2021

Politics, Identity, and Pleading Decisions on the U.S. Courts of Appeals

Stephen B. Burbank
University of Pennsylvania Law School

Sean Farhang
University of California - Berkeley

Repository Citation

ARTICLE

POLITICS, IDENTITY, AND PLEADING DECISIONS ON THE U.S. COURTS OF APPEALS

STEPHEN B. BURBANK & SEAN FARHANG[†]

We report the results of an empirical study of appeals from rulings on motions to dismiss for failure to state a claim under Federal Rule of Civil Procedure 12(b)(6) after the Supreme Court's decisions in Twombly and Iqbal. We first describe the role that pleading was intended to play in the original (1938) Federal Rules of Civil Procedure, review the Court's decisions in Twombly and Iqbal, and offer a brief discussion of common themes in normative scholarship that is critical of Twombly and Iqbal, including the claim that they threaten to amplify ideological and subjective decision-making, particularly in civil rights cases.

We then empirically examine the extent to which the party (of appointing president), gender, and racial composition of panels are associated with their disposition of 12(b)(6) appeals across all policy areas pooled, also separately analyzing discrimination claims, all "other civil rights" claims, and non-civil rights claims. We separately analyze a random sample of (predominantly non-precedential) cases and a set of only precedential cases.

In our random sample of cases, we find that panels with women and non-white judges are substantially more likely to rule in favor of a plaintiff reaching discovery in other civil rights claims, an important and cross-cutting civil rights category amounting to a quarter of 12(b)(6) appeals in our data, but that race and gender are insignificant outside that substantive area. Party is insignificant across the board in the random sample.

The results are different when the panel is making law. In precedential cases, we find that Democratic panels were significantly more likely to decide in favor of plaintiffs in non-civil rights claims. We also find that panels with one woman were more likely to decide precedential other civil rights claims in favor of plaintiffs, and that panels with two women (but not one) were more likely to do so in non-civil rights claims.

Our results for gender contradict conventional wisdom in the literature that women judges' preferences differ from men's only in cases implicating discrimination. They add to evidence suggesting the possibility that procedural law affecting access to justice may itself be a policy domain in which women have different (more pro-access) preferences that extend beyond discrimination claims. Gender, alone among the judge characteristics we study, is significant in both random sample and precedential-only models, and in both civil rights and non-civil rights models, revealing a distinctive propensity among women on the Courts of Appeal to support plaintiffs' access to discovery.

Finally, significant variation in our results across the random sample and precedential cases highlights the risk of error in drawing general inferences from either significant or null results in precedential cases—general inferences that are widespread in the literature on the Courts of Appeals.

INTRODUCTION

In recent work we have sought to illuminate the extent to which federal court decisions in the realm of procedure reflect non-legal influences on judicial behavior that have been shown to affect decisions about substantive law in policy areas of high salience. We started by studying the decisions of the Supreme Court on issues implicating private enforcement of federal law, such as standing, attorney's fees, arbitration of federal claims, and the interpretation of pertinent Federal Rules of Civil Procedure.[1] We found that, as a group, these decisions are more influenced by ideology than are the Court's decisions on merits issues, and that the Court's Federal Rules decisions are more influenced by ideology than either.[2]

Motivated by these findings to extend the scope of the inquiry beyond the Supreme Court and beyond ideology, we compiled an original comprehensive data set that includes precedential Court of Appeals decisions on issues of class certification under Rule 23 from 1967 through 2017, together with nonprecedential decisions since 2002. In the first article based on these data, we explored the roles of ideology, race and gender in class certification decisions.[3] We found that, at the Court of Appeals level, ideology is strongly associated with class certification decisions, playing a role akin to that found

---

[1] *See* STEPHEN B. BURBANK & SEAN FARHANG, RIGHTS AND RETRENCHMENT: THE COUNTERREVOLUTION AGAINST FEDERAL LITIGATION 130-91 (2017). For a full description of the data set, see *id*. at 143-44. *See also* Stephen B. Burbank & Sean Farhang, *Class Actions and the Counterrevolution against Federal Litigation*, 165 U. PA. L. REV. 1465 (2017).

[2] *See* BURBANK & FARHANG, *supra* note 1, at 180-81.

[3] *See* Stephen B. Burbank & Sean Farhang, *Politics, Identity, and Class Certification on the U.S. Courts of Appeals*, 119 MICH. L. REV. 231 (2020) (hereinafter *Politics & Identity*). A second article using this data set has recently been published. *See* Stephen B. Burbank & Sean Farhang, *Class Certification in the U.S. Courts of Appeals: A Longitudinal Study*, 84 LAW & CONTEMP. PROBS. 73 (2021).

in cases raising some of the most controversial substantive law issues of the day.[4] To those who understand the power of procedure in general and the catalytic power of the class certification decision in particular, and who have followed the increasing role that ideology has played in Court of Appeals appointments in recent decades, this was not a surprise.

Reflecting on the centrality of the class certification decision to court access and the importance of class actions in the struggles for racial and gender equality caused us to study whether identity characteristics such as race and gender may also play a role in class certification decisions. Here, unlike our investigations of the role of ideology, we were not merely extending prior panel-effects studies of judicial behavior, which have neglected procedural decisions. We were working in the context of an emerging consensus that Court of Appeals "judges' gender and race are associated with variation in preferences only in a narrow band of cases presenting issues of substantive law that directly and explicitly implicate discrimination and inequality."[5] We found that the presence of one African American on a panel, and the presence of two women (but not one), is associated with pro-certification outcomes even in cases not involving civil rights claims. For many, we think, this was a surprise.

Although neither our data nor prior panel-effects scholarship enabled us to identify the reasons for the different preferences of women and African Americans as to class certification, we offered several suggestions to guide further study:

> As transsubstantive procedural law, the Federal Rules of Civil Procedure apply across substantive domains and can enable or constrict access to justice. A controlling interpretation of a Federal Rule in an antitrust case, for example, will carry over into its application in a voting rights case. One important insight of this Article is that the transsubstantive nature of the Federal Rules can also convey the substantive effects of diversity across the landscape of American regulatory law. Court of Appeals judges understand that the Federal Rules are transsubstantive, as are the effects of some Federal Rules (importantly including Rule 23) on the enforcement of substantive law. As strategic actors, it would be rational for them to take into consideration how class-certification doctrine in a case that does not implicate issues on which they have strong preferences might affect certification in cases that do. Alternatively, or in addition, our results may be the first evidence that

---

4 *See* Burbank & Farhang, *Politics & Identity*, *supra* note 3, at 258 (Table 4). "We suspect that many will be surprised that the outcome changes associated with moving from unified Democratic to Republican panels in certification decisions is larger than, for example, such changes in obscenity, capital punishment, employment discrimination, desegregation, and abortion cases." *Id*. at 257.

5 *Id.* at 236.

transsubstantive procedural law affecting access to justice is itself a policy domain in which women and African Americans have distinctive preferences.[6]

In this article, we report the first results of a study designed to gain additional insight on the extent to which ideology, race and gender are associated with judges' decisions on procedural issues critical to court access. For this purpose, we chose to study federal appeals challenging rulings on motions to dismiss under Federal Rule of Civil Procedure 12(b)(6) after the Supreme Court's decisions in *Bell Atlantic Corp.* v *Twombly*[7] and *Ashcroft* v. *Iqbal*.[8] We created two original data sets, one consisting of precedential (published) decisions after *Iqbal* was decided in 2009 through 2019, and the other a random sample of decisions in the same period, a substantial majority of which are nonprecedential (unpublished) decisions.

In Section I we first describe the role that pleading was intended to play in the original (1938) Federal Rules of Civil Procedure, stressing the views of the original Advisory Committee that (1) pleading was an inferior means to uncover the facts grounding a plaintiff's claims, a role better left to discovery, and (2) judicial gate-keeping at the pleading stage often arbitrarily and unfairly deprived plaintiffs of access to information that was available only through discovery.

We then discuss strategies that emerged in response to an enormous increase in federal litigation beginning in the late 1960s, which sought more effectively to curb wasteful litigation behavior and identify and weed out claims lacking sufficient support to warrant a trial (or elicit a settlement). One such strategy was to restore fact pleading by amending the Federal Rules, but it never gained traction, probably because the Advisory Committee recognized that the effort would embroil the rulemaking process in political controversy, putting at risk the major source of the judiciary's control of procedure by legislative intervention in that domain.

Although the rulemaking process *was* used repeatedly to rein in the costs of discovery, the Supreme Court invoked discovery's costs as a reason to change the requirements that the pleading rules impose. We describe the Court's decisions in *Twombly* and *Iqbal*, stressing that the Court's reinterpretations of the pertinent rules introduced a new gatekeeping strategy, rooted in denying a presumption of truth to allegations deemed to be conclusory, and dismissing claims deemed to be implausible based on "judicial experience and common sense." We conclude Section I with a brief

---

[6] *Id.* at 238-39.

[7] 550 U.S. 544 (2007).

[8] 556 U.S. 662 (2009).

discussion of common themes in normative scholarship that is critical of *Twombly* and *Iqbal*, including the claim that they threaten to amplify subjective decision-making.

In Section II, we review the empirical literature on the association between outcomes and Court of Appeals judges' party (of appointing president), gender and race. That literature has overwhelmingly ignored transsubstantive procedural law. Empirical results in the literature are something of a patchwork, with scholars sometimes detecting significant relationships where they expected to find them, but sometimes not, in ways that are difficult to explain based upon general theories about salience or preferences. Still, an apparent consensus has emerged that gender and race are associated with voting only on some types of claims based on discrimination or inequality.

Finally, in Section III we describe our data and results. Critics of *Twombly* and *Iqbal* worry that the new 12(b)(6) standard introduces excessive subjectivity and ideology into disposition of 12(b)(6) motions, with a particular concern about civil rights cases. Although we do not compare pre- and post-*Iqbal* decision-making, we examine the extent to which the party, gender and race of panel members are associated with their disposition of 12(b)(6) appeals in cases brought by individuals against business or government since *Iqbal*. In addition to analyzing all policy areas pooled, we separately analyze discrimination claims, all "other civil rights" claims, and non-civil rights claims. We also separately analyze a random sample of (predominantly non-precedential) cases, and a set of only precedential cases. Our results vary across identity characteristic, policy area, and random sample versus precedential cases. The patchwork character of our results reflects the judicial behavior literature that we contribute to.

In our random sample of cases, we find that judges' gender and race are associated with outcomes in other civil rights claims (excluding discrimination). This broad and varied civil rights category amounts to one in four 12(b)(6) appeals in our data, overwhelmingly made up of constitutional claims against governmental actors, commonly arising in such areas as policing, prisons, and public employment. Panels with one woman or one non-white judge have more than double the likelihood of rendering a decision in favor of the plaintiff as compared to all-male and all-white panels. In the same models, party of appointing president is clearly insignificant. We are aware of no prior Court of Appeals study to find that gender and race are associated with outcomes when party is insignificant. Gender and race are not simply amplifying ideology as measured by party; they are consequential where party is not. With eighty-six percent of claims in the data decided by panels with some degree of gender or racial diversity, diversity on the Courts

of Appeals has a broad and favorable impact on plaintiffs' ability to reach discovery in other civil rights claims. The remaining judge characteristic variables in the other random sample models—all policy areas pooled, discrimination claims, and non-civil rights claims—are insignificant.

However, when panels are in the posture of making law, the results are different. In precedential cases we find that Democratic panels were significantly more likely to rule for plaintiffs when all non-civil rights claims are pooled. All-Democratic panels are more than twice as likely to allow the plaintiff to proceed to discovery when compared to Republican-majority ones. This set of cases asserting non-civil rights claims, pooled across many policy areas, amounts to about half the precedential cases. The most common are consumer, contract, labor, personal injury, antitrust, and securities (in that order).

Party remains insignificant, however, in precedential discrimination and other civil rights claims, as it was in the random sample. Contrary to expectations, party appears to matter least (or not at all) to pleading decisions in the policy area where many scholars (including us) thought it would matter most: civil rights. Although many studies of the Courts of Appeals have found party consequential to votes and outcomes in civil rights cases, we find that this is not so when the question is narrowed to whether a plaintiff has stated a claim sufficient to proceed to discovery.

Gender is significant across more policy areas in precedential cases. Panels with one woman were more likely to decide for the plaintiff in other civil rights claims, although they were not more likely to do so in discrimination claims. Panels with women in the majority, but not panels with one woman, were more likely to decide for the plaintiff in non-civil rights claims—more than twice as likely. One woman panel-affects (changes the votes of) male majorities in civil rights but not non-civil rights claims. These results contradict conventional wisdom in the literature that judge gender on the Courts of Appeals is only consequential in certain discrimination cases. They add further evidence, along with our class certification study discussed below, to the conclusion that women have more pro-access preferences that are more broad-ranging. Among the three judge characteristics that we study, only gender is significant in both random sample and precedential-only models, and in both civil rights and non-civil rights models. Gender is distinctively associated with panels' propensity to allow plaintiffs to proceed to discovery.

Race is insignificant in all precedential models, including other civil rights claims, an area in which it was highly significant with a substantively large association in the (mostly non-precedential) random sample. This highlights an unexpected and important lesson. A judge characteristic may not be significantly associated with outcomes in precedential cases—which are the basis of nearly all Court of Appeal studies—while a significant association

actually exists in the full universe. One potential explanation is that white majorities are less likely to give precedential status to decisions in which they make concessions to (are panel affected by) non-white judges. These findings show that null results for gender and race in precedential cases cannot support the inference that gender and race are not consequential to outcomes in the full universe—an inference widespread in the literature.

I. PLEADING, DISCOVERY AND ACCESS TO COURT UNDER THE FEDERAL RULES OF CIVIL PROCEDURE

A. *The Original Understanding*

Law prescribing how detailed and persuasive the complaint commencing a lawsuit must be has important implications for the ability of potential plaintiffs to pursue adjudication of disputes on the merits, including their ability to discover relevant information from defendants in order to prove their allegations. It thus also has important implications for the ability of those who have been injured to use litigation in order to secure compensation, and the ability of government to use private litigation for the enforcement of public law.

From the perspective of potential defendants, pleading law affects the ease with which they can be summoned to court and forced to incur costs in defending against, or settling, what may be meritless claims. Finally, from the (self-interested) perspective of the judiciary, pleading law affects the volume of civil litigation and the types of litigation activity that filed cases produce, both of which affect the allocation of resources by court systems.

The original (1938) Federal Rules of Civil Procedure were designed to facilitate litigation on the merits, including litigation enforcing public law. In this they reflected the jurisprudential and social commitments of the individuals who were responsible for drafting them. The way that those individuals approached pleading and discovery made these procedural features critical pillars of the regime they created, and it is thus not surprising that they have been important sites of contestation.

The Advisory Committee responsible for drafting the 1938 Federal Rules designed pleading rules that were simple and flexible. The Committee was in part reacting to the existence in many states of pleading law—applicable in federal courts in those states—that required the plaintiff to allege facts supporting each cause of action relied on. The Committee objected to this type of "fact pleading" because it entailed arbitrary distinctions among, and wasteful disputes about, "facts," "conclusions," and "evidence."[9] More

---

[9] As Edgar Tolman, who bore major responsibility for explaining the proposed new Federal Rules to Congress, put it in his 1938 House testimony:

fundamentally, vast changes in social and economic life since the mid-nineteenth century (when fact pleading was introduced) had made it harder for many people suffering injuries—especially those without resources to conduct an extensive pre-filing investigation – to know what the facts were. The drafters believed that pleading is an inferior method to find out what actually happened.[10]

The original Advisory Committee opted instead for "notice pleading," under which a plaintiff's complaint was at risk of dismissal (for merits-related reasons) only if, under Rule 12(b)(6), it failed to state a claim that was legally tenable or if, under Rule 12(e), it failed to give the defendant fair notice of what that claim was. Federal Rule 8 required that a complaint include only "a short and plain statement of the claim showing that the pleader is entitled to relief." In discussing one of the forms that the Committee created to demonstrate acceptable pleading practice under Rule 8 at an Institute for members of the bar, Dean Charles Clark, the Committee's Reporter, observed:

> [A]n allegation which says simply that the defendant did injure the plaintiff through his negligence is too general and would not stand, for really that tells you no differentiating features about the case whatsoever, except the very broad word "negligence"; while on the other hand . . . the statement of the act in question in a general way, and with a characterization that it is negligent, is sufficient. That is the allegation in this form (Form 9). Here, instead of saying defendant's negligence caused the injury, you say that defendant negligently drove his automobile against the plaintiff, who was then crossing the street, and you have then the case isolated from every other

---

> I want you now to consider this provision in Rule 8, as to what you have to put into your paper. You used to have the requirement that a complaint must allege the "facts" constituting the "cause of action." I can show you thousands of cases that have gone wrong on dialectical, psychological, and technical argument as to whether a pleading contained a "cause of action"; and of whether certain allegations were allegations of "fact" or were "conclusions of law" or were merely "evidentiary" as distinguished from "ultimate" facts.

Rules of Civil Procedure for the District Courts of the United States: Hearings Before the H. Comm. on the Judiciary, 75th Cong. 94 (1938) (statement of Edgar B. Tolman, Secretary of the Advisory Committee on Rules for Civil Procedure Appointed by the Supreme Court).

[10] Again, Tolman explained:

> One important consideration should be emphasized as to the method by which, under these rules, the opponents may be adequately advised as to the real matter in controversy. The simplified pleadings provided for . . . which give a general view of the controversy are supplemented by the provisions for depositions, discovery and pretrial practice . . . which enable each side by the examination of witnesses, documents, and other evidence, to ascertain in advance of the trial, precise knowledge as to the nature of the case.

Id. at 98.

type of case of the same character, really from every other case, as a pedestrian or collision case. At the pleading stage, in advance of the evidence, before the parties know how the case is going to shape up, that is all, in all fairness, you can require.[11]

Implementing the view that pleading should play a minor role in litigation—and that common law trials were inefficient because the parties often were in the dark about the issues—required other means to ascertain facts prior to trial. To that end the architects of the 1938 Federal Rules wrote rules that afforded parties pre-trial authority to demand information from other parties (and non-parties) much greater than had been available under prior systems.[12] Such broad discovery appealed to the commitments of the Progressive movement in American law,[13] of which Edson Sunderland, the chief architect of the Federal Rules on discovery, had long been a proponent. Progressives contended that effective regulation was impossible without access to the facts concerning the regulated enterprise. As Sunderland wrote in 1925:

> The spirit of the times calls for disclosure, not concealment, in every field—in business dealings, in governmental activities, in international relations, and the experience of England makes it clear that the courts need no longer permit litigating parties to raid one another from ambush.[14]

Eliminating the gatekeeping role of fact pleading required some other means to prevent the trial of claims that lacked evidentiary support after adequate opportunity for discovery. For this purpose, with Sunderland again taking the lead, the original Advisory Committee drew on experience in England. They made available to both plaintiffs and defendants, and in all cases, what in England had been used primarily to enable plaintiffs to collect debts: the motion for summary judgment.[15] Discussing this new tool, for

---

[11] AMERICAN BAR ASSOCIATION, FEDERAL RULES OF CIVIL PROCEDURE: PROCEEDINGS OF THE INSTITUTE AT WASHINGTON, D.C. AND OF SYMPOSIUM AT NEW YORK CITY 241 (1938). *See also id.* at 308 ("What these rules do emphasize with respect to the contents of a pleading (as the forms in the Appendix show) is that any plain telling of the story that shows that the pleader is entitled to relief upon the grounds he states is sufficient to bring the pleader's cause into court. That the statement or averment includes a conclusion of law is no ground for a motion to strike or for a motion to make definite, merely because the statement or averment embodies a conclusion which might be elaborated by a more particularized detailing of the facts.") (George Donworth).

[12] *See supra* note 10.

[13] *See* Ken I. Kersch, *The Reconstruction of Constitutional Privacy Rights and the New American State*, in 16 STUDIES IN AMERICAN POLITICAL DEVELOPMENT 61, 84 (2002).

[14] Edson R. Sunderland, *An Appraisal of English Procedure*, 24 MICH. L. REV. 109, 116 (1925).

[15] *See* Stephen B. Burbank, *Vanishing Trials and Summary Judgment in Federal Civil Cases: Drifting from Bethlehem or Gomorrah?*, 1 J. EMP. LEG. STUD. 591, 594-603 (2004). As discussed there,

which the Advisory Committee had no domestic empirical evidence and about which many members harbored doubts, including doubts arising under the Seventh Amendment,[16] the Reporter told members of the bar:

> The great question about the motion for summary judgment is whether it may not be attempted in all sorts of cases, whereas it is only really going to perform its function in the simple case where there isn't much of a defense. It is quite possible that the motion . . . may be resorted to too much and may become an instrument of delay.[17]

B. *The Road to Retrenchment*

In the years following the advent of the Federal Rules, a number of Supreme Court decisions, including *Hickman v. Taylor*[18] and *Conley v. Gibson*,[19] embraced the concepts of notice pleading and broad discovery, while others seemed to constrain the ability of summary judgment to separate wheat from chaff.[20] Eventually, however, notice pleading, broad discovery (unleashed further by amendments to the Federal Rules in 1970), and a restrictive view of summary judgment assumed a different complexion. In an era of growing social and economic regulation, statutory incentives to litigate (e.g., a host of new federal statutes with pro-plaintiff fee-shifting provisions),[21] the modern class action, and a bar responsive to such incentives elevated the role of litigation in American governance. As federal courts began to labor under the weight of increasing caseloads, those responsible for federal procedure sought better means to curtail wasteful litigation behavior

---

Sunderland had drafted broad summary judgment provisions for Michigan a few years earlier, but there was no documented experience under them.

   [16] A member of the Committee cautioned:

> I am in favor of retaining trial by jury inviolate and not in any instance substituting trial by affidavit, whether the party is in good faith or not. I think this is one of the most serious rules in our whole group, and it will be the one subject to the most criticism unless you throw every safeguard around the man who wants his case tried by a jury. Trial by jury is the safeguard of the man who otherwise would not get a square deal. I am very much opposed to giving any color to the charge that these rules in any way encroach on that right.

Proceedings of Meeting of Advisory Committee on Rules for Civil Procedure of the Supreme Court of the United States (Feb. 20–25, 1936), *microformed on* Records of the U.S. Judicial Conference, Committees on Rules of Practice and Procedures, 1935–1988, Nos. CI-210-35, CI-210-37, CI 210-54, at 830–31 (Feb. 20, 1936) (Mr. Donworth).

   [17] AMERICAN BAR ASSOCIATION, FEDERAL RULES OF CIVIL PROCEDURE: PROCEEDINGS OF THE CLEVELAND INSTITUTE 225 (1938).

   [18] 350 U.S. 544 (1947).

   [19] 355 U.S. 41 (1957).

   [20] *See, e.g.*, Adickes v. Kress, 398 U.S. 144 (1970).

   [21] *See* SEAN FARHANG, THE LITIGATION STATE (2010).

and to identify cases ripe for early termination. Through rulemaking and judicial decisions, they enlisted a variety of tools for these purposes, from sanctions to case management to discovery reform, and from the law of summary judgment to the law of pleading.

Pleading was an early candidate for retrenchment. In the 1950s, whether from affection for the old or fear of the new, some federal judges sought a return to fact pleading through amendments to the Federal Rules. They had no success with the Advisory Committee, and the timing of *Conley v. Gibson* was such that it could be viewed as a rebuke of the effort.[22] Several subsequent attempts similarly ran aground in the Advisory Committee.[23] The Committee likely recognized that, if pursued, such proposals would embroil the rulemaking process in partisan political controversy, endangering the main source of the federal judiciary's control of procedure by threatening legislative intervention in that domain. Moreover, their attitude in that regard did not change even after, twice within a decade (1993-2002), the Supreme Court reversed lower courts for imposing heightened pleading requirements in civil rights and employment discrimination cases through judge-made law, insisting that such changes must come from legislation or amendments to the Federal Rules.[24]

But pleading was not the main target of the long-running campaign against litigation that has been waged on behalf of business. Over the last fifty years, the greatest source of complaints about federal litigation has been its cost, with the primary culprit said to be the cost of discovery, particularly document discovery, most of which is borne by the party from which discovery is sought and cannot be shifted from the winner to the loser. At the same time, however, thoughtful scholars and judges have pointed out the potential costs of cutting back on discovery. Thus, Paul Carrington, former Reporter of the Advisory Committee, observed:

> We should keep clearly in mind that discovery is the American alternative to the administrative state . . . every day, hundreds of American lawyers caution their clients that an unlawful course of conduct will be accomplished by serious risk of exposure at the hands of some hundreds of thousands of lawyers, each armed with a subpoena power by which misdeeds can be uncovered. Unless corresponding new powers are conferred on public

---

[22] *See* Richard L. Marcus, *The Puzzling Persistence of Fact Pleading*, 76 TEX. L. REV. 1749, 1750 (1998).

[23] *See id.* at 1751–52.

[24] *See* Swierkiewicz v. Sorema N.A., 534 U.S. 506, 515 (2002); Leatherman v. Tarrant County Narcotics Intelligence and Coordination Unit, 507 U.S. 163, 168 (1993). *See also* Crawford-El v. Britton, 523 U.S. 574, 594–95 (1998). Apparently, the message was lost on, or simply unacceptable to, some lower federal courts, as the practice persisted even after *Swierkiewicz. See, e.g.,* Perry v. Southeastern Boll Weevil Erad. Fund, Inc., 154 Fed. Appx. 467, 472 (6th Cir. 2005); Danley v. Allen, 540 F.3d 1298, 1313-14 (11th Cir. 2008).

officers, constricting discovery would diminish disincentives for lawless behavior across a wide spectrum of forbidden conduct.[25]

The rulemakers have responded to complaints about discovery with round after round of amendments designed to discipline the discovery process,[26] the cumulative effect of which has been to render that process more complex. Yet, empirical research conducted over many decades has not demonstrated that discovery is a problem—is disproportionately expensive—in more than a small slice of litigation. Instead, study after study has found that discovery is a problem in precisely the types of cases that one would expect—high-stakes, complex cases.[27] These are the types of cases entailing the problems that have preoccupied the rulemakers in recent decades. Because the Federal Rules are transsubstantive and make few distinctions according to perceived procedural needs, the solutions the rulemakers have devised for complex cases—and the added expense those solutions can entail—are usually applicable in all cases. That which may be a cure in high-stakes, complex cases may also be a curse in simpler cases of modest stakes.[28]

C. *The New Regime*

1. The Supreme Court's Pleading Decisions

The Supreme Court invoked pervasive discovery abuse and crushing discovery expense, together with the asserted inability of federal judges to manage discovery, as reasons to change federal procedural law—but not the law that governs discovery. They did so by (1) resuscitating distinctions

---

[25] Paul D. Carrington, *Renovating Discovery*, 49 ALA. L. REV. 51, 54 (1997). Judge Patrick Higginbotham, former Chair of the Advisory Committee on Civil Rules, also emphasized the relationship of discovery to the ability of "private attorneys-general" to enforce congressional statutes, observing that "[c]alibration of discovery is calibration of the level of enforcement of the social policy set by Congress." Patrick Higginbotham, *Foreword*, 49 ALA. L. REV. 1, 5 (1997).

[26] For a discussion of the Committee's discovery proposals in historical context, see Letter from Stephen B. Burbank to Committee on Rules of Practice and Procedure 3-10 (Feb. 10, 2014), available at https://www.regulations.gov/docketBrowser?rpp=25&so=DESC&sb=commentDueDat e&po=0&s=burbank&dct=PS&D=USC-RULES-CV-2013-0002. As noted there, the Committee's 2013 proposals, which, as revised became effective in 2015, "represent[ed] the seventh set of (non-stylistic) proposed reforms since 1980." *Id.* at 10.

[27] *See, e.g.,* EMERY G. LEE III & THOMAS E. WILLGING, FED. JUDICIAL CTR., PRELIMINARY REPORT TO THE JUDICIAL CONFERENCE ADVISORY COMMITTEE ON CIVIL RULES (2009), *available at* http://www.fjc.gov/public/pdf.nsf/lookup/dissurv1.pdf/$file/dissurv1.pdf; Danya Shocair Reda, *The Cost and Delay Narrative in Civil Justice Reform: Its Fallacies and Functions,* 90 OR. L. REV. 1085, 1103–11 (2012) (discussing 2009 FJC study); *id.* at 1111 ("Nearly every effort to quantify litigation costs and to understand discovery practice over the last four decades has reached results similar to the 2009 FJC study.").

[28] *See* Stephen B. Burbank, *The Complexity of Modern American Civil Litigation: Curse or Cure?*, 91 JUDICATURE 163 (2008).

between "facts" and "conclusions" that the drafters of the 1938 Federal Rules had rejected when crafting Rule 8, and (2) transforming the 12(b)(6) motion from a vehicle for testing the viability of the plaintiff's legal theory into a means to weed out complaints that, shorn of conclusions, do not set forth sufficient facts to make the plaintiff's claim "plausible."

In order that defendants in a massive antitrust case might be spared impositional discovery,[29] in *Bell Atlantic Corp. v. Twombly,*[30] the Supreme Court made it more difficult for plaintiffs in such cases to survive a motion to dismiss for failure to state a claim under Rule 12(b)(6). The case involved an antitrust conspiracy complaint brought as a class action under Section 1 of the Sherman Act against the regional telecommunications service providers that remained after the breakup of AT&T. Reversing a panel of the Second Circuit, the Court "retired" the language in *Conley v. Gibson* that "a complaint should not be dismissed . . . unless it appears beyond doubt that the plaintiff can prove no set of facts . . . which would entitle him to relief."[31] Agreeing, however, with *Conley* that a complaint must give "fair notice of what the . . . claim is and the grounds upon which it rests,"[32] and emphasizing Rule 8's requirement that the statement of claim "show[] that the pleader is entitled to relief," the Court interpreted the former as requiring "more than labels and conclusions,"[33] and the latter as requiring that its "[f]actual allegations must be enough to raise a right to relief above the speculative level[.]"[34] The Court then held that for a Section 1 Sherman Act claim these standards "require[d] a claim with enough factual matter (taken as true) to suggest that an agreement was made."[35] Disregarding direct allegations of conspiracy as conclusory, the Court held that the plaintiffs' claims were not plausible because they rested on allegations of "parallel conduct and not on any independent allegation of actual agreement among [defendants]." [36]

In decisions after *Conley v. Gibson*, the Court had insisted that pleading practice is not the appropriate way to challenge the factual sufficiency of a

---

[29] Judge Easterbrook deployed the concept of "impositional discovery" in an article on which the *Twombly* Court relied heavily. *See* Frank H. Easterbrook, *Discovery as Abuse*, 69 B.U. L. REV. 633, 637–38 (1989) (defining "an impositional request" as "one justified by the costs it imposes on one's adversary rather than by the gains to the requester derived from the contribution the information will make to the accuracy of the judicial process"); Bell Atl. Corp. v. Twombly, 550 U.S. 544, 559 (2007).

[30] *Twombly*, 550 U.S. at 559.

[31] 355 U.S. at 45–46; *see Twombly*, 550 U.S. at 563 ("[T]his famous observation has earned its retirement").

[32] *Twombly*, 550 U.S. at 555 (*quoting Conley*, 355 U.S. at 47).

[33] *Id.*

[34] *Id. See id.* n.3; *id.* at 557 ("The need at the pleading stage for allegations plausibly suggesting (not merely consistent with) agreement reflects the threshold requirement of Rule 8(a)(2) that the 'plain statement' possess enough heft to 'sho[w] that the pleader is entitled to relief.'").

[35] *Id.* at 556.

[36] *Id.* at 564.

plaintiff's claim, which is a task for summary judgment, after the plaintiff has had the opportunity for discovery.[37] In retrojecting the gate-keeping function to the pleading stage, the *Twombly* Court appears to have been influenced by the experience it had developed in policing inferences in antitrust conspiracy cases at later stages. Whether at trial (via a motion for judgment as a matter of law), or on a pretrial motion for summary judgment, the question whether plaintiff had a plausible claim of a proscribed agreement could be tested (1) under substantive law making clear that evidence of parallel conduct by itself is not enough to sustain a claim under Section 1 of the Sherman Act, and (2) against a factual record.[38] Moreover, although the majority and dissent in *Twombly* disagreed about the discovery-management ability of district judges,[39] *Twombly* was undoubtedly a case in which discovery could be very expensive. These and other considerations prompted speculation that *Twombly's* domain might be limited, whether only to antitrust cases or to cases portending similarly costly discovery.[40]

Not for long. Two years later, in *Ashcroft v. Iqbal*,[41] the Court again expressed concern about the costs of discovery—but this time the costs of diverting the time and attention of high government officials asserting

---

[37] *See, e.g.,* Swierkiewicz v. Sorema N.A., 534 U.S. 506, 512-13 (2002); Leatherman v. Tarrant County Narcotics Intelligence and Coordination Unit, 507 U.S. 163, 168-69 (1993).

[38] *See Twombly,* 550 U.S. at 554 ("Accordingly, we have previously hedged against false inferences from identical behavior at a number of points in the trial sequence);" *id.* at 561 n.7; Theatre Enterprises, Inc. v. Paramount Film Distributing Corp., 346 U.S. 537 (1954); Monsanto Co. v. Spray-Rite Service Corp., 465 U.S. 752 (1984); Matsushita Elec. Industrial Co. v. Zenith Radio Corp., 475 U.S. 574 (1986).

[39] *See Twombly*, 550 U.S. at 559 (noting "common lament that the success of judicial supervision in checking discovery abuse has been on the modest side"). *See id.* at 596-97 (Stevens, J., dissenting) ("[I]t is only a lack of confidence in the ability of trial judges to control discovery, buttressed by appellate judges' independent appraisal of the plausibility of profoundly serious allegations, that could account for this stark break from precedent.")

[40] In dissent, Justice Stevens observed that "[w]hether the Court's actions will benefit only defendants in antitrust treble damages cases, or whether its test for the sufficiency of a complaint will inure to the benefit of all civil defendants, is a question that the future will answer." *Id.* at 596 (Stevens, J., dissenting). Contemporary commentary asserted:

> More probably, *Twombly* is an exercise in strategic ambiguity that empowers the lower federal courts to tighten pleading requirements in cases or categories of cases that augur similar discovery burdens (or are otherwise disfavored), while preserving deniability in the Court through the use of its discretionary docket to correct perceived excesses (as in *Erickson*).

Editorial, *The Devil in the Details*, 91 JUDICATURE 52 (2007). The reference is to Erickson v. Pardus, 551 U.S. 89 (2007), a case decided a few weeks after *Twombly* (without argument and per curiam) in which the Court reversed the Tenth Circuit's affirmance of a judgment dismissing a prisoner's complaint under Rule 12(b)(6).

[41] 556 U.S. 662 (2009).

qualified immunity.[42] The case involved some of the claims brought by a citizen of Pakistan whom federal officials arrested after the 9/11 attacks and who was subsequently transferred to the (federal) Metropolitan Detention Center in Brooklyn, New York (MDC), pending trial on charges of fraud in connection with identification documents to which he ultimately pleaded guilty, leading to his removal to Pakistan. The complaint alleged that Iqbal's seven-month confinement in highly restrictive conditions at MDC resulted from unlawful racial and religious discrimination. It asserted that Robert Mueller, the Director of the F.B.I., and John Ashcroft, the Attorney General of the United States, adopted and/or approved policies and directives pursuant to which Iqbal was confined, policies and directives that purposefully discriminated on the basis of religion and race.[43]

According to the dissent of four justices—including the author of the Court's opinion in *Twombly* and another justice who joined that opinion—the Court in *Iqbal* inconsistently treated some of the complaint's assertions as factual allegations and others as conclusions.[44] The Court also disregarded direct allegations of intentional discrimination, notwithstanding Rule 9(b)'s assurance that "[m]alice, intent, knowledge and other conditions of a person's mind may be alleged generally."[45] That move enabled the Court to assess the plausibility of the inferential basis for the theory of the plaintiff's case.[46] Relying on "judicial experience and common sense,"[47] the Court found the complaint implausible. Because the Federal Rules are transsubstantive, the Court was constrained to make clear that its approach applies across the board—that *Twombly* cannot be confined to its substantive context (antitrust) or according to some other criterion (e.g., cases with heavy discovery burdens).[48]

2. Evaluating the Court's Decisions: Normative Concerns

There is a vast normative literature on the Court's decisions in *Twombly* and *Iqbal*. It started shortly after the former was decided, when the extent of the Court's changes in pleading doctrine was unclear, and proliferated after the latter

---

[42] *See id.* at 685–86. *But see id.* at 700 (Breyer, J., dissenting) ("Neither the briefs nor the Court's opinion provides convincing grounds for finding these alternative case-management tools inadequate, either in general or in the case before us.").

[43] *See* Iqbal v. Hasty, 490 F.3d 143, 147-49, 165, 174-76 (2d Cir. 2007), *rev'd*, Ashcroft v. Iqbal, 556 U.S. 662 (2009).

[44] *See Iqbal*, 556 U.S. at 698–99 (Souter, J., dissenting).

[45] FED. R. CIV. P. 9(b). *See Iqbal*, 556 U.S. at 686–87. One line of criticism was that the Court's interpretation of Rule 9(b) betrays the original understanding and is otherwise "patently unsupportable." A. Benjamin Spencer, *Pleading Conditions of the Mind Under Rule 9(b)*, 41 CARDOZO L. REV. 1015, 1016 (2020).

[46] *See Iqbal*, 556 U.S. at 681–83.

[47] *Id.* at 679.

[48] *See id.* at 684.

was decided, when the fears of *Twombly's* critics and the dreams of its supporters seemingly had been confirmed. It appears to us that critics far outnumber supporters among academics (as opposed to practitioners) writing about these decisions and that the criticisms include a number of identifiable themes.[49]

One theme focuses on the Court's embrace of distinctions—among "facts," "conclusions," "legal conclusions," and "threadbare allegations"—akin to those that had bedeviled code pleading. In the view of critics, changes in the influences affecting federal litigation that have occurred since 1938 do not include the logical counterrevolution[50] that would have been necessary to rescue such distinctions from the original Advisory Committee's indictment of arbitrariness or from the more serious indictment that they unfairly impede access to justice.[51]

Two other themes concern the Court's turn to plausibility pleading. Once *Iqbal* affirmed that the new regime was transsubstantive, and with the reinterpretation of Rule 9(b) in that case, it became clear that in many cases plausibility pleading would require courts to assess inferences without the benefit of either substantive law rules or a factual record. As a result, many critics

---

[49] *See* David L. Noll, *The Indeterminacy of* Iqbal, 95 GEO. L.J. 117, 118–20 (2010). A common theme we do not discuss is a critique based on the illegitimacy of changing pleading law via judicial decision, given the requirements of the Rules Enabling Act as interpreted in past decisions, and on the epistemic shallowness of the Court's decisions in *Twombly* and *Iqbal*. *See, e.g.*, Stephen B. Burbank, *Pleading and the Dilemmas of "General Rules*," 2009 WIS. L. REV. 535, 548–49; Stephen B. Burbank, *Pleading and the Dilemmas of Modern American Procedure*, 93 JUDICATURE 109, 116, 118–20 (2009) (hereinafter *Dilemmas of Modern American Procedure*).

[50] *See* Walter Wheeler Cook, *Statements of Fact in Pleading Under the Codes*, 21 COLUM. L. REV. 416, 417 (1921) (arguing that "there is no logical distinction between statements which are grouped by courts under the phrases 'statements of fact' and 'conclusions of law'"); Robert G. Bone, *Plausibility Pleading Revisited and Revised: A Comment on* Ashcroft v. Iqbal, 85 NOTRE DAME L. REV. 849, 862–70 (2010).

[51] *See* Noll, *supra* note 49, at 120 (discussing critique "that when the defendant controls critical private information, *Iqbal* creates an apparent Catch-22 for plaintiffs, requiring them to plead information they do not know but denying them a means of discovering that information"). As one of us wrote shortly after *Iqbal* was decided:

> The architecture of *Iqbal's* mischief . . . is clear. The foundation is the power the Court claimed to parse a complaint, accepting allegations of fact as true and ignoring conclusory allegations, including legal conclusions. In *Twombly*, the Court ignored allegations of conspiracy; in *Iqbal*, notwithstanding Rule 9(b), it ignored allegations of discriminatory intent. Yet, an important reason why the drafters of the 1938 Federal Rules rejected fact pleading is that one person's "factual allegation" is another's "conclusion." The discretionary power of the judge to follow his or her personal preferences in deciding the plausibility of a complaint is enlarged to the extent that direct allegations of liability-creating conduct can be thus disregarded.

Burbank, *Dilemmas of Modern American Procedure*, *supra* note 49, at 115.

expressed concern that another layer of subjectivity had been added to that inherent in the enterprise of distinguishing between "facts" and "conclusions."[52]

Relatedly, critics of *Iqbal* regard with skepticism the idea that "judicial experience and common sense" can serve as a reliable guide in determining whether a complaint states a plausible claim. Given the lack of a factual record and, in many cases, of rules that adjudicate inferential force, critics maintain that the formula portends judicial decision-making that is vulnerable to unconscious bias. Most prominently mentioned among cases in which such a dynamic may be in play are those involving civil rights claims, including claims of employment discrimination.[53]

One reason may be that some discrimination and other civil rights decisions are vulnerable to what Professors Kahan, Hoffman and Braman call "cognitive illiberalism"[54] in an article on the dangers of summary adjudication exemplified by the Supreme Court's decision in *Scott* v. *Harris*.[55] That is because in some discrimination and other civil rights cases one would expect "Americans [to] interpret th[e] facts against the background of competing subcommunity understandings of social reality,"[56] making them strong

---

[52] "The assessment of likely trial success that the thick screening model requires is an all-things-considered prediction based on what the complaint tells the judge about the facts and what the judge knows from her experience about how facts like the ones alleged are usually proved in similar cases." Bone, *supra* note 50, at 873. *See also* Elizabeth M. Schneider, *The Changing Shape of Federal Civil Pretrial Practice: The Disparate Impact on Civil Rights and Employment Discrimination Cases*, 158 U. PA. L. REV. 517, 527–36 (2010).

[53] *See, e.g.,* Schneider, *supra* note 52; Victor D. Quintanilla, *Beyond Common Sense: A Social Psychological Study of Iqbal's Effect on Claims of Race Discrimination*, 17 MICH. J. RACE & L. REV. 1, 26 (2011) (arguing that, according to aversive racism theory, "[w]ith no evidence, and relying on 'common sense,' courts are more likely to be influenced by automatic stereotypes and implicit associations about race," while, lay theories of racism predict that "[u]nder *Iqbal's* more rigorous plausibility standard . . . many federal judges will likely fail to perceive subtle discrimination as plausibly suggesting unlawful discrimination" and that "White and Black judges will decide motions to dismiss in ambiguous cases differently"); Jerry Kang et al., *Implicit Bias in the Courtroom*, 59 UCLA L. REV. 1124, 1160 (2012); Elizabeth M. Schneider & Nancy Gertner, *"Only Procedural": Thoughts on the Substantive Law Dimensions of Preliminary Procedural Decisions in Employment Discrimination Cases*, 57 N.Y.L. SCH. L. REV. 767, 775 (2012–2013) (stating that "there is little difference between the 'common sense' and 'plausibility' standards that *Iqbal* and *Twombly* encourage and the very cognitive processes that social scientists have identified as producing bias"); Joseph A. Seiner, *The Discrimination Presumption*, 94 NOTRE DAME L. REV. 1115, 1126-28 (2019). For recent work that recognizes a possible role for "rules that adjudicate inferential force"—presumptions—in addressing the problem of implicit bias, see Elizabeth Thornburg, *(Un)Conscious Judging*, 76 WASH. & LEE L. REV. 1567, 1652-53 (2019) (suggesting presumptions for issues where "inferences are particularly likely to be unreliable because of the limits of generalizations or social stereotypes"). Even without reference to the problem of unconscious bias, critics saw in *Iqbal's* aggressive screening approach a particular threat to civil rights cases. *See, e.g.,* Bone, *supra* note 50, at 879.

[54] Dan M. Kahan, David A. Hoffman, and Donald Braman, *Whose Eyes Are You Going to Believe? Scott v. Harris and the Perils of Cognitive Illiberalism*, 122 HARV. L. REV. 838 (2009). *See id.* at 896.

[55] 550 U.S. 372 (2007).

[56] Kahan et al., *supra* note 54, at 887.

candidates for the operation of cognitive biases that judges may recognize in others but not in themselves.[57] Both plaintiffs and jurors in discrimination and other civil rights cases will often have "recognizable identity-defining characteristics" that might cause them to dissent from a view of plausibility grounded in a judge's cultural predispositions.[58]

In his book, *The Death of the American Trial*, Professor Robert Burns observes:

> Common sense very rarely confronts the level of detailed factual development that the trial provides. Every time the lawyer says, "Generally and for the most part . . . ." the other lawyer is likely to say, "Yes, but not where . . . ." Each new case requires a genuine insight, what Peirce called an "abduction," that must seek out the intelligibility inherent in these particular facts. Paradoxically, by giving particularity and empirical truth their due, the trial provides a strong critique of commonsense generalizations . . . The trial provides a self-criticism of the overgeneralized "scripts" with which much of our common sense is stored.[59]

According to some critics of *Iqbal*, "judicial experience and common sense" is subject to no such critique.

II. IDEOLOGY AND IDENTITY ON THE COURTS OF APPEAL: THE EMPIRICAL LITERATURE

In this section we review existing empirical research on the relationship between votes or outcomes in Court of Appeals decisions and judges' ideology (usually proxied by party of appointing president), gender, and race. The literature has a patchwork character, with salient and apparently ideologically divisive issues sometimes the locus of significant associations between judge characteristics and outcomes, and sometimes not. The literature is largely bereft of theory explicating the relationship between specific characteristics and preferences over outcomes across policy domains, particularly as to judges' gender and race. As a result, it has proceeded inductively, mapping the universe of relationships in fields of law that researchers consider important.

---

57 *See id.* at 843 (noting that "[w]e thus simultaneously experience overconfidence in the unassailable correctness of the factual perceptions we hold in common with our confederates and unwarranted contempt for the perceptions associated with our opposites"). *See also* Russell M. Robinson, *Perceptual Segregation*, 108 COLUM. L. REV. 1093, 1153 (2008) (introducing proposals to change Title VII jurisprudence in order to "respond in one way or another to judicial intuitions that (1) discrimination is rare and (2) most outsiders who claim to have suffered discrimination are either paranoid or strategic").

58 *See* Kahan et al., *supra* note 54, at 898–99.

59 ROBERT P. BURNS, THE DEATH OF THE AMERICAN TRIAL 33, 35 (2009).

A. *Party*

By now it will surprise few people that measures of judge ideology based, in whole or in part, on party of appointing president are associated with decision-making on the Courts of Appeals. The empirical literature establishing this has focused heavily on civil rights and liberties, and studies cover a lot of that landscape. They have found that Democratic appointees to the Courts of Appeals, on average, are more likely to decide in the liberal direction than Republican appointees in the areas of affirmative action, employment discrimination, sex discrimination, desegregation, disability rights, abortion, campaign finance, freedom of expression,[60] some types of religious liberty cases,[61] voting rights,[62] search and seizure, and obscenity.[63] There has been much less work outside of civil rights and liberties, and the results are more mixed. Researchers have found that judges appointed by Democrats or with more liberal ideology measures, on average, are more likely to decide in the liberal direction in the areas of labor, communications,[64] and an aggregation of cases between individuals and business.[65] However, studies have found no such differences in some areas that may be regarded as ideologically salient, such as takings of property rights, punitive damages, standing, and Commerce Clause challenges to national legislation.[66]

---

[60] *See* CASS R. SUNSTEIN, DAVID SCHKADE, LISA M. ELLMAN & ANDRES SAWICKI, ARE JUDGES POLITICAL?: AN EMPIRICAL ANALYSIS OF THE FEDERAL JUDICIARY 149 (2006) (evaluating all of these issues); Jonathan P. Kastellec, *Racial Diversity and Judicial Influence on Appellate Courts*, 57 AM. J. POL. SCI. 167 (2013) (evaluating affirmative action cases); Sean Farhang & Gregory Wawro, *Institutional Dynamics on the U.S. Court of Appeals: Minority Representation Under Panel Decision Making*, 20 J.L. ECON. & ORG. 299 (2004) (evaluating employment discrimination cases); Jennifer L. Peresie, Note, *Female Judges Matter: Gender and Collegial Decisionmaking in the Federal Appellate Courts*, 114 YALE L.J. 1759 (2005) (evaluating employment discrimination cases).

[61] *See* Gregory C. Sisk, Michael Heise & Andrew P. Morris, *Searching for the Soul of Judicial Decisionmaking: An Empirical Study of Religious Freedom Decisions*, 65 OHIO ST. L.J. 491, 595–96 (2004); Sepehr Shahshahani & Lawrence J. Liu, *Religion and Judging on the Federal Courts of Appeals*, 14 J. EMPIRICAL LEGAL STUD. 716 (2017).

[62] *See* Adam B. Cox & Thomas J. Miles, *Judging the Voting Rights Act*, 108 COLUM. L. REV. 1 (2008).

[63] *See* Donald R. Songer, Sue Davis & Susan Haire, *A Reappraisal of Diversification in the Federal Courts: Gender Effects in the Courts of Appeals*, 56 J. POL. 425 (1994).

[64] *See* SUNSTEIN ET AL., *supra* note 60, at 151.

[65] *See* SUSAN B. HAIRE & LAURA P. MOYER, DIVERSITY MATTERS: JUDICIAL POLICY MAKING IN THE U.S. COURTS OF APPEALS 47-48 (2015) (For Haire and Moyer, "distributive politics" is defined as suits between individuals and business, and they exclude discrimination cases). Other work finds Court of Appeals judges' votes to be associated with party of the appointing president when cases are pooled across a large number of policy areas including both civil rights and non-civil rights cases. *See* LEE EPSTEIN, WILLIAM M. LANDES, AND RICHARD A. POSNER, THE BEHAVIOR OF FEDERAL JUDGES: A THEORETICAL AND EMPIRICAL STUDY OF RATIONAL CHOICE ch. 4 (2013).

[66] *See* SUNSTEIN ET AL., *supra* note 60, at 149, 151.

B. *Gender and Race*

The study of gender and racial diversity on the federal bench is often tied to issues of representation. Pitkin's distinction between descriptive and substantive representation highlights two goals that diversity on the bench can serve.[67] Descriptive representation is concerned with whether an institution of governance mirrors, in salient respects, the composition of the community that it governs. Substantive representation, in contrast, is concerned with whether government actors, in their decision-making, actually represent the distinctive preferences or interests of a community that they are associated with.[68]

Advocates of gender and racial diversity on the bench have long argued that a value of judicial diversity is to create a bench that descriptively reflects the polity, which itself can promote the judiciary's appearance of impartiality and enhance its democratic legitimacy.[69] But they have also argued that women and members of racial minorities have distinct preferences in at least some policy domains, and that in particular they are, on average, more sensitive than white men to issues of discrimination and inequality.[70] The primary reason given for this view is that women and members of racial minorities are more likely to have seen or been subjected to discrimination, and these life experiences make the judges more likely to believe a plaintiff's claims of discrimination or other status-based injury and to empathize with such plaintiffs.[71]

1. Gender

The empirical literature on the Courts of Appeals suggests that judges' gender is consequential to their decision-making in a much narrower range of cases than may have been expected by some advocates of gender diversity on the bench. A series of studies has found that women judges vote in a more

---

[67] HANNA FENICHEL PITKIN, THE CONCEPT OF REPRESENTATION (1967).

[68] *See id.* at 80, 184.

[69] *See* HAIRE & MOYER, *supra* note 65; Allison P. Harris & Maya Sen, *Bias and Judging*, 22 ANN. REV. POL. SCI. 241, 247 (2019); Susan Moloney Smith, Comment, *Diversifying the Judiciary: The Influence of Gender and Race on Judging*, 28 U. RICH. L. REV. 179, 198 (1994); Carl Tobias, Commentary, *The Gender Gap on the Federal Bench*, 19 HOFSTRA L. REV. 171, 177 (1990).

[70] *See* HAIRE & MOYER, *supra* note 65, at 6, 13–114; Theresa M. Beiner, *What Will Diversity on the Bench Mean for Justice?*, 6 MICH. J. GENDER & L. 113 (1999); Christina L. Boyd, Lee Epstein & Andrew D. Martin, *Untangling the Causal Effects of Sex on Judging*, 54 AM. J. POL. SCI. 389 (2010); Sherrilyn A. Ifill, *Racial Diversity on the Bench: Beyond Role Models and Public Confidence*, 57 WASH. & LEE L. REV. 405 (2000).

[71] *See, e.g.*, HAIRE & MOYER, *supra* note 65, at 32, 48; Boyd et al., *supra* note 70; Ifill, *supra* note 70; Joy Milligan, *Pluralism in America: Why Judicial Diversity Improves Legal Decisions About Political Morality*, 81 N.Y.U. L. REV. 1206 (2006).

pro-plaintiff direction, and/or the presence of a woman on a three-judge panel is associated with more pro-plaintiff outcomes as compared to all male panels, in employment discrimination cases.[72]

Other than the discrete policy area of employment discrimination and our study of class certification (which we discuss later in this section), the literature on the Courts of Appeals has generally found that gender is not associated with Court of Appeals judges' votes or panel outcomes. Researchers have reported that gender is not associated with votes or outcomes in the areas of environmental protection, federalism, piercing the corporate veil, the Contracts Clause, or the Takings Clause.[73] They have likewise found that gender is not associated with Court of Appeals judges' voting behavior when one pools a large set of policy areas that can be characterized on a left-right spectrum.[74] Null results with respect to judges' gender also extend into domains of civil rights and liberties without explicit gender content, including campaign finance, capital punishment, disability rights, race discrimination in employment,[75] affirmative action,[76] voting rights,[77] religious liberty,[78] search and seizure, and obscenity.[79]

Finally, the null results with respect to judges' gender extend into some domains with quite salient gender content, including one notable form of employment discrimination. One study finds that gender is not associated with votes or outcomes in sexual harassment cases (which occur primarily in the employment context), or abortion.[80] In sum, the employment discrimination studies revealing gender differences in Court of Appeals decision-making are islands in a sea of null results.[81] This is the dominant view in the literature.[82]

---

[72] *See* Boyd et al., *supra* note 70; Farhang & Wawro, *supra* note 60; Peresie, *supra* note 60; HAIRE & MOYER, *supra* note 65; Songer et al., *supra* note 63.

[73] *See* Boyd et al., *supra* note 70.

[74] *See* HAIRE & MOYER, *supra* note 65, at 47; SUNSTEIN ET AL., *supra* note 60, at 167, 171, 185, 197.

[75] *See id.*

[76] Kastellec, *supra* note 60, at 178.

[77] *See* Cox & Miles, *supra* note 62, at 43–45.

[78] *See* Sisk et al., *supra* note 61, at 593.

[79] *See* Songer et al., *supra* note 63, at 433.

[80] *See* Boyd et al., *supra* note 70*; but see* Peresie, *supra* note 60 (reaching a contrary conclusion with respect to sexual harassment cases).

[81] *See* Burbank & Farhang, *Politics & Identity*, *supra* note 3*.*

[82] Surveying the literature on gender and judging, and reporting the results from a large-scale study in which they find gender associated with outcomes in employment discrimination claims based on gender, Haire and Moyer conclude that "issues of sex discrimination" are "[t]he single exception" to the general rule that "women judges . . . decide cases similarly to their male colleagues." HAIRE & MOYER, *supra* note 65, at 48. *See also* Christina Boyd & Adam Rutkowski, *Judicial Behavior in Disability Cases: Do Judge Sex and Race Matter?*, 8 POL., GROUPS, AND IDENTITIES 834, 837-38 (2020) ("[A] relatively large number of empirical studies . . . have failed to find evidence that female and male judges decide cases differently from one another, particularly

Why is employment discrimination different? Scholars have been left to speculate. Boyd, Epstein and Martin, in their noted study covering thirteen issue areas but finding gender significant in only one (employment discrimination based on gender), hypothesize that women vote differently, and influence men on three-judge panels, in domains in which they "possess unique and valuable *information emanating from shared professional experiences*."[83]

We have cautioned that none of these studies reporting null results for gender separately analyzed panels in which women were in the majority. As a result, it is not possible to be confident whether null results in any given policy domain (or aggregation of them) indicate that women do not have different preferences than men, or that women in the minority (where the bulk of their votes are cast) are not influencing panel outcomes on majority-male panels and are suppressing their dissents.[84] In a study of class certification, we find that one woman on a panel is not associated with increasing the probability of certification, but two women are, showing that in some domains women judges' preferences become visible, and are differentially associated with outcomes, only when they are in the majority.[85]

2. Race

Studies finding Court of Appeals judges' race to be significantly associated with votes and outcomes have occurred across a notably broader range of issue areas than gender, and such findings have often occurred in studies reporting null results for gender in the same data. Researchers have reported that African American judges were associated with pro-plaintiff voting and/or outcomes in voting rights cases,[86] affirmative action cases,[87] employment discrimination claims based on race,[88] religious liberty claims,[89] and death penalty cases.[90] Not all studies focused on civil rights have found

---

outside of issue areas that are not closely related to 'women's issues' like sex discrimination."); Jonathan P. Kastellec, *Race, Context and Judging on the Courts of Appeals: Race-based Panel Effects in Death Penalty Cases*, 41 JUST. SYSTEM J. 1, 2 (2020) ("[T]he addition of a woman to a panel increases the probability that men will vote for the plaintiff in sex discrimination cases, but makes no difference in cases not related to gender. . . . . In sum, judges tend to influence each other in areas of the law where we would expect such influence to occur.").

    83 Boyd et al., *supra* note 70, at 391–92, 398, 401; *see also* Christina Boyd, *Representation on the Courts? The Effects of Trial Judges' Sex and Race*. 69 POL. RES. Q. 69, 789-90 (2016); Boyd & Rutkowski, *supra* note 82.

    84 *See* Burbank & Farhang, *Politics & Identity*, *supra* note 3, at 250, 267–68.

    85 *See id.* at 261–63.

    86 *See* Cox & Miles, *supra* note 62, at 30, 43 (also reporting null results for gender).

    87 *See* Kastellec, *supra* note 60 (also reporting null results for gender).

    88 *See* HAIRE & MOYER, *supra* note 65, at 28–32.

    89 *See* Sisk et al., *supra* note 61 (also reporting null results for gender).

    90 *See* Kastellec, *supra* note 82.

race effects, however. Studies have found that non-white judges are not associated with more pro-plaintiff rulings in employment discrimination cases where the data were not restricted to race claims.[91] Two find that African American or non-white judges do not vote more liberally when all cases are pooled across a wide range of civil issues that can be classified on a liberal-conservative spectrum.[92] Still, as compared to gender, scholars have observed that the scope of the empirically discernable association between judges' race and their decision-making appears to be broader, traversing more issue areas.[93]

C. *Procedural Law and Access to Justice*

The judicial behavior literature on the Courts of Appeals has almost totally ignored access to justice issues, by which we mean rules affecting opportunities and incentives to enforce substantive rights through litigation. That is a problematic omission. If it is important to understand the relationship between ideology and diversity among Courts of Appeals judges and disposition of cases asserting rights of high public salience, it is equally important to understand the impact of the same factors on whether plaintiffs seeking to enforce those rights will have effective access to court in order to do so.

An empirical study examining the relationship between access to justice and judicial behavior entails a shift in how to conceptualize the unit of interest. In a conventional study of Title VII cases, for example, the researcher identifies a sample of cases in which the plaintiff is asserting a Title VII claim. Scholars are rarely clear regarding what cases qualify for inclusion other than the presence of the requisite type of claim. In particular, they often don't make clear whether procedural or other threshold issues, which can be dispositive, are included.[94] The key point is that the unit of analysis in most existing studies is defined by the claim (Title VII), and the data aggregate appeals of decisions rendered by the trial court throughout the stages of litigation (dismissal, summary judgment, post-trial motions, etc.).

---

[91] *See* Farhang & Wawro, *supra* note 60; Peresie, *supra* note 60.

[92] *See* HAIRE & MOYER, *supra* note 65, at 28–32; EPSTEIN ET AL., *supra* note 65, at 167, 171, 185, 197.

[93] *See* Boyd & Rutkowski, *supra* note 85, at 837–38 (theory and evidence predict that African American judges will be more pro-civil rights in general, and more likely to support disadvantaged groups, including disability benefits claimants); Boyd, *supra* note 83, at 789-90 (discussing theoretical accounts of case types that may be associated with judges' gender and race); Kastellec, *supra* note 82.

[94] For an exception in which researchers included procedural rulings and were explicit about it, see Shahshahani & Liu, *supra* note 61, at 723 ("Our object is to gauge the attitude of judges to religious liberties claims, and there is no reason to think that this attitude becomes uninteresting or entirely different when the issue before the court is procedural. Quite the opposite, we know that judges use procedural doctrines to achieve substantive outcomes they desire."). This study, however, provided no separate analysis of procedural issues, and thus it does not allow inferences about whether their outcomes were associated with judge characteristics.

When studying access to court, the criterion for inclusion is the nature and posture of the issue presented in a way that crosscuts substantive claims. In our prior study, the question was whether to certify a class under Federal Rule of Civil Procedure 23. In this study the question is whether, under Federal Rule of Civil Procedure 12(b)(6), the case should be dismissed for failure to state a claim without the opportunity to gather evidence in discovery. In both studies, data are aggregated across substantive claims. This research design allows one to evaluate whether judge characteristics such as party of the appointing president, gender and race are associated with voting behavior and outcomes in the discrete access to justice issue studied.

The conventional approach to claim aggregation contains an implicit assumption that procedural posture does not matter. Our approach takes seriously the idea that, independently of judges' policy preferences over fields of policy (e.g., environmental, anti-trust, voting rights), the nature of the procedural question may have a distinctive relationship to judicial preferences. The fact that two types of judges do not differ on whether to overturn jury verdicts in some policy domain(s) doesn't mean that they will not differ on whether to dismiss claims before discovery.

Our recent study of class certification illustrates how this approach can reveal relationships that are absent from the picture painted by the empirical literature on judges' ideology, race and gender that aggregates claims across all procedural postures, reviewed above. We analyzed Court of Appeals panel decisions addressing whether or not to certify a class under Rule 23. We found a very strong association between the political party of the appointing president and certification votes and outcomes, with all-Democratic panels yielding pro-certification outcomes at nearly triple the rate of all-Republican panels over about the past twenty years. [95]

The study also shows that racial and gender diversity on panels is consequential to certification, although we discern important differences between the race and gender dynamics on panels. The presence of a single African American on a panel, relative to none, increased the probability that the panel would yield a pro-certification outcome.[96] In notable contrast, the presence of a single woman on a panel, relative to none, was not associated with an increased probability of a pro-certification outcome, but the presence of two women was.[97] Panels addressing whether a class should be certified appeared to operate on a majoritarian basis with respect to gender but not race.

In addition, we found that the higher levels of pro-certification outcomes on panels with one African American and with two women were not driven

---

[95] *See* Burbank & Farhang, *Politics & Identity*, *supra* note 3, at 260–61.

[96] *See id*. at 264-65.

[97] *See id*. at 264-65.

by discrimination claims or civil rights claims more broadly. They remained present when analysis was restricted to non-civil rights claims.[98] Contrary to the dominant view in the literature, gender and race were associated with outcomes outside the area of discrimination and civil rights more broadly. These results motivate us to further explore the effects of diversity on transsubstantive procedural law affecting access to court. We now turn to an empirical analysis of pleading decisions.

III. Data, Models, and Analysis

A. *The Data*

We collected cases in which the Courts of Appeals reviewed district court decisions on whether to dismiss for failure to state a claim under Rule 12(b)(6). We excluded pro se cases and cases in which the court applied a heightened pleading standard required by rule or statute. We drew a random sample of 700 cases from the full universe of cases between the *Iqbal* decision and the end of 2019.[99] In this sample 36 percent of cases asserted a civil rights claim. In order to have sufficient data for separate analysis of discrimination claims, and all other civil rights clams, we took a random oversample of an additional 206 civil rights cases (including both discrimination and other civil rights), and an additional 130 cases asserting discrimination claims.

In the random sample of 700 cases, 35 percent were precedential. Because we wanted sufficient data for separate analysis of precedential cases, we collected an additional 942 precedential cases (without any policy area restrictions) between the *Iqbal* decision and the middle of 2020. This was all precedential cases that received general Westlaw headnotes for 12(b)(6) motions that were not already captured by our random samples. In total, our

---

[98] *See id.* at 265-67.

[99] The E-Government Act of 2002 required that federal circuits make even non-precedential opinions publicly available, allowing them to be included in commercial databases. *See* Andrew T. Solomon, *Making Unpublished Opinions Precedential: A Recipe for Ethical Problems & Legal Malpractice?*, 26 Miss. C. L. Rev. 185, 205–06 (2007) ("By 2005, every federal circuit released the full-text of its unpublished opinions."). However, recent work finds fewer of certain types of appeals on commercial databases than the number reported by the Administrative Office of the U.S. Courts, casting doubt on the U.S. Courts of Appeals' full compliance with the E-Government Act of 2002. *See* Merritt E. McAlister, *Missing Decisions*, 169 U. Pa. L. Rev. 1101 (2021). McAllister examined the nature of missing cases only in the First Circuit, where 67% were criminal and 49% were pro se. *Id.* at 1144. Although she does not report the percentage of missing cases with counseled civil plaintiffs, the forgoing percentages are consistent with the number being zero or miniscule. Although McAlister's data does not show non-compliance with the E-Government Act with respect to counseled civil plaintiffs, such noncompliance cannot be foreclosed without more evidence. Future empirical investigation will be necessary to reach confident conclusions. Although we have no basis to conclude that such data is missing, to the extent that it is, it would be missing from our sample. Finally, we note that this issue is not pertinent to what we report on published cases.

dataset includes 1,978 cases. We describe the data in further detail in the Appendix (Part I) and delineate which sets of cases above are included in which regressions reported below.

Although the idea is not new, there is growing awareness among scholars that, when studying the relationship between judge characteristics and their decision-making, we stand to learn different things from studies of random samples of all cases and studies of precedential cases.[100] The random sample of mostly nonprecedential decisions will allow us to assess the average effect of judge characteristics on disposition of the mine-run of 12(b)(6) appeals. It will answer the question of what the consequences are, on average, of the identity of the panel for the probability of a plaintiff prevailing on a 12(b)(6) appeal.

Precedential Court of Appeals decisions differ from nonprecedential decisions in important respects. On average, they are likely to raise more salient and non-routine legal issues, and they may not be representative of all litigated cases in other ways.[101] In addition to the possible unrepresentativeness of precedential decisions with respect to judicial behavior, there may be other important selection processes at play when analyzing only precedential decisions. The same judges that render precedential decisions also decide whether the opinion will be precedential, threatening to confound inferences about the relationship, in general, between judge characteristics and case outcomes when one studies only precedential decisions.[102] Studies finding a statistically significant relationship between a judge characteristic and an outcome in precedential cases may be explained by differences in voting that are confined to salient cases, or differences in publication behavior, or both, rather than differences in votes on outcomes in the full universe.

While recognizing this, we are interested, in part, in the influence of the ideology and identity characteristics of judges, if any, on the creation and development of law. Precedential Court of Appeals opinions are the vehicle through which circuits create and develop law that is binding on all subsequent panels and on all district courts in the circuit, while nonprecedential decisions are not binding. Even if some judge characteristic

---

[100] *See* Keith Carlson, Michael A. Livermore, and Daniel N. Rockmore, *The Problem of Data Bias in the Pool of Published U.S. Appellate Court Opinions*, 17 J. EMPIR. LEG. STUD. 224, 256 (2020).

[101] *See id.* at 230, 238–254; Sean Farhang, Jonathan P. Kastellec, and Gregory J. Wawro, *The Politics of Opinion Assignment and Authorship on the U.S. Court of Appeals: Evidence from Sexual Harassment Cases*, 44 J. LEG. STUD 59, 70 (2015); Theodore Eisenberg & Stewart J. Schwab, *What Shapes Perceptions of the Federal Court System?*, 56 U. CHI. L. REV. 501 (1989); David S. Law, *Strategic Judicial Lawmaking: Ideology, Publication, and Asylum Law in the Ninth Circuit*, 73 U. CIN. L. REV. 817 (2005); Deborah Jones Merritt & James J. Brudney, *Stalking Secret Law: What Predicts Publication in the United States Courts of Appeals*, 54 VAND. L. REV. 71 (2001).

[102] *See* Burbank & Farhang, *Politics & Identity*, *supra* note 3, at 251; Farhang et al., *supra* note 101, at 70–71; Carlson et al., *supra* note 100, at 230–31.

is not associated with outcomes in the full universe of assigned cases, it may still influence the direction of law's development through precedential opinions in ways that will affect future cases (including mine-run ones). Or the opposite may be true. Some judge characteristics may be consequential only when *not* making law. Thus, if we detect a relationship between a judge identity characteristic and outcomes in our random sample of 12(b)(6) motions, only with a parallel model of precedential cases will it be possible to evaluate whether it extends into the domain of lawmaking.

Finally, including analysis of precedential decisions contributes to comparison of our results with the leading works in the Court of Appeals literature on panel effects and diversity, which is based almost entirely on precedential decisions.[103] Unlike that literature, however, we present parallel models with a random sample of decisions to demonstrate whether and how results differ across the two sets of cases. We are not aware of any prior study to analyze both a random sample of cases and precedential cases covering the same issues, in the same court, and over the same period, allowing direct comparison. We find that in important respects the results across the two populations of cases do differ, sometimes in ways that are complex and unexpected.

A number of trial court studies seeking to evaluate the impact of *Twombly* and *Iqbal* either excluded or separately analyzed decisions turning on the legal sufficiency of a claim as distinguished from its factual sufficiency. The former, in their view, are unremarkable applications of pleading law that has been with us since 1938. The latter, they believe, are what plausibility pleading is all about. And since plausibility pleading was what interested them, they sought to segregate or exclude decisions grounded on legal insufficiency. These

---

[103] *See, e.g.,* Farhang & Wawro, *supra* note 60, at 310–311 (analyzing precedential employment discrimination cases); Boyd et al., *supra* note 70 (analyzing precedential cases across numerous policy areas; they analyze data described in SUNSTEIN ET AL., *supra* note 60, at 18); Cox & Miles, *supra* note 62, at 3 (analyzing precedential voting rights cases); Kastellec, *supra* note 60, at 173 (analyzing precedential affirmative action cases); Songer et al, *supra* note 63, at 430 (analyzing precedential employment discrimination, obscenity, and search and seizure cases); Sisk et al., *supra* note 61, at 534–35 (analyzing precedential religious liberty cases). *But see* Kastellec's recent paper on habeas death penalty decisions on the U.S. Court of Appeals, *supra* note 82, which draws on data from Jeffrey Fagan and James Liebman, *Processing and Outcome of Death Penalty Appeals After Furman v. Georgia, 1973-1995: [United States]* (ICPSR 3468) (2006), which endeavored to collect all such appeals. Carlson et al., *supra* note 100, at 230-3, observe that reliance on precedential opinions is dominant in the Court of Appeals literature on judicial behavior in general, not just in studies of diversity. This reliance on precedential cases by scholars doing work on periods before about 2002 was largely a function of necessity. Before the E-Government Act of 2002, neither complete nor representative samples of nonprecedential cases were accessible on electronic databases. *See* Solomon, *supra* note 99, at 203–215. Thus, studies including whatever nonprecedential cases were on electronic databases before around 2002 also suffer a significant selection threat analogous to that of publication. *See* Carlson et al., *supra* note 100, at 256; Peresie, *supra* note 60, at 1767 (including in her study precedential and nonprecedential employment discrimination cases from 1999-2001 that were available on Westlaw).

studies tend to regard legal sufficiency issues, at least in part, as threshold questions that could lead to dismissal without the court evaluating the merits of the complaint's factual allegations under governing substantive law.[104]

We elect not to take this approach and to simply analyze all 12(b)(6) appeals. We are not seeking to evaluate whether *Twombly* and *Iqbal* changed the dismissal behavior of judges, and thus isolating factual sufficiency issues is not necessary to our objectives. Further, our review of a large number of cases leads us to conclude that the distinction between legal and factual sufficiency is often highly ambiguous. Some legal sufficiency issues, we believe, are not threshold issues of the sort excluded by prior studies. And some threshold issues excluded by prior studies may in fact turn on factual pleadings that implicate *Twombly* and *Iqbal*. We are doubtful that one can systematically and objectively operationalize a legal sufficiency code in a data set as large as ours.

Further, when studying the relationship between judges' characteristics and their decisions, random assignment is critical to facilitate comparison of different judge types by allowing us to assume that they are deciding comparable claims. If the researcher filters out some cases after the random assignment has occurred based on disposition of threshold issues, selection may be introduced into the data. If the judge characteristics being studied are associated with how judges decide threshold legal sufficiency issues, then when one analyzes only cases in which judges evaluate the factual sufficiency of the pleadings, there is risk that the two types of judges are not deciding a comparable set of cases, confounding an inference that the judge characteristic explains observed difference in voting. Thus, we analyze all 12(b)(6) appeals after *Iqbal* and leave for future work analysis of the cases at a more granular level of reasoning.

Our unit of analysis is the claim, not the case. Cases often contain multiple claims; motions to dismiss are made with respect to claims, and both district courts and Courts of Appeals regularly conclude that a motion should be granted with respect to some claims and denied with respect to others. In order to code it, the court's Rule 12(b)(6) analysis in each decision was read in full. The random sample of 700 cases contained evaluation of 1,136 claims under Rule 12(b)(6) (about 1.6 claims per case). The oversample of 942 precedential cases contained evaluation of 1,794 claims under Rule 12(b)(6) (about 1.9 claims per case).

---

[104] *See* Scott Dodson, *A New Look at Pleading in Federal Civil Cases*, 96 JUDICATURE 127, 131 (2012); Alexander Reinert, *Measuring the Impact of Plausibility Pleading*, 101 VA. L. REV. 2117, 2139 (2015); *see also* Raymond H. Brescia & Edward J. Ohanian, *The Politics of Procedure: An Empirical Analysis of Motion Practice in Civil Rights Litigation under the New Plausibility Standard*, 47 AKRON L. REV. 329, 335 (2014); Quintanilla, *supra* note 53, at 33–34.

Our dependent variable is whether a decision is pro- or anti-plaintiff. We code a decision as anti-plaintiff (=0) if the Court of Appeals affirms the trial court's grant of a motion to dismiss or reverses the trial court's denial of a motion to dismiss. We code a decision as pro-plaintiff (=1) if the Court of Appeals reverses the trial court's grant of a motion to dismiss or affirms the trial court's denial of a motion to dismiss.

The task of measuring how a judge or panel characteristic may influence lawmaking is difficult. The most clearly observable manifestation of influence is a change in the probability that the appellant will prevail. However, much bargaining and deliberation among judges focuses on how to frame or justify a decision once it has been determined which party will prevail.[105] Such decisions about framing and justification can have important ramifications for the actual policy consequences of an opinion for future cases. Although we believe that our dependent variable captures much that is important to the development and application of the law governing dismissal under Rule 12(b)(6), we readily acknowledge its limits, which we regard as one cost of a large-N empirical study, as compared to a qualitative study that examines not just outcomes but also the scope and implications of reasoning.

This measurement constraint limits the inferences we can make from our data. If we find that some judge or panel characteristic is not associated with either pro- or anti-plaintiff decisions, we cannot conclude that the judge or panel characteristic has no directional influence on opinion content. On the other hand, flipping an outcome is a very strong form of influence, and thus to the extent that we find that a judge or panel characteristic is associated with a decision on dismissal in a particular direction, that characteristic is likely influencing opinion content in the same direction in more subtle ways.

For each case, we identified the party,[106] gender, and race of each judge using the Federal Judicial Center's biographical database.[107] With respect to race, we compare non-white judges to white judges.[108] The inferences we draw from the party, gender and race variables are based on the assumption that case assignment to panels is random, or "as-if" random, regarding the relationship between panel composition and the merits of the motion to dismiss.[109] We incorporate a battery of control variables that include a variety

---

105 *See* LEE EPSTEIN & JACK KNIGHT, THE CHOICES JUSTICES MAKE 9–14 (1998).

106 We use party of the appointing president as a proxy for judges' ideological preferences.

107 *Biographical Directory of Article III Federal Judges, 1789-Present*, FED. JUD. CTR., https://www.fjc.gov/history/judges [https://perma.cc/8HHF-JYZ9].

108 We are limited by sample size in our ability to analyze smaller racial subsets, such as African Americans and Hispanics. This is in part because we disaggregate the full data into smaller policy area subsets, and in part because the low rate of plaintiff wins in the random sample (15%) makes it difficult to estimate the effects of smaller racial subsets.

109 By "as-if" random we mean that departures from true random assignment of cases are inconsequential with respect to anything that would affect the outcomes studied. Levy and Chilton

of case characteristics, which are detailed in the Appendix (Part II). The models also contain circuit fixed effects and year fixed effects, the significance of which is also discussed in the Appendix (Part II).

B. *Policy Distribution of the Claims*

For all policy areas comprising 2% or more of the data, Table 1 shows the policy areas of the 1,136 claims underlying the motions to dismiss in our random sample of 700 cases, and of the 2,184 claims underlying the motions to dismiss in precedential cases occurring in the random sample and the oversample of precedential cases combined. The table excludes claims oversampled in the area of civil rights and discrimination in order that it be representative of the courts' 12(b)(6) docket. Because we round the percentages, the specific policy categories do not sum to 100 percent.

Other civil rights claims (excluding discrimination) will prove important in our empirical analysis below, and we thus provide additional information on them. They comprise 25 percent of the claims in the random sample and 31 percent of claims in precedential cases. When the data are collapsed down to the case level, 25 percent of cases in the random sample assert an other civil rights claim, and 28 percent of published cases do so. They overwhelmingly are constitutional claims against government. The three largest types of claims in this category are (1) policing, (2) public employment, and (3) prisoner. In the large residual category under other civil rights, the next largest seven areas are (4) judicial or prosecutorial misconduct, (5) education, (6) guns, (7) speech and religion, (8) family relations (primarily constitutional claims to parental rights), (9) voting and elections, and (10) privacy.[110] These ten policy areas are 89 percent of claims in the other civil rights category. Seventy-one percent of other civil rights

---

report the results of an empirical study finding small differences in the frequency with which circuits constituted panels with particular partisan configurations, such as panels with one Republican, or panels with two Democrats, relative to a scenario in which all panels were constituted purely by random draws from the circuit's slate of eligible judges. They suggest that such departures from randomness may arise from considerations of workload or judges' scheduling needs. *See* Adam S. Chilton & Marin K. Levy, *Challenging the Randomness of Panel Assignment in the Federal Courts of Appeals*, 101 CORNELL L. REV. 1, 5 (2015). The validity of inferences from our judge-characteristic variables do not depend on the assumption of random assignment of judges to panels, but rather on the assumption of random assignment of cases to panels once they are constituted. That is, we require the assumption that, for example, panels with two women, or panels with three Democrats, are not more likely to be assigned cases with stronger bases to deny the motion to dismiss. *See* Deborah Beim, Tom S. Clark & Benjamin E. Lauderdale, Random Assignment to Death 4 (Jan. 29, 2019) (unpublished manuscript).

   [110] Freedom of speech and religious liberty claims made by prisoners and public employees were coded as prisoner and public employment. Many public employment claims assert First Amendment rights.

claims are section 1983 damages actions. This is a large, cross-cutting, and important segment of the Court of Appeals 12(b)(6) docket.

Table 1: Policy Areas of Claims Underlying 12(b)(6) Decisions

	Random	Precedential
Civil Rights		
All Anti-Discrimination		
(race, gender, age, etc.)	13%	11%
Employment discrimination	10%	6%
Education discrimination	___	2%
Other discrimination		
(housing, voting, etc.)	2%	2%
All Other Civil Rights	25%	32%
Policing	9%	11%
Public Employment	5%	3%
Prisoner	2%	5%
Other	9%	12%
Non-Civil Rights	62%	57%
Consumer	15%	10%
Contract	12%	8%
Labor	7%	9%
Personal Injury	7%	6%
Antitrust	2%	3%
Securities	___	3%
Insurance	2%	2%
Intellectual Property	2%	2%
Other	14%	14%

C. *Panel Effects*

Our approach to the design of the statistical models presented in the next section is to assess the relationship between panel characteristics and claim outcomes rather than the relationship between an individual-level judge

characteristic and that judge's votes. There is a dissent rate of 2% in our cases, and thus outcomes are extremely highly correlated with votes. The literature shows that when Court of Appeals judges' party, gender, and race are associated with votes, their primary explanatory power is at the panel level, meaning that the composition of the panel often explains more variation in judges' votes than their own individual characteristics. For example, in many policy areas a Democratic appointee votes more liberally when sitting with two other Democratic appointees as compared to when sitting with two Republican appointees. The key point is that Court of Appeals judges' preferences (measured by characteristics) may influence outcomes by the way they influence the votes of co-panelists.[111]

The theoretical literature seeking to explain panel effects is built on the empirical fact that Court of Appeals panels are overwhelmingly unanimous even while we observe significant variation in case outcomes associated with panel composition.[112] On one account, unanimity may be driven by dissent avoidance by panel-minority judges who disagree with panel majorities but do not dissent because of workload pressures, strong norms against dissent, or the loneliness of dissent. These factors could lead to suppression of dissents on panels on which there is sincere disagreement, and the panel-majority view prevails without being influenced by the panel minority.[113] We use the phrase "panel minority" to refer to a minority position on a panel that has divided preferences, regardless of whether the judge is in a majority or minority group on the circuit.

Alternatively, the literature teaches, unanimity may be driven by panel minorities not dissenting because they *are* able to affect decisions. Mechanisms of influence include deliberation and bargaining, which allow panel minorities to change the preferences and/or votes of panel majorities.[114] As applied to minority-group judges, this view yields more positive normative implications than if they were suppressing dissents. It would allow minority-group preferences, when they differ systematically from majority-group preferences, to shape the application and development of law even when they are in the panel minority. By "minority group" we refer to groups of judges that are a numerical minority on the U.S. Courts of Appeals, such as women and nonwhites, regardless of their numbers in the general population. Multiple studies focused on civil rights cases have found that a single woman or racial minority can influence the votes of men and whites.[115]

---

[111] *See* Burbank & Farhang, *Politics & Identity*, *supra* note 3, at 240, and sources cited therein.

[112] *See id.* at 243, and sources cited therein.

[113] *See id.* at 244–-45, and sources cited therein.

[114] *See id.* at 246-49, and sources cited therein.

[115] *See id.* at 249, and sources cited therein.

Our recent study of class certification, however, shows that this is not always true. We found that one woman on a panel had no statistically discernable effect on the likelihood of a pro-certification outcome, but two women had a large effect in the pro-certification direction.[116]

D. *Statistical Models*

In our panel-level outcome model, the unit of analysis is the claim. In each case, we measure panel effects with dichotomous variables indicating whether the panel contained zero, one, two, or three Democrats; zero, one, two, or three women; and zero, one, two, or three racial minorities. Panels with three Democrats, panels with three men, and panels with three white judges are the reference categories for the party, gender, and race panel variables. This allows us to evaluate, for example, whether panels with one, two, or three Republicans have a statistically distinguishable probability of pro-plaintiff outcomes from an all-Democratic panel (the reference category), and if so, by what margin. All of the logistic regression models of pro-plaintiff outcomes reported in the Appendix (Part VI) contain the full set of these party, gender, and race panel variables and all of the control variables enumerated in the Appendix (Part II).

In the models presented in this section we impose the restriction of analyzing only cases in which at least one plaintiff is an individual person (or class of them), and at least one defendant is a business or governmental entity.[117] Seventy-eight percent of the claims in our random sample of 700 cases meet this criterion—95% of civil rights claims and 68% of non-civil rights claims. Most commonly, claims not meeting this criterion are business against business, or business against government. The dominant concerns about *Twombly* and *Iqbal* have focused on the pleading challenges faced by individuals suing business (as in *Twombly*) and government (as in *Iqbal*). Such cases are at the heart of our wider empirical investigation of judicial behavior and access to justice in the federal system. Further, this party structure restriction creates a much more plausible basis for testing preferences arrayed on a liberal/conservative continuum. For example, business against business commercial disputes are not often associated with expectations that judges' ideology, gender, or race will matter.

We report models: (1) combining all policy areas, and of (2) discrimination claims, (3) all other civil rights claims, and (4) non-civil rights claims. For each of these policy groupings, we run separate models for a

---

116 *See id.* at 261.

117 In the Appendix (Part III) we discuss models of all cases regardless of the party structure of the suit. Suits against individual persons for conduct undertaken on behalf of a business/government were coded business/government defendant cases.

random sample of (mostly non-precedential) cases, and for precedential cases. Table 2 shows plaintiffs' win rate in 12(b)(6) appeals with an individual suing business or government in each of the four policy groupings.

Table 2: Claim-Level Plaintiff Win Rate in 12(b)(6) Appeals, with Individual Suing Business or Government

	Random Sample	Precedential Only
All Cases	15%	31%
Civil Rights, Discrimination	16%	37%
Civil Rights, Other	15%	30%
Non-Civil Rights	14%	28%

In the random sample of claims by individuals against business or government, the plaintiff is the appellant 97% of the time. They win rarely—only 15% of the time.[118] Cases decided for the plaintiff are significantly more likely to be precedential, with about double the plaintiff win rate. In the random sample, we see no meaningful variation in win rates across the subcategories of discrimination, other civil rights, and non-civil rights. In precedential cases, plaintiffs are most successful in discrimination claims.

Before turning to the results, we note one important limitation. In many of our models we lack sufficient data to reach confident conclusions about the association between outcomes and panels with a *majority* of women or non-white judges. This problem is common in various policy subsets of the data (with fewer observations), particularly in the random sample. The relatively few observations of majority-women and non-white panels, combined with the low frequency of plaintiff wins in the random sample (15%), provide an insufficient basis for confident conclusions about such panels in many of the policy subset models. At the same time, a number of our models do have sufficient data to evaluate two-woman and two-racial minority panels—particularly models pooling all policy areas. Further, all the models can support strong inferences with respect to a single woman and single non-white because such panels are prevalent in the data. One-woman panels decide 47% of the claims in our random sample, and 45% of precedential claims. One-minority panels decide 43% of the claims in the random sample, and 38% of precedential claims.

The limited number of majority-woman and non-white panels in our data requires an interpretive caution. When panels with one woman or one non-

---

[118] The numbers are nearly identical (changing by less than one percentage point) when the full random sample is examined, without party restrictions.

white judge are not statistically distinguishable from all-male or all-white panels, and we lack sufficient data to confidently evaluate majority-woman and non-white panels, there remains an important indeterminacy. This pattern is consistent with gender and race being unassociated with preferences, or with women and non-whites having different preferences but failing to affect outcomes when they are in a minority. Only with sufficient data to evaluate the preferences of women and non-white judges when in the majority can we empirically adjudicate among these two possibilities.[119] Researchers may disagree about how many claims of a particular panel type in a model are sufficient to make a null result meaningful, and thus we provide those details in the Appendix (Part VI).[120]

In addition to the panel outcome-level models that we discuss below, in the Appendix (Part V) we also present parallel vote-level models. We only report as significant results that are robust across the outcome-level and vote-level models and that satisfy other robustness checks described in the Appendix (Parts IV-V). The Appendix (Part I) also details which samples are used in each model.

1. Party

RANDOM SAMPLE. All of the party variables are insignificant in the random sample of cases. This is true for all policy areas combined,[121] discrimination claims,[122] all other civil rights claims,[123] and all non-civil rights clams.[124] Notwithstanding all of the controversy about *Twombly* and *Iqbal*, and the prospect that they would introduce greater subjectivity and ideology into disposition of 12(b)(6) motions, we detect no party association with outcomes in the random sample of cases, even when comparing all-Republican to all-Democratic panels.[125] Party is consequential, however, in precedential cases.

PRECEDENTIAL CASES. In the model of precedential cases spanning all policy areas, some party variables are significant.[126] Because logit coefficients are not directly interpretable, we compute predicted probabilities that specific panel-types will render a pro-plaintiff outcome. All-Democratic panels have a 38% probability of a pro-plaintiff outcome.[127] Adding one Republican

---

119 *See* Burbank & Farhang, *Politics & Identity*, *supra* note 3, at 250, 267–68.

120 Table A-10.

121 Table A-1, Model C.

122 Table A-2, Model A.

123 Table A-2, Model C.

124 Table A-2, Model E.

125 Party remains insignificant in alternative specifications of the models reported in notes 121 to 124 that substitute a dichotomous variable comparing Democratic to Republican majority panels.

126 Table A-1, Model D.

127 Table A-5, Model D.

(DDR) is not statistically distinguishable.[128] The Democratic majority does not give ground to one Republican. When Republicans form a majority on RRD panels, however, the probability of a plaintiff win declines to 24%.[129] The variable measuring all-Republican panels is negative and approaches but does not achieve significance.[130] In light of the significance of the DDR variable, it appears to us that the RRR variable is insignificant because we have many fewer RRR than RRD panels in the data. In an alternative specification of an otherwise identical model, we combine RRR and RRD panels into a Republican majority variable, which is statistically significantly different from all-Democratic panels (p=.011) with a 26% probability of ruling for the plaintiff. Subsetting the data by policy area will reveal how this party association with outcomes is distributed across policy areas.

The party variables are not significant in either the discrimination model[131] or the other civil rights model[132] of precedential cases, consistent with their insignificance in both civil rights models in the random sample. We also examined a random sample model, as well as a precedential model, of all civil rights claims (combining discrimination and other civil rights), significantly increasing the sample size. All the party variables remain clearly insignificant in both models.[133]

Together, these null results for party in civil rights claims are contrary to our expectations, and, we think, contrary to what the literature on *Iqbal* seems to anticipate. They are also contrary to what the judicial behavior literature on the Courts of Appeals (which ignores procedural posture) would lead one to expect. That literature has found widespread associations between party and case outcomes across the landscape of civil rights, and we do not doubt those results. However, it appears that when the question is narrowed to whether civil rights plaintiffs have stated a claim sufficient to proceed to discovery, Democratic and Republican appointees decide similarly. This is true even when the data are restricted to more salient cases that make law (precedential decisions).

---

[128] Table A-1, Model D.

[129] Table A-1, Model D; Table A-5, Model D.

[130] Table A-1, Model D.

[131] Table A-2, Model B. We have the least data in our discrimination models. In an alternative specification, we pooled discrimination claims in the random sample and the precedential model into an omnibus discrimination model, yielding 614 claims arising in 368 cases. In that model all of the party variables remained insignificant.

[132] Table A-2, Model D.

[133] The random sample of combined civil rights claims (discrimination and other) contained 739 claims occurring in 424 cases, and the set of combined precedential civil rights claims contained 915 claims in 465 cases. We examined alternative specifications of these models substituting a dichotomous variable comparing Democratic to Republican majority panels, and it was consistently insignificant.

In contrast, in the model of precedential non-civil rights claims some party variables are significant.[134] This non-civil rights category covers about half the data, with the most common policy areas, in order, being consumer, contract, labor, personal injury, antitrust, and securities. All-Democratic panels have a 48% probability of a pro-plaintiff outcome.[135] Adding one Republican (DDR) is not statistically distinguishable.[136] The Democratic majority does not give ground to one Republican. However, when Republicans form a majority on RRD panels, the probability of a plaintiff win declines sharply, to 16%.[137] As in the precedential model spanning all policy areas, the all-Republican panel variable is negative and approaches but does not achieve significance, which we attribute to the limited number of RRR panels. In an alternative specification in an otherwise identical model, we combine RRR and RRD panels into a Republican majority variable, which is statistically significant (p=.002) with a 20% probability of ruling for the plaintiff. Thus, there is a decline of 28-percentage points in the probability of a pro-plaintiff outcome moving from an all-Democratic to a majority-Republican panel.

In sum, party is strongly associated with outcomes in precedential non-civil rights claims, but not precedential civil rights claims. In the random sample, party is not associated with outcomes in the disposition of 12(b)(6) appeals either in aggregate or in any of our policy subsets.

2. Gender

RANDOM SAMPLE. Controlling for the partisan and racial composition of the panel, and the variables listed in the Appendix (Part II), panels with one woman are significantly more likely to rule for the plaintiff in the random sample of other civil rights claims (excluding discrimination).[138] All-male panels have a 10% probability of a pro-plaintiff outcome.[139] The addition a single woman panel-affects the two males and increases the probability to 21%,[140] about doubling it. This cross-cutting set of civil rights claims comprises 1 in 4 cases on the Courts of Appeals' 12(b)(6) docket. Forty-eight percent of claims in the random sample are decided by panels with one woman, and thus this result reflects a broad impact of gender on plaintiff's

---

134   Table A-2, Model F.
135   Table A-6, Model F.
136   Table A-2, Model F.
137   Table A-2, Model F; Table A-6, Model F.
138   Table A-2, Model C. As discussed in the Appendix (Part IV), the result is robust in a model designed for rare event outcomes.
139   Table A-6, Model C.
140   Table A-6, Model C.

access to discovery when the defendant has moved to dismiss in other civil rights cases.

The variable measuring two-woman panels is insignificant in the model of other civil rights claims,[141] but we lack sufficient data to be confident in this null result.[142] All of the gender variables are insignificant in the other random sample models—all policy areas combined,[143] discrimination claims,[144] and non-civil rights clams.[145] We do have sufficient data to be confident in the null results for panels with two women in the larger model of all policy areas pooled, but not in the much smaller policy subsets.[146]

PRECEDENTIAL CASES. Turning to the models of precedential cases, in other civil rights claims (excluding discrimination) panels with one woman are again significantly more likely than all-male panels to rule for the plaintiff, associated with a growth in the probability of a pro-plaintiff ruling from 24% to 37%.[147] The variable measuring panels with two women is insignificant in this model, but we again have insufficient data to be confident in the null result.[148]

Gender is also significant in precedential non-civil rights claims, but only when women are in the majority. Panels with one woman do not have a significantly different probability of ruling in a pro-plaintiff direction as compared to all-male panels, while panels with two and three women do. All-male panels have a 21% probability of a pro-plaintiff outcome. When women form a majority sitting with one man, the probability grows substantially to 41%, and on all-woman panels it grows to 58%.[149] The model contains only 19 claims decided by all-women panels, and thus we put no stock in this result. In contrast, the model contains 115 claims decided by panels with two women. In an alternative specification we created a majority-woman variable (134 claims, arising in 79 cases) and substituted it for the variables measuring two and three women panels. It was highly statistically significant (p=.004), and such panels had a 43% probability of a pro-plaintiff ruling, more than double that of panels with three men.

To evaluate the robustness of the result, we randomly oversampled an additional 25 cases with non-civil rights claims asserted by individuals against

---

141 Table A-2, Model C.

142 *See supra* notes 119-120 and accompanying text; Table A-10.

143 Table A-1, Model C.

144 Table A-2, Model A.

145 Table A-2, Model E. Panels with two women barely crosses the .1 threshold with a negative sign, but this result is insignificant in the vote-level model (Table A-4, Model E), and we treat as significant only results that are significant across our alternative specifications.

146 *See supra* notes 119-120 and accompanying text; Table A-10.

147 Table A-2, Model D; Table A-6, Model D.

148 *See supra* notes 119-120 and accompanying text; Table A-10.

149 Table A-6, Model F.

business or government, and decided by majority-woman panels. With the oversample, the model contains 164 claims arising in 104 cases decided by majority-woman panels. In the claim-level model with the oversample, the significance level (.004) and probability of a pro-plaintiff outcome (43%) for majority-woman panels were the same as those reported in the last paragraph. In our vote-level analysis yielding the same substantive results, the model contains 492 votes cast on majority-woman panels.[150]

All of the gender variables are insignificant, or lack robust significance, in the other precedential models—when all policy areas are combined,[151] and in the subset with only discrimination claims.[152] We do have sufficient data for confidence in the null result for panels with two women in the larger model of all policy areas, but not in the much smaller discrimination subset.[153]

The structure of gender panel effects differs moving from other civil rights to non-civil rights claims in the precedential models. In the civil rights context (excluding discrimination), we observed that a single woman changed the votes of her male colleagues relative to how they vote on all-male panels. In non-civil rights claims, a single woman on a panel has no discernable impact on the voting behavior of male colleagues. This is so even though women judges have much more pro-plaintiff preferences (revealed when they are in the majority).

It is unclear why one woman affects a male majority in civil rights cases but not non-civil rights cases, although the literature suggests several potential (and speculative) explanations. One focuses on intensity of preferences. Sunstein et al. argue that the impact, if any, of judges in the preference-minority on a three-judge panel will vary across policy domains and is less likely to occur when judges have intense preferences, such as capital punishment and abortion, undercutting the ability of panel-minorities to influence majorities.[154] This view suggests that the presence or magnitude of panel effects will be negatively associated with the panel-majority's

---

[150] *See* Appendix (Part V).

[151] Table A-1, Model D; Table A-5, Model D. Although the variable measuring panels with one woman is significant, it is not robust in all alternative specifications discussed in the Appendix (Part V), in particular the vote-level model with standard errors clustered on case. When the regressions are disaggregated by policy area (Table A-2), it seems clear that these suggestive results are driven mainly by other civil rights claims. And while panels with three women are significant across all models, this is based on only 25 claims arising in 12 cases, and thus the result is not credible.

[152] Table A-2, Model B. All gender variables remain insignificant when discrimination claims from the random sample and precedential models are pooled. *See supra* note 131.

[153] *See supra* notes 119-120 and accompanying text; Table A-10.

[154] *See* SUNSTEIN ET AL., *supra* note 60, at 62–63, 69–70. Using different methods and improved data, Professor Fischman finds that there are panel effects in capital punishment and abortion cases. Joshua B. Fischman, *Interpreting Circuit Court Voting Patterns: A Social Interactions Framework*, 31 J.L. ECON. & ORG. 808, 829 (2015). Our point here is conceptual and not about Sunstein et al.'s specific empirical claims.

preference intensity—more intense preferences will be associated with more majoritarian panel behavior.

Our results for party, discussed above, show that party composition of panels is not significantly associated with 12(b)(6) dispositions in either discrimination or other civil rights cases. In contrast, our results for party in precedential non-civil rights cases show that they are the subject of significant ideological voting. This is some evidence that many judges have more intense preferences in non-civil rights cases than in civil rights cases when disposing of 12(b)(6) appeals. One possibility is that one woman panel-affects male judges in other civil rights cases because the male judges lack intense preferences, while a single woman is unable to discernably panel-effect male majorities in non-civil rights cases, where preference intensity is higher. This would be consistent with the fact (noted above) that one Republican does not panel-effect two Democrats in the same non-civil rights model.

Another mechanism offered to explain panel effects is "cue taking." Cue taking is a dynamic whereby some judges, seeking an efficient path to rendering a decision, show greater deference to other judges in issue domains in which they are perceived to be more credible or expert.[155] Studies finding race and gender panel effects in civil rights cases have proposed cue taking, or something akin to it, as a possible explanatory mechanism.[156] The structure of gender panel effects that we observe could be explained if male judges perceive women as more credible or expert in deciding 12(b)(6) appeals in the area of civil rights, but not more generally.[157]

Among non-civil rights claims, policy areas comprising more than 2% of our cases, in order, are consumer, contract, labor, personal injury, antitrust, securities, and intellectual property. Like our class certification study, the results contradict conventional wisdom in the literature that women judges'

---

[155] DAVID E. KLEIN, MAKING LAW IN THE UNITED STATES COURTS OF APPEALS 31 (2002).

[156] *See, e.g.,* Peresie, *supra* note 60, at 1783–84 (suggesting cue taking as an explanation for one-woman panel effect); Kastellec, *supra* note 60, at 171–72 (observing that the mere presence of an African American in an affirmative action case, independent of the content of deliberations, and independent of her vote, may influence the behavior of whites on the panel); Boyd et al., *supra* note 70, at 392 n.8 (suggesting the same possibility with respect to gender, and likening this to cue taking).

[157] The panel effects literature's invocation of cue taking in the context of gender is undertheorized, lacking clear criteria for specifying when a male judge would perceive a woman as possessing expertise worthy of deference. The cross-cutting set of other civil rights claims is heterogeneous and does not include claims expressly based on discrimination. To the extent that gender-based cue taking explains one-woman panel effects in other civil rights cases, the cue taking is occurring at a very high level of generality, which does not strike us as very likely. In our class action study we further invoked ideas from the literature on the gender gap in deliberative decision-making when considering possible explanations for why one African American panel-affected white/other judges in the majority, but in the same set of cases women influenced outcomes only when in the majority. *See* Burbank & Farhang, *Politics & Identity*, *supra* note 3, at 269–72. That pattern is not present in our pleading data.

preferences differ from men's only in cases implicating discrimination (or even civil rights more broadly). By shifting the focus to procedural posture we learn that when making law in non-civil rights claims, women in the majority exercise gatekeeping powers differently than men. They are more likely to make law that gives plaintiffs the opportunity to gather evidence in an effort to prove their claims.

In sum, in other civil rights claims (excluding discrimination), panels with one woman are materially more likely to rule in a pro-plaintiff direction in both the random sample and in the precedential model. Majority-woman panels, but not panels with one woman, are also substantially more likely to do so in non-civil rights claims in precedential cases, where they are more than twice as likely to rule for the plaintiff as compared to all-male panels.

3. Race

RANDOM SAMPLE. Controlling for the partisan and gender composition of the panel, and the variables listed in the Appendix (Part II), in the random sample of other civil rights cases the variable measuring panels with one non-white is highly statistically significant (p=.001), and its magnitude is large.[158] The result is consistent with the worries of some of *Iqbal*'s critics. All-white panels have an 11% probability of rendering a pro-plaintiff outcome.[159] The probability more than doubles, growing to 25%, when there is one non-white judge on the panel.[160] Viewed conversely, as compared to panels with one non-white judge, assignment of an all-white panel cuts plaintiffs' chances of reaching discovery by more than half. This important cross-cutting category of civil right cases represents a quarter of the federal appellate 12(b)(6) docket. Forty-three percent of claims in the random sample are decided by panels with one non-white, and thus this result reflects a broad impact of race on plaintiff's access to discovery when the defendant has moved to dismiss in other civil rights cases. The variable measuring panels with two non-white judges is insignificant, but we lack sufficient data for confidence in the null results for this panel type.[161]

All of the race variables are insignificant, or lack robust significance, in the remaining random sample models—discrimination claims[162] (the area in

---

[158] Table A-2, Model C. As discussed in the Appendix (Part IV), the result is robust in a model designed for rare event outcomes.

[159] Table A-6, Model C.

[160] *Id.*

[161] Table A-2, Model C; *see supra* notes 119-120 and accompanying text; Table A-10.

[162] Table A-2, Model A. Although the two and three non-white variables are significant in the claim-level model, neither result is credible. The three non-white variable is based on only six causes of action in three cases, and thus the result is meaningless. The two non-white result is based on only 21 cases/26 claims with two non-whites in the model, and this is clearly insufficient to produce

which some expected race to be especially salient under *Iqbal*), non-civil rights claims, and all policy areas pooled.[163] With respect to panels with two non-whites, only the model of cases pooled across all policy areas has sufficient data to allow confidence in the null results.[164]

PRECEDENTIAL CASES. In the models of precedential claims, all of the race variables are insignificant in every model—claims pooled across all policy areas,[165] discrimination claims,[166] other civil rights claims,[167] and non-civil rights claims.[168] We have insufficient data to be confident in the null results for panels with two non-whites in any of these models.[169]

The clear insignificance of the variable measuring panels with one non-white judge in the precedential model of other civil rights claims alongside its strong significance in the random sample is striking. The conventional wisdom is, we believe, that if voting or outcome variation exists that is associated with judge characteristics in the full universe of cases, it will be present in precedential cases in light of their higher salience and legal consequences. This is an important assumption because, although nonprecedential cases far outnumber precedential cases,[170] the judicial behavior literature on the Courts of Appeals is based almost entirely on precedential cases.[171] Null results in these studies are regularly understood to mean that white versus racial minority judges, or men versus women, or Democrats versus Republicans, are not associated with directionally different votes or outcomes. Our results show that the assumption underlying this inference is sometimes false. They highlight a significant risk of inferential

---

a reliable estimate. *See supra* notes 119-120 and accompanying text; Table A-10. Further, the result becomes insignificant in the vote-level models, which is a key robustness check that we discuss in the Appendix (Part V). Throughout the paper we report as significant only results that are significant in both the claim-level and vote-level models.

163 Table A-1, Model C (all cases); Table A-2, Model E (non-civil rights cases). In the original random sample of 700 cases, panels with two non-whites are significant in these models. However, these results are based on only 39 cases/62 claims and 24 cases/41 claims with two non-whites, respectively. To evaluate the robustness of the result we randomly oversampled cases with two non-white judges so that we had a total of 115 cases/196 claims and 76 cases/136 claims, respectively. The two non-white judge variables became clearly insignificant in both models.

164 *See supra* notes 119-120 and accompanying text; Table A-10. When the oversample is added (discussed in the last footnote), the null result for two non-whites in non-civil rights cases also looks reliable.

165 Table A-1, Model D.

166 Table A-2, Model B. All race variables remain insignificant when discrimination claims from the random sample and precedential models are pooled. *See supra* note 131.

167 Table A-2, Model D.

168 Table A-2, Model F.

169 *See supra* notes 119-120 and accompanying text; Table A-10.

170 *See* Law, *supra* note 101; Merrit & Brudney, *supra* note 101.

171 *See supra* note 103 and accompanying text.

error when drawing general inferences from null results in precedential cases, as is widespread in the Court of Appeals literature.

Although this result may strike some as counterintuitive, recent empirical research on differences between precedential and nonprecedential cases suggests a possible explanation. Carlson, Livermore, and Rockmore find that single-party panel opinions (all-Democratic or all-Republican) are overrepresented in the universe of published decisions.[172] Grunwald offers a strategic theory of publication that is consistent with this pattern: "when the judges cannot reach a compromise [on opinion content], they may exchange publication for unanimity."[173] On average, panels with heterogeneous preferences, coupled with the norm of unanimity on the Courts of Appeals, are characterized by greater need to bargain and compromise,[174] and the question of precedential status may be one element in that compromise.[175] A panel majority may give ground to a panel minority to attain unanimity, drawing opinion content away from their preferences and making the opinion less attractive as a vehicle to make law.[176]

In contrast, when a panel has homogeneous preferences, on average panel decisions will be closer to the preferences of all members, presenting a better opportunity to make law. This dynamic would provide an explanation for Carlson et al.'s finding that single-party panels are overrepresented in precedential cases. It also offers a plausible explanation for why we observe panel effects for one-nonwhite sitting in a random sample of (mostly non-precedential) civil rights cases, but they disappear in precedential cases.

In sum, race is significant in only one model. However, that model is an important one—the random sample of all other civil rights claims. The presence of one non-white on a panel more than doubles a plaintiff's chances of success. For reasons that are unclear, non-white judges' pro-plaintiff influence in the random sample of these cases does not extend to precedential cases.

CONCLUSION

Critics of *Twombly* and *Iqbal* worried that the new 12(b)(6) standard would introduce excessive subjectivity and ideology into disposition of 12(b)(6) motions. With our data we cannot compare pre-*Twombly* to post-*Iqbal* decision-making, but we can evaluate the degree to which the party (of

---

[172] *See* Carlson *et al.*, *supra* note 100, at 239–241; *see also* Ben Grunwald, *Strategic Publication*, 92 TUL. L. REV. 744 (2018).

[173] *See* Grunwald, *supra* note 172, at 759.

[174] *See* Burbank & Farhang, *Politics & Identity*, *supra* note 3, at 242–44, 246–49.

[175] *See* Grunwald, *supra* note 172, at 759.

[176] *See id.* at 760.

appointing president), gender and race of panel members are associated with their disposition of 12(b)(6) appeals in the age of *Iqbal*. Like the judicial behavior literature writ large, especially as applied to race and gender, our results are a patchwork that varies across identity characteristic, policy area, and random sample versus precedential cases.

In our random sample of cases, judges' race and gender were associated with outcomes in other civil rights claims (excluding discrimination), which are overwhelmingly constitutional claims against governmental actors, commonly arising in such areas as policing, prisons, and public employment. In this cross-cutting civil rights category, amounting to a quarter of 12(b)(6) appeals, panels with one racial minority or one woman are more than twice as likely, as compared to all-white and all-male panels, to render a decision favoring a plaintiff's opportunity to gather evidence in support of her claim. Party is clearly insignificant in these models. This is the first Court of Appeals study that we are aware of in which gender and racial panel composition are associated with outcomes but party composition is insignificant. Gender and race are not simply amplifying ideology as measured by party; they are consequential where party is not. With eighty-six percent of claims in the data having some degree of gender or racial diversity on the panel, and with influence on outcomes occurring when women and non-whites are in the panel-minority, diversity on the Courts of Appeals has a broad impact on plaintiffs' ability to reach discovery in other civil rights claims. The remaining judge characteristic variables in the other random sample models are insignificant, with the important caveat that we lack sufficient data to evaluate panels with two women or non-whites in many of the regressions.

The results in precedential cases differ in important respects. Deciding the case at hand does not exhaust the ways that judges' preferences shape case outcomes, and it may not even be the most important way. Law plays an important role in determining how routine appeals are decided. Indeed, the substance of law is part of what makes a claim routine—it is unambiguously covered by a rule. And we found somewhat more associations between judge characteristics and lawmaking in a pro-plaintiff direction when disposing of 12(b)(6) appeals through precedential decisions.

Democratic panels were significantly more likely to decide in favor of plaintiffs in precedential non-civil rights cases even though this was not true in the random sample. In fifty-five percent of precedential cases there are non-civil rights claims asserted, and thus this party result covers about half of the data, pooled across many policy areas. Party was insignificant, however, in precedential discrimination claims and other civil rights claims. Party therefore was not associated with outcomes in civil rights claims across the board, in both the random sample and precedential models. Party matters

least (or not at all) in the policy area in which many scholars (including us) expected it to matter most: civil rights.

Panels with women in the majority were more likely to decide precedential non-civil rights claims in favor of plaintiffs, although panels with one woman were not. In contrast, panels with one woman were more likely to decide other civil rights claims (excluding discrimination) in favor of plaintiffs. Thus one woman panel-affected the male majority in other civil rights claims, but not in non-civil rights claims. Race was insignificant in every precedential model. In the decade after *Iqbal,* during which the Courts of Appeals were grappling with elaboration of notoriously ambiguous pleading doctrine, judges' ideology and gender mattered to how they made law. Our models do not detect any race associations with lawmaking under *Iqbal*, but we do not claim that our data forecloses their existence. We lacked the data to evaluate non-white majorities in precedential civil rights cases (discrimination or other), and we cannot speak to many more specific issue areas.

Our results for gender, in combination with similar results in our class certification study, contradict conventional wisdom in the literature that women judges' preferences differ from men's only in cases implicating discrimination. They add further evidence to the possibility, noted in our class action study, that transsubstantive procedural law affecting access to justice may itself be a policy domain in which women have different (more pro-access) preferences.

Our results also offer some significant lessons about the relationship between judicial behavior in precedential versus random sample cases, although the lessons are hard to decipher. What can be said with confidence is that researchers should be cautious in drawing general inferences about judicial behavior on the Courts of Appeals from precedential decisions. This lesson is important because the Court of Appeals literature on judge attributes is based overwhelmingly on precedential cases, which, in aggregate, are a small fraction of Court of Appeals decisions.

What is (to us) unexpected is the inconsistency of the relationship. Party and gender were significant in precedential non-civil rights cases but not in the random sample. Sometimes a statistically significant relationship exists in precedential cases when in fact there is no relationship in the full universe. As a result, existing studies finding significant relationships between judge characteristics and case outcomes in precedential cases do not warrant the inference that the relationship exists in the full universe. Such results may be measuring patterns in salient cases, or publication behavior, or some combination. We suspect that few will be surprised by this result.

More surprisingly (at first blush), the presence of one non-white on a panel is strongly associated with plaintiff wins in the random sample of other civil

rights cases, but clearly insignificant in the precedential model. A judge characteristic may be insignificant in published cases when in fact a relationship exists in the full universe. We are aware of no work in the judicial behavior literature that anticipates this result. A simple and plausible explanation is that publication is partly a strategic decision to embed legal views in law, and the white majority may be less likely to give precedential status in cases in which they make concessions to (are panel affected by) non-white judges.

This result highlights an important caution in interpreting existing Court of Appeals studies reporting null results for gender and race based on precedential cases (virtually the entire literature). Such studies have routinely inferred from null results that no relationship exists between a judge characteristic and votes or outcomes. They are the basis of the widely held view that race and gender rarely matter on the Court of Appeals. If we had collected only precedential cases, we would have reported that race was clearly insignificant across the board. The results in our random sample of other civil rights cases show that this would have been error. Null results in studies of precedential cases cannot support the inference that race and gender are not associated with decisions in the full universe of cases, although they have been widely so interpreted. Viewing all the results together, we conclude that the existing literature, based overwhelmingly on precedential cases, leaves many more questions unanswered than previously understood.

APPENDIX

I. SAMPLES USED IN EACH REGRESSION

In Part III.A we describe the samples of 12(b)(6) appeals we collected. Below we describe which samples were used in each regression reported, and how many cases the sample contained. This indicates the number of cases underlying the number of causes of action (the unit of analysis) listed as the sample size in the regression tables.

- The models designated "Random, All Policy Areas" in Table A-1 (Model A), and "Random, Non-Civ Rts" in Table A-2 (Model E), are based on the random sample of 700 cases. The former model uses all such cases, and the latter model uses 271 cases.
- The model designated "Random, Other Civ Rts" in Table A-2 (Model C) is based on other civil rights claims in (1) the random sample of 700 cases, and (2) the random oversample of an additional 206 civil rights cases. The model uses 302 cases.
- The model designated "Random, Discrim" in Table A-2 (Model A) is based on discrimination claims in (1) the random sample of 700 cases, (2) the random oversample of an additional 206 civil rights cases, and (3) the random oversample of an additional 130 cases asserting discrimination claims. The model uses 238 cases.
- The models designated "Preced., All Policy Areas, All Parties" in Table A-1 (Model B), "Preced., All Policy Areas, Pty Rest" in Table A-1 (Model D), and "Preced., Non-Civ Rts" in Table A-2 (Model F), are based on (1) the Westlaw headnote precedential cases, and (2) precedential decisions in the random sample of 700 cases. There are 1,178 cases in the first model, 852 cases in the second, and 472 cases in the third.
- The model designated "Preced., Other Civ Rts" in Table A-2 (Model D) is based on other civil rights claims in (1) the Westlaw headnote precedential cases, (2) precedential cases in the random sample of 700 cases, and (3) precedential cases in the random oversample of an additional 206 civil rights cases. The model uses 352 cases.
- The model designated "Preced., Discrim" in Table A-2 (Model B) is based on discrimination claims in (1) the Westlaw headnote precedential cases, (2) precedential cases in the random sample of 700 cases, (3) precedential cases in the random oversample of an additional 206 civil rights cases, and (4) precedential cases in the random oversample of an additional 130 discrimination cases. The model uses 151 cases.

II. CLAIM-LEVEL MODEL SPECIFICATIONS

The unit of analysis in our primary models, discussed in the body of the Article, is the claim. We ran logit models with standard errors clustered on case because multiple claims within the same case are not independent of one another. In all of the statistical models reported below, the following control variables are included:

- *TRIAL COURT OUTCOME*: Indicator variable reflecting whether the trial court granted or denied the motion to dismiss on the claim that is under consideration by the Court of Appeals.
- *TRIAL JUDGE SITTING BY DESIGNATION*: Indicator variable recording whether there was a trial judge sitting by designation on the panel.
- *DEFENDANT TYPE*: Non-mutually exclusive indicator variables measuring whether there was a federal defendant, state defendant, business defendant, or other type of defendant.
- *LAW TYPE:* Mutually exclusive indicator variables measuring whether the claim was under federal law, state law, or both.
- *POLICY AREA:* Mutually exclusive indicator variables reflecting policy areas comprising 2% or more of the data. Policy areas comprising less than 2% of the data were aggregated into an "other" policy category.
- *CIRCUIT FIXED EFFECTS*: Circuit fixed effects (dichotomous variables for each circuit) account for any time-varying covariates that take the same value for each judge on a panel within the circuit.
- *YEAR FIXED EFFECTS*: Year fixed effects (dichotomous variables for each year) account for any time-varying covariates that take the same value for each judge on a panel within the year.

Circuit fixed effects account for any variables that change across circuits and that would take the same value for each judge on a panel within that circuit, such as circuit doctrine that may have a pro- or anti-dismissal slant and variation in the size and content of caseloads across circuits. Year fixed effects account for any variables that change over time and that would take the same value for each judge on a panel within that year, such as national trends in caseload, the evolution of Supreme Court doctrine, changing composition of the Supreme Court, changes in Federal Rules, and salient features of the partisan or political environment, such as an anti-litigation posture in a party agenda. They also account for trends over time in attitudes among male and white judges toward co-panelists who are women and racial minorities, which may affect the extent to which the former are influenced by the latter. The circuit and year fixed-effects approach leverages only variation in the relationship between panel characteristics and outcomes *within* circuit and year. This approach allows us to estimate the effects of panel characteristics most effectively because it controls for the influence of any

variables that would take the same value for each panel in the same circuit and each panel in the same year.

III. Model Interpretation

In Part III.D of the Article we discuss interpretation of all models except Models A and B in Table A-1. Those are models, which we discuss here, of all policy areas combined, but with no party restriction imposed. In the random sample, all of the party, gender, and race variables are insignificant or not robustly significant.[177] Though the variable measuring panels with two non-whites is significant, as discussed in Part III.D.3, it is based on insufficient data to allow confidence in the result, and when we randomly oversample such panels to address this concern, the variable becomes clearly insignificant.[178]

In the precedential cases, two of the party variables are significantly distinguishable from all-Democratic panels.[179] All-Democratic panels have a 38% probability of a pro-plaintiff outcome.[180] Panels with one Republican approach but do not achieve statistically significant difference from all-Democratic panels. However, majority-Republican panels (RRD and RRR) are significantly different from all-Democratic panels, and both are associated with a 14-percentage point reduction in the probability of a pro-plaintiff outcome, which they render with a probability of 24%.[181]

Panels with one woman and panels with three women are both significantly more likely to render pro-plaintiff outcomes than all-male panels, but the variable measuring panels with two women is insignificant.[182] However, the result for all-woman panels is not meaningful because it is based on only 39 causes of action arising in 15 cases. All-male panels have a 27% probability of a pro-plaintiff outcome, and the addition of a single woman panel-affects the two males and increases the probability modestly, to 33%.[183] All of the race variables are insignificant in this model.[184]

IV. Firth Models for Rare Events

Because plaintiff wins in our random sample are relatively infrequent (15%), we evaluated the robustness of the significant results in these models

---

177 Table A-1, Model A.
178 *See supra* note 164.
179 Table A-1, Model B.
180 Table A-5, Model B.
181 Table A-1, Model B; Table A-5, Model B.
182 Table A-1, Model B.
183 Table A-5, Model B.
184 Table A-1, Model B.

using Firth logit modes designed for "rare events."[185] Those significant results were panels with one woman, and panels with one non-white, in the random sample of other civil rights claims. They are robustly significant at comparable levels in the Firth models.

V. VOTE-LEVEL MODELS

We only characterize results as significant in the body of the paper if they are also present in individual vote-level models. An individual vote-level model of panel effects requires that we disaggregate the party, gender, and race variables. Our approach is to create variables that capture the identity of the voting judge and the characteristics of her colleagues on the panel. This requires that each variable measuring a characteristic (party, gender, race) be disaggregated into six variables. Table A-9 defines each of the six variables associated with party, gender, and race panel effects. Although the table is labored, the information is necessary in order to understand the textured information conveyed by the regression models. For each set of indicator variables, we designate the reference category with the "reference" parenthetical.

In the vote-level models reported, we ran logit models with standard errors clustered on claim because multiple votes on the same claim are not independent of one another. We also examined alternative specifications with standard errors clustered on case, and on judge, and obtained consistent results for results found to be significant across the outcome-level models and our vote-level models with standard errors clustered on claim.

Because the vote-level models are consistent with the claim-level models, we do not rehash substantive interpretation of all models. Instead, we interpret only the significant gender and race results in order to illustrate how to read the tables. For the gender variables, a man voting on an all-male panel is the reference category. In the random sample of other civil rights claims, the variables measuring the votes of a male judge sitting with one man and one woman, and the votes of a woman sitting with two men, are both statistically significant.[186] These are votes occurring on panels with one woman. A man voting on an all-male panel has a 12% predicted probability of ruling for the plaintiff. A man serving with one woman and one man has a 23% probability, and a woman serving with two men has a 27% probability—about triple that of a man serving on an all-male panel.[187]

---

[185] *See* David Firth, *Bias Reduction of Maximum Likelihood Estimates*, 80 BIOMETRIKA 27–38 (1993).
[186] Table A-4, Model C.
[187] Table A-8, Model C.

This one-woman panel effect in other civil rights claims is also present in the precedential model.[188] A male serving on an all-male panel votes in favor of the plaintiff with a probability of 24%. The probability grows to 38% for males serving with one man and one woman, and to 41% for one woman serving with two men.[189]

Moving to precedential non-civil rights cases, men and women voting on MMW panels are not statistically distinguishable from men voting on all-male panels.[190] However, when women assume a majority in MWW panels, both the men and women judges are significantly more likely to vote for the plaintiff as compared to men on all-male panels.[191] Men voting on all-male panels vote in favor of the plaintiff 21% of the time. For men sitting with two women the probability is 40%, and for women sitting with one man and one woman it is 39%.[192] A woman serving on an all-woman panel votes in a pro-plaintiff direction with a probability of 54%, but as noted in Part III.D.2 there are too few such cases (19) to credit this result. In an alternative specification we pool women's votes on MWW and WWW panels and find that when women are in the majority they have a 50% probability of voting in a pro-plaintiff direction in precedential non-civil rights cases—about two and a half times the probability for a male on an all-male panel. This result is based on 402 votes on 134 claims arising in 79 cases decided by majority-woman panels.

To evaluate the robustness of the result, we randomly oversampled an additional 25 cases with non-civil rights claims asserted by individuals against business or government and decided by majority-woman panels. With the oversample the model contains 492 votes on 164 claims arising in 104 cases decided by majority-woman panels. The result is robust in the vote-level model with the oversample, where the variable measuring a woman voting on majority-woman panels is significant (p=.007) with a predicted probability of 49%.

For the race variables, a white judge voting on an all-white panel is the reference category. In the random sample of other civil rights claims, the variables measuring the votes of a white judge sitting with one white and one non-white judge, and the votes of one non-white sitting with two white judges, are both statistically significant.[193] These are votes occurring on panels with one non-white judge. A white judge voting on an all-white panel has a 12% predicted probability of ruling for the plaintiff. A white serving with one white and one non-white judge has a 26% probability, and a non-white judge

---

188 Table A-4, Model D.

189 Table A-8, Model D.

190 Table A-4, Model F.

191 *Id.*

192 Table A-8, Model F.

193 Table A-4, Model C.

serving with two white judges has a 29% probability—more than triple that of a white judge serving on an all-white panel.[194]

---

[194] Table A-8, Model C.

VI. TABLES

TABLE A-1: LOGIT MODEL OF CLAIM-LEVEL PANEL EFFECTS IN 12(B)(6) OUTCOMES
POOLING ACROSS POLICY AREAS

	Model A Random All Policy All Parties	*Model B* Preced All Policy All Parties	*Model C* Random All Policy Pty Rest	*Model D* Preced All Policy Pty Rest
PARTY				
3 Democrats (reference)				
1 Rep, 2 Dems	-.16 (.35)	-.28 (.22)	-.09 (.41)	-.23 (.25)
2 Reps, 1 Dem	-.17 (.39)	-.52** (.22)	-.27 (.46)	-.60** (.26)
3 Reps	.04 (.45)	-.47* (.28)	.20 (.54)	-.40 (.34)
GENDER				
3 Men (reference)				
1 Wom, 2 Men	.22 (.23)	.25* (.15)	.24 (.26)	.34* (.17)
2 Wom, 1 Man	-.24 (.34)	.11 (.20)	-.41 (.38)	.37 (.25)
3 Wom	-.89 (.82)	.79* (.45)	-1.24 (.87)	1.41** (.59)
RACE				
3 White (reference)				
1 NW, 2 White	.18 (.22)	.02 (.15)	.42 (.27)	.11 (.18)
2 NW, 1 White	.81** (.36)	.01 (.26)	1.09** (.45)	.10 (.31)
3 NW	——	-.37 (1.19)	——	-.16 (1.33)
All models include circuit fixed effects, year fixed effects, and independent variables measuring policy area, direction of the trial court outcome, trial judge sitting by designation, defendant type (federal government, state government, business, other), and law type (federal law, state law, both). Random sample models additionally contain a variable indicating whether the case was published. "Pty Rest" in Models C & D indicates that the models contain only claims in which an individual (or class or them) sues a business or government defendant.				
N=	1,117	2,171	852	1605
Pseudo R²=	.13	.10	.15	.13
***p < .01; **p < .05; *p < .1 Standard errors in parentheses, clustered on case.				

TABLE A-2: LOGIT MODEL OF CLAIM-LEVEL PANEL EFFECTS IN 12(B)(6) OUTCOMES BY
POLICY SUBSET, IN CLAIMS WITH INDIVIDUAL PLAINTIFFS (INCLUDING CLASSES) SUING
BUSINESS OR GOVERNMENT

	Model A Random Discrim-ination	Model B Preced Discrim-ination	Model C Random Other Civ Rts	Model D Preced Other Civ Rts	Model E Random Non-Civ Rts	Model F Preced Non-Civ Rts
PARTY						
3 Democrats (reference)						
1 Rep, 2 Dems	.16 (.56)	-.08 (.52)	-.12 (.53)	.11 (.47)	-.09 (.49)	-.39 (.33)
2 Reps, 1 Dem	.17 (.66)	-.01 (.54)	-.17 (.59)	-.20 (.44)	-.15 (.57)	-1.17*** (.38)
3 Reps	-1.23 (.86)	.27 (.81)	.63 (.65)	.26 (.54)	-.16 (.66)	-.63 (.46)
GENDER						
3 Men (reference)						
1 Wom, 2 Men	-.35 (.45)	.07 (.50)	1.07** (.47)	.71** (.29)	-.23 (.40)	.27 (.26)
2 Wom, 1 Man	.17 (.50)	-.18 (.63)	.03 (.51)	-.10 (.40)	-.83* (.49)	.81** (.35)
3 Wom	——	.97 (1.93)	1.50 (1.03)	.40 (1.72)	1.07 (1.05)	1.59** (.65)
RACE						
3 White (reference)						
1 NW, 2 White	.07 (.39)	.32 (.47)	1.39*** (.43)	.11 (.31)	.22 (.38)	.23 (.24)
2 NW, 1 White	1.00* (.54)	.09 (.66)	-.30 (.92)	-.69 (.51)	1.00* (.55)	.15 (.48)
3 NW	2.28* (1.17)	——	——	——	——	-.18 (1.46)
All models include circuit fixed effects, year fixed effects, and independent variables measuring policy area, direction of the trial court outcome, trial judge sitting by designation, defendant type (federal government, state government, business, other), and law type (federal law, state law, both). Random sample models additionally contain a variable indicating whether the case was published. All models contain only claims in which an individual (or class or them) sues a business or government defendant.						
N=	406	246	507	702	422	789
Pseudo R^2=	.24	.17	.30	.19	.14	.15
***p < .01; **p < .05; *p < .1 Standard errors in parentheses, clustered on case.						

TABLE A-3: LOGIT MODEL OF VOTE-LEVEL PANEL EFFECTS IN 12(B)(6) OUTCOMES
POOLING ACROSS POLICY AREAS

	Model A Random All Policy All Parties	Model B Preced All Policy All Parties	Model C Random All Policy Pty Rest	Model D Preced All Policy Pty Rest
PARTY				
D with 2Ds (All Democrats) (reference)				
D with 1D & 1R	-.19 (.31)	-.19 (.17)	-.15 (.36)	-.12 (.20)
R with 2Ds	-.23 (.31)	-.39** (.17)	-.18 (.36)	-.41** (.21)
D with 2Rs	-.26 (.33)	-.40** (.18)	-.36 (.38)	-.46** (.22)
R with 1R & 1D	-.19 (.33)	-.60*** (.18)	-.26 (.38)	-.70*** (.22)
R with 2Rs (All Reps)	-.05 (.40)	-48** (.21)	.12 (.48)	-.39 (.26)
GENDER				
M with 2 Ms (All Men) (reference)				
M with 1M & 1Wm	.19 (.21)	.19* (.11)	.17 (.24)	.24* (.14)
Wm with 2Ms	.20 (.21)	.19 (.11)	.20 (.24)	.27** (.14)
M with 2Wm	-.44 (.33)	.04 (.17)	-.53 (.36)	.31 (.21)
Wm with 1Wm & 1M	-.19 (.31)	.03 (.16)	-.35 (.34)	.22 (.20)
Wm with 2Wm (All Women)	-.94 (.95)	.66* (.37)	-1.30 (1.50)	1.19** (.50)
RACE				
W with 2 Ws (All White) (reference)				
W with 1W & 1NW	.10 (.21)	-.03 (.11)	.28 (.24)	.07 (.14)
NW with 2Ws	.23 (.21)	.01 (.12)	.43* (.24)	.07 (.14)
W with 2NWs	.79** (.33)	-.19 (.22)	1.02*** (.40)	.14 (.28)
NW with 1NW & 1W	.74 (.32)	.05 (.21)	.97** (.39)	.20 (.26)
NW with 2NWs (All Non-White)	—	-.34 (1.15)	—	-.10 (1.28)

	Model A Random All Policy All Parties	Model B Preced All Policy All Parties	Model C Random All Policy Pty Rest	Model D Preced All Policy Pty Rest
All models include circuit fixed effects, year fixed effects, and independent variables measuring policy area, direction of the trial court outcome, trial judge sitting by designation, defendant type (federal government, state government, business, other), and law type (federal law, state law, both). Random sample models additionally contain a variable indicating whether the case was published. "Pty Rest" in Models C & D indicate that the models contain only claims in which an individual (or class or them) sues a business or government defendant.				
N=	3,351	6,512	2,556	4,814
Pseudo R²=	.13	.10	.15	.12
***p < .01; **p < .05; *p < .1 Standard errors in parentheses, clustered on claim.				

TABLE A-4: LOGIT MODEL OF VOTE-LEVEL PANEL EFFECTS IN 12(B)(6) OUTCOMES BY
POLICY SUBSET IN CLAIMS WITH INDIVIDUAL PLAINTIFFS (INCLUDING CLASSES) SUING
BUSINESS OR GOVERNMENT

	Model A *Random* *Discrimination*	*Model B* *Preced* *Discrimination*	*Model C* *Random* *Other Civ Rts*	*Model D* *Preced* *Other Civ Rts*	*Model E* *Random* *Non-Civ Rts*	*Model F* *Preced* *Non-Civ Rts*
PARTY						
D with 2Ds (All Democrats) (reference)						
D with 1D & 1R	.37 (.57)	.08 (.55)	-.15 (.47)	.29 (.39)	-.01 (.46)	-.28 (.29)
R with 2Ds	.09 (.56)	-.13 (.55)	-.37 (.51)	-.14 (.40)	-.14 (.46)	-.50* (.29)
D with 2Rs	.55 (.60)	.28 (.54)	-.02 (.49)	.06 (.37)	-.13 (.55)	-1.04*** (.34)
R with 1R & 1D	.23 (.62)	.05 (.55)	-.37 (.51)	-.31 (.37)	-.06 (.54)	-1.28*** (.34)
R with 2Rs (All Reps)	-1.24* (.72)	.36 (.74)	.21 (.56)	.28 (.43)	-.21 (.66)	-.64* (.39)
GENDER						
M with 2 Ms (All Men) (reference)						
M with 1M & 1Wm	-.33 (.42)	.09 (.44)	.83** (.39)	.62*** (.23)	-.20 (.38)	.20 (.21)
Wm with 2Ms	-.70 (.43)	-.03 (.44)	1.05*** (.39)	.68*** (.23)	-.21 (.38)	.27 (.21)
M with 2Wm	.07 (.47)	-.27 (.64)	-.24 (.49)	.01 (.35)	-.75 (.48)	.70** (.31)
Wm with 1Wm & 1M	.12 (.46)	-.20 (.63)	-.07 (.45)	-.32 (.35)	-.62 (.46)	.70** (.31)
Wm with 2Wm (All Women)	⎯	.88 (2.00)	1.16 (1.08)	.40 (1.41)	.94 (1.00)	1.43*** (.53)
RACE						
W with 2 Ws (All White) (reference)						
W with 1W & 1NW	-.13 (.35)	.23 (.45)	1.09*** (.39)	.05 (.23)	.06 (.35)	.17 (.20)
NW with 2Ws	-.11 (.35)	.40 (.45)	1.23*** (.39)	-.06 (.23)	.11 (.35)	.18 (.21)
W with 2NWs	.89 (.61)	-.10 (.70)	-.51 (.81)	-.78* (.46)	.95* (.51)	.08 (.43)
NW with 1NW & 1W	.88 (.57)	.30 (.70)	-.53 (.81)	-.23 (.38)	.94* (.51)	-.11 (.42)
NW with 2NWs (All Non-White)	1.95 (1.31)	⎯	⎯	⎯	⎯	-.18 (1.37)

	Model A Random Discrimination	Model B Preced Discrimination	Model C Random Other Civ Rts	Model D Preced Other Civ Rts	Model E Random Non-Civ Rts	Model F Preced Non-Civ Rts
All models include circuit fixed effects, year fixed effects, and independent variables measuring policy area, direction of the trial court outcome, trial judge sitting by designation, defendant type (federal government, state government, business, other), and law type (federal law, state law, both). Random sample models additionally contain a variable indicating whether the case was published. All models contain only claims in which an individual (or class or them) sues a business or government defendant.						
N=	1,218	737	1,519	2,115	1,266	2,367
Pseudo R^2=	.24	.17	.28	.17	.13	.15
***p < .01; **p < .05; *p < .1 Standard errors in parentheses, clustered on claim.						

2186 Random

TABLE A-5: CLAIM-LEVEL PREDICTED PROBABILITIES OF 12(B)(6) OUTCOMES POOLING
ACROSS POLICY AREAS FOR PARTY, GENDER, AND RACE PANEL COMBINATIONS

	Model A *Random* *All Policy* *All Party*	*Model B* *Preced* *All Policy* *All Party*	*Model C* *Random* *All Policy* *Pty Rest*	*Model D* *Preced* *All Policy* *Pty Rest*
"___" indicates that panel type is not statistically distinguishable from the reference category (in italics), or not significant in both individual and vote-level models. ** indicates no cases in model.				
PARTY				
3 Dems	14%	38%	13%	38%
1 Rep, 2 Dems	___	___	___	___
2 Reps, 1 Dem	___	24%	___	24%
3 Reps	___	24%	___	___
GENDER				
3 Men	13%	27%	13%	27%
1 Wom, 2 Men	___	33%	___	34%
2 Wom, 1 Man	___	___	___	___
3 Wom	___	46%	**	59%
RACE				
3 White	11%	30%	10%	30%
1 NW, 2 White	___	___	___	___
2 NW, 1 White	___	___	___	___
3 Non-White	**	___	**	___

TABLE A-6: CLAIM-LEVEL PREDICTED PROBABILITIES OF 12(B)(6) OUTCOMES IN POLICY
SUBSETS FOR PARTY, GENDER, AND RACE PANEL COMBINATIONS

	Model A Random Discrim- ination	*Model B Preced Discrim- ination*	*Model C Random Other Civ Rts*	*Model D Preced Other Civ Rts*	*Model E Random Non- Civ Rts*	*Model F Preced Non- Civ Rts*
"___" indicates that panel type is not statistically distinguishable from the reference category (in italics), or not significant in both individual and vote-level models. ** indicates no cases in model.						
PARTY						
3 Dems	14%	38%	17%	29%	16%	48%
1 Rep, 2 Dems	___	___	___	___	___	___
2 Reps, 1 Dem	___	___	___	___	___	16%
3 Reps	___	___	___	___	___	___
GENDER						
3 Men	18%	36%	10%	24%	17%	21%
1 Wom, 2 Men	___	___	21%	37%	___	___
2 Wom, 1 Man	___	___	___	___	___	41%
3 Wom	**	**	___	___	___	58%
RACE						
3 White	12%	34%	11%	34%	10%	26%
1 NW, 2 White	___	___	25%	___	___	___
2 NW, 1 White	___	___	___	___	___	___
3 NW	___	**	**	**	**	___

TABLE A-7: VOTE-LEVEL PREDICTED PROBABILITIES OF 12(B)(6) OUTCOMES POOLING
ACROSS POLICY AREAS FOR PARTY, GENDER, AND RACE PANEL COMBINATIONS

	Model A Random All Policy All Parties	*Model B Preced All Policy All Parties*	*Model C Random All Policy Pty Rest*	*Model D Preced All Policy Pty Rest*
"___" indicates that panel type is not statistically distinguishable from the reference category (in italics), or not significant in both individual and vote-level models. ** indicates no cases in model.				
PARTY				
D with 2Ds (All Democrats)	17%	37%	17%	41%
D with 1D & 1R	___	___	___	___
R with 2Ds	___	24%	___	24%
D with 2Rs	___	24%	___	23%
R with 1R & 1D	___	22%	___	21%
R with 2Rs (All Reps)	___	23%	___	___
GENDER				
M with 2Ms (All Men)	13%	28%	14%	27%
M with 1M & 1Wm	___	32%	___	33%
Wm with 2Ms	___	___	___	35%
M with 2Wm	___	___	___	___
Wm with 1Wm & 1M	___	___	___	___
Wm with 2Wm (All Women)	___	43%	___	54%
RACE				
W with 2 Ws (All White)	11%	30%	10%	30%
W with 1W & 1NW	___	___	___	___
NW with 2Ws	___	___	___	___
W with 2NWs	___	___	___	___
NW with 1NW & 1W	___	___	___	___
NW with 2NWs (All Non-White)	**	___	**	___

TABLE A-8: VOTE-LEVEL PREDICTED PROBABILITIES OF 12(B)(6) OUTCOMES IN POLICY SUBSETS FOR PARTY, GENDER, AND RACE PANEL COMBINATIONS

	Model A Random Discrim- ination	Model B Preced Discrim- ination	Model C Random Other Civ Rts	Model D Preced Other Civ Rts	Model E Random Non- Civ Rts	Model F Preced Non- Civ Rts
"___" indicates that panel type is not statistically distinguishable from the reference category (in italics), or not significant in both individual and vote-level models. ** indicates no cases in model.						
PARTY						
D with 2Ds (All Democrats)	13%	32%	19%	25%	16%	48%
D with 1D & 1R	___	___	___	___	___	___
R with 2Ds	___	___	___	___	___	21%
D with 2Rs	___	___	___	___	___	15%
R with 1R & 1D	___	___	___	___	___	14%
R with 2Rs (All Reps)	___	___	___	___	___	20%
GENDER						
M with 2Ms (All Men)	18%	34%	12%	24%	17%	21%
M with 1M & 1Wm	___	___	23%	38%	___	___
Wm with 2Ms	___	___	27%	41%	___	___
M with 2Wm	___	___	___	___	___	40%
Wm with 1Wm & 1M	___	___	___	___	___	39%
Wm with 2Wm (All Women)	**	___	___	___	___	54%
RACE						
W with 2 Ws (All White)	16%	34%	12%	30%	15%	26%
W with 1W & 1NW	___	___	26%	___	___	___
NW with 2Ws	___	___	29%	___	___	___
W with 2NWs	___	___	___	___	___	___
NW with 1NW & 1W	___	___	___	___	___	___
NW with 2NWs (All Non-White)	___	**	**	**	**	___

TABLE A-9: VOTE-LEVEL PANEL VARIABLES FOR PARTY, GENDER, AND RACE

Variable Description	Variable Name
Democratic Majority Panels	
Democrat voting with 2 other Democrats (reference)	**D with 2Ds**
Democrat voting with 1 Democrat & 1 Republican	D with 1D & 1R
Republican voting with 2 Democrats	R with 2Ds
Republican Majority Panels	
Democrat voting with 2 Republicans	D with 2Rs
Republican voting with 1 Republican & 1 Democrat	R with 1R & 1D
Republican voting with 2 other Republicans	R with 2Rs
Male Majority Panels	
Man voting with 2 other Men (reference)	**M with 2Ms**
Man voting with 1 Man & 1 Woman	M with 1M & 1Wm
Woman voting with 2 Men	Wm with 2Ms
Female Majority Panels	
Man voting with 2 Women	M with 2Wm
Woman voting with 1 Woman & 1 Man	Wm with 1Wm & 1M
Woman voting with 2 other Women	Wm with 2Wm
White/Other Majority Panels	
White voting with 2 other Whites (reference)	**W with 2Ws**
White voting with 1 White & 1 Non-White	W with 1W & 1NW
Non-White voting with 2 Whites	NW with 2Ws
Non-White Majority Panels	
White voting with 2 Non-Whites	W with 2NWs
Non-White voting with 1 Non-White & 1 White	NW with 1NW & 1W
Non-White voting with 2 other Non-Whites	NW with 2NWs

TABLE A-10: NUMBER OF CLAIMS AND CASES WITH A MAJORITY OF WOMEN OR NON-WHITE JUDGES

Regressions	2 Wm COAs	3 Wm COAs	2 Wm Cases	3 Wm Cases	2 NW COAs	3 NW COAs	2 NW Cases	3 NW Cases
Table A-1, Model A: Random All, No Party Restrictions (PR)	206	22	128	13	82	10	55	3
Table A-1, Model B: Preced All, No PR	329	39	185	15	134	3	88	3
Table A-1, Model C: Random All, PR	156	10	94	8	62	9	39	2
Table A-1, Model D: Preced All, PR	219	25	125	12	81	2	53	2
Table A-2, Model A: Random, Discrim, PR	83	7	42	6	26	6	21	3
Table A-2, Model B: Preced, Discrim, PR	34	2	22	2	15	1	14	1
Table A-2, Model C: Random, Other Civil Rights, PR	88	6	48	6	34	0	24	0
Table A-2, Model D: Preced, Other Civil Rights, PR	95	5	53	4	29	0	18	0
Table A-2, Model E: Random, Non-Civil Rights, PR	83	3	61	3	41	9	24	2
Table A-2, Model F: Preced, Non-Civil Rights, PR	115	19	71	8	48	2	30	2

*   *   *   *   *