Cleveland State Law Review

Volume 69 | Issue 2 Article

3-15-2021

The Elastics of Snap Removal: An Empirical Case Study of Textualism

Thomas O. Main
William S. Boyd School of Law

Jeffrey W. Stempel
William S. Boyd School of Law

David McClure
William S. Boyd School of Law

Recommended Citation

Thomas O. Main, Jeffrey W. Stempel, and David McClure, *The Elastics of Snap Removal: An Empirical Case Study of Textualism*, 69 Clev. St. L. Rev. 289 (2021)
available at https://engagedscholarship.csuohio.edu/clevstlrev/vol69/iss2/6

THE ELASTICS OF SNAP REMOVAL: AN EMPIRICAL CASE STUDY OF TEXTUALISM

THOMAS O. MAIN, JEFFREY W. STEMPEL & DAVID MCCLURE*©

ABSTRACT

This Article reports the findings of an empirical study of textualism as applied by federal judges interpreting the statute that permits removal of diversity cases from state to federal court. The "snap removal" provision in the statute is particularly interesting because its application forces judges into one of two interpretive camps—which are fairly extreme versions of textualism and purposivism, respectively. We studied characteristics of cases and judges to find predictors of textualist outcomes. In this Article, we offer a narrative discussion of key variables, and we detail the results of our logistic regression analysis. The most salient predictive variable was the party of the president who appointed the judge. Female judges and young judges were also more likely to reach textualist outcomes. Cases involving torts were substantially more likely to be removed even though the statute raises a pure legal question upon which the subject matter of the case should have no bearing. Our most surprising finding was the impact of a judges' undergraduate and legal education: the eliteness of the educational institution was positively correlated with removal for judges appointed by Republicans, but negatively correlated for judges appointed by Democrats. This disordinal interaction was especially striking since there was no party effect among judges who attended non-elite institutions. In addition to the aforementioned variables which were significant, several variables that were not predictive are also discussed; these include race, seniority, state court experience, and the prospect of multi-district case consolidations.

* Respectively, William S. Boyd Professor of Law, Doris S. & Theodore B. Lee Professor of Law, and Associate Professor & Associate Director of the Wiener-Rogers Law Library, William S. Boyd School of Law, University of Nevada Las Vegas,

Thanks to organizers and participants of the 2019 Civil Procedure Workshop, especially Brooke Coleman, Bob Bone, Liz Page and David Marcus, as well as to Dean Dan Hamilton, Associate Dean Ian Bartrum, Associate Dean Sara Gordon, Randy Maniloff, Ann McGinley, Jean Sternlight, Sean Farhang, Howard Wasserman, Rocky Rhodes, Jamie Leavitt, Adam Sopko, Laura Rios, Kyle Hill, Aden Kebede, Petya Pucci, Donghyun (Don) Lee, Krysa Cieszko, Zahava Liberman, McKay Holley, and others.

CONTENTS

I. INTRODUCTION

Scholars and judges are growing weary of the debate between textualism and purposivism.[1] Numerous studies have disputed the narrative that judges belong

---

[1] Abbe R. Gluck & Richard A. Posner, *Statutory Interpretation on the Bench: A Survey of Forty-Two Judges on the Federal Courts of Appeals*, 131 HARV. L. REV. 1298, 1300–01 (2018) (referring to the debate as "boring").

exclusively to one tribe or the other.[2] And when reporting the results of their recent survey of appellate judges, Professor Abbe Gluck and Judge Richard Posner observed that their subjects assiduously avoided referring to themselves as either textualists or purposivists.[3] There may be an emerging consensus that most judges fulfill their judicial role pragmatically, drawing sensibly from legislative history and other indicia of intent—with some judges invoking interpretive tools more frequently and liberally than others.[4] Indeed, the bench, bar, and academy all have suggested that pragmatism, contextualism, soft brands of intentionalism, and other hybrids have replaced the extreme interpretive methodologies of textualism and purposivism.[5]

But some of that literature does more than merely eulogize what they label a false dichotomy. It suggests that "statutory interpretation doctrine [i]s a way to express

---

2 *See, e.g.*, Anita S. Krishnakumar, *Backdoor Purposivism*, 69 DUKE L.J. 1275 (2020); Michael C. Mikulic, *The Emergence of Contextually Constrained Purposivism*, 91 NOTRE DAME L. REV. 128 (2016); Richard M. Re, *The New Holy Trinity*, 18 GREEN BAG 2D 407, 417 (2015).

3 Gluck & Posner, *supra* note 1, at 1302 ("None of the 42 judges [in their survey] . . . was willing to associate himself or herself with textualism [or purposivism] without qualification.").

4 *See* Jonathan R. Siegel, *Textualism and Contextualism in Administrative Law*, 78 B.U. L. REV. 1023, 1057 (1998) ("In a significant sense, we are all textualists now."); Glen Staszewski, *Introduction to Symposium on Administrative Statutory Interpretation*, 2009 MICH. ST. L. REV. 1, 4 (2009) (observing that symposium articles "suggest that 'we are all purposivists again'"); James J. Brudney & Lawrence Baum, *Protean Statutory Interpretation in the Courts of Appeals*, 58 WM. & MARY L. REV. 81, 722 (2017) (measuring liberal invocation of legislative history by textualists). This phenomenon is not new:

> Do not expect anybody's theory of statutory interpretation, whether it is your own or somebody else's, to be an accurate statement of what courts actually do with statutes. The hard truth of the matter is that American courts have no intelligible, generally accepted, and consistently applied theory of statutory interpretation.

HENRY M. HART JR. & ALBERT M. SACKS, THE LEGAL PROCESS: BASIC PROBLEMS IN THE MAKING AND APPLICATION OF LAW 1169 (Eskridge & Frickey eds., Found. Press 1994) (1958).

5 *See* Gluck & Posner, *supra* note 1, at 1302–03; Brudney & Baum, *supra* note 4, at 751–52 (describing the interpretive approach of most appellate judges as "versatile," "functional," "practical," "eclectic," and "protean"); Victoria F. Nourse, *A Decision Theory of Statutory Interpretation: Legislative History by the Rules*, 122 YALE L.J. 70, 90 (2012) ("Today, we are all textualists and we are all purposivists."); Jonathan T. Molot, *The Rise and Fall of Textualism,* 106 COLUM. L. REV. 1, 2 (2006) ("Textualism has outlived its utility as an intellectual movement. . . . Textualists have been so successful discrediting strong purposivism, and distinguishing their new brand of 'modern textualism' from the older, more extreme 'plain meaning' school, that they no longer can identify, let alone conquer, any remaining territory between textualism's adherents and nonadherents."); Henry Paul Monaghan, *Supremacy Clause Textualism*, 110 COLUM. L. REV. 731, 733 (2010) ("Textualists have made an important contribution by forcing statutory interpreters to take the enacted text seriously; but they have persuaded few, if any, trained in the old school that, as the directive force of the statutory text attenuates, one can dispense with a comprehensive consideration of legislative purpose in determining statutory meaning.").

results in opinions, rather than as a tool that actually decides cases."[6] The real question, then, is what makes judges reach the outcome-decision that, in turn, leads them to use the tools of textualism and/or purposivism to justify that decision?[7] We pick up the baton and answer that question. We examined hundreds of published and unpublished opinions authored by federal judges about the availability of so-called "snap removal." We personally hand-coded dozens of characteristics about each of the cases and the judges who decided them, and then incorporated certain variables into a logit model that predicts case outcomes no matter the court's justification.[8]

The empirical analysis is not our only contribution. Our study also offers a counterexample to those who find the textualism/purposivism debate passé. This section of the removal statute does not lend itself to the hybrid interpretive methodologies that allow a judge to be (or to purport to be) in some middle ground between the extremes of textualism and purposivism. Rather, application of the snap removal statute necessarily requires a strong commitment to one of those extremes.[9]

---

6 Gluck & Posner, *supra* note 1, at 1314; *see also id.* at 1330, 1333; Michael Abramowicz et al., *Citation to Legislative History: Empirical Evidence on Positive Political and Contextual Theories of Judicial Decision Making*, 38 J. LEGAL STUD. 419, 423 (2009); Morell E. Mullins, Sr., *Tools, Not Rules: The Heuristic Nature of Statutory Interpretation*, 30 J. LEGIS. 1, 21 (2003); Philip P. Frickey, *From the Big Sleep to the Big Heat: The Revival of Theory in Statutory* Interpretation, 77 MINN. L. REV. 241, 251 (1992).

7 In this Article we take a wide arc around the fine distinctions between and among textualism, originalism, original intent, plain meaning, original meaning, and the like; by "textualism" we simply mean the various schools of interpretation that emphasize text over purpose. Likewise, when referring to purposivism we mean any and all of the various schools of interpretation that invoke extrinsic interpretive aids in the prioritization of legislative purpose over legislative text when the two arguably conflict.

8 Regression analysis measures the predictive value of one or more independent variables on a dependent or outcome variable. A logistic regression model differs from the more-familiar linear regression model in that the outcome variable (here: removal) in logistic regression is binary or dichotomous. This difference is reflected both in the form of the model and its assumptions. With these adjustments, the same general principles apply.

9 As with any legal issue, there are nuances in even the most seemingly clear dichotomy. As noted above, the interpretative debate between permitting snap removal and rejecting it divides neatly into textualist and purposivist camps. But in a prior article, we criticized the practice of snap removal not only because we reject textual literalism (arguably hyper-literalism in this context) and support purposivism as an interpretative tool, but also because various historical, structural, practical, and functional factors auger against snap removal, including the asymmetry of allowing an entity that is not subject to the control of the court (i.e., the unserved defendant) to utilize federal judicial power to divest the plaintiff of a chosen state court, which also has implications for federalism. *See* Jeffrey W. Stempel et al., *Snap Removal: Concept; Cause; Cacophony; and Cure*, 72 BAYLOR L. REV. 423 (2020).

One might in fact describe opposition to snap removal as based on a functionalist jurisprudence (rather than a purely purposive approach) that rejects formalism and textualism. *See* JEFFREY W. STEMPEL ET AL., PRINCIPLES OF INSURANCE LAW 104 (5th ed. 2020) (distinguishing legal formalism and legal functionalism).

For purposes of this empirical examination of judicial traits and decision making, snap removal nonetheless presents a stark illustration of the choice between embracing a literal

Accordingly, our study provides something of a census for the status of each of those methodologies in the courts—even if only for this narrow legal issue.

A very short primer on removal, diversity jurisdiction, and the forum defendant rule will suffice for an understanding of snap removal. Fights about snap removal are about whether certain cases will proceed in state court or in federal court. The distinction matters because, *inter alia*, win rates can vary;[10] in the modern era, defendants—especially corporate defendants—tend to prefer federal court.[11] Defendants in civil matters involving state law claims may remove cases from state to federal court when there is complete diversity between the parties.[12] Federal diversity jurisdiction is premised on the notion that state courts may be biased (or perceived as biased) against out-of-state parties.[13] That premise, in turn, cues an exception to the

---

reading of clear text and consideration of non-textual indicia of meaning, which is commonly described as a purposive approach even though it might better be deemed a functional or instrumental approach.

10 *See generally* Kevin M. Clermont & Theodore Eisenberg, *Do Case Outcomes Really Reveal Anything About the Legal System? Win Rates and Removal Jurisdiction*, 83 CORNELL L. REV. 581, 593 (1998) (noting overall plaintiff win rate of 58 percent for cases initially filed in federal court compared to win rate of 37 percent for removed cases); Marc Galanter, *Planet of the APs: Reflections on the Scale of Law and Its Users*, 53 BUFF. L. REV. 1369, 1389–91 (2006) (summarizing studies).

11 Alexandra D. Lahav & Peter Siegelman, *The Curious Incident of the Falling Win Rate: Individual vs. System-Level Justification and the Rule of Law,* 52 U.C. DAVIS L. REV. 1371 (2019); Brooke D. Coleman, *One Percent Procedure*, 91 WASH. L. REV. 1005, 1017–18, 1023 & 1027–28 (2016); Gillian K. Hadfield, *Exploring Economic and Democratic Theories of Civil Litigation: Differences Between Individual and Organizational Litigations in the Disposition of Federal Civil Cases*, 57 STAN. L. REV. 1275, 1322 (2005); *see also* Erwin Chemerinsky, *Parity Reconsidered: Defining a Role for the Federal Judiciary*, 36 UCLA L. REV. 233, 233–34 (1988); *see also* EDWARD A. PURCELL, JR., BRANDEIS AND THE PROGRESSIVE CONSTITUTION (2000); EDWARD A. PURCELL, JR., LITIGATION AND INEQUALITY (1992) [hereinafter PURCELL, LITIGATION AND INEQUALITY].

12 Complete diversity exists when the states of citizenship of all plaintiffs do not overlap with any of the states of citizenship of all defendants. *See* 28 U.S.C. § 1332 (providing original jurisdiction over actions between "citizens of different States."); Strawbridge v. Curtiss*,* 7 U.S. (3 Cranch) 267 (1806) (requiring complete diversity).

13 *See* THE FEDERALIST NO. 80, at 537–38 (Alexander Hamilton) (Jacob E. Cooke ed., 1961) (articulating the bias rationale); Bank of the U.S. v. Deveaux*,* 9 U.S. (5 Cranch) 61, 87 (1809) (observing national courts as forum to address "possible fears and apprehensions"). *See generally* Graham C. Lilly, *Making Sense of Nonsense: Reforming Supplemental Jurisdiction,* 74 IND. L.J. 181, 190 (1998); Scott Dodson, *Beyond Bias in Diversity Jurisdiction*, 69 DUKE L.J. 267 (2019) (reviewing traditional basis and collecting sources); John P. Frank, *The Case for Diversity Jurisdiction,* 16 HARV. J. LEGIS. 403, 407 (1979) (recognizing bias as a danger). The empirics and the wisdom of the underlying premise have long been questioned. *See, e.g.*, Henry Friendly, *The Historic Basis of Diversity Jurisdiction*, 41 HARV. L. REV. 483, 493–95 (1928) (interrogating the existence of bias); RICHARD A. POSNER, THE FEDERAL COURTS: CRISIS AND REFORM 142–43 (1985) (disputing the prevalence of bias); (suggesting that, instead, maintaining "federal jurisdiction over issues and interests . . . [of] national importance" has been the Court's touchstone); PURCELL, LITIGATION AND INEQUALITY, *supra* note 11, at 254–

right of removal: preventing removal of a diversity case by a citizen of the forum state;[14] after all, a local defendant needs no escape from their home court.

This exception—often called the forum defendant rule—has been a part of the jurisdictional schema of removal since the federal courts were established.[15] The statute has been amended on several occasions and currently provides that a case "may not be removed if any of the parties in interest *properly joined and served* as defendants is a citizen of the State in which such action is brought."[16] The italicized language was added in 1948. Although there is no legislative history directly on point, the obvious purpose of the "joined and served" language was to prevent gamesmanship by plaintiffs. Such gamesmanship was made possible by the Supreme Court's *Pullman Co. v. Jenkins* decision in 1939.[17] The reasoning in that decision created the opportunity for plaintiffs to exploit (a blander version of) the forum defendant rule by naming but not serving a forum defendant.[18] The 1948 amendment thus added language to prevent plaintiffs from nominally joining a forum defendant (as a sham) and then never serving them, thereby thwarting removal by the out-of-state defendant even though the plaintiff was never seriously pursuing the forum defendant.[19]

---

56 (suggesting that, instead, maintaining "federal jurisdiction over issues and interests . . . [of] national importance" has been the Court's touchstone).

14 28 U.S.C. § 1441(b)(2).

15 For a meticulous account of the history, see Hawkins v. Cottrell, Inc., 785 F. Supp. 2d 1361, 1373–78 (N.D. Ga. 2011); Bowman v. PHH Mortg. Corp., 423 F. Supp. 3d 1286, 1291 & n.5 (N.D. Ala. 2019).

16 28 U.S.C. § 1441(b)(2) (emphasis added).

17 Pullman Co. v. Jenkins, 305 U.S. 534 (1939).

18 *Id.* at 540.

19 The basic purpose of the language is agreed upon by commentators, although they may differ in their views of the propriety of snap removal. *See, e.g.*, Saurabh Vishnubhakat, *Pre-Service Removal in the Federal Defendant's Arsenal*, 47 GONZ. L. REV. 147 (2011-2012) (finding snap removal appropriate notwithstanding the purpose of the language); Matthew Curry, Note, *Plaintiff's Motion to Remand Denied: Arguing for Pre-Service Removal Under the Plain Language of the Forum-Defendant Rule*, 58 CLEV. ST. L. REV. 907 (2010) (same); Jordan Bailey, Comment, *Giving State Courts the Ol' Slip: Should a Defendant be Allowed to Remove an Otherwise Irremovable Case to Federal Court Solely Because Removal Was Made Before Any Defendant is Served?*, 42 TEX. TECH. L. REV. 181 (2009) (criticizing snap removal as inconsistent with overall structure of removal and diversity jurisdiction).

Judges likewise acknowledge the obvious purpose of the language.

> Despite the lack of [legislative history], the purpose of the "properly joined and served" requirement of section 1441(b) is abundantly clear in light of the historical development of the policy of the remand provisions, the practical application of the "joined and served" provision by district courts in recent decades, and common sense.

Sullivan v. Novartis Pharms. Corp., 575 F. Supp. 2d 640, 644 (D.N.J. 2008); *see also* Brown v. Organon Int'l Inc., No. 07-3092, 2008 WL 2833294, at *5 (D.N.J. July 21, 2008) (remanding

Ironically, paradoxically even, the forum defendant rule now facilitates gamesmanship by defendants. Because of the quoted, italicized language added by the 1948 amendment, forum defendants claim that the forum defendant rule does not apply, provided defendants file their notice of removal prior to service; after all, the forum defendant rule is an exception to the right of removal, and that exception applies only to forum defendants who have been "properly . . . served."[20] Because this tactic works only if the notice of removal is filed in the brief time period between filing and service, it is often referred to as "snap removal."[21] And, as we document below, more than half of judges allow it. To be clear, with this tactic, forum defendants are removing diversity cases to federal court for the first time since 1789.

In this Article, we are agnostic on questions about how the statute should be applied. We have argued elsewhere that allowing snap removal requires an excessively textualist reading of the statute that derogates the history, purpose, and logic of both diversity jurisdiction and removal.[22] We also proposed a fix—one that textualists might even welcome.[23] But our focus in this Article is not how it should be applied or fixed, but rather how in fact it is applied. Simply stated, we interrogate the questions: what are the traits of a judge who allows snap removal under the current statute? And what are the traits of a judge that rejects snap removal?

---

rather than "violate the clear purpose of the legislative provision"); Marsh v. Monster Beverage Corp., No. ED CV 15-2205, 2016 WL 11508266, at *2 (C.D. Cal. Jan. 4, 2016) ("District courts across the country have concluded that the rule serves to 'prevent plaintiffs from improperly joining forum defendants that they do not actually intend to pursue in order to prevent removal from state court,' which was a 'pervasive problem' at the time the 'properly joined and served' language was adopted by Congress in 1948." (quoting Standing v. Watson Pharma., Inc., No. CV09-0527, 2009 WL 842211, at *3 (C.D. Cal. Mar. 26, 2009) (collecting cases))); Saratoga Advantage Tr. Tech. & Comms. Portfolio v. Marvell Tech. Grp., No. 15-cv-04881, 2015 WL 9269166, at *2 (N.D. Cal. Dec. 21, 2015) (allowing snap removal and denying motion to remand, "[n]otwithstanding the rule's purpose").

20 By referring to the forum defendant rule as an "exception" we have accepted the characterization typically employed by defendants. Alternatively, the forum defendant rule is a limitation on the right to remove.

21 While the term "snap removal" appears to have taken hold among many courts, practitioners, and academics, others have referred to the practice by different monikers, such as "pre-service removal," Gibbons v. Bristol-Myers Squibb Co., 919 F.3d 699, 705 (2d Cir. 2019); "early removal," Vivas v. Boeing Co., 486 F. Supp. 2d 726, 733 (N.D. Ill. 2007); "jack rabbit removal," Beavers v. Medtronic, Inc., 41 F. Supp. 3d 633, 639 (W.D. Ky. 2014); "race to remove," Ayala-Castro v. GlaxoSmithKline (*In re* Avandia Mktg.), 624 F. Supp. 2d 396, 409–10 (E.D. Pa. 2009); "preemptive removal," Bailey, *supra* note 19, at 182; "snatch and remove," Paul D. Rheingold, *Removal, Remand, and Transfer, in* LITIGATING MASS TORT CASES § 5.24 (May 2020 update); and "wrinkle removal" (because the practice "capitalize[s] on a 'wrinkle' in the language of 28 U.S.C § 1441(b)(2)"), Rachel B. Weil, *Remand Denied in Two Shingles Vaccine Cases out of the District of New Jersey*, DRUG & DEVICE L. BLOG (Feb. 1, 2019), https://www.druganddevicelawblog.com/2019/02/nd-denied-in-two-shingles-vaccine-cases-out-of-the-district-of-new-jersey.html.

22 Stempel et al., *supra* note 9, at 468.

23 *Id.* at 468–69.

Our questions lend insight into the evergreen debate of textualism versus purposivism, because this statute has an unusual quality that forces judges into one interpretive camp or the other. Indeed, applications of this statute require either a hyper-literal reading that flouts Congressional intent or a purposive reading that evades crystal-clear text; there is no middle ground.

To reject snap removal requires an especially strong brand of purposivism because the text unambiguously limits the scope of the forum defendant rule to those who have been served. This is not a scrivener's error. Nor is it patently absurd to allow forum defendants access to a federal court in diversity cases; forum plaintiffs, for example, have access to a federal court in diversity cases.[24] The statutory schema refers broadly to "defendants" in other parts, but qualifies the scope of the forum defendant rule to defendants who have been served.[25] The plain meaning of the statute suggests no need to consult legislative history. And worse, there is no legislative history to directly contradict the plain meaning. The historical context of the statute confirms a different purpose (to prevent gamesmanship by plaintiffs), but that purpose is irrelevant or at most orthogonal to the plain meaning of the text; it is not directly contradictory.[26] All this is to say that there is not even a strained reading of the text for the purposivist to use as a fig leaf to hide behind.[27]

---

24 There is no forum plaintiff rule that prevents in-state plaintiffs from filing a diversity action in a federal district court in their home state.

25 *Cf.* 28 U.S.C. § 1441(a).

26 *See, e.g.*, *Gibbons*, 919 F.3d at 699.

27 Searches for the middle ground may be noble and well-intentioned, but are unavailing. In an effort to bolster our purposivist understanding of the forum defendant rule with text, we have suggested that a named-but-unserved party may not be a "defendant". Because only defendants can remove, this reading would prevent some of the most egregious instances of snap removal. We also observe that this analysis is problematic in that it "runs counter to a significant body of case law and commentary." Stempel et al., *supra* note 9, at 489.

Another "formalist" objection to snap removal can be raised in certain unique circumstances. For example, when a forum defendant is the only defendant, snap removal could be denied on the basis that the absence of any joinder of parties means that the "joined and served" language of section 1441(b) cannot be satisfied. But without resorting to the *purpose* of the forum defendant rule, failure to trigger the "joined and served" criterion would seem to authorize removal (because the limitation upon or exception to removal would not apply). Ditto efforts that treat "mischief" as a tool for textualists. *See, e.g.*, Howard M. Wasserman, *Mischief and Snap Removal*, JOTWELL (June 3, 2020), https://courtslaw.jotwell.com/mischief-and-snap-removal/ (reviewing Samuel Bray, *The Mischief Rule*, 109 GEO. L.J. (forthcoming 2021) (available at SSRN)).

Another "formalist" objection to snap removal could leverage the federalism canon. *See generally* Gregory v. Ashcroft, 501 U.S. 452, 460–61 (1991). Pursuant to this canon, the snap removal statute should be interpreted to avoid interfering with state court jurisdiction over matters involving forum defendants. Many opinions in our dataset referenced the notion that removal statutes should be construed narrowly–"in order to promote the goals of federalism, restrict federal court jurisdiction, and support the plaintiff's right to choose the forum." Torchlight Loan Servs., LLC v. Column Fin., Inc., No. 12 Civ. 8579, 2013 WL 3863887, at *2 (S.D.N.Y. July 24, 2013). Our point here is that none of these opinions refers to this as a

Yet to allow snap removal requires an especially strong brand of textualism because the text is inconsistent with tradition, principle, and common sense. Congress did not intend the availability of a federal forum to turn on the timing of service. Congress did not intend to elevate the idiosyncrasies of state service practice and procedure into something that is outcome-determinative. Congress could not have foreseen that advances in technology would allow defendants to know that they had sued before they had been served. And worse, snap removal gives comparatively wealthy defendants another procedural weapon to disarm comparatively impoverished plaintiffs; even someone who is not bothered by favoring this lot could be troubled by the optics. All this is to say that there is not a shred of congressional intent for the textualist to use as a fig leaf to hide behind.[28]

We emphasize this binary choice because its starkness served to clarify the data that we classified and analyzed. We could safely assume that: (1) judges who allowed snap removal are textualists (at least for this particular issue); and (2) judges who disallowed snap removal are purposivists (at least for this particular issue). We had to apply the labels because, as Gluck and Posner observed through their surveys, judges do not label themselves.[29] Although these opinions exhibit much reverence for text, the word "textualist" appears in only one opinion in our dataset,[30] and "textualism" not even once. "Purposivism" and "purposivist" do not appear at all; though, naturally, "purpose" is a popular referent.

"canon," which is the currency of textualism. *See generally* William N. Eskridge, Jr., *The New Textualism and Normative Canons*, 113 COLUM. L. REV. 531 (2013) (reviewing ANTONIN SCALIA & BRYAN GARNER, READING LAW (2012)); Anita S. Krishnakumar, *Dueling Canons*, 65 DUKE L.J. 909 (2016).

Another purportedly formalist-like objection to snap removals distinguishes snap removals initiated by forum defendants from snap removals initiated by out-of-state defendants. *See, e.g.*, Carpenter v. Apotex Corp., No. 08-60526-CIV, 2008 WL 11332029, at *3 (S.D. Fla. Aug. 8, 2008). But this is not a textual argument. Rather it draws on the purpose of removal, to-wit: "to protect out-of-state defendants from possible prejudices in state court." *Id.* at *2.

28 Searches for the middle ground may be noble and well-intentioned, but are unavailing. In an effort to bolster its textualist result by genuflecting to a hypothetical purpose, the U.S. Court of Appeals for the Second Circuit suggested that "Congress may [have intended to] provide a bright-line rule keyed on service, which is more clearly more easily administered than a fact-specific inquiry into a plaintiff's intent or opportunity to actually serve a home-state defendant." *Gibbons*, 919 F.3d at 706 (citing no relevant authority); *see also* Tex. Brine Co. v. Am. Arb. Ass'n, 955 F.3d 482, 486 (5th Cir. 2020) (same). That "a reasonable person could intend the results of [this] plain language," *id.* at 486, is surely more damning than persuasive. For the origins of this reasonable person standard, see SCALIA & GARNER, *supra* note 27, at 574.

29 Gluck & Posner, *supra* note 1, at 1302–03.

30 *See* Breitweiser v. Chesapeake Energy Corp., No. 15–CV–2043, 2015 WL 6322625, at *5 (N.D. Tex. Oct. 20, 2015) (observing that "[e]ven the most ardent textualist" will consider legislative intent and history to avoid an absurd result). The word "text" appears in only 54 of the opinions; references to the language of the statute are much more popular.

We coded the courts' reasoning in every removal case.[31] In cases allowing snap removal, judges invoked the refrains of textualism to justify their results even if not also naming the hymnal from which they were drawn. Typical examples from judicial opinions include the following:

- "The plain meaning of Section 1441(b)(2) is clear and unambiguous."[32]
- "Read literally, the forum defendant rule only precludes removal when a forum defendant has been 'properly joined and served.'"[33]
- "[T]he language of the statute is plain, and, thus, adherence to the plain language is required."[34]
- "This court must apply the statute as it is written, and not as plaintiffs maintain it is intended."[35]

And in every case that rejected snap removal, judges sampled the familiar rhymes of purposivism to justify their results even if not also crediting their source material. Typical examples from judicial opinions include the following:

- "[Applying t]he plain meaning of the statute . . . would eviscerate the purpose of the forum defendant rule. The court may not adopt a plain language interpretation of a statutory provision that directly undercuts the clear purpose of the statute."[36]
- "[G]iving effect to the plain language of the forum defendant rule . . . would result in an outcome at odds with the intentions of its drafters. . . . There [is] no evidence that 'Congress . . . intended to create an arbitrary

31 References such as "every removal case" include cases that were ultimately remanded after snap removal. To be clear, removal occurs upon filing by the defendant; it is not a motion and it does not require permission of the court. The propriety of a removal is tested by the filing, by the plaintiff, of a motion to remand.

32 Dechow v. Gilead Scis., Inc., 358 F. Supp. 3d 1051, 1054 (C.D. Cal. 2019).

33 Selective Ins. Co. of S.C. v. Target Corp., No. 13 C 5910, 2013 WL 12205696, at *1 (N.D. Ill. Dec. 13, 2013) (citing 28 U.S.C. § 1441(b)).

34 Poznanovich v. AstraZeneca Pharms. LP, No. 11–4001, 2011 WL 6180026, at *4 (D.N.J. Dec. 12, 2011).

35 Johnson v. Precision Airmotive, LLC, No. 07CV1695, 2007 WL 4289656, at *6 (E.D. Mo. Dec. 4, 2007).

36 Mass. Mut. Life Ins. Co. v. Mozilo, No. 12–CV–03613, 2012 WL 11047336, at*2 (C.D. Cal. June 28, 2012).

means for a forum defendant to avoid the forum defendant rule simply by filing a notice of removal before the plaintiff is able to effect process.'"[37]

- "A literal reading of Section 1441(b)(2) . . . is clearly at odds with Congress's intent to eliminate gamesmanship [by plaintiffs]."[38]

- "This is surely not what Congress intended. . . . [T]he courts 'must look beyond the plain meaning of the statutory language.'"[39]

Our aim with these quotes is merely to convey that removing and remanding judges make the noises of textualism and purposivism, respectively. There is no opinion in our dataset where the judge relies principally on the statutory text to remand or on the statute's purpose to remove.[40]

The statute is also somewhat unique in that its applications are almost unreviewable as a matter of appellate practice and procedure.[41] Accordingly, more than 95% of the opinions in our dataset were authored by district court judges. This has numerous benefits. First, it helps answer the call of those who have hailed the

37 Williams v. Daiichi Sankyo, Inc., 13 F. Supp. 3d 426, 430–31 (D.N.J. 2014) (quoting Sullivan v. Novartis Pharms. Corp., 575 F. Supp. 2d 640, 645 (D.N.J. 2008)).

38 Penn v. Amazon.com, Inc., No. CV 17-02682, 2017 WL 8229625, at *2 (C.D. Cal. May 24, 2017).

39 Prather v. Kindred Hosp., No. 14–0828–CV, 2014 WL 7238089, at *2–3 (W.D. Mo. Dec. 17, 2014) (citing Perez v. Forest Lab'ys, Inc., 902 F. Supp. 2d 1238, 1245 (E.D. Mo. 2012)).

40 *See supra* notes 27–28 and accompanying text.

41 An order of remand is unreviewable. 28 U.S.C. § 1447(d) (prohibiting review of a district court's order remanding a case to state court subject to exceptions not relevant here); Thermtron Prods., Inc. v. Hermansorfer, 423 U.S. 336, 355–56 (1976) (limiting prohibition of appellate review to grounds specified in 28 U.S.C. § 1447(c), namely (i) lack of subject matter jurisdiction, and (ii) any defect other than lack of subject matter jurisdiction); Lively v. Wild Oats Mkts., Inc., 456 F.3d 933 (9th Cir. 2006) (collecting cases); Holmstrom v. Peterson, 492 F.3d 833, 838–39 (7th Cir. 2007) (same).

The denial of a motion to remand is reviewable, but must await a final judgment. So the issue seldom reaches the Courts of Appeals. *But see* Tex. Brine Co. v. Am. Arb. Ass'n, 955 F.3d 482, 482 (5th Cir. 2020) (plaintiff appealed after district court granted defendant's motion for judgment on the pleadings); Gibbons v. Bristol-Myers Squibb Co., 919 F.3d 699, 699 (2d Cir. 2019) (plaintiff appealed after district court granted defendant's motion to dismiss for failure to state a claim); Encompass Ins. Co. v. Stone Mansion Rest. Inc., 902 F.3d 147 (3d Cir. 2018) (plaintiff appealed after district court granted defendant's motion to dismiss for failure to state a claim).

See Adam Sopko, Swift *Removal*, 12 FED. CTS. L. REV. (forthcoming 2020) (manuscript at 16 n.105) (on file with authors) (citing Curry, *supra* note 19, at 931–32); Arthur Hellman et al., *Neutralizing the Strategem of "Snap Removal": A Proposed Amendment to the Judicial Code*, 9 FED. CTS. L. REV. 103, 106–07 (2016); Valerie M. Nannery, *Closing the Snap Removal Loophole*, 86 U. CIN. L. REV. 541, 557–58 (2018); Tempe D. Smith, *Snap Removal: What Is It and What Do You Do If It Happens to You*, 39 ALA. ASS'N JUST. J. 55, 57 (2019).

"merits of turning more scholarly attention away from the Supreme Court and instead to the everyday decisionmakers in the system."[42]

More substantively, district court opinions do not have the obfuscations that plague studies of appellate opinions. For example, we can avoid the uncertainty regarding whether the content of an opinion is the true expression of the judge or is instead the product of a compromise to obtain the support of judicial colleagues.[43] Further, the (published and unpublished) opinions of district court judges are usually written with a degree of candor and directness because they are not written for an extrajudicial audience.[44]

To be sure, district court opinions are frequently short—and also lean on their justification and reasoning. But that shortcoming is no handicap in a study that is focused principally on matters that are upstream from the decision, such as the demographics of the judge, rather than downstream, such as the justification for that decision.

The availability of snap removal also presents a pure question of law with crisp resolutions that simplified coding and analysis. Further, this is not the sort of legal question that judges might strategically recast as a fact question or a mixed law-and-fact question in order to shield a judge's policy preferences from appellate review.[45] The legal question requires no fact finding and is, thus, usefully naked. Shielding the decision from appellate review is also not a concern when appellate review is unlikely in any event.

We anticipated that our study's focus primarily on district court opinions would also present challenges. For example, compared with their appellate counterparts, district judges have a more demanding caseload, less law clerk and library support, burdensome efficiency mandates, and inferior briefs from parties.[46] One might predict that this would bias a district judge toward textualism since text requires comparatively less of an investment than what is required to unearth legislative history. But with regard to snap removal there is no buried treasure; there was little to

42 Gluck & Posner, *supra* note, 1 at 1306 (referring to Appeals Court judges as everyday decisionmakers) (citing RICHARD POSNER, HOW JUDGES THINK (2008)).

43 *See, e.g.*, Abramowicz et al., *supra* note 6, at 420 (explaining how the choice of what to put in an opinion "may be driven more by collegiality or strategic calculation in an effort to persuade or assuage the other judges and courts than by an authoring judge's own preferences"); Brudney & Baum, *supra* note 4, at 744 (explaining possible influence of collegiality and strategy in judicial behavior).

44 *See generally* Brudney & Baum, *supra* note 4, at 759.

45 *Cf.* Abramowicz et al., *supra*, note 6 at 423–24 (citing Emerson H. Tiller & Pablo T. Spiller, *Strategic Instruments: Legal Structure and Political Games in Administrative Law*, 15 J.L. ECON. & ORG. 349 (1999); Joseph L. Smith & Emerson H. Tiller, *The Strategy of Judging: Evidence from Administrative Law*, 31 J. LEGAL STUD. 61 (2002); Max Schanzenbach & Emerson H. Tiller, *Strategic Judging Under the United States Sentencing Guidelines: Positive Political Theory and Evidence*, 23 J.L. ECON. & ORG. 24 (2007).

46 *See, e.g.*, '*Staggering Workloads' for district judges*, *in* THE MONTANA LAWYER 5, 5 (June–July 2007).

learn from an interpretive standpoint, and the legislative context has long been clear, widely available, and persuasive to the persuadable.[47]

Another challenge when studying district court opinions is the hierarchy (or principal-agent) effect. A trial judge's decision (and/or the justification of that decision) may be a prediction of the appellate court's preference, rather than a sincere expression of the trial judge's own belief.[48] But yet again, because a judge's decision on a motion to remand is practically unreviewable, there is less reason for a trial judge to defer-and-predict rather than simply to reason-and-rule.

The near-unreviewability of applications of this statute also means that we have three decades of essentially unchecked variation among district courts. There was no binding precedent on snap removal anywhere prior to the decision of the U.S. Court of Appeals for the Third Circuit in the fall of 2018.[49] In the state court where that suit began, the Pennsylvania defendant's lawyer had agreed orally to accept service on his client's behalf, but then removed to federal court rather than return the form that would have constituted service.[50] The dubious nature of these circumstances deterred neither the federal district court nor the appeals court from endorsing snap removal.[51] Since the Third Circuit's decision in *Encompass Ins. Co.*, two more circuit courts have

---

47 *See, e.g.*, Hawkins v. Cottrell, Inc., 785 F. Supp. 2d 1361, 1373–78 (N.D. Ga. 2011); Sullivan v. Novartis Pharms. Corp., 575 F. Supp. 2d 640, 644 (D.N.J. 2008).

48 *See generally* Tracey E. George & Albert H. Yoon, *The Federal Court System: A Principal-Agent Perspective*, 47 ST. LOUIS U. L.J. 819, 822 (2003); Lee Epstein, *Some Thoughts on the Study of Judicial Behavior*, 57 WM. & MARY L. REV. 2017, 2071 (2016); Pauline T. Kim, *Beyond Principal-Agent Theories: Law and the Judicial Hierarchy*, 105 NW. U. L. REV. 535, 538–39 (2011); Pauline T. Kim, *Lower Court Discretion*, 82 N.Y.U. L. REV. 383, 403–04 (2007).

49 Encompass Ins. Co. v. Stone Mansion Rest. Inc., 902 F.3d 147, 147 (3d Cir. 2018).

There are earlier Appeals Court decisions that discuss snap removal more or less directly, but none is resolved exclusively on the snap removal issue, and none is treated as binding on the issue by lower courts in those circuits. *See* McCall v. Scott, 239 F.3d 808 (6th Cir. 2001); Clarence E. Morris, Inc. v. Vitek, 412 F.2d 1174 (9th Cir. 1969); Goodwin v. Reynolds, 757 F.3d 1216, 1220–21 (11th Cir. 2014); *see also* Little v. Wyndham Worldwide Ops., Inc., 251 F. Supp. 3d 1215, 1219 (M.D. Tenn. 2017) ("*McCall* is not controlling"); Loewen v. McDonnell, No. 19-cv-00467, 2019 WL 2364413, at *7 n.8 (N.D. Cal. June 5, 2019) (distinguishing *Vitek* as addressing the question whether the citizenship of unserved defendants should be considered in evaluating whether diversity jurisdiction exists); Timbercreek Asset Mgmt., Inc. v. De Guardiola, No. 19-CV-80062, 2019 WL 947279, at *2 (S.D. Fla. Feb. 27, 2019) (referring to the *Goodwin* court's discussion of snap removal as "dicta").

50 *Encompass Ins. Co.*, 902 F.3d at 149.

51 *Id.* at 153–54 ("[W]e conclude that the language of the forum defendant rule in section 1441(b)(2) is unambiguous. . . . Permitting removal on the facts of this case does not contravene the apparent purpose [of the statute, to-wit:] fraudulent joinder by a plaintiff. . . . [T]his result may be peculiar in that it allows Stone Mansion to use pre-service machinations to remove a case that it otherwise could not . . . ,[but t]here are simply no grounds upon which we would substitute Encompass' interpretation for the literal interpretation. Reasonable minds might conclude that the procedural result demonstrates a need for a change in the law; however, if such change is required, it is Congress—not the Judiciary—that must act.").

likewise endorsed snap removal: in early 2019, the U.S. Court of Appeals for the Second Circuit decided *Gibbons v. Bristol-Myers Squibb Co.*,[52] and on April 7, 2020, the U.S. Court of Appeals for the Fifth Circuit decided *Texas Brine Co. v. American Arbitration Ass'n*.[53] Textualism is winning.

Now we turn to our key findings which we summarize here. Perhaps unsurprisingly, the best single predictor of case outcomes was the party of the president who appointed the judge. Following the familiar stereotype, judges appointed by Republicans were significantly more likely (than their colleagues appointed by Democrats) to reach a textualist outcome.

Other significant variables included the subject-matter of the action: torts cases were more likely to be snap-removed than contract cases. Because the statute itself is context-independent, the courts' appetite for tort cases hinted at some ideological effect. We found more evidence of such effects in the rates at which judges published their opinions.

Women allowed snap removal at higher rates than men. Younger judges, too, were substantially more likely than their senior colleagues to reach textualist outcomes.

We found that the eliteness of the institution where a judge received his or her undergraduate and legal education was not only a strong predictor of case outcomes, but this eliteness variable had disordinal effects on Republican and Democratic appointees.[54] Moreover, we observed that Democratic appointees who attended non-elite institutions had the same removal rate as Republican appointees who attended non-elite institutions. All party effects, then, were manifest among those Democratic and Republican judges who attended elite institutions. This was a striking finding about party polarization within the judiciary.

The data also rejected several theories that we tested: the prospect for multidistrict litigation in complex cases, for example, had no observable effect on snap removal rates. Similarly, the race or ethnicity of the judge was not predictive. Nor was the amount of prior judicial experience, including whether the judge previously served as a state court judge.

II. METHODOLOGY

Our data derive from repeated and redundant searches of Lexis and Westlaw databases of published and unpublished federal court opinions at the district and circuit court levels. We identified cases from 1960 through 2019, utilizing targeted keyword searches with Boolean connectors, key number and headnote searches, and case citators.[55] Although this approach may admit occasional errors, we have no reason to

---

52 Gibbons v. Bristol-Myers Squibb Co., 919 F.3d 699, 705 (2d Cir. 2019).

53 Tex. Brine Co. v. Am. Arb. Ass'n, 955 F.3d 482, 486 (5th Cir. 2020).

54 A disordinal interaction describes a circumstance where one variable has significant positive *and* negative influence on the outcome variable, depending on the presence of some other variable. In our study, eliteness of education was positively correlated with removal by Republican appointees, but negatively correlated with removal by Democratic appointees.

55 One of our first steps was to run keyword searches in the Lexis and Westlaw case databases for "snap removal" and its various nicknames. *See supra* note 21 and accompanying text. However, not all snap removal cases (particularly older ones) use a consistent term or phrase to describe the practice, so we utilized other methods to find true snap removal cases. Among the

believe that those errors would have any bias. One or more of the named authors then read each case in this pool of 337 opinions. Because our search terms erred on the side of being overly inclusive, many of the cases in that pool were principally about fraudulent joinder or lack of diversity jurisdiction and did not allow or require the judge to rule specifically on the issue of snap removal.

Nearly 300 cases involving snap removal were entered (by hand, by the authors) into a database. The data set included, for each opinion, data identifying the case name, citation, whether the opinion was published or unpublished, year, the state where the court was located, type(s) of plaintiff(s), type(s) of defendant(s), which defendant filed the notice of removal, subject matter of the case, disposition, stated rationale, length of discussion, key authorities relied upon, the identity of the judge, and miscellaneous notes. We then supplemented the dataset with additional research in order to code whether the action was a prospect for consolidation.[56]

We also researched and coded demographic characteristics about each judge: date of judicial appointment, appointing president, gender, race, year of birth, prior judicial experience, other work experience, and details of their undergraduate and legal education. This data about the judges was drawn from the Federal Judges Biographical Database, Wikipedia, and various websites including those maintained by district and circuit courts.[57]

We excluded cases that were decided by magistrate judges. This decision was informed by the lack of reliable demographic information about all magistrate judges; we did not want to include some of their opinions and not all. The authority of magistrate judges also varies considerably by district,[58] and it was often unclear

---

techniques we used was reviewing case citator reports from Shepard's and KeyCite for numerous opinions, particularly the published circuit court decisions available at the time (*Stone Mansion* and *Gibbons*) and those that were frequently cited in other cases and in the secondary literature. We also ran keyword searches and generated case citator reports of 28 U.S.C. § 1441(b)(2) limited by Boolean searches of frequently appearing terms, such as "PLAIN /S MEANING" OR ABSURD OR GAMESMANSHIP. We reviewed cases listed in the citator reports, identified cases cited in those cases, and continued repeating variations of the process until we consistently ran across citations to the same published and unpublished opinions.

56 In some instances consolidation was apparent from the opinion in our dataset. *See, e.g.*, *In re* Lipitor (Atorvastatin Calcium) Mktg., Sales Pracs. & Prods. Liab. Litig., No. 2:14-mn-02502, 2016 WL 7338594 (D.S.C. Nov. 21, 2016); Valido-Shade v. Wyeth, LLC, 875 F. Supp. 2d 474 (E.D. Pa. 2012). In other instances, we consulted the Judicial Panel on Multidistrict Litigation to determine whether the action was a candidate for consolidation. *See Judicial Panel on Multidistrict Litigation*, U.S. CTS., https://www.jpml.uscourts.gov/ (last visited Jan. 12, 2021).

57 *See Biographical Directory of Article III Federal Judges, 1789–Present*, FED. JUD. CTR., https://www.fjc.gov/history/judges; en.wikipedia.org (last visited June 1, 2020); *see, e.g.*, *Boston Judge Information*, U.S. DIST. CT. D. MASS., www.mad.uscourts.gov/boston/boston-judgeinfo.htm (last visited June 1, 2020).

58 *See generally* Philip M. Pro, *United States Magistrate Judges: Present but Unaccounted for*, 16 NEV. L.J. 783, 787 (2016); Douglas A. Lee & Thomas E. Davis, *"Nothing Less than Indispensable": The Expansion of Magistrate Judge Authority and Utilization in the Past Quarter Century*, 16 NEV. L.J. 845, 928 (2016). In addition, of course, Magistrate Judges are not selected by Republican or Democratic Presidents in the same sense as are Article III judges in that the Article I magistrate judges are as a practical matter selected by the Article III judges

whether the available opinion was only a report and recommendation or was, instead, a matter being resolved by a magistrate judge upon consent of the parties under 28 U.S.C. § 636. Note, however, that several opinions included in our dataset were decisions by a district judge adopting the report and recommendation of a magistrate judge.[59]

We also excluded cases decided by district judges after the issuance of binding precedent by a circuit court. Our empirical study assumed that judges had some freedom to choose between the textualist and purposivist approaches to the forum defendant rule. Once the Third Circuit issued its opinion in *Encompass Ins. Co.*,[60] district court judges in Delaware, New Jersey, and Pennsylvania lacked that freedom. And of course the same logic applies for federal district court judges in those states encompassing the Second and Fifth Circuits, after those circuit courts issued binding precedent in *Gibbons* and *Texas Brine*, respectively.[61] We included the decisions of the nine circuit judges who formed the panels on those appellate cases.

Our dataset thus differs in scope from other scholars who have examined snap removal from different perspectives and with different objectives. Readers who are especially interested in matters regarding the frequency of snap removals and the identity of the snap removers are strongly encouraged to read our colleagues' work.[62]

We uploaded our dataset into Stata, a command-driven software package for statistical analysis. Because the principal focus of our study was judicial decisions, we clustered the standard errors for judges who wrote more than one opinion in our dataset.[63] We did not want to misrepresent the magnitude of our dataset, and we did not want to distort any of our findings by counting echoes as distinct voices.

One important exception to this protocol involved judges who issued more than one decision when those decisions were not consistently for or against snap removal. Five judges drifted from one side of the snap-removal divide to the other—all five moved from allowing snap removals to denying them. After considering various options, we decided to include these five judges in both of their respective iterations, treating their changed heart as a new person. This had the virtue of inclusion. Unfortunately, it also sabotaged predictive models by putting the same five judges in both camps, ensuring five erroneous predictions no matter the model.

Net of these decisions, our core dataset included 193 unique cases of first impression. A dataset of 193 judges is a robust size for a study of federal judges. One

---

of the relevant district. *See* Pro, *supra*, at 848 (citing Wellness Int'l Network, Ltd. v. Sharif, 575 U.S. 665 (2015)).

59 *See, e.g.*, Higgins v. City of Savannah, No. CV417-257, 2018 WL 2424138, at *1 (S.D. Ga. Mar. 5, 2018) (Smith, M.J.), *adopted by* 2018 WL 2418550 (S.D. Ga. May 29, 2018) (Moore, J.).

60 Encompass Ins. Co. v. Stone Mansion Rest. Inc., 902 F.3d 147 (3d Cir. 2018).

61 Gibbons v. Bristol-Myers Squibb Co., 919 F.3d 699, 699 (2d Cir. 2019); Tex. Brine Co. v. Am. Arb. Ass'n, 955 F.3d 482 (5th Cir. 2020).

62 *See* Sopko, *supra* note 41; Nannery, *supra* note 41, at 541.

63 *See, e.g.*, Gentile v. Biogen Idec, Inc., 934 F. Supp. 2d 313 (D. Mass. 2013) (Woodlock, J.); Howard v. Genentech, Inc., No. 12–11153, 2013 WL 680200 (D. Mass. Feb. 21, 2013) (Woodlock, J.).

useful reference point is that there are currently 677 authorized district court judgeships.[64] But because our dataset spans three decades of caselaw, only ninety-four (or 49%) of the cases in our dataset were authored by judges who are *currently* active Article III judges. Another sixty-nine (or 36%) of the cases in our dataset were authored by judges who are currently on the bench with senior status.

Although we made efforts to ensure that our dataset was the entire population of the scope of cases that we have described, we acknowledge that our dataset was but a sample of a larger population that included cases beyond our reach.[65] Accordingly, we were sensitive to concerns about the representativeness of our dataset. To that end, we observed that forty-six cases (or 24%) were authored by female judges, and forty-four (or 23%) were authored by judges who are persons of color. These numbers correlated fairly closely with the gender and racial diversity among current federal district court judges which is 33% and 30%, respectively.[66] Our numbers were heavy with white and with male district judges, but diversity lags in federal courts and our study spanned a 30-year window of judicial decision-making.

Of the cases in our dataset, 92 (or 48%) were authored by judges who were appointed by Democratic Presidents and 101 (or 52%) by judges who were appointed by Republican Presidents. These numbers correlate closely with the current federal district court bench which, according to one recent estimate, is divided 52%/48% with the slight majority being appointed by Democratic Presidents.[67]

A separate issue of representativeness regards the potential for distortion because of selection bias that could attend the decision to publish an opinion in the official reporters or to make it available for the online databases to upload it.[68] We were

---

64 *See U.S. District Courts—Additional Authorized Judgeships*, U.S. CTS., https://www.uscourts.gov/sites/default/files/districtauth.pdf (last visited May 12, 2020) (detailing Congressional authorization for 667 permanent and 10 temporary district court judgeships).

65 Valerie Nannery reviewed snap removal cases over a three-year period from 2012 to 2014 compiling a data set of 221 cases. She utilized Westlaw, Google Scholar, and electronic court records and administrative case documents of the federal district courts. *See* Nannery, *supra* note 41, at 541. That study was recently updated by Adam Sopko. *See* Sopko, *supra* note 41. Our study differs in that it utilizes Westlaw and Lexis to review unpublished and published opinions over a longer period of time. In order to achieve this broader perspective spanning several decades, an in-depth review of every federal docket to identify snap removal cases was not feasible. The opinions we have gathered constitute a robust and substantial data set, but we join Ms. Nannery in acknowledging many limiting factors that "make it next to impossible to identify every snap removed case in the federal district courts." *See* Nannery, *supra* note 41, at 561 n.118.

66 *See Diversity on the Bench*, FED. JUD. CENT., https://www.fjc.gov/history/judges/diversity-bench (last visited May 12, 2020); *Diversity of the Federal Bench*, AM. CONST. SOC., https://www.acslaw.org/judicial-nominations/diversity-of-the-federal-bench/ (last visited May 12, 2020).

67 *See United States District Court*, BALLOTPEDIA, https://ballotpedia.org/United_States_District_Court (last visited July 3, 2020).

68 *See infra* text accompanying notes 85–91 (describing selection bias by judges and commercial legal publishers).

mindful of this risk, but our concerns faded with increased investigation. A judge's decision on a motion to remand a snap-removed case is dispositive of an important question of law about the court's subject matter jurisdiction; it is not a mine-run matter that merely requires another exercise of judicial discretion. It is also a threshold issue that is unlikely to get lost in the fog of litigation or rolled into some other issue.

Nor is it the sort of motion that a judge would be looking to shield from disclosure because, say, its resolution is unusually political, embarrassing, or complicated. Judges had plenty of cover and company no matter their decision, because snap-removal decisions were about evenly divided between removals and remands (53% and 47%, respectively).

Finally, the issue with representativeness is not that our dataset included every one of those orders, but rather that the opinions and the orders that were available and included were not systematically different than the opinions and the orders that were unavailable and not included.

III. FINDINGS

We divide the discussion of our empirical findings into three sections. In Section A, we discuss general findings about snap removal. These findings describe the phenomenon of snap removal generally, such as when and where it is happening, and what is published. In Section B, we examine specific variables that correlate with snap removal—which is to explore the who and why. In Section C, we describe a predictive model that integrates four key variables.

A. General Discussion

This Section is divided into three subsections. In the first subsection, we trace the recent history of snap-removal rates. In the second subsection, we map the geographic areas where the issue is arising and where snap removals are allowed. These areas are concentrated, and we offer various explanations for these patterns. In the third subsection, we chart the rate at which opinions raising the snap removal issue are published in the official reporters (as opposed to unpublished but available in commercial databases); these numbers are lower than one might expect.

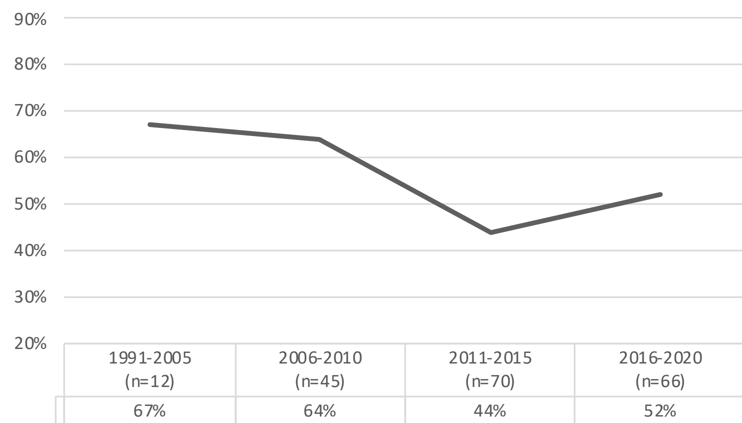
1. Historical Rates of Removal

Our dataset of 193 opinions includes published and publicly available unpublished opinions by judges who are offering their first impressions on snap removal. The cases are almost evenly divided with a total aggregate successful removal rate of 53% for the past thirty years. The annual rates have varied from 0% to 100% during these three decades, albeit with extremely small populations in some years.[69] In Figure 1 we clustered the removal rates to more elegantly convey the trend.

---

69 As recently as 2006, the removal rate was 100% (n=2); and as recently as 2003, the removal rate was 0% (n=1).

Figure 1. Snap Removal Rate Over Time (Clustered)



We remind the reader that our dataset excluded snap removal cases authored by district judges once they were bound by circuit precedent. Naturally, once that happened the removal rate in each of those jurisdictions was 100%. Even in our dataset, however, removal rates are rising. After five consecutive years with annual removal rates at or below 50% nationwide, the last two years were notably higher. The removal rates in 2018 and 2019 were 63% (n=16) and 56% (n=16), respectively. For the narrow slice of 2020 included in our dataset the removal rate is 100% (n=3).

2. Geography

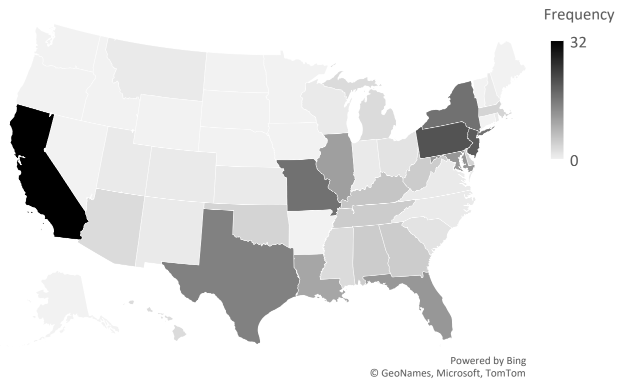
Snap removal is not a nationwide phenomenon. As depicted in Figure 2, the issue was relatively dormant in nearly half of all states.[70] Six states account for 53% of the cases in our dataset; of these six high-frequency states, four are located in circuits that now have binding precedent on the propriety of snap removal.[71] Figure 2, below, captures the frequency at which the issue of snap removal was raised, not the rate at which it was granted.

---

70 Our dataset includes no cases from Alaska, Arkansas, Connecticut, District of Columbia, Idaho, Iowa, Maine, Minnesota, Nebraska, Nevada, North Dakota, Oregon, South Dakota, Vermont, Washington, and Wyoming. And there is only one case from each of Colorado, Indiana, Kansas, Montana, North Carolina, New Hampshire, New Mexico, Rhode Island, Utah, Virginia, and Wisconsin.

71 California, Missouri, New Jersey, New York, Pennsylvania, and Texas account for 122 of the 232 cases depicted in Figure 2.

Figure 2. Geographic Distribution of Snap Removal Cases



When reporting our data, we occasionally are able to use one of our larger datasets. In Figure 2, for example, we have included not only the 193 cases that form the core of our dataset, but an additional thirty-nine opinions that were authored by judges who authored another opinion that was already included within the core dataset. Although the additional thirty-nine cases were not first impressions, they are properly part of a metric that demonstrated how frequently the issue surfaced.[72]

There are several possible contributing explanations for a concentration of snap removal cases in certain geographic areas and an absence of such cases in other areas. First, a state's service practice and procedure can affect the likelihood of snap removal. Specifically, in some states, a civil action commences upon service, rather than filing.[73] This distinction minimizes the trap for the plaintiff's lawyer who is unwary of the snap removal phenomenon. In these states, an action is not removable until some defendant is served;[74] so, provided the forum defendant is served first or is served contemporaneously with non-forum defendants, there is no window for a snap removal. By contrast, in states where a civil action commences upon filing, the action is removable from the moment of filing and continues to be removable until the forum defendant is served.

State practice and procedure can also make snap removal more likely. In New Jersey, for example, actions commence upon filing, and even savvy plaintiffs' lawyers cannot serve the defendant immediately to prevent snap removals.[75] Rather, plaintiffs' lawyers must first wait up to ten days for a Track Assignment Notice to be issued by

---

72 These 232 cases do not include snap removal cases that were decided by district judges under binding circuit court precedent.

73 *See, e.g.*, MINN. R. CIV. P. 3.01(a); N.D.R. CIV. P. 3; WASH. CIV. R. 3. Note that there are also states where actions commence upon filing and yet snap removals are rare. *See, e.g.*, ALASKA R. CIV. P. 3; ARK. R. CIV. P. 3(a); IOWA R. CIV. P. 1.301. *See also* Figure 2, *supra*.

74 Federal practice and procedure allow removal of an action, but there is no action to be removed until it has commenced (according to state law).

75 *See* N.J.R. CIV. P. 4:5A-2.

the clerk.[76] Similarly in jurisdictions like Pennsylvania that require process to be served by the sheriff, plaintiffs' lawyers cannot use private process-servers to minimize the time-window for a snap removal.[77]

Second, geographic variance with respect to snap removal may be explained, in part, by a corresponding geographic variation in certain types of litigation and litigants. For example, more than 30% of the cases in our dataset involve defendants who manufacture pharmaceuticals or medical devices.[78] Many of these companies are located in New Jersey, Pennsylvania, and California, and thus are frequently defending suits there.[79] Conferences and publications that target this cohort of repeat-players in litigation may have popularized snap removal.[80]

---

76 *See id.*; *see also* N.Y. CIV. P.L. & R. 304, R. 306 (clerk must issue an "index number" before process may be served).

77 PA. R. CIV. P. 400(a) (stating that "original process shall be served with the Commonwealth only by the sheriff," who has 30 days to effect service); s*ee also* LA. CODE CIV. P. ART. 1291 (noting "service shall be made by the sheriff").

78 These defendants include Abbot Lab'ys, Actavis Pharma, Alza Corp., Amgen Inc., Apotex Inc., AstraZeneca PLC, Bayer AG, Biogen Inc., Bio-Reference Labs., Inc., Bristol-Myers Squibb Co., ConMed Linvatec, Daiichi Sankyo Co., DuPont de Nemours & Co., Eli Lilly & Co., ExeGi Pharma, LLC, Eon Labs, Inc., Gilead Sciences, Inc., GlaxoSmithKline PLC, F Hoffmann-Laroche AG, Janssen Pharmaceutica, Johnson & Johnson, MDS Inc., McKesson Corp., Medtronic PLC, Merck & Co., Inc., Novartis Intl. AG, Organon Pharmaceuticals, Pfizer Inc., Schering-Plough Corp., Wright Med. Grp. N.V., and Wyeth, LLC.

See also Sopko, *supra* note 41 (manuscript at 29 n.201) (citing James M. Beck, *What's up with Removal Before Service?*, DRUG & DEVICE L. (May 26, 2011), https://www.druganddevicelawblog.com/2011/05/whats-up-with-removal-before-service.html [https://perma.cc/GM8M-SP28] (observing in a Westlaw search of snap removal cases that "more than half of the cases . . . didn't involve drugs and devices," which benefits corporate defendants since significant use of the tact by corporations outside the pharmaceutical industry makes it "harder for the other side to characterize it as some sort of procedural gimmick that shouldn't be allowed")). Sopko's article also includes an appendix that catalogues which law firms are employing snap removal. Of the 274 snap removal cases in his dataset, more than 75% of the defendants were represented by one of three firms, namely Fox Rothschild (125), Riker Danzig (68), or Husch Blackwell (20). *See* Sopko, *supra* note 41 (manuscript at 66 tbl. 6).

79 For example, Bristol-Myers Squibb Co., Merck & Co., Novartis Int'l AG, and many other companies have a substantial presence in New Jersey. AstraZeneca PLC, Pfizer Inc., Johnson & Johnson, and others are in Pennsylvania. Amgen Inc., Bayer AG, Gilead Scis., Inc., GlaxoSmithKline PLC, and others are in California.

80 *See, e.g.*, John P. Lavelle, Jr. & Erin E. Kepplinger, *Removal Prior to Service: A New Wrinkle or a Dead End?*, 75 DEF. COUNS. J. 177 (2008); Zach Hughes, *A New Argument Supporting Removal of Diversity Cases Prior to Service*, 79 DEF. COUNS. J. 205, 212 (2012); James M. Beck, *Making it Snappy in New York*, DRUG & DEVICE L. BLOG (July 3, 2020, 4:25 PM), https://www.druganddevicelawblog.com/2020/05/making-it-snappy-in-new-york.html; James M. Beck, *How'd the World Miss This?* Thomson v. Novartis, DRUG & DEVICE L. BLOG (Nov. 25, 2007, 4:27 PM), https://www.drug anddevicelawblog.com/2007/11/howd-world-miss-this-thomson-v-novartis.html; Katie A. Fillmore, *Third Circuit Recognizes Viability of Snap Removal by in-State Defendant,* 30 PRODS. LIAB. 15 (2019); John Polzer & Derek Carson, *"Oh Snap!" A Loophole Around the Forum-Defendant Rule*, DALLAS BAR ASS'N (Sept. 26, 2018), https://www.dallasbar.org/book-page/%E2%80%9Coh-snap%E2%80%9D-loophole-

A third and somewhat similar possible explanation for geographic variance is different legal cultures. Snap removal is something of a trick: it elevates form over substance, and it is something that, practically speaking, only rich defendants do. That is not a recipe for success in some regions of the country. Informal conversations with practitioners confirm this. One senior partner at a major law firm in Boston told us, for example, that he would never try snap-removing a case there because such hijinks would infuriate the judge.[81] There are four snap removal cases from Massachusetts in our expanded dataset, and in all four instances the cases were remanded.[82] Naturally, the earliest cases on a subject like snap removal could reflect *or create* a culture that discourages tactics like snap removal.[83] In this vein, it may be noteworthy that there was no Massachusetts snap removal case (in our dataset) after 2016.[84]

---

around-forum-defendant-rule; Lindsay C. Omolecki, *Safe at Home: Preservice Removal and the Forum Defendant Rule,* 29 PRODS. LIAB. 1 (2018); Molly E. Flynn & Christopher D. Liwski, *Looking Beyond Removal: Procedural Tactics in Plaintiff-Friendly Jurisdictions When Removal Is Not an Option*, IN-HOUSE DEF. Q., Summer 2015, at 21; Alicia A. Baiardo, *Snap Removals—An Often Unconsidered Path to Federal Court*, MCGUIRE WOODS (Jan. 16, 2019), https://www.classactioncountermeasures.com/2019/01/articles/strategy/snap-removals-an-often-unconsidered-path-to-federal-court/; Paul Werner & Abraham Shanedling, *Second Circuit Affirms "Snap" Removal Practice*, SHEPPARD MULLIN (Mar. 29, 2019), https://www.classactiondefensestrategy.com/2019/03/articles/federal-class-action/snap-removal-mdl/.

Adam Sopko has collected many more examples of law firms boasting about their effective use of snap removal. *See* Sopko, *supra* note 41 (manuscript at 28–29).

81 This attorney has requested that they remain unnamed.

82 Gentile v. Biogen Idec, Inc., 934 F. Supp. 2d 313 (D. Mass. 2013) (Woodlock, J.); Howard v. Genentech, Inc., No. 12–11153, 2013 WL 680200, at *1 (D. Mass. Mar. 19, 2013) (Woodlock, J.); Adams v. Beacon Hill Staffing Grp., No. 15-CV-11827, 2015 WL 6182468, at *5 (D. Mass. Oct. 21, 2015) (Burroughs, J.); Rizzi v. 178 Lowell St. Operating Co., 180 F. Supp. 3d 66, 68 (2016) (Gorton, J.).

83 The first two opinions on snap removal in Massachusetts were authored by Judge Douglas Woodlock in 2013. Judge Woodlock was then (and remains) a highly regarded and experienced judge. Query whether his rejection of snap removal had ripple effects beyond his courtroom; that is, if Judge Woodlock wouldn't allow snap removal, the perception might have been that no judge in that courthouse would. After all, Judge Woodlock was an elite-educated, former big-firm lawyer and Assistant U.S. Attorney, who was appointed by President Reagan.

84 Similarly, in Arizona, there are three snap removal cases and, as in Massachusetts, all were remanded. The most recent of these decisions was in 2011. *See* Traslavina v. MDS Pharma Servs. Inc., No. CV–11–0742, 2011 WL 2132880, at *2 (D. Ariz. May 27, 2011) (remanding action to state court); Rodriguez v. Cooper Tire & Rubber Co., No.: 09-CV-00415, 2010 WL 11519175, at *3 (D. Ariz. Jan. 22, 2010) (same); Ibarra v. Protective Life Ins. Co., No. CV–09–049, 2009 WL 1651292, at *4 (D. Ariz. June 12, 2009) (same). In New Mexico, the only opinion on snap removal was published in 2013. *See* Lone Mountain v. Santa Fe Gold, 988 F. Supp. 2d 1263, 1269 (D.N.M. 2013) (remanding action to state court).

However, there are also districts where the only decisions have allowed snap removal, yet there are no recent cases. In Montana, for example, the only case on the subject allowed snap removal in 2012. *See* Mahana v. Enerplus Res. U.S.A. Corp., No. CV–12–31, 2012 WL

A fourth explanation for the geographic variation could be selection bias by judges and online legal publishers that influence which published and unpublished opinions appear in commercial legal databases. Many choices that judges make, particularly at the district court level, impact the likelihood that a particular opinion or order will later appear in a commercial legal database.[85] Such choices include: deciding whether the decision will be a formal, written opinion, a memorandum opinion, a summary order, or an oral announcement from the bench;[86] labeling an opinion as a "Written Opinion" in PACER; designating it for publication; and deciding whether to send the opinion to Westlaw or Lexis.[87] Local court practices and court culture are among the factors that can influence the form of judicial determinations and lead to geographic variations.[88] Such determinations in turn impact the universe of cases available to commercial publishers who frequently rely on designations, such as "Written Opinion" in PACER and opinions recommended by judges.[89] On top of this, commercial publishers also make strategic, discretionary choices to include cases in

---

1947101, at *1 (D. Mont. May 30, 2012) (denying remand motion); s*ee also* Chace v. Bryant, No. 10–CV–85, 2010 WL 4496800, at *2 (E.D.N.C. Nov. 1, 2010) (denying remand motion).

85 *See* Elizabeth Y. McCuskey, *Submerged Precedent*, 16 NEV. L.J. 515, 536 ("In their role as decision-makers, judges must decide whether to write, how much to write, what form their writing should take, and whether to broadcast their writing, all of which respond to the circumstances of individual cases and invoke judges' discretion to manage cases. . . . The initial choice to hold hearings or decide an issue on paper has some bearing on the submergence of any decision resulting from this choice.").

86 FED. JUD. CTR., JUDICIAL WRITING MANUAL 3 (2d ed. 2013); Christina Boyd, *Opinion Writing in the Federal District Courts*, 36 JUST. SYS. J. 254, 254 (2015).

87 *See* McCuskey, *supra* note 85, at 537.

88 *See, e.g.*, *id.* at 538–39 (discussing the decreased availability in Westlaw of certain opinions from the Northern District of Illinois due to the use of a particular order form utilized in that district from 2005 to 2014); Brian Lizotte, *Publish or Perish: The Electronic Availability of Summary Judgments by Eight District Courts*, 2007 WIS. L. REV. 107, 143 ("Although difficult to quantify, local norms may create categorical regional differences in publication practice . . . ."); Pauline T. Kim et al., *How Should We Study District Judge Decision-Making?*, 29 WASH. U. J.L. & POL'Y 83, 97 (2009) (noting how publication decisions "may depend upon formal rules, court culture, personal predilections, or strategic considerations"); Nancy Leong, *The Saucier Qualified Immunity Experiment: An Empirical Analysis*, 36 PEPP. L. REV. 667, 700–01 (2009) ("Different jurisdictions, however, have markedly different practices with respect to publication . . . ."). There are other influential factors that are discussed extensively elsewhere. *See, e.g.*, McCuskey, *supra* note 85, at 535 (discussing "judges' managerial discretion, the technology used to collect and access law, the structure of the substantive law applied, a decision's appealability, and the sophistication of the parties"); Lizotte, *supra*, at 138–45 (discussing several "biases in the publication decision").

89 *See* McCuskey, *supra* note 85, at 537–38, 539; *see also Submission Guidelines for Court Opinions*, THOMSON REUTERS, https://legal.thomsonreuters.com/en/solutions/government/court-opinion-submission-guidelines (last visited July 31, 2020) (providing instructions and guidelines for judges for forwarding opinions for print and online publication).

their databases.[90] Commercial publishers may be influenced by the needs and demands of certain legal markets that "may drive publication channels to give some issues and courts greater attention."[91] All of these factors play important roles in the overall availability of written legal opinions and also contribute to geographic variations.

Importantly, the concentration of cases into states and regions was not necessarily correlated with countenance of snap removal. Nor do the judges within a state or judicial district tend to agree about the propriety of snap removal. The following examples are offered to illustrate both the variance of removal rates across jurisdictions and the lack of unanimity within most jurisdictions.[92]

Figure 3. Snap Removal Rate in Selected States

Arizona	0% (n=3)
Kentucky	25% (n=4)
Maryland	29% (n=7)
Missouri	36% (n=14)
New York	36% (n=14)
New Jersey	47% (n=15)
Florida	63% (n=8)
Illinois	63% (n=8)
California	64% (n=25)
Texas	79% (n=14)
West Virginia	80% (n=5)
Louisiana	100% (n=8)

Again, our core dataset did not include cases decided after the issuance of binding circuit precedent. Accordingly, the numbers reported for New York, New Jersey, Texas, and Louisiana in Figure 3 include only district court cases decided before the Second, Third, and Fifth Circuits, respectively, allowed snap removal.[93]

3. Publication

Our dataset included snap removal opinions that were published, as well as opinions that were formally "unpublished" though published by online commercial vendors.[94] The rate at which opinions are (formally) published is, at minimum, intriguing trivia. But we also observed evidence of ideology at work.

---

90 *See* McCuskey, *supra* note 85, at 537–38; *see also* Leong, *supra* note 88, at 701 (discussing Westlaw's selection process and varying rates of including cases across jurisdictions).

91 *See* Lizotte, *supra* note 88, at 143.

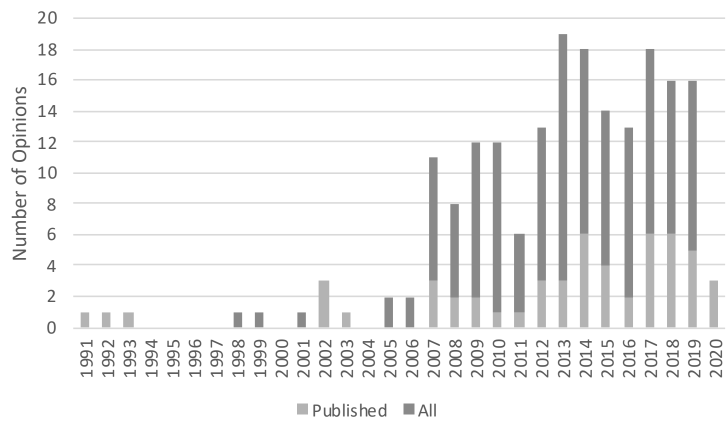
92 In Figure 3 we have returned to the core dataset (of 193 cases), because a state's removal rate is more accurately captured by excluding multiple cases decided by a single judge.

93 *See supra* text accompanying notes 60–61.

94 To be clear, "snap-removal opinions" include those where snap removal was allowed and those where it was disallowed.

Approximately 28% of the opinions in our dataset were (formally) published. Figure 4 presents raw annual data for published and unpublished opinions.

Figure 4. Annual Published and Unpublished Snap Removal Opinions



The criteria for publishing an opinion vary by jurisdiction but often reflect standards recommended in a 1973 report of the Federal Judicial Center's Advisory Council on Appellate Justices.[95] An opinion should not be published unless it "lays down a new rule of law, or alters or modifies an existing rule . . . , involves a legal issue of continuing public interest . . . , criticizes existing law . . . , [or] resolves an apparent conflict of authority."[96] Similarly, Westlaw encourages federal judges to submit opinions for publication that are "of general interest and importance to the bench and bar, such as those that: deal with an issue of first impression; establish, alter, modify, or explain a rule of law; provide a review of the law; involve unique factual situations; present a unique holding; [or] involve newsworthy cases."[97]

---

95 *See* Robert J. Martineau, *Restrictions on Publication and Citation of Judicial Opinions: A Reassessment*, 28 U. MICH. J.L. REFORM 119, 125 (1994).

96 A Report of the Committee on Use of Appellate Court Energies of the Advisory Council on Appellate Justice, Standards for Publication of Judicial Opinions, FJC Research Series No. 73–2, at 15–17 (1973). For circuit courts, the criteria for publishing an opinion are often prescribed in their local rules. *See* Joseph L. Gerken, *A Librarian's Guide to Unpublished Judicial Opinions*, 96 L. LIBR. J. 475, 480 (2004); *see also Criteria for Publication*, *in* 2A FEDERAL PROCEDURE, L. ED. § 3:916 (June 2020 update). For district courts, "the decision to publish is entirely in the judge's discretion," FED. JUD. CTR., *supra* note 86, at 7, but criteria similar to those in the Federal Judicial Center report and Westlaw's publication submission guidelines likely guide publication practices for district courts, as well. *See* Hillel Levin, *Making the Law: Unpublication in the District Courts*, 53 VILL. L. REV. 973, 987 n.76 (2008); Karen Swenson, *Federal District Court Judges and the Decision to Publish*, 25 JUST. SYS. J. 121, 121 (2004).

97 *Submission Guidelines for Court Opinions*, *supra* note 89.

In light of such publication guidelines, one might fairly expect a greater percentage of these opinions to have been slated for publication. After all, the interpretation of the forum defendant rule is a pure question of law. Broadly circulating the ruling, then, would be useful and could have positive externalities—possibly to help spread an ideology, but that would not be the only reason. The issue of snap removal is important: the interpretation of this statute resolves a substantial question about the availability of a federal forum. There was no binding precedential authority in any of the cases (in our dataset); so this was not a routine matter that was unworthy of broader circulation. Yet, in nearly three out of every four cases, the judge's order or opinion was unpublished and hidden behind an online paywall. In addition to cost, another consequence of an opinion being unpublished is that it may be treated as less persuasive authority, and depending on the date it was issued, may not be cited at all in some jurisdictions.[98]

One possible explanation for this practice is collegiality and deference. Publishing an opinion that agrees with a colleague's interpretation might seem unnecessary, and publishing an opinion that disagrees might seem discourteous. These suppositions seem unlikely given the independence of judges and the absence of (horizontal) stare decisis. Yet, it is at least conceivable that this might explain why, for example, a federal judge in Hawaii could decide not to publish his decision rejecting snap removal when two of his colleagues had, in prior years, allowed it.[99] However, in many jurisdictions, even the first opinion on snap removal was not published.[100] Collegiality cannot readily explain this mystery.

We explored the possibility that publication rates were low because district judges within a district would coalesce around either a pro- or anti- snap removal position, and would thus view the matter as routine and therefore unworthy of publication. But there is little evidence for this in light of the lack of consensus within jurisdictions, as illustrated in Figure 3 above. Indeed, Louisiana is the only state where both of the following conditions were true: (1) there were at least four opinions from the state; and (2) all of the state's opinions concurred on the availability vel non of snap removal.

---

98 The adoption of Federal Rule of Appellate Procedure 32.1 helped provide some uniformity across circuits in allowing citations to unpublished federal judicial opinions issued on or after January 1, 2007. However, Rule 32.1 is very limited in scope: "It does not prescribe rules for determining when a decision should be published. It does not require courts to permit citation of unpublished opinions issued prior to 2007. And it does not prescribe the precedential value, if any, of unpublished opinions. Circuits take varying approaches to all these questions, and practitioners should be sure to consult the local rules of the relevant circuit." *Citing Unpublished Opinions*, *in* 16AA CHARLES ALAN WRIGHT ET AL., FEDERAL PRACTICE & PROCEDURE § 3978.10 (Apr. 2020 update).

99 U.S. Bank Nat'l Ass'n v. Martin, No. 15–00061, 2015 WL 2227792 (D. Haw. Apr. 23, 2015); Deutsche Bank v. Hagan, No. 15–00376, 2015 WL 7720465 (D. Haw. Nov. 27, 2015); Watanabe v. Lankford, 684 F. Supp. 2d 1210 (D. Haw. 2010) (allowing removal).

100 *See, e.g.*, Maple Leaf Bakery v. Raychem, No. 99 C 6948, 1999 WL 1101326 (N.D. Ill. Nov. 29, 1999); Johnson v. Precision Airmotive, LLC, No. 07-CV-1695, 2007 WL 4289656 (Dec. 4, 2007); FTS Int'l Servs., LLC v. Caldwell-Baker Co., No. 13–2039, 2013 WL 1305330 (D. Kans. Mar. 27, 2013). *But see* Stan Winston Creatures, Inc. v. Toys "R" Us, Inc., 314 F. Supp. 2d 177 (S.D.N.Y. 2003).

With respect to the issuance of a published opinion, we saw no patterns. In jurisdictions where one judge published an opinion allowing snap removal, there were usually subsequent cases where judges rejected snap removal. In California, for example, there were two published (plus two unpublished) opinions allowing snap removal before the first (of nine) unpublished opinion(s) rejecting it.[101] Although less common, there were examples of the mirror situation: In Maryland there have been three published (plus two unpublished) opinions rejecting snap removal, yet two judges have authored unpublished opinions allowing it.[102] There were several districts with published opinions on each side of the divide,[103] and some with only unpublished opinions on each side.[104]

At first blush, the judges' decisions to publish appeared not to be correlated with the decision to allow or disallow snap removal. Figure 5 reveals nearly identical publication rates for remands and removals.

Figure 5. Publication Rate for Remands and Removals

x2 = 0.022; p = .882

Remands	27% (25/91)
Removals	28% (29/102)

The woefully small chi square value generated a very high p-value.[105] The p-value suggests that there is an 88.2% chance that the null hypothesis is true, to-wit: there is no statistical correlation between the decision to publish and the outcome in the case.

101 *See* Republic W. Ins. Co. v. Int'l Ins. Co., 765 F. Supp. 628 (N.D. Cal. 1991) (allowing snap removal); Cucci v. Edwards, 510 F. Supp. 2d 479 (C.D. Cal. 2007) (same); City of Ann Arbor Emps. Ret. Sys. v. Gecht, No. C-06-7453, 2007 WL 760568 (N.D. Cal. Mar. 9, 2007) (same); Waldon v. Novartis Pharms. Corp., No. C07-01988, 2007 WL 1747128 (N.D. Cal. June 18, 2007) (same); Mohammed v. Watson Pharms. Inc., No. CV09–0079, 2009 WL 857517 (C.D. Cal. Mar. 26, 2009) (remanding).

102 To be clear, only the first of the published opinions preceded the two unpublished opinions. Opinions rejecting snap removal were issued in 2002 (publ.), 2015 (publ.), 2016 (unpubl.), 2017 (publ.), and 2019 (unpubl.). The opinions allowing snap removal were issued in 2011 and 2015, respectively.

103 In the Northern District of Illinois, for example, there are two published opinions allowing snap removal and two published opinions rejecting it. *See* Vivas v. Boeing Co., 485 F. Supp. 2d 726, 735 (N.D. Ill. 2007) (remanding); *In re* Testosterone Replacement Therapy Prods. Lia. Litig., 67 F. Supp. 3d 952, 962 (N.D. Ill. 2014) (remanding); Graff v. Leslie Hindman Auctioneers, Inc., 299 F. Supp. 3d 928, 939 (N.D. Ill. 2017) (denying remand); DC by Cheatham v. Abbott Lab'ys, Inc., 323 F. Supp. 3d 991, 997 (N.D. Ill. 2018) (denying remand).

104 *See, e.g.*, Stefan v. Bristol-Myers Squibb, No. 13–1662, 2013 WL 6354588 (D. Del. Dec. 6, 2013) (remanding); Laugell v. Bell Helicopter Textron, Inc., No. 10–1080, 2012 WL 368220 (D. Del. Feb. 2, 2012) (remanding); Munchell v. Wyeth LLC, No. 12-906, 2012 WL 4050072 (D. Del. Sept. 11, 2012) (denying remand).

105 A chi square test tests the likelihood that an observed distribution is attributable to chance. It is designed to analyze categorical (as opposed to parametric or continuous) data. The test compares the observed data to a model that distributes the data according to the expectation (or null hypothesis) that the variables are independent. The chi square value is used in combination

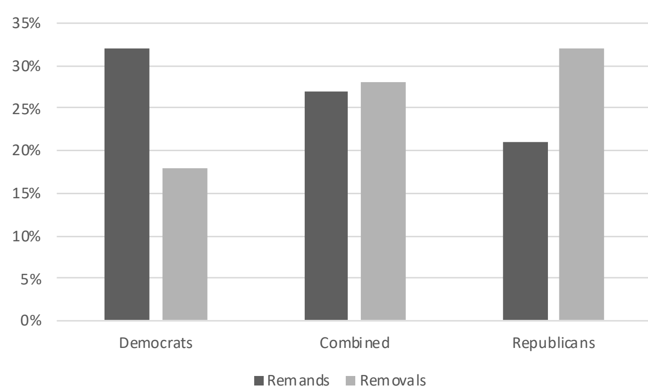
Yet Figure 5 does not tell the whole story—for a reason that is a harbinger of things to come in this Article. To be sure, the decision to publish was not correlated with case outcomes *when other variables were held constant*. Yet, there were interactions with other variables, as demonstrated in Figure 6, below. Here, we introduce the concept of party effects which we discuss throughout the paper. Party refers to the party of the President who appointed the judge. Using this proxy, we often refer to "Democratic judges" and "Republican judges" throughout.[106]

Figure 6. Publication Rate for
Remands and Removals, with Party Effects (Table)

	Democrats x2 = 3.338; p = .068*	Republicans x2 = 2.662; p = .103
Remands	32% (17/53)	21% (8/38)
Removals	18% (6/39)	37% (23/63)
	25%	31%

Although the publication rates for Democratic and Republican judges were substantially similar (25% vs. 31%), Figure 6 reveals what is referred to as a crossover (or disordinal) interaction. Figure 7, below, illustrates that interaction graphically by combining the data from Figures 5 and 6.

Figure 7. Publication Rate for
Remands and Removals, with Party Effects (Bar Graph)



In Figure 7, the middle set of bars represents the publication rate for remanded and removed cases (by Democratic and Republican judges combined); these are the numbers represented in Figure 5 where there was an 88.2% chance that there is no statistical correlation between the decision to publish and the outcome in the case. The sets of bars on the left and right sides are subgroups for Democratic and Republican

with the degrees of freedom to calculate a p-value. *See generally* DAVID R. ANDERSON ET AL., STATISTICS FOR BUSINESS & ECONOMICS (14th ed. 2019).

106 We discuss this conventional practice more thoroughly in Part III.B.1, *infra*.

judges, respectively. As reported in Figure 6, within each of these subgroups there was a strong correlation between the decision to publish and the outcome in the case. The p-values were .068 and .103, respectively. With the relatively small sample size of our dataset, we flag p-values below .10 with one star to indicate statistical significance, p-values below .05 with two stars, and p-values below .01 with three stars.[107]

The decision to publish was thus correlated with case outcomes, but that correlation ran in opposite directions for Republicans and Democrats, respectively.[108] Specifically, although Democrats published 25% of their opinions, they published 32% of their remands and only 18% of their removals. Republicans, on the other hand, published 31% of their opinions, but only 21% of their remands and 37% of their removals. This phenomenon likely reflects some desire to influence other judges and/or future litigants. Alternatively, the decision to publish may indicate a judge's higher level of satisfaction (whether visceral or ardent) with the opinion's reasoning or conclusion. Whether proselytizing or peacocking is more descriptive, ideology was correlated with a judge's decision to publish.

B. Specific Variables

In this Section we report the results of our empirical study of several variables as predictors of textualism in the context of snap removal. The most important variable was the party of the President who appointed the judge. This variable proved important in two different respects. First, its predictive power was higher than any other variable we studied. That analysis appears in the first subsection. Second, and much more subtly, we observed this party effect in several of the other variables that we analyze in subsections two through seven of this Section. Whether some of these other variables had predictive value was conditioned upon whether the judges were appointed by Republicans or Democrats. This is the "party effect" first observed in the context of publication decisions, as discussed in Part III.A.3, *supra*.

1. Party of President Who Appointed Judge

The best single predictor of a textualist case outcome was the party of the president who appointed the judge. Unsurprisingly, judges appointed by Republican presidents were more likely to allow snap removal than judges appointed by Democratic presidents. We say unsurprisingly because textualism is strongly associated with

---

107 Null hypothesis testing typically sets the decision-making threshold of statistical significance at .05. But .10 can be an acceptable tolerance for small sample sizes like ours. The risk of a higher tolerance is the possibility of more Type 1 errors. *See* Joseph F. Mudge et al., *Setting an Optimal α That Minimizes Errors in Null Hypothesis Significance Tests* 6, PLoS ONE (Feb. 2012), https://doi.org/10.1371/journal.pone.0032734 (describing disadvantages in consistently relying on the traditional but arbitrary significance level of .05).

108 The correlation between the decision to publish and the outcome in the case for all judges (combined) was not statistically significant because the correlation was masked by a cancellation of the oppositional forces of the two subgroups—one establishing correlation between publication and remand, and the other establishing correlation between publication and removal.

Republicans (since President Reagan),[109] and purposivism with Democratic judges.[110] Using the party of the appointing president as a proxy for the ideological preference of judges is a routine practice.[111] And it is a routine practice because it has been empirically substantiated in many studies.[112] That said, our study adds something new to that literature. Although it is generally accepted that ideology can help explain judicial outcomes at the appellate level, many have suggested—and verified—that there are fewer or no party effects at the trial court level.[113] To the extent that we demonstrate party effects here and throughout this Article, we offer a counterexample to that conventional wisdom.

Figure 8 below, presents the basic tally. This difference in the rate of removal was statistically significant at the highest level.

Figure 8. Snap Removal Rate by Party of Appointing President

x2 (1, *n*=193) = 7.717; p = .005***

Republicans	62% (63/101)
Democrats	42% (39/92)

109 *See, e.g.*, Abramowicz et al., *supra* note 6, at 436; Orin S. Kerr, *Shedding Light on* Chevron*: An Empirical Study of the* Chevron *Doctrine in the U.S. Courts of Appeals*, 15 YALE J. ON REG. 1, 28 (1998); Steven R. Greenberger, *Civil Rights and the Politics of Statutory Interpretation*, 62 U. COLO. L. REV. 37, 51–52 (1991); Jeffrey W. Stempel, *The Rehnquist Court, Statutory Interpretation, Inertial Burdens, and a Misleading Version of Democracy*, 22 U. TOL. L. REV. 582, 591 (1991); Nicholas S. Zeppos, *Justice Scalia's Textualism: The "New" New Legal Process*, 12 CARDOZO L. REV. 1597, 1599 (1991).

110 Kerr, *supra* note 109, at 28; Stempel, *supra* note 109, at 564; Abramowicz et al., *supra* note 6, at 436.

111 *See, e.g.*, Abramowicz et al., *supra* note 6, at 423; Cass R. Sunstein et al., *Ideological Voting on Federal Courts of Appeal: A Preliminary Investigation*, 90 VA. L. REV. 301, 305 (2004); Daniel R. Pinnello, *Linking Party to Judicial Ideology in American Courts: A Meta-Analysis*, 20 JUST. SYS. J. 219 (1999); Richard L. Revesz, *Environmental Regulation, Ideology, and the D.C. Circuit*, 83 VA. L. REV. 1717, 1718 n.6 (1997).

112 *See, e.g.*, Richard J. Pierce, Jr., *Two Problems in Administrative Law: Political Polarity on the District of Columbia Circuit and Judicial Deterrence of Agency Rulemaking*, 1988 DUKE L.J. 300, 302 ("[D]emocratic D.C. Circuit judges are more likely to reverse agency policies at the behest of individuals, and republican D.C. Circuit judges are more likely to reverse agency policies challenged by business interests."); *see supra* notes 109–11 and accompanying text; *see infra* note 113 and accompanying text.

113 *See, e.g.*, Orley Ashenfelter et al., *Politics and the Judiciary: The Influence of Judicial Background on Case Outcomes*, 24 J. LEGAL STUD. 257, 281 (1995) (finding that judges' political preferences do not affect the outcome of summary judgment motions in civil rights cases); Brian T. Fitzpatrick, *An Empirical Study of Class Action Settlements and Their Fee Awards*, 7 J. EMPIRICAL LEGAL STUD. 811, 840–41 (2010) (finding no effect of the judges' political party on class action settlements and their fee awards); Laura Beth Nielsen et al., *Individual Justice or Collective Legal Mobilization? Employment Discrimination Litigation in the Post Civil Rights United States*, 7 EMPIRICAL LEGAL STUD. 175, 193 (2010) (finding that the "party of the deciding judge bears no relation to outcome").

An odds ratio is sometimes a more useful way to analyze correlation amid categorical outcomes.[114] The odds of removal by a Republican judge were 2.25 times greater than the odds of removal by a Democratic judge.[115]

Moreover, there is evidence that party effects may be getting stronger. In Figure 9, below, we report the removal rates for judges when sorted by their appointing President. The darker shading distinguishes Republicans from Democrats.

Figure 9. Snap Removal Rate by Appointing President

L. Johnson	* (1/1)
R. Nixon	* (3/4)
G. Ford	* (1/1)
J. Carter	25% (2/8)
R. Reagan	68% (13/19)
G.H.W. Bush	48% (13/27)
W. Clinton	49% (20/41)
G.W. Bush	65% (30/46)
B. Obama	38% (16/42)
D. Trump	* (3/4)

Judges appointed by President Obama had lower removal rates (38%) than judges appointed by the Democratic president who immediately preceded him (49%). And judges appointed by President George W. Bush had higher removal rates (65%) than judges appointed by the Republican president who immediately preceded him

114 Odds ratios are particularly useful in logistic regression—and with dichotomous variables more generally—because they approximate how much more likely or unlikely (in terms of odds) it is for the outcome to be present among those subjects where the dummy variable of *x* equals 1 compared to those where it equals 0.

Also, for polychotomous or continuous independent variables (*see infra* notes 146–47 and accompanying text [Part III.B.5]), an odds ratio represents the constant effect of an independent variable on a predicted outcome. If the effect is represented as a probability (rather than as odds) the effect is not constant. *See also infra* note 115 and accompanying text.

115 For the uninitiated, odds are not probabilities. Probability is the ratio between the number of events favorable to some outcome and the total number of events. By contrast, odds are the ratio between probabilities: the probability of an event favorable to an outcome and the probability of an event against the same outcome. Probability is constrained between zero and one, while odds are constrained between zero and infinity. An odds ratio is the ratio between odds.

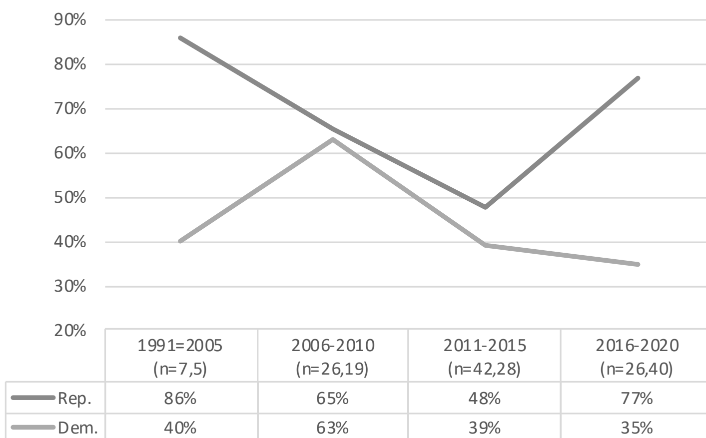
The odds of removal by a judge who was appointed by a Republican president were 63/38 = 1.658. The odds of removal by a judge who was appointed by a Democratic president were 39/53 = .736. The ratio of these odds is 1.658 / .736 = 2.25.

Odds can be converted into probabilities by dividing the odds by (1 + odds). Thus the *probability* of a Republican judge removing is 1.658 / (1 + 1.658) = .62.

(48%).[116] These trends portend more polarization as the Obama and the younger Bush judges (plus the Trump judges) decide more of the forthcoming snap removal cases (and as the Clinton and elder Bush judges decide fewer). We expect that to happen soon since the turnover on the District Court bench occurs rapidly.[117] Given the large number of judicial appointments from President Trump, we would expect the rate of successful snap removals to steadily increase in the short term. We discuss this more thoroughly in Part III.C, *infra*.[118]

The fairly recent party polarization on the issue of snap removal is apparent also in Figure 10, below. This graph is the same as Figure 1 except that Figure 1 combined Republican and Democratic judges into a single line graph. Figure 10 shows the compositional subgroups.

Figure 10. Snap Removal Rate Over Time (Clustered), with Party Effects



116 This polarization in judicial ideology, which has been documented in numerous contexts, is generally thought to have begun with President Reagan. *See generally* RICHARD A. POSNER, THE FEDERAL COURTS: CHALLENGE AND REFORM 18 (1996) (describing President Reagan's aggressive efforts to rebalance the bench after "notably liberal" judges appointed by President Carter). *But see* Timothy B. Tomasi & Jess A. Velona, *All the President's Men? A Study of Ronald Reagan's Appointment to the U.S. Courts of Appeals*, 87 COLUM. L. REV. 766 (1987) (examination of Reagan circuit court appointees suggest comparatively little difference in views relative to predecessors).

117 As of May 1, 2020, there were 260 active district court judges appointed by President Obama on the bench, compared to only 55 appointed by President Clinton. And there were 130 active district court judges appointed by President George W. Bush on the bench, compared to only 8 appointed by President George H.W. Bush. President Trump has appointed 145 active district court judges who have begun service; another 42 are in various pending stages. *See Law360's Guide to Trump's Judicial Picks*, LAW360.COM, https://www.law360.com/articles/963060/law360-s-guide-to-trump-s-judicial-picks (last updated July 28, 2020).

118 *See infra* note 168 and accompanying text.

Although there was only modest separation between Republican and Democratic removal rates for a decade, since 2016 the difference was glaring. And it is growing: since the start of calendar year 2019, the Republican removal rate was 100% (n=8) and the Democratic removal rate was 36% (n=11).

2. Gender of Judge

Next we consider gender effects in outcomes while controlling for all other variables. The gender of the judge appeared to have some predictive value with respect to textualist outcomes. As indicated in Figure 11, below, women judges were more likely than men to allow snap removal.

Figure 11. Snap Removal Rate by Gender

x2 = 1.559; p = .212

Women	61% (28/46)
Men	50% (74/147)

Again, odds ratios can illuminate relationships between and among categorical variables. The odds of a female judge permitting snap removal were about 1.5 (or 3:2) times greater than the odds of a male judge allowing it.[119] This difference does not, however, register as statistically significant. With a p-value of .212, there is roughly a 21% chance that this differential could be observed even if our null hypothesis were true, to-wit: that there is no correlation between gender and removal rates. To be sure, the relatively small number of opinions by female judges (n=46) made it harder to find statistical significance.[120]

There was also nuance to this gender effect. Or, put another way, there were two different gender effects, namely Republican women and Democratic women. Figure 12, below, tracks both of these cohorts.

Figure 12. Snap Removal Rate by Gender, with Party Effects (Table)

	Democrats x2 = .269; p = .604	Republicans x2 = 4.099; p = .043**
Females	46% (13/28)	83% (15/18)
Males	41% (26/64)	58% (48/83)

These data indicate that there was a very small gender effect among Democrats, but a strong and statistically-significant effect among Republicans. The odds of a Republican female judge allowing snap removal were 7.3 times greater than the odds of a Democratic male judge allowing it.[121] The odds of Republican female judge

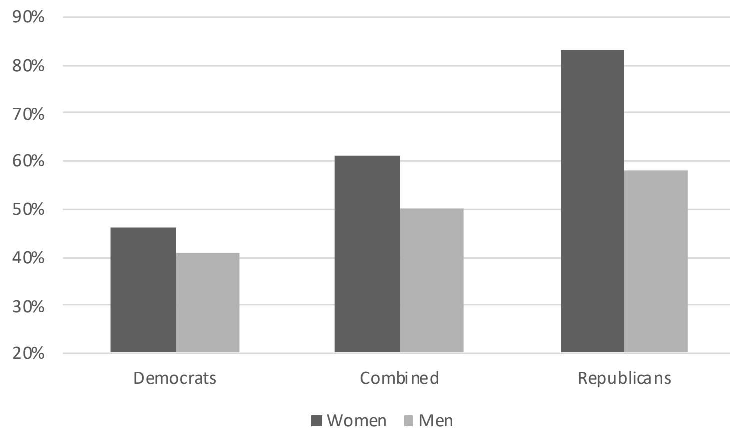
119 The odds of removal by a female judge were 28/18 = 1.56. The odds of removal by a male judge were 74/73 = 1.01. The ratio of these odds is 1.56/1.01 = 1.54.

120 By contrast, with the party effects variable, we had 92 and 101 datapoints in each group.

121 The odds of a Republican female judge allowing removal were 15 / 3 = 5. The odds of a Democratic male judge allowing removal were 26 / 38 = .684. The odds ratio is 5 / .684 = 7.310.

allowing snap removal were 3.6 times greater than the odds of a Republican male judge allowing it.[122] Figure 13 further elaborates this gender effect.

Figure 13. Snap Removal Rate by Gender, with Party Effects (Bar Graph)



The middle set of bars report the data presented in Figure 11. The sets of bars on the left and right sides are subgroups for Democratic and Republican judges, respectively, as reported in Figure 12. To be clear, this was not the sort of crossover interaction discussed in Part III.A.3, *supra*. Here, the winds of the gender effect blew in only one direction, with a breeze for Democrats and a gust for Republicans. The direction was toward more textualist outcomes.

Thus, even if not overwhelming, the variable of gender was important. Gender alone as a separate variable was compelling but not statistically significant. But gender with party effects revealed something more subtle. Democratic women were only slightly more likely than Democratic men to remove. But Republican women were substantially more likely than Republican men to remove.

3. Race of Judge

Testing for race effects required the admittedly crass act of combining all of the non-White races into a single cluster in order to assemble a sufficiently large sample. This gave us a set of 44 cases, a number worthy of meaningful study. The data are presented in Figure 14, below.

Figure 14. Snap Removal Rate by Race (Clustered)

$x2 = .063; p = .802$

Non-White	51% (22/43)
White	53% (80/150)

[122] The odds of a Republican female judge allowing removal were 15 / 3 = 5. The odds of a Republican male judge allowing removal were 48 / 35 = 1.371. The odds ratio is 5 / 1.371 = 3.647.

This was decidedly uninteresting from a statistical perspective.

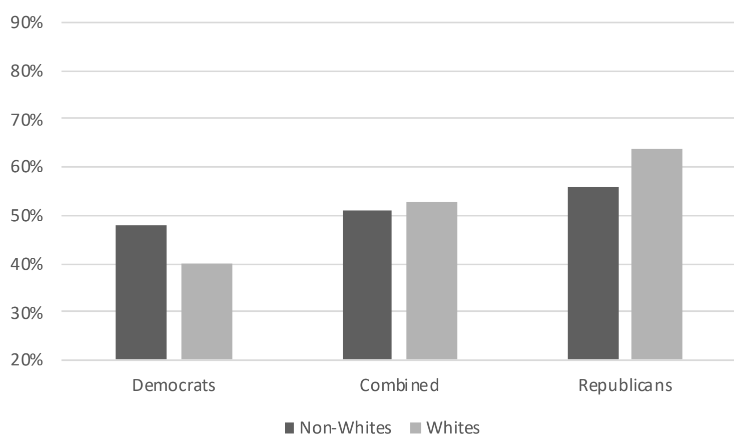
There was some intriguing evidence, however, that racial diversity moderated party effects. As indicated in Figures 15 and 16, below, non-white Democrats were more inclined to allow snap removal than other Democrats, and non-white Republicans were less inclined to allow snap removal than other Republicans.

Figure 15. Snap Removal Rate by Race (Clustered), with Party Effects (Table)

	Democrats x2 = .442; p = .506	Republicans x2 = .434; p = .510
Non-Whites	48% (12/25)	56% (10/18)
Whites	40% (27/67)	64% (53/83)

This was another crossover interaction, where the main effect of the race variable was neutralized by opposing party effects.[123]

Figure 16. Snap Removal Rate by Race (Clustered), with Party Effects (Bar Graph)



This moderating effect on both parties suggests that, with time, the increasing diversification of the federal judiciary will (or would, as the case may be[124]) decrease

123 "Main effect" (or "simple effect") refers to the effect that the independent variable (i.e., race) has on the dependent variable (i.e., snap removal). "Interactions" refer to the combined effect of two independent variables (i.e., race and party) on the dependent variable (i.e., snap removal).

124 The gender diversity of President Trump's appointees to the district court bench are slightly higher than the percentage of female judges in our dataset, but considerably lower than the gender diversity of President Obama's judges. The racial diversity of President Trump's appointees to the district court bench are substantially lower than the percentage of non-white judges in our dataset, and substantially lower than the percentage of non-white judges appointed by President Obama. *See infra* note 168; *see also Diversity of the Federal Bench*, *supra* note 66 (reporting gender and racial diversity of President Obama's appointees as 42% and 36%, respectively); John Gramlich, *How Trump Compares with Other Recent Presidents in*

the level of polarization among Republicans and Democrats. However, what we observed did not register as statistically significant. Because there was no strong evidence that the race of the judge influenced outcomes, we do not further discuss this variable.

4. Judicial Experience

We tested several hypotheses regarding correlations between judicial experience and case outcomes. Our dataset allowed us to interrogate experience from different angles, including age, years of experience as a federal judge, senior/active status, and prior experience as a state judge.

a. Age

We first considered the age of the judge at the time the opinion was issued. Age is a tricky variable for a couple of reasons. First, as a linear variable, the data gets sliced too thinly for a meaningful logistic regression: our population of 193 cases was spread across a range of 55 years (from age 39 to 94).[125] That problem was easily solved by treating age as a categorical variable and clustering the population into age groups. But there was a second and more fundamental limitation about using age as a variable. Lest one be p-hacking, use of the age variable needed to be constrained to plausible explanatory hypotheses.[126]

There were two plausible hypotheses with respect to a judge's age and case outcomes. The first hypothesis, floated most recently by Gluck and Posner, was that young judges were more likely to reach textualist outcomes than their middle-aged and elder peers.[127] A second hypothesis presumably would examine old judges. But what is the explanatory theory behind that? Perhaps being more or less textualist is correlated with advanced age (and emerging senility) because of a lack of rigor or pronounced stubbornness or an outsourcing of work to law clerks. We were dubious of such theories and we discuss them further below. The immediate point of emphasis

---

Appointing Federal Judges, PEW RSCH. CTR. FACTTANK (July 15, 2020), https://www.pewresearch.org/fact-tank/2020/07/15/how-trump-compares-with-other-recent-presidents-in-appointing-federal-judges/. *See generally* Carl Tobias, *President Donald Trump's War on Federal Judicial Diversity*, 54 WAKE FOREST L. REV. 531 (2019); Stacy Hawkins, *Trump's Dangerous Judicial Legacy*, 67 UCLA L. REV. DISC. 20 (2019).

125 The rule of thumb for logistic regression is a minimum of ten cases with the least frequent outcome for each independent variable in the model.

Further, a linear experience variable (like a linear age variable) has a long tail that makes correlations with a categorical dependent variable like removal unlikely. Confidence or wisdom (or for whatever age is serving as the proxy here) likely does not intensify in a linear fashion: the effect of moving from year 2 to 3 is not the same as from 14 to 15 or from 26 to 27.

126 Researchers should not use the data to create the hypotheses, because the data inevitably produce correlations that are pure chance. As explored more fully in Part III.C, *supra*, there is a danger in creating a model that is, in fact, a *prediction of the past*. The danger is crafting a theory that is so tied to idiosyncrasies in past data that it would not be validated by more (say, incoming future) data of the same sort.

127 Gluck & Posner, *supra* note 1, at 1311–12; *see also id.* at 1300, 1302, 1305, 1319, 1326, 1331, 1342, 1351 & 1353.

with respect to p-hacking, however, is not the young nor the old, but rather the middle aged: there was no theoretical basis for exploring a difference in removal rates between, say, sixty-one- and sixty-eight-year-old judges.[128]

But a theory about eager young judges was plausible. Gluck and Posner described the influence of the emergence of an intellectually vibrant field of statutory interpretation (that featured textualism) beginning in the 1980s.[129] That movement ultimately found its way into legal education where it displaced legal process theory.[130] This, in turn, has caused a generational shift such that contemporary younger judges "do not seem to focus on the big picture questions about the judicial role and inherent authority that we heard emphasized by the older judges—indeed, they seem more insecure than the older judges that they even have . . . authority [to fill gaps in laws, to correct Congress's mistakes, or to make law.]"[131]

We investigated—and to some extent confirmed—Gluck and Posner's hypothesis about the young judge.[132] We established a dummy variable, Young, that separated the young from the not young. Of course, there is something arbitrary about choosing

---

128 The data might suggest, for example, that judges who are in the first three years of their 50s, 60s, 70s, and 80s might have higher removal rates than judges who are in the last three years of their 50s, 60s, 70s, and 80s. Or even that such a pattern holds for the first three of those decades but then reverses for judges in their 80s. All this is easy to incorporate into a model to improve its predictive capacity. But absent some meaningful explanation, it is noisy nonsense.

For the hopelessly curious, we share the following data:

Figure 16.1. Snap Removal Rate
by Age (Clustered), with Party Effects (Table)

	Democrats	Combined	Republicans
≤ 50	50% (2/4)	69%	78% (7/9)
51–59	45% (13/29)	57%	80% (12/15)
60–69	38% (14/37)	46%	52% (25/48)
≥ 70	45% (10/22)	56%	66% (19/29)

The two takeaways here were (1) the relative stability of the Democratic removal rate across age groups, compared to Republicans; and (2) the loping U-shape of the removal rate across age groups.

129 Gluck & Posner, *supra* note 1, at 1304 (citing Frickey, *supra* note 6, at 242) (practitioners in the 1970s had no scholarship to draw upon); John F. Manning, *Legal Realism & The Canons' Revival*, 5 GREEN BAG 2D 283, 290 (2002); Eskridge, *supra* note 27, at 532–33 (reviewing SCALIA & GARNER, *supra* note 27).

130 Gluck & Posner, *supra* note 1, at 1312 (citing sources).

131 *Id.* at 1311.

132 Gluck and Posner's survey was limited to appellate judges.

any particular age for a characteristic of a young judge, but we chose 55.[133] The average age-at-appointment for judges in our dataset was 50.[134] About one-fifth of the judges in our dataset fell into this young judge category. To be clear, however, our young variable was dynamic and, thus, included a judge who was young at the time they issued their decision in, say, 1992.[135] Even with that important qualification, almost all of our young judges were in fact on the young side of the generational shift. Among the cases in our dataset decided by young judges, thirty-four of thirty-nine were decided in 2006 or later.[136] The typical young judge in our dataset was born in 1964 and graduated from law school in 1990.[137] All thirty-nine of the young judges in our dataset assumed the bench well after Justice Scalia was already on the Supreme Court.[138]

Figure 17, below, contrasts the removal rates for young and over-55 judges. The distinction was very nearly statistically significant at the .10 level.

Figure 17. Snap Removal Rate
by Age (Clustered)
x2 = 2.484; p = .115

Young	64% (25/39)
Over 55	50% (77/154)

Once again we sub-grouped the data for a deeper understanding of this variable and its possible interaction with the party variable. That data is presented in Figure 18.

Figure 18. Snap Removal Rate
by Age (Clustered), with Party Effects (Table)

	Democrats x2 = 1.844; p = .175	Republicans x2 = 2.332; p = .127
Young	54% (13/24)	80% (12/15)
Over 55	38% (26/68)	59% (51/86)

133 Using the age of 60 as the cut-off delivered substantially similar results.

134 The mean and the median were both 50. The mode was 48.

135 For example, in Wensil v. E.I. DuPont de Nemours & Co., 792 F. Supp. 447 (D.S.C. 1992), a 39-year-old then-District Judge Dennis Shedd allowed snap removal. Judge Shedd graduated from law school in 1978. In 2020, now-Circuit Judge Shedd is 67 years old.

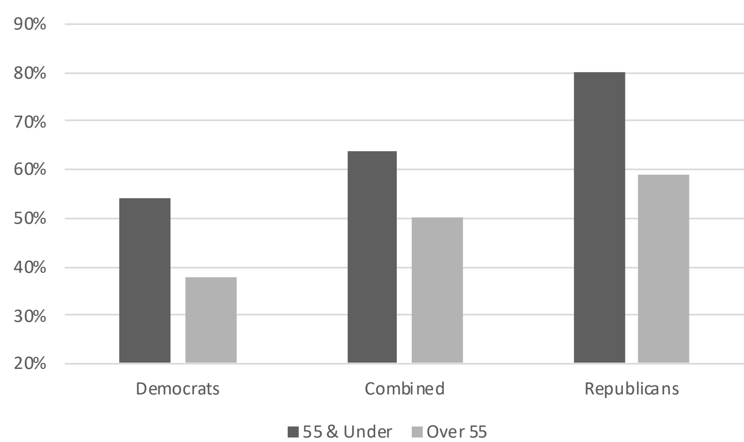
136 The five exceptions were decided in 1991, 1992, 1998, 2002, and 2003, respectively.

137 Only two were born before 1951—in 1946 and 1949, respectively.

138 The earliest appointment to the federal bench among this group of young judges occurred in 1990. *See generally* William N. Eskridge, Jr., *The New Textualism*, 37 UCLA L. REV. 621, 623 (1990) (describing how Justice Scalia inspired a "new" textualism); Nicholas S. Zeppos, *Justice Scalia's Textualism: The "New" New Legal Process*, 12 CARDOZO L. REV. 1597, 1598 (1991) (discussing Justice Scalia's singular role); ANTONIN SCALIA, A MATTER OF INTERPRETATION (1997).

We did not observe much of a party effect: among both Democrats and Republicans, their young judges had higher removal rates. The odds of removal by a young Democratic judge were 1.91 times that of the non-young Democratic judge. The odds of removal by a young Republican judge were 2.75 times that of the non-young Republican. The p-value for each of these sub-groups was less impressive than the p-value for the combined group (as captured in Figure 17); this is attributable to the smaller population size of the sub-groups. Figure 19, below, offers a graphic illustration of the same data.

Figure 19. Snap Removal Rate by Age (Clustered), with Party Effects (Bar Graph)



The party effect was modest, so the more compelling finding was the higher removal rate as a main effect of the young variable.

We also interrogated the second hypothesis regarding an "old judge" effect. We struggled with the definition of who was old, and more fundamentally, we lacked conviction in possible hypotheses that explored something beyond the "not young" judge, which we have already (if indirectly) addressed above. In any event, we settled upon something with a theoretical grounding, but reserve discussion of that effect until Part III.B.4.c, *infra*.

b. Article III Experience

We calculated the number of years of experience as an Article III judge that each judge had accrued when their opinion was issued. We observed less effect on this alternative measure of experience than we observed on our examination of age in Part III.B.4.a, *supra*.

We followed the same methodological approaches that we used for the age variable by converting experience into categorical data, and avoiding reliance on atheoretical correlations.[139] Using three or fewer years of experience as the definition of a new judge, we found essentially no difference in removal rates between the two cohorts.

Figure 20. Snap Removal Rate by Art. III Experience (Clustered)
x2 = .007; p = .934

0-3 Years	54% (15/28)
More than 3	53% (87/165)

This suggests that Gluck and Posner were correct in suggesting that textualism was a "young judge" phenomenon, as opposed to a "new judge" phenomenon.[140] Indeed even when we used four years (instead of three) as the definition of a "new judge," we reached similar results (with 52% and 53% removal rates, respectively).[141] That is to say that the "young judge" and "new judge" cohorts were not coextensive.

As we observed with the "young judge" cohort, the "new judge" cohort is more heavily populated with Democrats. The interaction of the "new judge" and party effect variables is demonstrated in Figure 21 below.

139 The atheoretical forays were not especially productive in any event. But lest we disappoint the data-dredging enthusiast, we share the following data:

Figure 19.5. Snap Removal Rate by Years of Experience (Clustered), with Party Effects (Table)

	Democrats	Combined	Republicans
0 – 3	45% (9/20)	54%	75% (6/8)
4 – 10	39% (12/31)	44%	50% (13/26)
11 – 20	50% (11/22)	71%	83% (25/29)
> 20	37% (7/19)	47%	53% (20/38)

The two takeaways here are (1) the relative stability of the Democratic removal rate across experience clusters, compared to Republicans; and (2) the matched patterns (of fall, rise, and fall) of the removal rates with increased experience among Democrats and Republicans.

140 We do not mean to suggest that Gluck and Posner rejected the "new judge" label in favor of their "young judge" label. They are silent on this point.

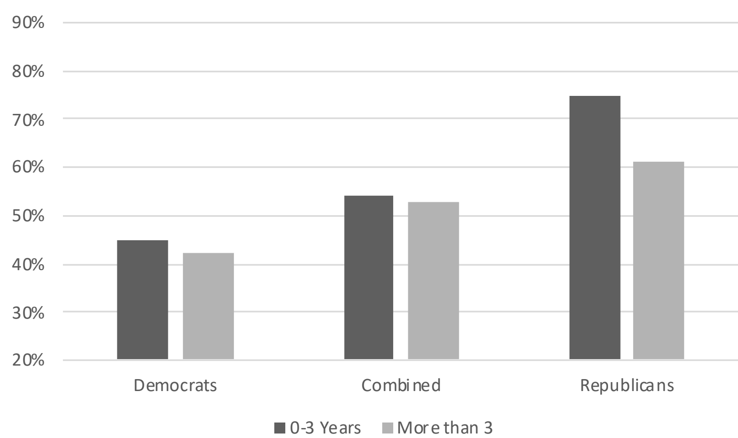
141 Removals by judges with up to four years of experience: 17 / 33 = 52%. Removals by judges with more than four years of experience: 85/160 = 53%. This calculation with the four-year cutoff had a sample size (n) of 33, compared to the n of 28 for the "young judge" calculation.

Figure 21. Snap Removal Rate by Art. III Experience (Clustered), with Party Effects (Table)

	Democrats x2 = .071; p = .790	Republicans x2 = .059; p = .442
0-3 Years	45% (9/20)	75% (6/8)
More than 3	42% (30/72)	61% (57/93)

With these subgroups, there was modest evidence that being a new judge was more predictive of removal for Republicans than Democrats. Yet, it was not statistically significant—especially with such a small sample size. Figure 22 illustrates the magnitude of the effect graphically.

Figure 22. Snap Removal Rate by Art. III Experience (Clustered), with Party Effects (Bar Graph)



For both Democrats and Republicans, the spread between "new judge" and more-experienced judge was substantially smaller than it was on the measure of age, which is reflected in Figures 18 and 19, above.

At the other end of the experience continuum, of course, are judges with extensive experience. We discuss them in the next subpart.

c. Senior Status

We have previously alluded to hypotheses about the removal rate of elder or especially-experienced judges. Thrashing for a defensible variable to explore, we settled upon Senior Judge status. To qualify for senior status, a judge must satisfy two conditions: (i) be at least 65 years of age and (ii) have a sum of age plus years of federal court service that is at least 80.[142] Senior status is voluntary, and election of this status does not necessarily mean any diminishment of the judge's caseload, commitment, or capacity. But a Senior Judge does not participate in en banc panels, suggesting some nominal detachment. And many Senior Judges do transition to lighter caseloads. Though undoubtedly both underinclusive and overinclusive of elder

142 28 U.S.C. § 371(c).

statesmanship, electing senior status is an objective indication of advanced age and experience. Query then whether it was also indicative of an intensified fondness for—or resistance to—the formalism represented in textualist outcomes.

More than a quarter of the cases in our dataset were decided by Senior Judges.[143] The average age of a Senior Judge in our dataset (measured at the time of the decision) was 75.2 years, with 24.8 years of experience.[144] As demonstrated in Figure 23, their rates of removal were similar to their active peers.

Figure 23. Snap Removal Rate by Senior and Active Judges

$x2 = .107; p = .744$

Senior	51% (27/53)
Active	54% (75/140)

The interaction of the senior status and party variables demonstrated in Figure 24 was more interesting, however.

Figure 24. Snap Removal Rate by Senior and Active Judges, with Party Effects (Table)

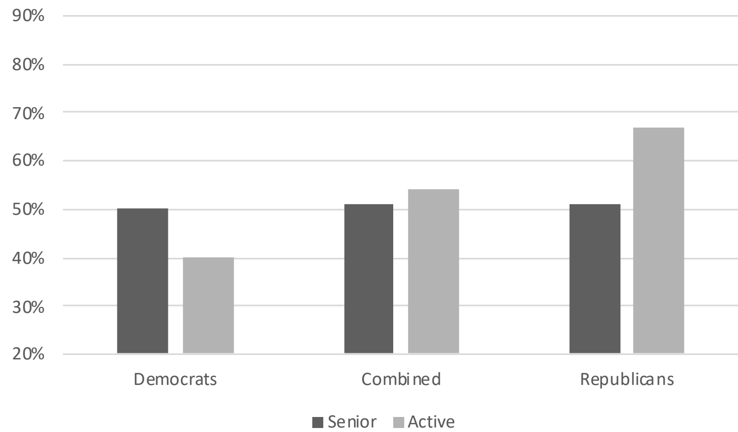
	Democrats $x2 = .685; p = .408$	Republicans $x2 = 2.208; p = .137$
Senior	50% (11/22)	51% (16/31)
Active	40% (28/70)	67% (47/70)

Here we observed senior status playing a role similar to that which we observed in the context of race. Senior status had a crossover interaction with the parties that had a moderating influence on both subgroups: raising the rate of removal by Democrats and lowering the rate of removal by Republicans. The effect within the Republican subgroup was nearly statistically significant at .10. Figure 25 graphically illustrates the crossover interaction. The removal rate among senior judges was essentially 50% regardless of party.

143 This datum is no surprise to those who have documented the incredible workload borne by the judiciary's Senior Judges. *See* Stephen B. Burbank et al., *Leaving the Bench, 1970–2009: The Choices Federal Judges Make, What Influences Those Choices, and Their Consequences*, 161 U. PA. L. REV. 1, 93 (2012) (observing that, in 2009, senior district judges accounted for 21.2% of case terminations and 26.8% of all trials).

144 The medians for these measures were 74 and 23, respectively.

Figure 25. Snap Removal Rate
by Senior and Active Judges, with Party Effects (Bar Graph)



The moderating influence of senior judges (on their more polarized, active peers) was an attractive narrative, to be sure. And the near-statistical significance of the Republican subgroup justified a hunt. But we ultimately abandoned the chase when we could not corroborate these findings with enough data to allay our concerns with using it. The strongest argument in favor of using it was that senior status is something of a compromise variable between two other purported measures of advanced age and experience. Figure 26, below, lays out the three measures with the ordinal headings for each of them.

Figure 26. Effect of Different Variables for Age and Experience
on Snap Removal Rates, with Party Effects (Table)

	Democrats	Combined	Republicans
Senior Status	Increases	No effect	Decreases
75+ Years Old	Increases	Increases	Increases
20+ Years Experience	Decreases	Decreases	Decreases

The middle row suggests that old judges had higher rates of removal than their younger colleagues. The bottom row suggests that highly experienced judges had lower rates of removal than their less experienced colleagues. And of course, the first row suggests that senior judges had the same rates of removal as their (generally younger, generally less experienced) active colleagues—and had a mollifying effect on the removal rates of their fellow Democrats and Republicans. Statistical noise is the most conservative explanation for these disparate results. Drawing conclusions about this data would likely be as dubious as the explanatory theory that would need to accompany it.

d. State Court Experience

Approximately one-third of the cases in our dataset were authored by judges who served as state court judges before their appointment to the federal bench. We speculated that that exposure or affinity might lead a judge to be more inclined to remand a snap-removed case, since the judge would know first-hand that it was an unbiased forum with competent judges. Accordingly, we investigated whether prior service as a state court judge helped predict outcomes. We measured this as a binary categorical variable, rather than as a linear variable that segmented state court experience by years of service.[145]

As demonstrated in Figure 27, below, service as a state court judge was another statistically-uninteresting variable.

Figure 27. Snap Removal Rate
by Whether Any State Court Judicial Experience
$x2 = .312; p = .576$

Experience	50% (32/64)
No Experience	54% (70/129)

Looking again for distinctive interactions with party, the removal rates for each of these subgroups appear in Figure 28, below.

Figure 28. Snap Removal Rate by Whether
Any State Court Judicial Experience, with Party Effects

	Democrats $x2 = .260; p = .610$	Republicans $x2 = .065; p = .798$
Experience	39% (12/31)	60% (20/33)
No Experience	44% (27/61)	63% (43/68)

The effect of this variable was uniform but negligible across both subgroups. Accordingly, we concluded that there was no evidence that prior state court judicial service influenced a judge's decision to allow snap removal, and we do not further discuss this variable. We found this surprising in that we had initially posited that a federal judge with state court service would be more deferential to state court proceedings and more likely to remand snap removed cases. Apparently not.

5. Education

In this Section we present our most surprising—and probably the most important—finding of our study. The party effect which was observed in the parsing of so many variables discussed above was, in fact, driven almost entirely by judges with elite educations. In other words, judges who did not attend elite institutions had essentially the same removal rates, regardless of whether they were Republican or Democratic appointees. That is, party effect is a feature of elite-educated Republican judges who pushed removal rates up, and of elite-educated Democratic judges who pulled removal rates down.

145 *See supra* text accompanying note 125.

a. Undergraduate

We investigated and coded the undergraduate education of each of the judges who authored opinions in our dataset. We then grouped the colleges and universities into three categories of eliteness.[146] We were curious whether the eliteness of the judge's undergraduate education had any predictive value in determining that judge's treatment of snap removal.

As indicated in Figure 29 below, there was no observable effect in the eliteness of the undergraduate education variable.

Figure 29. Snap Removal Rate
by Eliteness of Undergraduate Education (Clustered)
$x2 = .299$; $p = .861$

Top 4	48% (14/29)
Top 20	54% (25/46)
Others	53% (63/118)

However, we observed that this undergraduate education variable was being distorted by its interaction with party. Specifically, the eliteness of a judge's undergraduate education had an observable negative correlation with snap removal for Democrats, but a positive correlation for Republicans. This is depicted in Figure 30, below.

Figure 30. Snap Removal Rate by
Eliteness of Undergraduate Education (Clustered),
with Party Effects (Table)

	Democrats $x2 = 1.842$; $p = .241$	Republicans $x2 = 4.231$; $p = .121$
Top 4	33% (5/15)	64% (9/14)
Top 20	32% (8/25)	81% (17/21)
Others	50% (26/52)	56% (37/66)

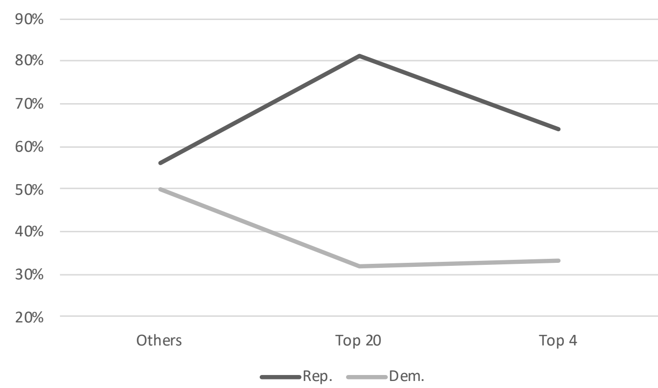
This was another crossover interaction where the opposing forces of the party variable had a cancellation effect with respect to the main variable. Put another way, there was evidence that eliteness matters, but it mattered differently—indeed, oppositionally—for the party subgroups.

A point worth emphasizing is the relatively modest difference between the removal rates of Republicans and Democrats who attended non-elite undergraduate institutions. Among that cohort, there was only a modest difference in removal rates:

146 The "Top 4" includes Yale, Harvard, Stanford, and Princeton Universities. The "Top 20" includes the other top twenty national universities and also the top twenty liberal arts colleges, according to U.S. News & World Report's 2020 College Rankings. *See Best National University Rankings*, U.S. NEWS & WORLD REP., https://www.usnews.com/best-colleges/rankings/national-universities (last visited May 1, 2020); *National Liberal Arts Colleges*, U.S. NEWS & WORLD REP., https://www.usnews.com/best-colleges/rankings/national-liberal-arts-colleges (last visited May 1, 2020). Our model thus assumes a certain linearity between the two steps of this hierarchy.

50% for Democrats and 56% for Republicans. But the difference in the respective parties' removal rates spread as each attended more elite institutions. This fact is illustrated well by the line graph in Figure 31, below.

Figure 31. Snap Removal Rate by Eliteness of Undergraduate Education (Clustered), with Party Effects (Line Graph)



As reported in Figure 30, the effect of eliteness of undergraduate education on removal was more pronounced for Republican judges; it was nearly statistically significant at the .10 level.

b. Law School

We also explored the legal education of each of the judges who authored opinions in our dataset. We then grouped the law schools into four categories of eliteness.[147] We were curious whether the eliteness of the judge's legal education had any predictive value in determining that judge's treatment of snap removal.

As indicated in Figure 32 below, there was an observable effect with removal rates declining with the increase in eliteness. Yet a main effect was not, by itself, statistically significant.

Figure 32. Snap Removal Rate by Eliteness of Legal Education (Clustered)

x2 = 3.873; p = .276

Top 3	41% (12/29)
Top 14	47% (20/43)
Top 50	54% (26/48)
Others	61% (44/73)

147 The "Top 3" includes Yale, Harvard, and Stanford. The "Top 14" and "Top 50" schools were ranked according to U.S. News & World Report's 2020 Law School Rankings. *See Best Law Schools*, U.S. NEWS & WORLD REP., https://www.usnews.com/best-graduate-schools/top-law-schools/law-rankings (last visited May 1, 2020). Our model thus assumes a certain linearity among the three steps of this hierarchy.

We investigated and discovered yet another interaction with party effect. The eliteness of a judge's education had a strong negative correlation with snap removal for Democrats, and a slight positive correlation for Republicans. This was another crossover interaction, as depicted in Figure 33, below.

Figure 33. Snap Removal Rate by Eliteness of Legal Education (Clustered), with Party Effects (Table)

	Democrats x2 = 8.528; p = .036**	Republicans x2 = .497; p = .919
Top 3	24% (4/17)	67% (8/12)
Top 14	30% (7/23)	65% (13/20)
Top 50	44% (11/25)	65% (15/23)
Others	63% (17/27)	59% (27/46)

The variable was statistically significant for the subgroup of Democratic judges at an impressive .05 level.

A line graph can effectively convey a crossover interaction, and Figure 34, below, does not disappoint.

Figure 34. Snap Removal Rate by Eliteness of Legal Education (Clustered), with Party Effects (Line Graph)

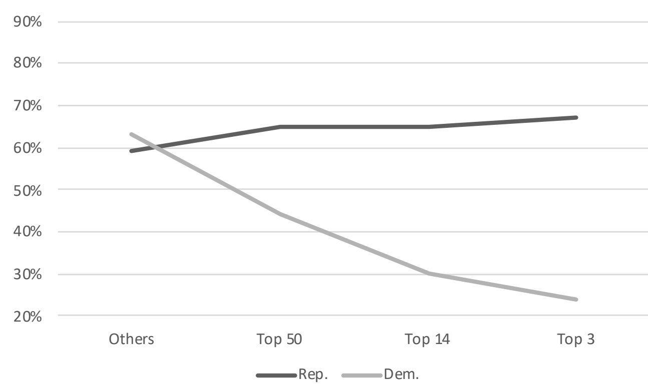


Figure 34 casts critically important light on nearly every finding in this paper. The party effects which we have discussed *ad nauseum* were, in fact, localized to judges (Republicans and Democrats) who attended elite institutions. Figure 34 demonstrates that the removal rates for Republican and Democratic judges who attended non-elite institutions were nearly the same—actually, the rate of removal was slightly *higher* among Democrats in that cohort.

The two eliteness-of-education variables sound similar notes. With both undergraduate education and legal education, eliteness increased the removal rate for Republicans and decreased it for Democrats. The undergraduate education variable was more statistically indicative for Republicans than Democrats, and the legal education variable was statistically significant for Democrats.

6. Subject Matter of Case

We also considered the theory that the characteristics of the case, rather than merely the judge, could be predictive of textualist outcomes. To be clear, the answer to the naked legal question presented by the forum defendant rule is not formally shaped by the subject matter of the case. Yet, we coded whether each case involved matters of tort or contract. We speculated that judges might systematically be more sympathetic to snap removal in cases where defendants were exposed to tort liability. The data confirmed that suspicion.

As an initial matter, we noticed that the *issue* of snap removal arose in a disproportionate number of torts cases. Well over 80% of the cases in our dataset were torts cases, with contracts cases accounting for the remaining 16.6%.[148] As a reference point, the percentage of torts cases among federal diversity cases generally for that same time period was consistently below 60% until the year 2003, below 70% until 2008, and has never reached the 80% plateau.[149] So, assuming our dataset is representative, defendants in torts cases were disproportionately more likely than defendants in contract cases to attempt a snap removal. This observation made our next finding all the more compelling.

Judges also allowed snap removal in a disproportionate number of torts cases. The numbers are reported in Figure 35, below.

Figure 35. Snap Removal Rate by Case Subject-Matter

x2 = 2.301; p = .129

Contracts	40% (13/32)
Torts	55% (89/161)

Even with this relatively small number of contracts cases, the correlation was nearly statistically significant at the .10 level. An odds ratio is also illustrative: the odds that a judge would remove a torts case were 1.81 times greater than the odds that a contract case would be removed.[150]

Of course we also tested for interactions with party effects. Both Democratic and Republican judges were more likely to remove torts cases. Data for the subgroups are presented in Figure 36.

148 Our categorization of torts cases includes a very small number of civil rights and discrimination suits. Our categorization of contracts cases includes a small number of breach of fiduciary duty suits, wrongful foreclosures, some employment matters (non-competes), and labor suits.

149 *See Table 4.9—U.S. District Courts–Civil Judicial Facts and Figures*, U.S. CTS. (Sept. 30, 2019), https://www.uscourts.gov/statistics/table/49/judicial-facts-and-figures/2019/09/30.

150 The odds for removal of a contract case are 13 / 19 = .684. The odds for removal of a torts case are 89 / 72 = 1.236. The ratio of these odds is 1.236 / .684 = 1.807.

Figure 36. Snap Removal Rate by Case Subject-Matter, with Party Effects

	Democrats	Republicans
	$x2 = .430; p = .512$	$x2 = 1.853; p = .173$
Contracts	35% (6/17)	47% (7/15)
Torts	44% (33/75)	65% (56/86)

The subject-matter effect was more pronounced among Republicans, but the p-value was stronger in the combined data (in Figure 35) than it was for either of the subgroups. This was attributable to the parsing of the already-small population of contracts cases. The combined data in the main effect, then, was the stronger statistical measure.

7. Efficiency Considerations

Finally, we contemplated that a judge's allowance of snap removal could be motivated in part by enthusiasm for consolidation of related cases in federal court. A case is not eligible for multidistrict litigation ("MDL") consolidation, for example, unless it is properly in federal court.[151] We hypothesized that the lure of efficiency and uniformity that is (correctly or incorrectly) associated with consolidation could influence the decision whether to allow snap removal.

With binary coding, we generously included within the category of MDL cases, all opinions that involved former, current, or prospective MDL matters.[152] With this classification protocol, fifty-nine (or 31%) of the opinions in the dataset were MDL cases. As a reference point, in recent years, MDL cases have constituted approximately 30% of the total civil caseload of the federal courts.[153]

An overwhelming majority—although not all—MDL consolidations involve torts, not contracts.[154] The prospect of collinearity between this MDL variable and the tort (subject-matter) variable deserved careful attention: if the MDL variable was highly

151 28 U.S.C. § 1407.

152 Indeed, any inter-district consolidation of matters was treated as an "MDL" case for our purposes. The availability of snap removal plays the same role with respect to the (real or perceived) gains of efficiency and uniformity through consolidated treatment of related matters.

153 *See Table S-19—Cases Transferred by Order of Judicial Panel on Multidistrict Litigation*, U.S. CTS. 1 (Sept. 30, 2019), https://www.uscourts.gov/sites/default/files/data_tables/jb_s19_0930.2019.pdf (reporting 134,462 as the total number of pending consolidated cases); *Table C—U.S. District Courts—Civil Cases Commenced, Terminated, and Pending*, U.S. CTS. 1 (Sept. 30, 2019), https://www.uscourts.gov/sites/default/files/data_tables/jb_c_0930.2019.pdf (reporting 357,566 as the total number of pending cases). *See generally* Linda S. Mullenix, *Reflections of a Recovering Aggregationist*, 15 NEV. L.J. 1455 (2015); Linda S. Mullenix, *Aggregate Litigation and the Death of Democratic Dispute Resolution*, 107 NW. U. L. REV. 511 (2013); Judith Resnik, *From Cases to Litigation,* 54 L. & CONTEMP. PROBS. 5 (1991); Thomas Metzloff, *The MDL Vortex Revisited*, 99 JUDICATURE 36 (2015).

154 *See Calendar Year Statistics January Through December 2019*, U.S. JUD. PANEL ON MULTIDISTRICT LITIG. 11, https://www.jpml.uscourts.gov/sites/jpml/files/JPML_Calendar_Year_Statistics-2019_1.pdf (last visited June 23, 2020) (reporting contract cases as constituting 2.1% of current caseload).

predictive of removal, then the prospect of MDL consolidation could have been exaggerating the effect of the tort variable discussed in Part III.B.6, *supra*.

But the MDL variable was not strongly correlated with the allowance of snap removal. In fact, as reported in Figure 37, below, an MDL case was slightly less likely to be removed than a non-MDL case.

Figure 37. Snap Removal Rate by Whether Subject to Consolidation

$x2 = .137; p = .712$

MDL	51% (30/59)
Non-MDL	54% (72/134)

Finding nothing statistically significant in MDL consolidation as a main variable, we turned our attention to interactions with our variables. Again we observed a crossover interaction with the party variable: the prospect of consolidation made Democrats a little more likely to remove and Republicans a little less likely to remove. This data is presented in Figure 38, below.

Figure 38. Snap Removal Rate by Whether Subject to Consolidation, with Party Effects (Table)

	Democrats	Republicans
	$x2 = .147; p = .702$	$x2 = .452; p = .501$
MDL	45% (14/31)	57% (16/28)
Non-MDL	41% (25/61)	64% (47/73)

Relative to their respective baselines, Democratic judges appear to have had the more sympathetic response to the instrumental arguments for snap removal. But even when divided into subgroups, these data did not approach statistical significance. Accordingly, we concluded that there was no evidence that the prospect of MDL consolidation influences a judge's decision to allow snap removal, and we do not further discuss this variable.

C. Logistic Regression Model

Our logit model is expressed in only four variables. The model reflects our balancing of the competing goals of formal parsimony, theoretical grounding, and predictive accuracy.[155] We make no pretense about the level of precision in this mode of inquiry. We are not under the illusion that judicial decision-making is a product of certain independent variables that combine to produce a dependent variable. But we are interested in general trends, defensible associations, and demonstrable correlations.

As an initial matter, it is important to appreciate that regression models predict the past rather than the future. Accordingly, only creativity and dubious ethics are necessary to produce a model with 100% accuracy. Cynics have described linear

155 *See generally* DAVID W. HOSMER ET AL., APPLIED LOGISTIC REGRESSION 89 (2013) ("Successful modeling of a complex data set is part science, part statistical methods, and part experience and common sense.").

regression as "torturing" the data until it "confesses."[156] The constraining principle on sound model design is to create a model that is also plausibly predictive of the future. Paradoxically, overfitting the model to predict the past compromises the model's capacity to predict the future, and vice versa. We preferred a forward-focused model, and we look forward to validating it with data in upcoming years. In the meantime, the ability to offer a theoretical explanation for each variable is a meaningful substitute.

Three of the four variables in our integrated model are main-effect variables that were strongly correlated with removal: Female, Young, and Tort.[157] The fourth is the interaction variable that included Democratic judges and the eliteness of their legal education.[158] A logistic regression of these variables produced the results in Figure 39, below.

Figure 39. Logistic Regression of Removal, Results

LR $chi^2(4) = 26.04$; Prob > $chi^2 = .000$***

	Coef.	Std. Err.	P>|z|
Female	.711	.385	.064*
Young (≤55)	.965	.420	.022**
Tort	.641	.433	.139
Elite LS + Democrat	- .498	.122	.000***
_cons	-.228	.442	.606

Variables with positive coefficients are correlated with removal, and variables with negative coefficients are correlated with remand. The magnitude of each coefficient indicates its relative prominence in the model's predictions.

The coefficients in the table in Figure 39 can be used to make predictions about whether a case is likely to be removed.[159] Every prediction begins with the coefficient of the constant (_cons), which is -.228. When predicting the outcome of a case, add the coefficients for any and all of the variables that are triggered by that scenario. In other words, if the judge is female, add .711 to the constant. Or (And) if the judge is young, add (another) .965. Or (And) if the case is a tort, add (another) .641. (The interaction variable requires something slightly different, and we will discuss that shortly.) The coefficients are the log odds of removal. Log odds are not an intuitively

---

156 *See* Gordon Tullock, *A Comment on Daniel Klein's 'A Plea to Economists Who Favor Liberty*,*'* 27 E. ECON. J. 203, 205 (2001) (quoting Nobel-Prize-Winning Economist Ronald Coase as saying, "if you torture the data long enough it will confess").

157 As a reminder the p-values for female, young, and tort were respectively: .212, .115, and .129. *See supra* Figures 11, 17, 35. For a discussion of the importance of using higher significance levels as a screening criterion for initial variable selection, see Ruth M. Mickey & Sander Greenland, *A Study of the Impact of Confounder-Selection Criteria on Effect Estimation*, 129 AM. J. EPIDEMIOLOGY 125 (1989) (showing that the use of a more demanding level of significance has distortion effects).

158 The p-value for this variable was .036**.

159 The estimated coefficients for the independent variables represent the slope (i.e., rate of change) of a function of the dependent variable per unit of change in the independent variable.

meaningful number.[160] But log odds are the logarithm of the (more intuitive) odds, and one can convert log odds into odds by raising the exponential constant to the power of those odds, using e^x. And further still, if one prefers probabilities to odds, that requires only one more step: dividing odds by (1 + odds).[161]

It may seem that a variable for Republican judges is conspicuously absent from the list of variables since this Article has made clear that, as a first approximation, Republican judges are the removers. But notice what happens with respect to predictions in cases involving Republican judges in each of the following scenarios that trigger (only) one of the listed variables:

(1) The Republican judge is a woman:
-.228 + .711 = .483. $e^{.483}$ = odds = 1.621. (Probability of removal = 62%)

(2) The Republican judge is young:
-.228 + .965 = .737. $e^{.737}$ = odds = 2.090. (Probability of removal = 68%)

(3) The Republican judge is hearing a tort case:
-.228 + .641 = .413. $e^{.413}$ = odds = 1.511. (Probability of removal = 60%)

In all three of the above scenarios, the odds of removal are greater than 1 (and the probability greater than 50%). The model thus predicts removal for all three. In fact, the model always predicts removal by a Republican judge unless the scenario presented triggers none of the three listed variables:

(4) The Republican judge is a male, 65 years old, in a contract case:
-.228. $e^{-.228}$ = odds = .796. (Probability of removal = 44%)

Of course, the odds and probabilities get much higher if the scenario triggers not just one, but two (or, hypothetically, all three) of the listed variables. Imagine, for example:

(5) The Republican is a woman, 63 years old, in a tort case:
-.228 + .711 + .641 = 1.124. $e^{1.124}$ = odds = 3.077. (Probability of removal = 75%)

Even without a separate Republican variable, the correlation of removal with Republican judges was captured in the reference category (i.e., the constant). All

---

160 Readers familiar only with linear regression must appreciate that if we try to predict the probabilities of a categorical variable like removal we have a problem of non-linearity, specifically the floor at 0 and the ceiling of 1 inherent in probabilities. But when the explanatory variables predict the log odds we do not have this problem.

161 The math conveys a certain precision that can be misleading. It can be useful to think of the coefficients not as the number reported, but rather as a range that hovers above and below that number. The size of that range is described by multiples of the standard error, which is also reported in the results in Figure 39. The standard error reflects the fragility of the coefficient, with that fragility defined by characteristics of the dataset. Notice that the coefficient for the Tort variable, while positive, is only (.641 / .433 =) 1.5 standard errors away from drifting into negative numbers. By contrast, the Young variable is (.965 / .420 =) 2.3 standard errors away from zero.

Democratic judges ineluctably vacate the reference category by invoking at least the interaction variable (and perhaps other listed variables as well). Accordingly, the reference category essentially incorporates the relevant correlation between Republicans and removals; adding a Republican variable would have been superfluous.[162]

The fourth variable in the table is different than the other variables in Figure 39 in two respects. First, it is an interaction variable rather than a main effect variable. As discussed in Part III.B.5.b, the variable for eliteness of legal education was not statistically significant as a main-effect variable because Republican and Democratic judges were pulling it in opposite directions. Incorporated here, then, is only the Democratic subgroup as identified in Figure 33. Every Democratic judge always triggers this variable and no Republican judge ever triggers this variable.

Second, this variable, unlike the other three variables in this model, is not binary. Rather, as depicted in Figure 33, there are four levels of eliteness: outside the Top 50, Top 50, Top 14, and Top 3. For variables like this that are not binary, we add multiples of the coefficient when doing the calculations allowed by the model. Thus, for Democratic judges in the model, the coefficient of -.498 is multiplied by 1, 2, 3, or 4, respectively, depending on the level of eliteness of the judge's legal education. Consider these scenarios involving Democratic judges:

(6) The Democratic judge is a woman, 63 years old, Yale Law School grad, in a tort case:

---

162 To illustrate the point, adding a Republican variable to our model produces the following results.

Figure 39.1. Logistic Regression of Removal,
Results of Alternative Model 1
LR chi2(5) = 26.50; Prob > chi2 = .000***

	Coef.	Std. Err.	P>|z|
Female	.705	.387	.068*
Young (≤55)	.964	.424	.023**
Tort	.600	.439	.171
Elite LS + Democrat	- .624	.225	.006***
Republican	- .393	.583	.500
_cons	. 148	.711	.835

In this alternative model, the log odds of removal by a Republican who does not trigger any other variable is .148 + (- .393) = - .245. e^{-245} = .783. (Probability of removal = 44%). Hence, it is the same result as in hypothetical (4) in the main text. This alternative model in Figure 39.1 predicts the dataset with 64.3% accuracy, with true positive and true negative rates of 78% and 48%, respectively.

-.228 + .711 +.641 + (4 * -.498) = -.868. $e^{-.868}$ = odds = .420. (Probability of removal = 30%)

(7) The Democratic judge is a male, 51 years old, Marquette grad, in a tort case: -.228 + .965 +.641 + (1 * -.498) = .880. $e^{.880}$ = odds = 2.41. (Probability of removal = 71%)

Throughout this Article we have implied that, to a first approximation, Democratic judges remand. That inclination is reflected in this model too, but remember that Democratic judges who attended non-elite law schools had removal rates of 63%,[163] which is also the removal rate for the typical Republican judge. This model's predictions are highly sensitive to the eliteness of the Democratic judge's legal education. For scenarios that trigger two of the main-effect variables (say, Tort and Young), the model predicts removal by a Democratic judge unless that judge has at least 3 (of the possible 4) levels of eliteness.[164] For scenarios that trigger all three of the main-effect variables, the model predicts removal by a Democratic judge even with all 4 levels of eliteness.[165]

Figure 40, below, presents a classification table for this model. It is a scorecard of the model's ability to predict the outcomes of all cases in the dataset—i.e., to predict the past. The model is accurate in 64.3% of the cases.[166]

Figure 40. Classification Table

		Observed	
		Removed	Remanded
Predicted	Removed	72 (71%)	39 (43%)
	Remanded	30 (29%)	52 (57%)
			= 64.3%

The upper-left quadrant and the lower-right quadrant represent the true positives and true negatives, respectively.[167] The model is better at predicting removals than at predicting remands. The upper-right quadrant reports the false positives ("Type I errors") and the lower-left quadrant the rate of false negatives ("Type II errors"). One might fairly criticize this model as erring on the side of predicting removals, rather than risk missing some. Erring on the side of over-predicting removals was a deliberate

163 *See supra* Figure 33.

164 The math for this scenario follows. -.228 + .965 + .641 + (3 * -.498) = -.116. $e^{-.116}$ = odds = .890. Probability of removal: 47%.

165 The math for this scenario follows. -.228 + .711 + .965 +.641 + (4 * -.498) = .097. $e^{.097}$ = odds = 1.102. Probability of removal: 52%.

166 As a reference point, a model based exclusively on party effect would generate a 60.1% level of accuracy. All models that we tested were an effort to improve on this baseline—while also maintaining our conservative commitments.

167 The true positive and true negative rates are often referred to as sensitivity and specificity measures, respectively.

decision. We wanted ballast in the model to make it more durable for future predictions. We have speculated that there will be more polarization and also more removals.[168] We also expect the Young judge variable to become less potent as judges

168 Our prediction of more removals is a function also of President Trump's appointments to the district court bench. As of June 27, 2020, there were 143 district judges who had been approved by the Senate. Figure 40.1 suggests that there is more gender diversity in this pool of new appointees relative to the judges in the dataset. Because gender (female) is positively correlated with removal (especially among Republican women, see Figure 12), this supports a prediction of more removals in the short term.

Figure 40.1. Gender Diversity Among Judges

		Judges in Dataset	
	Trump Appointees	Republicans	All
Female	25.9% (37)	17.8% (18)	23.8% (46)
Male	74.1% (106)	82.2% (83)	76.2% (147)

Figure 40.2 suggests that there is less racial diversity in the pool of new appointees relative to the judges in the dataset. Because race (non-white) was a disordinal variable that was negatively correlated with Republican removal rates and positive correlated with Democratic removal rates (see Figure 15), the lack of diversity among new appointees supports a prediction of more removals in the short term.

Figure 40.2. Racial Diversity Among Judges

		Judges in Dataset	
	Trump Appointees	Republicans	All
Non-White	12.6% (18)	17.8% (18)	22.3% (43)
White	87.4% (125)	82.2% (83)	77.7% (150)

Figure 40.3 suggests that the pool of new appointees attended substantially less elite undergraduate institutions relative to the judges in the dataset. Because Republican judges who attended non-elite undergraduate institutions had substantially lower removal rates than Republican judges who attended more elite institutions (see Figure 30), this datum does not support a prediction of more removals in the short term.

outgrow the age-restricted group, but do not outgrow their legal education and other factors that made that age group more inclined to remove relative to their elders.

The classification table in Figure 40 is not fine-grained to distinguish the accuracy of a prediction in light of its strength. In other words, predictions of removal are treated the same, whether the prediction was removal by a hair's breadth or removal by a mile. But getting both of those predictions right (or wrong) is not the same achievement (or problem). Goodness-of-fit tests, broadly speaking, assess whether the model's strongest predictions prove correct and the model's errors are clustered within the predicted close-calls.[169] These tests produce a p-value to gauge statistical significance, but it is the *null hypothesis* that is consistent with a well-specified distribution: an assumed model of independence is evaluated against the observed data. Put another way, when the p-value is statistically significant (at .05) the model does not fit the data.

Figure 40.3. Eliteness of Undergraduate Education Among Judges

		Judges in Dataset	
	Trump Appointees	Republicans	All
Top 4	2.0% (3)	13.9% (14)	15.0% (29)
Top 20	18.2% (26)	20.8% (21)	23.8% (46)
Others	79.8% (114)	65.3% (66)	61.1% (118)

Figure 40.4 suggests that the pool of new appointees attended law schools in proportions that approximate the proportions of judges in the dataset. This datum would suggest that the removal rate should be relatively stable. (See Figure 33.)

Figure 40.4. Eliteness of Legal Education Among Judges

		Judges in Dataset	
	Trump Appointees	Republicans	All
Top 3	11.2% (16)	11.9% (12)	15.0% (29)
Top 14	13.3% (19)	19.8% (20)	22.3% (43)
Top 50	34.3% (49)	22.8% (23)	24.9% (48)
Others	41.3% (59)	45.5% (46)	37.8% (73)

169 In linear regression, summary measures of fit are functions of a residual defined as the difference between the observed and fitted value. In logistic regression there are several possible ways to measure the difference between the observed and fitted values.

Our model fit the data very well. Using a Pearson Goodness-of-Fit Test, our significance level, with 28 degrees of freedom, was .907.

We were stubbornly uncompromising and conservative in conducting this study. All of the variables mentioned in this Article were identified before we coded our first case. We also did not include in our model geographical variables which could be statistically powerful predictors (and which we coded from the outset). By adding the state where the opinion was issued, the model could have leveraged that variable in states where the removal rate was especially high or low. In Louisiana, for example (where all eight cases were removed[170]), the model could have essentially encoded any case from Louisiana as a surefire removal, and there would be eight correct predictions in the model. To be sure, there is an argument for including geographic variables: there is a local legal culture in every state, and that culture might be a greater influence on case outcomes than variables that we include. Yet, we thought it antithetical to our goal of creating a model that could be predictive of the future as well as the past, and creating a model that is generalizable rather than being localized to a particular region.

Our final model did not include all of the consequential variables discussed in Part III.B.[171] We also considered adding an interaction variable for Republicans and eliteness of undergraduate education, for example, but this added complexity to the model without improving its predictions.[172]

---

170 *See supra* Figure 3.

171 *See generally* HOSMER ET AL., *supra* note 155, at 90 ("The rationale for minimizing the number of variables in the model is that the resultant model is more likely to be numerically stable, and is more easily adopted for use. The more variables included in a model, the greater the estimated standard errors become, and the more dependent the model becomes on the observed data.")

172 Adding an interaction variable for Republicans and eliteness of undergraduate education produced the following results.

Figure 40.5. Logistic Regression of Removal,
Results of Alternative Model 2
LR chi2(5) = 26.36; Prob > chi2 = .000***

	Coef.	Std. Err.	P>|z|
Female	.722	.385	.060*
Young (≤55)	.958	.420	.022**
Tort	.661	.435	.128
Elite LS + Democrat	- .436	.163	.007***
Elite UG + Republican	- .133	.238	.575
_cons	-. 415	.553	.453

This model predicts the dataset with 64.3% accuracy, with true positive and true negative rates of 71% and 57%, respectively.

IV. Conclusion

Permitting or rejecting snap removal under the current statute requires judges to make a choice between a strong brand of textualism and a strong brand of purposivism. Both options are popular, but textualism is winning. We thus posed and answered the question: Who are these textualists?

We analyzed data which showed that most of them were judges who were appointed by Republicans.[173] The odds of removal by a Republican judge were 2.25 times the odds of removal by a Democratic judge. Further, when answering a pure legal question about the availability of a federal forum, Republican judges were especially sympathetic of defendants facing tort liability. Further still, the publication rates of both parties suggested deliberate efforts to spread each party's competing ideology.

But party was hardly a complete explanation. Indeed, judges appointed by Democrats allowed almost 40% of all snap removals.[174] And Republicans were responsible for more than 40% of all remands.[175] Moreover, young judges of both parties were more inclined to remove than their senior colleagues. This finding aligned with other scholarship that has recognized the ascendance of a more constrained and technical understanding of the judicial process.[176] Such is the consequence of the last three decades' march toward a more formalist application of rules and statutes.

We speculated that party may be a better predictor of case outcomes in the future. This is a likely consequence if a recent trend toward party polarization in removal rates continues. This trend was apparent in party removal rates in the past 5 years compared to the decade that preceded it. We also demonstrated that the judges appointed by each

---

173 *See supra* Figure 8.

174 Judges appointed by Democratic Presidents allowed 39 of the 102 snap removals.

175 Judges appointed by Republican Presidents remanded 38 of the 91 remands.

176 *See* Gluck & Posner, *supra* note 1, at 1302–03. *See generally* Anton Metlitsky, *The Roberts Court and the New Textualism*, 38 CARDOZO L. REV. 671, 688 (2016) ("[T]he Court's insistence on the modern textualist method in every case provides a clear signal to lower courts--and especially the intermediate federal appellate courts--that they, too, must follow the modern textualist approach to reading statutes, meaning that they must give effect to clear text unless they can plausibly justify labeling the text to be "ambiguous." . . . [T]he courts of appeals are also likely to understand that open purposivism is verboten, which will likely mean fewer cases at the intermediate appellate level that depart from clear text in favor of statutory purpose."); Krishnakumar, *supra* note 2, at 1277 (challenging view that purposivism is dead, but also noting: "In the thirty-some years since the late Justice Scalia joined the U.S. Supreme Court and began waving the textualist flag, it has become in vogue to chronicle the Court's move toward a highly textualist approach to statutory interpretation. Scholars have, for example, noted a discernible decline in the rate at which the Court invokes legislative history, a marked increase in its use of dictionary definitions to interpret statutes, and a rise in its use of both linguistic and substantive canons of construction."); Lawrence Baum & James J. Brudney, *Two Roads Diverged: Statutory Interpretation by the Circuit Courts and Supreme Court in the Same Cases*, 88 FORDHAM L. REV. 823, 840 (2019) ("The circuit courts' lower frequency of reliance on purpose (like their higher reliance on ordinary meaning and agency deference) may reflect attention to case-management priorities.").

party's presidents were more extreme (in both directions) than the appointments by their predecessors in those parties.

At the same time, however, we observed that party polarization was a product exclusively of the elites within each of the parties. Increasing eliteness levels in education were positively correlated with removals for Republicans and were negatively correlated for Democrats. But Republicans and Democrats who attended non-elite educational institutions had the same removal rates. Because party effects are a staple of empirical study, this finding about the intersection of party effects and education should force a reassessment of the role of political affiliation in judging.

Women—especially Republicans—were more likely than men to reach textualist outcomes. Perhaps this finding was an anomaly of the relatively small sample size, but it registered as statistically significant and the inclusion of a gender variable improved the predictive capacity of all models we investigated.

Our study included many other variables that bore no correlation with removal rates. Of course, the lack of a finding is occasionally as telling and profound as the presence of significance. Variables that were statistically unimportant included race, the prospect of MDL consolidation, advanced age, advanced experience, senior status, and judicial experience on a state court.

In conclusion, we hope that this Article inspires further inquiry into the interaction between eliteness of education and ideology. Party effects are a staple of empirical study and, if our finding is typical of a broader phenomenon, then the insight that we have revealed should force a reassessment of the role of political affiliation in judging. Future studies could also explore whether elite institutions cause the polarization or, instead, tend to attract more students who are dogmatic.[177]

---

[177] Put another way, non-elite institutions might cause depolarization or, instead, attract more students who are pragmatic.