, stipulations of dismissal, orders granting notices or stipulations of dismissal, and orders disposing of the case with explicit reference to "voluntary dismissal" or Rule 41(a).
Other Dismissal	Entries that dismiss the case but do not fall into any of the other dispositive event categories are tagged with this label.

In simple terms, our workflow inputted the raw text of each docket entry in the SCALES corpus and applied a standardized litigation event label(s).[86]

To enable the model to work effectively on docket entry text, we further trained it on 11 million docket entries using a technique known as masked language modeling. This task involves randomly replacing approximately 15% of words or numbers, collectively known as "tokens," in docket entries with a special "mask" token and then training the model to predict the original text from the masked text. This technique is commonly used to adapt models pre-trained on general language corpora to more narrow domains, such as legal text.[87] This method improved our model's performance when

[86] Less simply, our starting point was a pre-trained base large language model, "large DeBERTaV3," a publicly available model developed by Microsoft researchers that has been exposed to (trained on) massive quantities of text, including BookCorpus, a dataset consisting of 11,038 unpublished books and all English-language Wikipedia entries. *See DeBERTaV3: Improving DeBERTa Using ELECTRA-Style Pre-Training with Gradient-Disentangled Embedding Sharing*, HUGGING FACE, https://huggingface.co/microsoft/deberta-v3-large [https://perma.cc/GS6G-8VU8]; PENGCHENG HE, JIANFENG GAO & WEIZHU CHEN, DEBERTAV3: IMPROVING DEBERTA USING ELECTRA-STYLE PRE-TRAINING WITH GRADIENT-DISENTANGLED EMBEDDING SHARING (Mar. 24, 2023), https://arxiv.org/pdf/2111.09543 [https://perma.cc/4WUM-Y4LA]; YUKUN ZHU, RYAN KIROS, RICHARD ZEMEL, RUSLAN SALAKHUTDINOV, RAQUEL URTASUN, ANTONIO TORRALBA & SANJA FIDLER, ALIGNING BOOKS AND MOVIES: TOWARDS STORY-LIKE VISUAL EXPLANATIONS BY WATCHING MOVIES AND READING BOOKS (June 22, 2015), https://arxiv.org/pdf/1506.06724 [https://perma.cc/8AKK-WCTS] (describing BookCorpus, which supplied training data for DeBERTa).

[87] For examples in other domains, see Jinhyuk Lee, Wonjin Yoon, Sungdong Kim, Donghyeon Kim, Sunkyu Kim, Chan Ho So & Jaewoo Kang, *BioBERT: A Pre-Trained Biomedical Language Representation Model for Biomedical Text Mining*, 36 BIOINFORMATICS 1234 (2020); Iz Beltagy, Kyle Lo & Arman Cohan, *SciBERT: A Pretrained Language Model for Scientific Text*, ARXIV (Sept. 10, 2019), https://arxiv.org/pdf/1903.10676 [https://perma.cc/UMP8-P3MK].

confronted with legal language that is used differently on docket sheets than in general usage—"settlement" is more strongly associated with the resolution of legal claims, for example, than the establishment of a community of people, and carries the same meaning as "agreement" or "resolution" within docket sheet entries.[88]

We then fine-tuned our model to build a set of binary classifiers that generate a prediction as to which litigation event label(s) applies to a given docket entry.[89] For each litigation event label, we generated between 1,000 and 4,000 manually labeled positive examples from among the docket entries and negative examples, which acted as the classifier model's training set.

We improved model performance iteratively by training the model, assessing labeling performance using a separate held-out validation set, identifying docket language where the model failed to apply the correct label, sampling additional examples with this "tricky" language to annotate and add to the training set, and repeating. Once the model reached satisfactory performance,[90] we computed litigation event labels for the full corpus of SCALES docket entries beyond the subset on which it was trained and validated.

We now turn to the particular disposition labels relevant to party resolution (shaded gray in Table 2): settlement and voluntary dismissal.

2. SCALES Party Resolution Labels: Settlement and Voluntary Dismissal

As introduced in Part I, a fundamental challenge in identifying party resolution from docket entry text stems from the extraordinary variation in the procedures that parties and judges use to memorialize out-of-court agreements. The list in Part II offers some examples of differing procedural configurations, from notices of settlement to consent decrees to joint motions

---

[88] *SCALES-OKN/Docket-Language-Model*, HUGGING FACE, https://huggingface.co/scales-okn/docket-language-model [https://perma.cc/HGL5-AHJ9].

[89] The final, publicly available version uses multi-label classification rather than multi-class classification, meaning that a single classifier will apply all relevant labels to any given docket entry rather than running each entry through multiple separate classifiers that apply nonexclusive binary labels.

[90] We defined "satisfactory performance" as achieving evaluation F1-scores above 0.96, with a crucial emphasis on the qualitative assessment of errors. Given our targeted approach to sampling, which focused on challenging or "tricky" language that was not representative of the overall distribution, a high F1-score indicated not just general performance in applying labels correctly but also the model's ability to handle complex cases in particular. This nuanced approach meant that even with high F1-scores, our primary criterion was whether the model's errors were reasonable and limited to edge-case (tricky) language within the context of our specialized dataset. Consequently, despite our training set's targeted nature, we expect the model's actual performance in broader applications to be substantially better. For a general explanation of classification-model performance assessment and F1-scores, see MAX KUHN & KJELL JOHNSON, APPLIED PREDICTIVE MODELING 247–74 (2013) (explaining "Measuring Performance in Classification Modeling").

to dismiss.[91] Further, even when describing the same procedure, language varies. "Total," "final," "complete," and "full" settlement may all describe an agreement that applies to all claims and parties. Likewise, when parties petition the court to dismiss a case after settling, they may file a "joint" or an "agreed" motion to dismiss. Judges similarly record settlement-related dismissals using a variety of phrases.

While our use of a large language model enabled us to handle the many lexical alternatives and synonyms used by the parties and the courts, we needed to identify a conceptual touchstone that would allow us, and by extension, our models, to differentiate between the docket entries that indicate settlement and those that do not.

We prioritized signals of bilateralism, mutuality, and agreement consistent with our definition of party resolution. Although peering into the hearts and minds of the parties is impossible, indicators of moral compromise do appear in docket text. Language such as "consent," "joint," "agreed," and "with prejudice" are examples.[92] Our settlement label, therefore, sweeps in settlement agreements, granted joint motions to dismiss with prejudice, consent judgments, and final agreed judgments—all explicit signals of consent and compromise. The settlement label also captures minute orders or notices from the court stating that the parties have reached a settlement or resolved the claims where the judge signals the party agreement.[93] We also

---

[91] *See supra* Section I.D.

[92] Dismissals with prejudice sacrifice the current lawsuit and bar future claims:

> It is settled that a judgment or order cannot act as a bar to a subsequent action between the same parties on the same cause of action unless it is rendered on the merits of the claim. However, it is frequently stated that a dismissal which recites that it is "with prejudice" is as conclusive of the rights of the parties as if the suit had been fully tried with a resultant judgment against the plaintiff.

David F. Ulmer, *Civil Procedure—Judgments—Res-Judicata Effect on Dismissal with Prejudice*, 50 MICH. L. REV. 600, 601 (1952).

[93] *See, e.g.*, Redman v. Keystone RV Co., No. 2:17-CV-44, 2018 BL 208687, at *1 (E.D.N.C. June 13, 2018) ("The court has been advised that the parties have settled all matters in controversy among them. Therefore, this matter is DISMISSED subject to the right of any party to file a motion to reopen the case should settlement not be consummated within 45 days hereof. The parties are directed to file their Stipulation of Dismissal with Prejudice on or before July 30, 2018. As there appears to be no further reason at this time to maintain the file as an open one for statistical purposes, this case is removed from the active docket.").

capture consent decrees[94] and accepted offers of judgment under Rule 68[95] as explicit signals of mutuality within our settlement label.

Our separate voluntary dismissal label is the one exception to the general "mutuality" rule. As explained above, voluntary dismissals are an inconclusive signal of settlement, as they might reflect *unilateral* decisions by plaintiffs to end a case.[96] However, some such filings mention settlement as the reason for dismissal. Further, we observed some U.S. district court practices that require parties to file voluntary dismissals under Rule 41 after notifying the court of a settlement.[97] In short, some courts use voluntary

---

[94] *See, e.g.*, Consent Decree at 1–3, Rocky Mountain Horse Ass'n, Inc. v. Rocky Mountain Horse Show Ass'n, Inc., No. 5:16-CV-00449 (E.D. Ky. Oct. 31, 2017), ECF No. 21 ("IT IS HEREBY ORDERED, ADJUDGED AND DECREED as follows": (1) "Absent prior written permission from RMHA to do so, Miller and Gean . . . are permanently ENJOINED and restrained from any use of Plaintiff's names and marks, namely, the words ROCKY MOUNTAIN HORSE, or ROCKY MOUNTAIN HORSE ASSOCIATION, or the initials RMHA, or any other marks or indicia confusingly similar to any of the foregoing marks including but not necessarily limited to any uses in association with any promotional materials, goods, services, printed materials, Internet websites, or programming"; (2) "Miller and Gean . . . are further permanently ENJOINED and restrained from any use of Plaintiff's names and marks, namely, the words ROCKY MOUNTAIN HORSE or ROCKY MOUNTAIN HORSE ASSOCIATION, or RMHA, or any other marks or indicia confusingly similar to any of the foregoing marks in association with promoting any horse shows or criteria for or reference to the nature or character of horses eligible to participate in same"; (3) "Miller and Gean . . . are further permanently ENJOINED and restrained from any actions tending to create any false or misleading representation tending to lead the trade or public erroneously to believe their goods or services have been produced, distributed, offered for distribution, advertised, promoted, displayed, officially sanctioned by, or otherwise licensed, sponsored, approved, or authorized by RMHA"; (4) "[P]arties . . . shall be responsible for their own costs and attorneys' fees"; (5) "This action is dismissed *without prejudice* as to Defendant Rocky Mountain Horse Show Association, Inc., a now legally dissolved corporation, and as to Defendant, Joe Miller"; and (6) "[T]his action is otherwise dismissed from the Court's active docket with prejudice as to Defendants Jeff Miller and Jane Gean.").

Because these two types of party resolution are distinct from others in that a judge enters a consent decree and retains jurisdiction to enforce, and the parties engage in a formal process dictated by Rule 68 of the Federal Rules of Civil Procedure, we created sublabels for consent decrees and accepted Rule 68 offers of judgment. The numbers were very small (398 consent decrees and 487 accepted Rule 68 offers in the full set), so we included them in the more general settlement-label counts in our analyses rather than calling them out separately.

[95] *See, e.g.*, Judgment Pursuant to Rule 68 at 2, Simmons v. City of New York, No. 1:16-CV-07306 (S.D.N.Y. Feb. 28, 2017), ECF No. 18 (finding in favor of Simmons against the City of New York for $1,501, "plus reasonable attorneys' fees, expenses and costs" for $3,000).

[96] *See, e.g.*, Notice of Voluntary Dismissal at 3, United States *ex rel.* Cole v. Barra, No. 5:16-CV-396 (N.D.N.Y. Jan. 19, 2017), ECF No. 13 (announcing settlement and providing the court with the parties' "Notice of Voluntary Dismissal").

[97] *See, e.g.*, Eckardt v. Eckardt, No. 2:16-CV-14211 (S.D. Fla. Dec. 21, 2016), ECF No. 53 ("ENDORSED ORDER. Plaintiffs have informed the Court that the Parties have settled this matter. The Parties shall submit closing documents by January 3, 2017. I advise the Parties that, contrary to Plaintiffs' suggestion in the Notice of Settlement [DE 52], a notice of dismissal is not an acceptable method of terminating an action, where, as here, the opposing party has already filed an answer [DE 26] and/or a motion for summary judgment [DE 47]. In order to properly dismiss the claims, the Parties must file a

dismissals to memorialize a party resolution and terminate a case. We, therefore, developed a separate voluntary dismissal label under our larger party resolution umbrella.

A docket entry can receive one or both settlement and voluntary dismissal labels in our labeling scheme. A plaintiff's bare-bones motion for voluntary dismissal, with no other description, would receive only the voluntary dismissal label. By contrast, a plaintiff-filed joint motion for voluntary dismissal with prejudice would receive both a voluntary dismissal label and a settlement label, as the "joint" and the "with prejudice" language act as signals of mutuality or compromise.[98] A "notice of settlement," in turn, receives only a settlement label.

Subcategorizing party resolutions at this granular level allows us to decide how widely to cast the empirical net, choosing how to treat (1) settlements that are described as such and have clear indicators of mutuality and bilateralism, (2) voluntary dismissals that are facially unilateral and have no indicators of mutuality, and (3) docket entries that have some combination of the two.[99]

Because our labeling scheme operates at the level of the docket entry, we can also capture the first indication of party resolution—a notice of settlement, for example—as well as the court's ultimate termination of a party, claim, or case via dismissal. Our methodology enables studying either or both of these litigation moments, which may be of separate interest to researchers for different reasons.[100] We think our labels indicate a "zone

---

joint stipulation of dismissal." (citing FED. R. CIV. P. 41(a)(1))); *see also* Cotropia et al., *supra* note 67, at 7 (finding that patent settlements were "typically accompanied with voluntary dismissals of the case by the court under Rule 41 of the FRCP").

[98] A close examination of Federal Rule of Civil Procedure 41, which governs voluntary dismissal, revealed a complication. Rule 41 requires a weakened form of consent—unopposed motions for voluntary dismissal—for unilateral action after an answer has been filed in a case. FED. R. CIV. P. 41. Because a true unilateral voluntary dismissal would still require "unopposed" status, we chose to exclude "unopposed" status as an indicator of mutuality. A plaintiff's unopposed motion for voluntary dismissal, without more, would only receive a voluntary dismissal label and not a settlement label.

[99] *See* Eisenberg & Lanvers, *supra* note 4, at 114 (noting the absence of a "single 'best' measure of the settlement rate" and suggesting that specific research questions should influence the definition of settlement). Granular settlement data allows researchers to categorize settlements in a way that is tailored to their individual research questions. For example, some researchers may want to exclude any possible false positives so that they may exclude voluntary dismissals. Other researchers may want to study the difference in docket proceedings between cases that indicate settlement and voluntary dismissal.

[100] There may be an important conceptual distinction between the first indication of settlement on a docket sheet and the court's ultimate action that terminates a party, claim, or case. For a discussion of settlement murkiness, see Hadfield, *supra* note 4, at 706–12. Substantial time and activity can intervene between these two events, including disputes about attorneys' fees and costs and future enforceability. Researchers interested in the hydraulics or drivers of civil settlement may be very interested in the timing of the agreement rather than the court's eventual ratification in the form of dismissal. Those scholars

of settlement" on a docket sheet, which we define as the period between the parties' out-of-court moral compromise and the effectuation of the legal compromise.

When combined with the other SCALES dispositive event labels listed in Table 2, our docket entry-level labeling also allows us to examine the sequencing and timing of party resolution activity with other dispositive actions in a case. This enables us to capture partial case dispositions, such as a partially granted defendant's motion for summary judgment and subsequent settlement of the remaining claims, which is a key point of departure between our docket sheet-based SCALES approach and the IDB. For a less granular view, one might also roll up all docket entry-level party resolution indicators to a single case-level indicator, capturing whether the case contained any party resolution, in full or in part.[101]

C. Limitations

Despite its flexibility and careful design, our methodology has some known limitations. The varied language and procedures that lawyers use to notify courts of a settlement—and that courts use in managing this stage of litigation—can muddy the demarcation line between mutual and unilateral case resolutions. We developed our approach to solve this, but errors may nevertheless occur. We might undercount because we missed some pocket of local party resolution procedure that is substantially different from the norm. For example, manual reviews of model outputs revealed that our labeling scheme missed docket entries in the Eastern District of Pennsylvania that cited a local rule involving settlement.[102] Another challenge was parties' use of relatively obscure terms, at least in federal practice, such as "praecipe."[103]

Further, settlement might be hiding in other places on the docket sheets. Our manual review found some docket entries labeled only "Motion to Dismiss," but the filing itself specified that the motion was prompted by a settlement. Without universal access to the underlying documents, differentiating between a unilateral motion to dismiss under Rule 12(b) and a bilateral motion to dismiss prompted by a settlement is impossible when

would likely focus on the closest proxy on the docket sheet: the parties' notification to the court. On the other hand, scholars interested in courts' caseload management practices would likely focus on removal of a case from the active caseload via a dismissal.

[101] Vacek & Schilder, *supra* note 4, at 210.

[102] E. DIST. PA. CIV. R. 41.1(b).

[103] *See, e.g.*, Plaintiffs' Praecipe to Withdraw Without Prejudice, Maull v. Strang Corp., No. 2:16-CV-02738 (E.D. Pa. July 28, 2016), ECF No. 3 (using the phrase praecipe). A praecipe can be a request for a court to issue a judgment or order, and so could act as a synonym for "motion." *Praecipe*, BLACK'S LAW DICTIONARY (11th ed. 2019).

the docket text lacks detail. Because most such motions that we reviewed were, in fact, Rule 12(b) motions, we grouped them with our SCALES dispositive label for Rule 12(b) dismissals rather than with our party resolution label. This likely caused us to miss some party resolutions, contributing to a known undercount.[104]

One more likely instance of undercounting stems from docket sheets in which our SCALES models detected no dispositive events. This set of "zero disposition" cases represents 3.8% of the full set of 346,916—or 13,270 cases. From our manual review of docket sheets, we suspect that in some of these, the parties abandoned the litigation—perhaps due to settlement—and the cases lay dormant, our models did not detect the court action that closed the case, or both.

However, the IDB disposition codes assigned to these cases support our hunch that they diverged from the common, more typical set of civil litigation outcomes and may include some party resolutions. A plurality (27%) had IDB code 14 for "other dismissal"; another 12% did not appear in the IDB at all, suggesting that they were perhaps opened in error and never recorded by the AOUSC; and 10% were listed with IDB code 18 for "statistical closing." A combined 23% of these SCALES zero-disposition cases had one of the three IDB codes most associated with settlement (5, 12, and 13 for consent judgments, voluntary dismissals, and settlements, respectively). Thus, after further manual review, some subset of these approximately 13,000 cases may receive a SCALES party resolution disposition label.

On the other hand, considering too many docket entries to be settlement signals may have resulted in overcounting. The main culprits here are the aforementioned bare-bones voluntary dismissals, which we addressed by labeling separately. However, docket entries (such as scheduling orders) discuss settlement *procedures* in great detail but do not actually state a party resolution event. We conducted extensive quality control to ensure that such entries did not receive either of our party resolution labels, but errors may nevertheless occur.[105]

---

[104] Without full access to the underlying documents, we cannot quantify the undercount's size. Our review of a sample of motions and the rest of the docket sheet entries in those cases suggested that most motions labeled "Motion to Dismiss" on the docket sheet without more were filed under Rule 12(b).

[105] We built quality controls throughout the research process by screening out docket entries that only mention settlement in scheduling orders. Before building the model, we studied the docket sheet language to identify the different uses of settlement and settlement-like language, such as "consent" and "by agreement." We built the initial training sets for the model using keyword combinations derived from our docket sheet review, such as "settlement" with "resolved," "dismissed," or "all claims." The researchers also reviewed model outputs at all stages of development to confirm the model's output. The

We also note some data limitations. Our docket sheet corpus captures cases filed during 2016 and 2017 and terminated within four years of filing. Excluded cases may be those that were especially complex and long-lasting. Other years of data may display different patterns around party resolution, particularly during and after the COVID-19 years. Yet as other troves of docket sheets become available, the SCALES labels can be applied to additional time periods; experiments are also underway to apply the SCALES labeling scheme to state court docket sheets.

Globally, we view our SCALES work as a complement to the IDB. Each data source helps to fill holes in and addresses the limitations of the other. The IDB crosswalk in Part IV expands on this theme, allowing us to address the limitations of our SCALES data and methods while improving on the current state of the art as reflected by the IDB.

IV. FINDINGS

We now present our party resolution findings on their own and then as a complement to the IDB's settlement numbers. We end by combining our SCALES-derived tallies of party resolution with the IDB's equivalent codes, offering a single definitive estimate of party resolution in federal civil litigation.

A. Party Resolution Results

We start with our full set of 346,916 docket sheets; 196,595—or almost 57%—include one of our SCALES party resolution labels as either the case's single disposition or one disposition among many.[106] Table 3 shows the distribution of the party resolution label options, dividing docket entries into those that have only unilateral, bare-bones voluntary dismissal language ("Voluntary dismissal only") and those that have some indication of mutuality and compromise, even if they also use the language of voluntary dismissal and Rule 41 ("Settlement").

TABLE 3: PARTY RESOLUTION DISPOSITIONS

Disposition Label	Frequency	Percent of All Cases (N=346,916)	Percent of Party Resolutions (N=196,595)
Settlement	105,431	30.4%	53.6%
Voluntary dismissal only	91,164	26.3%	46.4%

model was refined and iterated upon each stage of review for each label until we could reliably classify each of our party resolution docket entries.

[106] The approximately 13,000 cases still pending are in the 346,916 denominator.

A more conservative interpretation of the labeling might exclude the "Voluntary dismissal only" labels and only count those docket entries with mutuality language, as captured by the "Settlement" label. Under this approach, rather than about 57% of all cases, only about 30% of cases contain a party resolution.

The conservative approach offers the advantage of no false positives, but it almost certainly excludes party resolution events. For example, within the settlement scholarship cited in this Essay, the most frequently reported settlement rate is around 67%, with an unweighted average of 58.5%.[107] Apples-to-apples comparisons with previous studies are nearly impossible due to differences in definitions and data sources.[108] Substantial differences may also emerge across courts and case types. Those caveats aside, the scholarship suggests that more than half of all civil cases contain at least some party resolution, supporting a less conservative approach to combining our party resolution labels.

We now turn to party resolution in the context of single- and multiple-disposition cases. First, single-disposition cases represent about 84% of the cases in our scope or 292,911 cases. Of these, 162,329, or 55%, ended with one of the two party resolution labels as the single disposition. Table 4 below shows the breakdown of the party resolution labels. The more conservative approach would include only the "Settlement" labels and suggest that only 29% of cases ended with a full settlement. Regardless, even the 29% figure would represent a plurality compared to other single-disposition outcomes.

---

[107] Boyd & Hoffman, *supra* note 3, at 911 (finding that 77% of 585 cases in their database—all piercing the corporate veil cases—settled); Clermont & Eisenberg, *supra* note 18, at 136 (examining almost 260,000 federal civil cases and finding a settlement "one way or another" in at least 66.7%); Eisenberg & Lanvers, *supra* note 4, at 114–15 (estimating the aggregated settlement rate across all case categories in two districts to be 66.9% and finding variation by case type); Barkai et al., *supra* note 2, at 35 (finding that 84% of tort cases settle and 20% of foreclosure cases settle); Hadfield, *supra* note 4, at 730 (finding an overall settlement rate—including consent judgments—of 40.5%); Colvin, *supra* note 3, at 1, 6 ("In the AAA-CC dataset, 2,328 cases representing 59.1 percent of the sample were resolved by settlement."); Peter Grajzl & Katarina Zajc, *Litigation and the Timing of Settlement: Evidence from Commercial Disputes*, 44 EUR. J.L. & ECON. 287, 293 (2017) ("240 out of the 564 resolved cases (43%) were resolved via settlement."); Posner, *supra* note 13, at 424 (finding the "average probability of settlement [to be] 69 per cent").

[108] *See, e.g.*, Barkai et al., *supra* note 2, at 35 (finding that 84% of tort cases, 45% of contract cases, 20% of foreclosure cases, and 51% of "other" cases settle).

TABLE 4: SINGLE-DISPOSITION CASES

Disposition Label	Frequency	Single-Disposition Percent (N=292,911)	Party Resolution Percent (N=162,329)
Settlement	86,343	29.5%	53.2%
Voluntary dismissal only	75,986	25.9%	46.8%
Other dismissal	66,484	22.7%	-
Transfer or remand	24,777	8.5%	-
Rule 12(b) dismissal	17,079	5.8%	-
Summary judgment	11,950	4.1%	-
Default judgment	7,669	2.6%	-
Administrative closing	1,528	0.5%	-
Trial	1,095	0.4%	-

While the single-disposition findings above are roughly consistent with the previous literature on settlement, our multiple-disposition findings shed new light on settlement's role as a trimming tool. Of our full scope of 346,916 cases, our models detected multiple dispositions in about 12% or 40,735 cases. Of those, just over 84%, or 34,266, contained one of our party resolution labels. Our findings suggest that partial party resolutions, either in the form of settlements or voluntary dismissals, are a common and important winnowing process within litigation. This phenomenon is understudied, however, as research that relies on the IDB's disposition codes likely misses the presence of party resolutions among multiple dispositive events.

Table 5 lists the other dispositions that appear alongside party resolution in multi-disposition cases. Though work is ongoing to identify and analyze the particular disposition sequences in which party resolution appears and in which order, one can surmise from Table 5 that Rule 12(b) dismissals and default judgments likely appeared before party resolution, as they typically occur at the outset of civil litigation. Summary judgment is less clear, as is trial.[109] Nevertheless, these results—to be refined in future work—are fodder for further explorations of litigation hydraulics that push parties toward or away from settlement. Antecedent litigation events may be

[109] For a discussion on the efficiency of summary judgment and its effects on trials, see Edward Brunet, *The Efficiency of Summary Judgment*, 43 LOY. U. CHI. L.J. 689, 690 (2012), (arguing that summary judgment has several efficient effects, including fact clarification, early legal analysis, formal pretrial assessment of a case's strength, and a "settlement premium" that nonmoving parties gain when a motion for summary judgment is dismissed). *But see* John Bronsteen, *Against Summary Judgment*, 75 GEO. WASH. L. REV. 522, 527–30 (2006) (arguing that summary judgment is costly because it discourages early settlement and thus requires parties to go through pretrial litigation).

particularly powerful in shaping parties' attitudes toward compromise as other judicial dispositions winnow and trim their claims or liability exposure.

TABLE 5: OTHER DISPOSITIONS APPEARING WITH PARTY RESOLUTION IN MULTI-DISPOSITION CASES

Disposition Label	Frequency	Multi-Disposition Percent (N=34,266)
Rule 12(b) dismissal	14,127	41.2%
Summary judgment	6,189	18.1%
Default judgment	5,713	16.7%
Multiple other dispositions	4,350	12.7%
Transfer or remand	2,592	7.6%
Trial	1,000	2.9%
Other dismissal	295	0.9%

Consistent with previous settlement studies, we observed differences in party resolution distribution in both single- and multiple-outcome cases by NOS code, used as a proxy for claim type, district, and judge. For example, 75% to 80% of single-disposition cases involving claims under the Fair Labor Standards Act (NOS 710), Americans with Disabilities Act—Other (NOS 446), Copyrights (NOS 820), and Consumer Credit (NOS 480) had single-disposition party resolutions. In comparison, cases involving Prison Conditions (NOS 555) and some federal Personal Injury claims (NOS 367) had single-disposition party resolution percentages around or below 15%.[110] Districts, too, displayed different party resolution rates; for example, the Southern District of West Virginia resolved about 77% of its cases in our study via full party resolution, compared with only 27% in the Middle District of Tennessee.[111]

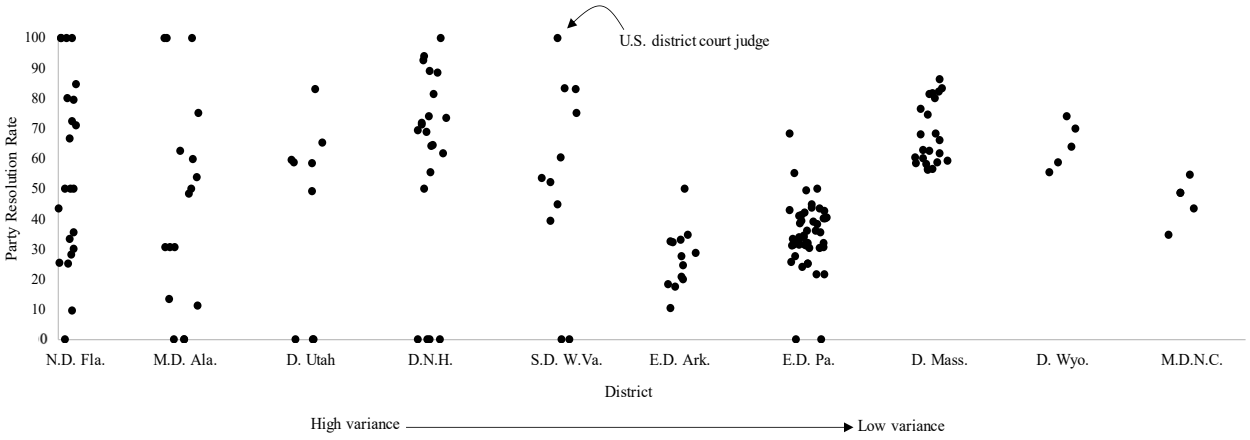
Finally, party resolution varied by individual judges within districts. Because case assignment to judges within federal judicial districts is generally random, there is no reason to expect that one judge would receive a relatively more or less tractable set of cases or litigants than another. In some districts, however, we observe substantial variation in per-judge party resolution rates (the percentage of cases in which any party resolution appeared, whether in single or multiple dispositions). For purposes of

[110] Online Appendix C lists the number and percentage of cases with party resolutions by NOS code.

[111] Online Appendix D lists the number and percentage of cases with party resolutions by U.S. district court.

illustration, Figure 1 shows the five districts with the highest and lowest variance by judge. Each judge is represented by a dot, grouped by district, with per-judge party resolution rates on the y-axis. This illustration is merely suggestive and raises many questions. How many cases did each judge resolve?[112] What was the size of each judge's caseload at the time? Are there interesting relationships between years on the bench or other judge-level characteristics and party resolution rates? Interestingly, the five lowest variance districts on the right of the graph differ across districts in their groupings of judges' party resolution rates, lending further support to the interdistrict variation observations above.

FIGURE 1: PARTY RESOLUTION RATES BY JUDGE—FIVE HIGHEST AND LOWEST VARIANCE DISTRICTS



We present the above findings not as definitive as to NOS, district, or judge distribution—some methodology or data-related phenomena might very well explain the differences noted above—but rather as suggestive of the types of analyses that might build on our SCALES-based approach to defining and counting party resolutions in federal court. We return to these avenues for further inquiry in Part V. Next, we move to the IDB.

[112] For instance, some judges with senior status may handle very few cases and settle all or none, producing a party resolution rate of 100% or 0%. Small denominators likely explain the very high and very low per-judge party resolution rates in Figure 1. A full analysis of all judges in all courts could include normalization strategies to account for low-caseload judges.

B. Party Resolution Comparison: IDB and SCALES

In this Section, we crosswalk our SCALES party resolution findings with the IDB's disposition codes. This exercise helps validate our SCALES party resolution methods and findings while also revealing some of the IDB's shortcomings as the sole data source on federal civil litigation pathways and outcomes.

1. Crosswalk Methodology

To match our SCALES dispositions to the codes recorded in the IDB, the SCALES team created a common key that would allow a crosswalk between the two data sets. Cases are recorded in the IDB using codes for "district," "office" or division within a judicial district, and "docket number."[113] Because each of our docket sheets in the SCALES corpus also has a value for each of these fields, we constructed a unique identifier for most cases in the SCALES corpus that tracked those three IDB codes.

Using this methodology, we could locate an IDB record for 98% of our full set of 346,916 cases. The remaining cases were likely missing from the IDB for various reasons, including errors in the multi-stage data-gathering process described in Section I.A above and per-district variations in reporting.[114]

An additional complexity in crosswalking to the IDB arises where there is more than one IDB row per case. Although the IDB is typically structured with one row and, therefore, one disposition code per case, sometimes multiple rows are generated when a case is closed and reopened.[115] In our analysis, when multiple rows for a case had the same disposition code, we consolidated them into a single case record. A small number of cases had multiple unique disposition codes and were not consolidated but labeled as "Multiple IDB Dispositions" in subsequent tables.[116]

For the next Sections' findings, we further removed the set of SCALES "zero disposition" cases that we described in Section III.C because they had no dispositive-event label on the SCALES side to match to a disposition code on the IDB side. This left a new scope of 328,869 cases with both an IDB and SCALES disposition code or codes, or 95% of the original scope.

---

[113] *Integrated Data Base Civil Documentation*, *supra* note 37.

[114] *See Crosswalking PACER to the IDB*, SCALES (Nov. 30, 2021), http://livingreports.scales-okn.org/#/idbCrosswalkReport [https://perma.cc/F6FV-KPAK] (detailing reasons for crosswalk failures).

[115] E-mails from Kristin Garri to Charlotte S. Alexander, *supra* note 38.

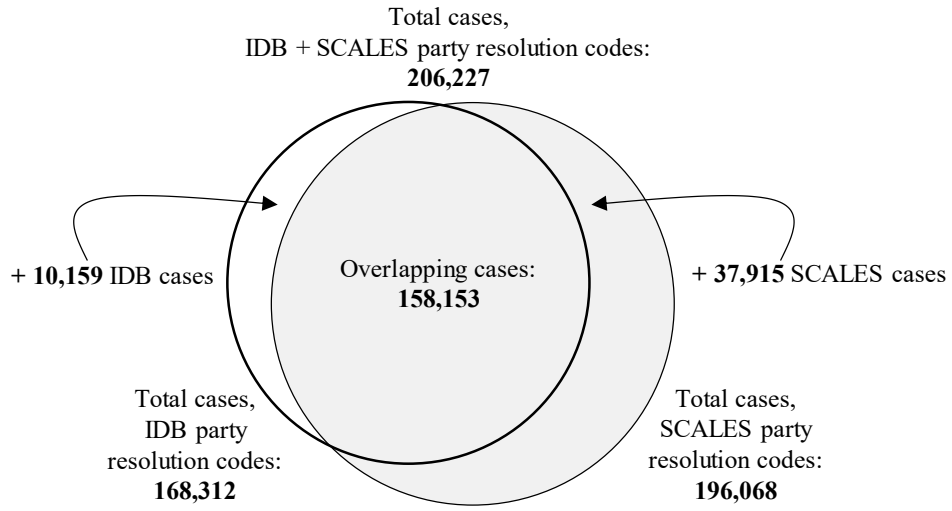
[116] This means that the "Multiple IDB Dispositions" category in later tables includes cases that could qualify under other disposition categories, thereby slightly influencing their overall distribution.

2. One Number to Rule Them All?

We start by harmonizing the SCALES and IDB data sources to generate a single settlement number that captures our concept of party resolution in both single and multiple disposition cases and all of its procedural manifestations.

Figure 2 below illustrates this process, showing the party resolution dispositions that both data sources captured (the bulk of cases), plus the approximately 10,000 cases that SCALES missed but the IDB captured, and the approximately 40,000 cases that the IDB missed but that SCALES identified. This analysis suggests that of the 328,869 cases in this scope, 206,227 cases, or 63%, included a party resolution of some type.

FIGURE 2: IDB AND SCALES PARTY DISPOSITION CODE HARMONIZATION



Next, we turn to comparisons of IDB-to-SCALES and SCALES-to-IDB data to provide an in-depth depiction of their overlap and variation.

3. IDB-to-SCALES Comparison

First, we compare the set of IDB disposition codes that appear most likely to indicate party resolution, marked in gray in Table 1: 5, or consent judgment; 12, or voluntary dismissal; and 13, or settled. Within the new scope above, 51%, or 168,312 cases, had one of those three IDB codes. Of those, our SCALES models applied our voluntary dismissal, settlement, or both labels to 158,153 cases, a 94% overlap. Our SCALES models missed 10,159 cases, or 6% of the cases.

Table 6 indicates that those approximately 10,000 cases primarily received two other SCALES labels instead of party resolution: "Other dismissal" and "Rule 12(b) dismissal." This is a known limitation of our SCALES modeling approach, discussed in depth in Section III.C, and can be refined in future modeling runs.

TABLE 6: SCALES DISPOSITIONS IN CASES NOT MATCHED TO IDB PARTY RESOLUTION DISPOSITIONS

SCALES Disposition	Frequency	Percent (N=10,159)
Other dismissal	6,983	68.7%
Rule 12(b) dismissal	2,020	19.9%
Summary judgment	357	3.5%
Transfer or remand	319	3.1%
Multiple dispositions	251	2.5%
Default judgment	194	1.9%
Trial	29	0.3%
Administrative closing	6	0.1%

4. SCALES-to-IDB Comparison

Second, we switch starting points and begin with the cases that SCALES identified as containing party resolution dispositions and investigate the dispositive codes they received in the IDB. Of the new 328,869-case scope, 60% (196,068 cases) had any of our SCALES party resolution labels. Of those, 158,153 cases (81%) also had one of the three IDB party resolution disposition codes, leaving 37,915 cases (or 19%) where our SCALES models picked up signals of party resolution from the docket sheet text, but where the IDB applied a different disposition code.

Four possible reasons may account for these nearly 40,000 SCALES–IDB mismatches:

(1) The IDB is coding party resolution but using another IDB disposition code to do so;
(2) The IDB is not coding party resolution but is instead coding some other dispositive event that is also present in the case;
(3) There is some other IDB coding error that undercounts party resolutions; or
(4) There is some other SCALES error that overcounts party resolutions.

While we addressed the risk of SCALES overcounts in Section III.C, we cannot opine further about the third possible reason, as we do not have enough insight into the various entry points for error in the IDB assembly process. Below, however, we explore the first two reasons centered on the IDB's disposition codes.

Table 7 shows the distribution of the 37,915 cases in which SCALES identified a party resolution, but the IDB did not. The gray rows together represent just over 80% of the full set of cases. We suspect that these are cases in which the clerks or other court personnel responsible for generating the data for inclusion in the IDB selected a disposition code other than 5, 12, or 13 to record party resolution.

Indeed, as discussed above, the parties might very well file a "Motion before trial" as a motion to approve a settlement agreement or seek "Other dismissal" of their claims due to an out-of-court compromise. Likewise, cases reported settled may receive a "Statistical closing" label while awaiting the filing of final dismissal documents or proceed to "Other judgment" to memorialize a consent decree. Professor Hadfield's previous analyses of the IDB's disposition codes revealed as much, finding settlements on the docket sheets of cases with the "Other dismissal" and "Motion before trial" IDB disposition codes.[117] We therefore suspect that the great majority of the almost 40,000 party resolutions that the IDB missed but that SCALES caught were merely using other IDB codes to capture party resolution, which is the first possibility in the list above.

TABLE 7: IDB DISPOSITIONS IN CASES NOT MATCHED TO SCALES PARTY RESOLUTION DISPOSITIONS

IDB Disposition Code	IDB Disposition Label	Frequency	Percent (N=37,915)
14	Other dismissal	19,697	52.0%
6	Motion before trial	5,169	13.6%
18	Statistical closing	2,783	7.3%
17	Other judgment	2,704	7.1%
0	Transfer to another district	1,962	5.2%
4	Default	1,475	3.9%
2	Want of prosecution	971	2.6%
-	Multiple IDB dispositions	815	2.1%
1	Remanded to state court	725	1.9%
7	Jury verdict	562	1.5%

[117] Hadfield, *supra* note 4, at 719–20.

IDB Disposition Code	IDB Disposition Label	Frequency	Percent (N=37,915)
3	Lack of jurisdiction	288	0.8%
-8	Missing	183	0.5%
9	Court trial	171	0.5%
10	Multi-district litigation transfer	152	0.4%
16	Stayed pending bankruptcy	130	0.3%
11	Remanded to U.S. Agency	43	0.1%
15	Award of arbitrator	39	0.1%
8	Directed verdict	34	0.1%
19	Appeal affirmed (magistrate judge)	7	0.02%
20	Appeal denied (magistrate judge)	5	0.01%

Looking beyond those four shaded IDB codes, we can explore the second possibility listed above: that party resolution was one of many dispositions in these cases, and the IDB's single-disposition structure privileged another disposition over a partial party resolution. In Online Appendix B, we reproduce the nongray rows from Table 7 above and report for each row the percentage of cases in which we detected multiple SCALES dispositions. For those multi-disposition SCALES cases, we then searched for nonparty resolution dispositions from among the SCALES labels that matched the IDB label applied to the case. For the cases in which the SCALES models detected party resolution *and* trial, for example, how often did the IDB apply the trial label? We hypothesize that the court personnel who assembled the data that eventually became the IDB might have applied an informal set of trumping rules, in which trial trumped party resolution or later-occurring dispositions trumped earlier ones.

Online Appendix B supports this hunch. As an example, the first two rows list the 733 cases in which SCALES detected party resolution, but the IDB assigned one of two trial-related IDB dispositions. SCALES also detected trial in almost all of these cases alongside a party resolution label, suggesting that the parties reached a settlement on some claims and proceeded to trial on others. Beyond trial, transfers, remands, defaults, and dismissals appear near the top of the table, suggesting that the IDB's disposition coding may privilege these outcomes over party resolution when both appear in the course of a case. This data structure disguises the role that settlement plays in narrowing the scope of a dispute and prevents the study of how party resolution and other court-driven adjudications interact with and influence one another.

We now turn to future work that we anticipate will stem from the SCALES project, moving from our focus on defining and counting party resolution to a broader consideration of federal civil litigation's myriad pathways and outcomes.

V. FUTURE WORK

This Essay offers a conceptual framework for defining settlement as party resolution and an integrated approach to counting it, drawing from both the SCALES project and the IDB. With this as a starting point, we turn to multiple areas of future research, including pathways, prompts, parties, case types, and judicial factors.

Pathways—Future work will create a topography of federal civil litigation by describing the different pathways that cases travel, settlement's placement along those paths, and the events that come before and after. Although party resolution is a common civil litigation outcome regardless of which data source or definition one uses, relatively little is known about its timing and antecedents in the litigation process.[118] Combining SCALES data with the IDB identifies the universe of cases that involve party resolution in full or in part. Adding both further litigation-event labels such as those available in the larger SCALES ontology (as shown in Table 2), as well as additional labels yet to be developed, such as tracking the beginning and end of discovery, situates party resolution within the lifecycle of litigation.[119]

Further, while the results reported herein group voluntary dismissal and settlement events together as two components of party resolution, there is even more granularity available in the SCALES labeling scheme. We separately labeled accepted offers of judgment under Rule 68 and consent decrees, which are unique procedural subsets of party resolution. One can study these party resolution events separately or in conjunction with general settlements. Our intuition is that consent decrees will have different pathways than voluntary dismissals, which will differ from settlement agreements. Identifying the different party resolution events facilitates a clearer window into the particular pathways of each.

Prompts—Accounting for prior litigation events will also shed light on the hydraulics of party resolution. Another area of study is to explore what prompts various party resolutions. What forces push a case, a claim, or a party toward a non-adjudicated resolution? What role do judicial actions

[118] *See supra* notes 4–10 and accompanying text.

[119] This line of future research is our most established and draws on our earlier case pathway work. *See* Charlotte Alexander, Khalifeh al Jadda, Mohammad Javad Feizollahi & Anne M. Tucker, *Using Text Analytics to Predict Litigation Outcomes*, *in* LAW AS DATA: COMPUTATION, TEXT, & THE FUTURE OF LEGAL ANALYSIS (Michael A. Livermore & Daniel N. Rockmore eds., 2019).

such as decisions on dispositive motions, judicial structures such as mandatory mediation, and judicial signaling such as comments on the record encouraging settlement have in prompting party resolution?[120] In our review of thousands of civil dockets, we have observed the full spectrum of court behavior vis-à-vis party resolution—from total silence, to nudges, to sending parties to mediation repeatedly.[121] We hope to expand on existing scholarship examining judicial influence in settlement outcomes to explore the relationship between judging and nudging parties to resolution.[122]

Parties—Who wins or loses in litigation predicts the success of future claims. But these outcomes also suggest whether systems are even-handed and fair or if they give some but not all an advantage. Understanding which parties resolve cases and when can help assess how settlement contributes to civil litigation outcomes.[123]

Party resolution, whether perceived as a win or a loss by a party, has positive economic consequences in reducing dispute resolution costs and in time savings.[124] Who settles and who goes to trial may also tell an important story about who has access to courts, who has access to the necessary legal

[120] *See, e.g.*, Lynch & Levine, *supra* note 9, at 241–51 (describing the role of judges in prompting settlement).

[121] In this project, *Judging and Nudging*, we empirically investigate if greater judicial management, through Rule 16 conferences and actions, is associated with faster case resolution times. Charlotte Alexander, Roger M. Michalski & Anne M. Tucker, Judging & Nudging (June 27, 2024) (unpublished manuscript) (on file with author); *see also* Sean P. Sullivan, *Why Wait to Settle? An Experimental Test of the Asymmetric-Information Hypothesis*, 59 J.L. & ECON. 497, 521 (2016) (finding that asymmetric information delays settlement but arguing that, given the high settlement rates, litigants are better served by divulging information sooner rather than later).

[122] *See, e.g.*, Boyd, *supra* note 18, at 204 (finding judge-gender effects on settlement outcomes). We are particularly interested in understanding if judicial prompts on the docket are associated with a party resolution outcome and, if so, which prompts facilitate which resolution events.

[123] For example, Professor Hadfield found differences in settlement rates when comparing individual versus organizational plaintiffs and whether attorneys were compensated hourly or by contingency fees. Hadfield, *supra* note 19, at 1314; *see also* James D. Cox, Randall S. Thomas & Lynn Bai, *There Are Plaintiffs and . . . There Are Plaintiffs: An Empirical Analysis of Securities Class Action Settlements*, 61 VAND. L. REV. 355, 367–74 (2008) (discussing how the type of plaintiff that heads a securities class action impacts the overall outcome of the case); Marc Galanter, *Why the "Haves" Come Out Ahead: Speculations of the Limits of Legal Change*, 9 L. & SOC'Y REV. 95, 97–119 (1974) (discussing the effects different kinds of parties may have on the workings of the legal system); Shari Seidman Diamond & Jessica M. Salerno, *Reasons for the Disappearing Jury Trial: Perspectives from Attorneys and Judges*, 81 LA. L. REV. 120, 121 (2020) (reporting that respondents ranked jury trials as "less predictable, slower, and less cost-effective" than alternative procedures that facilitate party-led resolutions).

[124] Hadfield, *supra* note 19, at 1311–12 ("Increased settlement rates are an important economic phenomenon in terms of the possible saving of dispute resolution costs, but the implications of this saving are different if we are talking about the resolution of a commercial contracting dispute than if we are talking about a civil rights dispute."); *see also* Paul Pecorino & Mark Van Boening, *An Empirical Analysis of the Signaling and Screening Models of Litigation*, 20 AM. L. & ECON. REV. 214, 217 (2018). Parties' repeat exposure to litigation, including settlement and representation by an attorney, are modeled to affect settlement outcomes positively. *Id.* at 239.

knowledge to navigate litigation successfully, and who does not.[125] Professor Hadfield's work advanced theories of legal haves and have-nots framed in the larger question of democratic ideals when claims brought by individuals were associated with higher litigation costs, longer delays, and less favorable outcomes.[126] More comprehensive data on party resolution, combined with party type, lawyer, and law firm characteristics—additional SCALES data features—can advance research into the winners and losers of settlement, civil litigation, and access to justice more generally.[127]

Case Types—Different case types have different settlement rates, a phenomenon that has been documented in several prior studies and confirmed in our NOS code findings reported in the Online Appendices.[128] Using these data, researchers interested in particular subsets of litigation such as patents,[129] employment,[130] securities,[131] commercial disputes,[132] or torts[133] can explore party resolution questions within their disciplines.

Judicial Factors—Party resolution is not adjudication,[134] but it happens in the context of litigation or in the shadow of the court.[135] As suggested by the findings in Section IV.A, but not fully explored in this project, party resolution rates vary—sometimes substantially—by district and judge.[136] This variation may indicate that the settlement temperature of the chambers or the court context in which party resolution negotiations occur influences party resolution rates.[137]

---

[125] Hadfield, *supra* note 19, at 1319 (referencing Galanter, *supra* note 123).

[126] *Id.* at 1321–22.

[127] *See generally* David Colarusso & Erika J. Rickard, *Speaking the Same Language: Data Standards and Disruptive Technologies in the Administration of Justice*, 357 SUFFOLK U. L. REV. 387, 388 (2017) (arguing that state trial court data is the biggest barrier to access to justice).

[128] *See supra* notes 110–111 and accompanying text.

[129] *See, e.g.*, Cotropia et al., *supra* note 4, at 108 (studying settlement in patent cases).

[130] *See, e.g.*, Colvin, *supra* note 3, at 3 (studying settlement in employment cases).

[131] *See, e.g.*, Cox et al., *supra* note 123, at 358 (examining settlement in securities cases).

[132] *See, e.g.*, Boyd & Hoffman, *supra* note 3, at 900 (studying settlement rates in nearly 600 piercing the corporate veil cases).

[133] Eisenberg & Lanvers, *supra* note 4, at 130 (examining settlement by case type, including torts).

[134] *See, e.g.*, *supra* notes 45–54 and accompanying text (comparing and contrasting party resolution and adjudication); *see also* Galanter, *supra* note 19, at 460–65 (describing the decline in civil trials).

[135] *See, e.g.*, Barkai et al., *supra* note 2, at 35–36 (examining settlement rates across case types). *See generally* Pauline T. Kim, Margo Schlanger, Christina L. Boyd & Andrew D. Martin, *How Should We Study District Judge Decision-Making?*, 29 WASH. U. J.L. & POL'Y 83, 83–112 (2009) (advocating for increased focus on the litigation process in studying judicial decision-making in federal district courts).

[136] *See supra* Figure 1 and accompanying text.

[137] *See, e.g.*, Grajzl & Zajc, *supra* note 107, at 312 (finding that the time to settlement decreases with participation in court mediation programs and early hearings in a sample of Slovenian commercial disputes).

Future work is needed to unpack the relationship(s) between party resolution and court- and judge-related factors, and other case-, party-, and lawyer-related variables suggested above.

CONCLUSION

This Essay has undertaken the task of defining and counting settlements in federal civil cases. Building on our review of the settlement literature, we first offered a flexible, empirically informed, and operationalizable definition of settlement as party resolution. This definition encompasses any resolution of claims, in part or in full, reached by the parties in the absence of those claims' adjudication by the court. Next, we counted party resolutions as they appeared on lawsuits' docket sheets using a set of machine learning classification models trained on 11 million docket sheet entries newly available via the SCALES–OKN project. We built our modeling approach to accommodate the wide lexical and procedural variations in how parties informed the court of their settlements (including plaintiff-filed voluntary dismissals) and that courts then used to dispose of the affected claim or claims.

From our docket entry-level labels, we generated a set of findings on the frequency and distribution of party resolutions. Notably, we discovered approximately 40,000 additional party resolutions that were missing from the main existing source of administrative data on federal litigation, the FJC's IDB. We also explored the co-occurrence of other dispositive events, such as trial or partially granted motions to dismiss or for summary judgment, with party resolution. Finally, we combined our new SCALES party resolution numbers with the IDBs, producing a single definitive estimate of the number of federal civil cases involving party resolution, in part or in full. We outlined multiple directions for future research to further explore party resolution's characteristics, predictors, and distribution. This research and the work of the larger SCALES collaboration will contribute to continuing scholarly examinations of litigation and the functioning of the courts, including questions around fairness and equity, efficiency, and access to justice.

ONLINE APPENDICES

The Online Appendices are hosted on Digital Commons and can be accessed at https://scholarlycommons.law.northwestern.edu/nulr/vol119/iss1/4/.