Justice-Level Heterogeneity in Certiorari Voting: U.S. Supreme Court October Terms 1939, 1968, and 1982.*

Gregory A. Caldeira† Daniel Lempert‡

March 18, 2021

Abstract

Though the literature on agenda-setting at the U.S. Supreme Court is sizable, justice-vote-level multivariate analyses of certiorari are almost exclusively limited to samples of discussed cases from 1986–1993. Moreover, these studies have done very little to explore justice-level heterogeneity on certiorari. Here, we present an initial effort to address these lacunae by analyzing the predictors of individual justices' cert votes on all paid dockets from the 1939, 1968, and 1982 terms. We find substantial justice-level heterogeneity in the weight that justices place on the standard factors shaping the cert vote. We also show that some of this heterogeneity is associated with justice experience and ideological extremism, largely in theoretically predicted ways. In closing, we sound a note of caution on drawing conclusions about effects of justice attributes, when the number of justices is relatively small.

*Revision of version presented at the 2020 Western Political Science Association Annual Conference, where we received thoughtful comments from Joanna Wuest and other attendees. We thank Larry Baum and Jeff Budziak for helpful feedback and Audra J. Smith for excellent research assistance. For use of data for OT 1982, we thank John Wright.

†Distinguished University Professor, Political Science, Ohio State University. email: caldeira.1@osu.edu.

‡(corresponding author) Associate Professor of Politics, SUNY Potsdam. email: dalempert@gmail.com.

Though the literature on agenda-setting at the U.S. Supreme Court is sizable, justice-vote-level analyses of certiorari are limited in several respects. First, multivariate justice-vote-level analyses have largely been based on samples of cases from 1986–1993 (Benesh, Armstrong II, and Wallander 2020 [625 cases]; Black and Boyd 2012*a* [305 cases]; Black and Boyd 2012*b* [447 cases]; Black and Owens 2009 [358 cases]; Johnson 2018 [c. 300 cases]). As Lane and Black (2017, 4) points out, this is because data from these terms are readily available as part of the Digital Archive of the Papers of Harry A. Blackmun (Epstein, Segal and Spaeth 2007).

The limited temporal scope of these analyses is concerning for two reasons. More obviously, there are questions about generalizability—whether the patterns demonstrated plausibly apply to other periods in the Court's history. This is not only of historical interest. Our lack of knowledge about the over-time stability of findings limits what we can say about how likely it is that the results demonstrated continue to hold *today*.[1]

There is a less obvious implication of agenda-setting studies' narrow temporal scope: difficulty disentangling justice *attributes* from justice *identities*. The concern is that the fewer justices there are included in the sample, the greater the risk that observed justice-level characteristics are confounded with unobserved, potentially idiosyncratic factors associated with individual justices. The risk of such confounding is particularly high for covariates that vary minimally or not at all *within justices*, during a given time period—for example, ideology and related variables. With only nine or so justices in a sample, effects attributed to these variables may be driven by a single justice; when this occurs, it is effectively impossible to determine whether one should credit the justice-specific covariate or simply a given justice's idiosyncracies. Increasing the temporal scope of an analysis, and thus the number of justices included, reduces—though, as we will see, does not eliminate—the chance that this ambiguity arises.

Perhaps even more important than the limited temporal scope of the modern literature on agenda-setting is that studies have done very little to explore justice-level heterogeneity

[1]Of course, it is near-impossible to study the contemporary Court's agenda-setting directly, given the essential role of justices' private papers—typically released years after retirement or death, if ever—as a data source.

on certiorari. Sorely lacking are systematic examinations of how the effects of factors known to influence the Court's collective cert decision vary across justices—whether as individuals or as a function of observed characteristics. While some research has assessed the impact of a few individual-level variables on the cert vote, little-to-no work has sought to explain how individual-level variables (including justice fixed effects) condition the effects of case-level variables on the cert vote. Our primary objective here is to determine the extent of justice-level variability; secondarily, we seek to explain this variability as a function of justice attributes.

Finally, we note another issue related to sampling: typically, in the studies mentioned above, cases are drawn from a set of dockets that have made the Discuss List—at the least, this raises potential inferential difficulties associated with selection bias. By considering all paid dockets in our analyses—whether discussed or not—we avoid these statistical difficulties, and in addition present a more complete picture of the Court's cert decision.

Background

The major historical shift in Supreme Court agenda-setting was the Judges' Bill of 1925, which gave the Court substantial control over its own docket for the first time. Thereafter, the Court promulgated formal rules about the criteria for granting certiorari, i.e., discretionary review of lower court decisions.

These criteria have remained largely unchanged over the terms we cover, 1939, 1968, and 1982. They are specified in Rule 10 today (Rule 38 in 1939), and are familiar. As Lane and Black (2017, 5) summarizes, the Court prefers cases involving "decisions issued by either state supreme courts or U.S. courts of appeals that conflict with either state courts of last resort, U.S. circuit courts, or the Supreme Court," as well as "decisions in which either state supreme courts or federal circuit courts decide an issue that has never been settled by the Supreme Court," and, finally, decisions wherein "one of these courts departs from the accepted and

usual course of judicial proceedings."[2]

Of course, scholars have long recognized that the formal rules do not constitute a complete explanation of certiorari. The Court denies petitions that appear to meet the rule's criteria and grants others, absent those criteria. Clearly, other, informal forces are at work. Nonetheless, these formal criteria guide petitioners, as suggested by the leading volume on Supreme Court practice (Stern and Gressman 1950; Shapiro, Geller, Bishop, Hartnett and Himmelfarb 2019; see also Robinson and Kirkham 1936).

Though unanimous denials are the most common outcome, there is often a fair amount of disagreement among justices about whether a given case is "cert-worthy" (Caldeira, Wright and Zorn 1999). These votes, however, are not made public, except in the rare cases where one or more justices publicly dissents from a denial of certiorari, and even then one cannot necessarily infer that all those who did not *publicly* dissent were in favor of a grant. Thus, as scholars we have access to individual votes only if preserved in justices' archived papers. Necessarily, this has led to the rather limited temporal scope of the research on Court agenda-setting.

[2]The relevant language of today's Rule 10 refers to cases where:

> (a) a United States court of appeals has entered a decision in conflict with the decision of another United States court of appeals on the same important matter; has decided an important federal question in a way that conflicts with a decision by a state court of last resort; or has so far departed from the accepted and usual course of judicial proceedings, or sanctioned such a departure by a lower court, as to call for an exercise of this Court's supervisory power;
>
> (b) a state court of last resort has decided an important federal question in a way that conflicts with the decision of another state court of last resort or of a United States court of appeals;
>
> (c) a state court or a United States court of appeals has decided an important question of federal law that has not been, but should be, settled by this Court, or has decided an important federal question in a way that conflicts with relevant decisions of this Court.

Similarly, in 1939, Rule 38 stated the criteria for review from federal courts as:

> "(1) a circuit court of appeals has rendered a decision in conflict with the decision of another circuit court of appeals on the same matter; (2) or has decided an important question of local law in a way probably in conflict with the weight of authority; (3) or has decided an important question of federal law which has not been, but should be, settled by this court; (4) or has decided a federal question in a way probably in conflict with applicable decisions of this court; or (5) has so far departed from the accepted and usual course of judicial proceedings, or so far sanctioned such a departure by a lower court, as to call for an exercise of this court's power of supervision."

Literature

The foundational piece in the modern literature on Supreme Court agenda-setting is Caldeira and Wright (1988). In a multivariate, Court-level analysis, it demonstrates that certiorari is more likely to be granted if (1) the United States is a petitioner, (2) the court below reverses the decision below *it*, (3) the legal question is associated with a live inter-circuit or other conflict enumerated in Rule 10, (4) the decision below is ideologically discordant with the Court[3], and (5) amicus briefs are present on certiorari—whether in favor of in favor of review or in opposition. Caldeira and Wright (1990) additionally demonstrates that a dissent in the court below increases the likelihood of a cert grant, though not if the analysis is limited to discussed cases. The effects of these "standard factors" have been confirmed repeatedly in subsequent research, some based on justice-vote level data (see in particular Schoenherr and Black (2019), but also, e.g., Black and Boyd (2012*b*), Black and Owens (2009), Black and Owens (2012), and McGuire and Caldeira (1993)).

Yet, as we noted above, the extent to which the effects of these standard factors may vary, that is, be heterogeneous, has not been systematically explored. The occasional study has assessed whether an isolated factor is conditional on a case-level variable (see Black and Boyd (2012*b*, 300), which argues, though does not formally test, that the effect of amicus briefs is conditioned by litigant status). But explorations of whether the factors influencing a justice's vote are heterogeneous as a function of a justice's identity or characteristics have been limited to studies of decisions on the merits (e.g., Bartels 2011; Collins 2008*b*; Segal and Spaeth 1996).

This is despite a growing (e.g., Rainey 2016)—though depending on one's perspective, still limited (Feller and Holmes 2009; Lam 2013)—literature in political science concerned with effect heterogeneity. Thus, in an initial effort to explore potential heterogeneity, we examine whether justices place different weights on the standard factors, using data on all paid cert petitions to the Court in October Terms (OT) 1939, 1968, and 1982.

[3]To be precise, Caldeira and Wright (1988) demonstrates that liberal decisions below are more likely to be reviewed, but observes that the Court at the time had a "decided [...] conservative ideological orientation."

Data

For the 1939 Term, for information on the relatively objective, descriptive features of cases, we relied on petitions and briefs in opposition in *Supreme Court Briefs and Records* (Hein), lower court opinions, and *United States Law Week*. For data on subjective indicators, we read and coded the "cert memoranda" in the papers of William O. Douglas (Library of Congress), primarily, and, in the absence of them, those in the papers of Stanley F. Reed (University of Kentucky).

For the 1968 Term, we gathered our data from *Records and Briefs of the Supreme Court* on microfiche (issue area, presence of amici curiae, the United States as a petitioner or respondent, reversals between or dissents in the lower courts), *United States Law Week* (dates of actions, resolutions of cases), and, for conflict and allegations of conflict, Thurgood Marshall's certiorari memoranda. The data for the OT 1982 term are from, and as described in, Caldeira and Wright (1990). Individual cert votes are as recorded in the docket books of Douglas (for OT 1939), Warren (1968), and Brennan (1982).

The operationalization of our variables accords with usual practice (see e.g., Caldeira and Wright 1990), and we describe the definitions concisely. Our outcome variable, *Grant* is whether a justice voted to grant(= 1) or deny (= 0) a cert petition.

U.S. Petitioner is coded 1 if the "United States," a federal agency, or its representative (in an official capacity) is one of the petitioners. *Intermediate Reversal* is coded 1 if the court immediately below (nearly always either a federal court of appeal or state supreme court) reversed the lower court (usually a trial court, less often an agency).

Actual Conflict is coded 1 if the case involved a square conflict between between two or more cases, occurring in different circuits, state supreme courts, or between the lower court and the Supreme Court—i.e., the conflicts enumerated in Supreme Court Rule 10 (formerly Rule 38). For OT 1939 and 1968, we relied on assessments in certiorari memoranda (primarily those of Douglas and Marshall); for OT 1982, we relied on the *New York University Law Review*'s Supreme Court Project (*NYU Law Review* Supreme Court Project 1984*a*; *NYU Law Review*

Supreme Court Project 1984*b*) to code this variable. *Alleged Conflict* is coded 1 if the clerk (OT 1939 and 1968) or the *NYU Law Review*'s Supreme Court Project (OT 1982) notes that a petitioner has alleged a conflict enumerated in Rule 10.

Civil Liberties is coded 1 if the primary substantive issue area of the case involves civil liberties (including criminal procedure and civil rights). If amicus curiae in support or opposed to certiorari appear in a case, we code *Amicus Present* as 1, and 0 otherwise. *Dissent Below* is coded 1 if, in the court immediately below the Supreme Court, one or more judges dissented.

Finally, *Incompatible Decision Below* is coded 1 if a conservative justice is reviewing a liberal decision below or a liberal justice is reviewing a conservative decision below. Otherwise, it is coded 0. We generally follow the coding rules from Spaeth (2001) to categorize the ideological direction of a lower court decision.[4] We classify a justice as liberal if, conditional on covariates, she was more likely to grant a conservative case below than a liberal one. Correspondingly, we classify a justice as conservative if, conditional on covariates, he was more likely to favor review of a liberal decision below than a conservative one. In practice, this operationalization means that the effect of Incompatible Decision Below in our analyses is constrained to be non-negative; of course, the effect's *magnitude*, which we are primarily interested in, is unaffected by the operationalization (i.e., can take on any possible value).

How Much Justice-Level Heterogeneity?

First, we examine the extent to which individual justices place different weights on the standard factors that have been shown to influence the Court's collective vote on certiorari. There is some reason to expect heterogeneity, as anecdotal accounts suggest that justices are not homogenous in how they approach the cert vote. Perhaps the most-cited example is Justice White's strong inclination to grant cases involving inter-circuit conflicts, which is well-known due to White's frequent public dissents from denial of certiorari (Broyde 1987; Sullivan 2002).

[4]See Caldeira and Lempert (2017) for a minor exception that applies to the 1939 term. Note also that a small number of cases included in our sample are categorized as non-ideological, i.e., neither conservative nor liberal.

More generally, accounts in Cordray and Cordray (2004) and Perry (1991) suggest that justices assess cert-worthiness in non-uniform ways. And to the extent that the merits vote is not independent of the cert vote (Caldeira, Wright and Zorn 1999), we might expect that certain justice idiosyncrasies demonstrated at the merits stage—e.g., late-career Justice Douglas' propensity to vote against the U.S. Government (see Wolfman, Silver and Silver 1973)—could influence their agenda-setting vote too.

Our modeling strategy for this exploratory analysis is straightforward: for each of the twenty-seven justices,[5] we estimate a regression where each covariate (U.S. Petitioner, Actual Conflict, Alleged Conflict, Intermediate Reversal, Dissent Below, Amicus Present, Incompatible Decision Below) is interacted with a binary justice indicator variable. This approach allows us to derive, for each justice-covariate combination, estimates of effect size on the probability of the individual justice voting to grant cert. Less important, but still interesting, it allows us to test formally whether a given justice differs, at the .05 level, in the weight he puts on a given covariate in deciding whether to grant cert, from the average weight for all other justices combined.[6] To facilitate comparison of effect sizes across justices, we calculate the change in predicted probabilities from a baseline of approximately .15 ($.15 \pm .05$) as the covariate in question goes from 0 to 1.[7]

Figure 1 show the effect sizes by justice for two interesting but not unrepresentative covariates, Actual Conflict and Incompatible Decision Below. For each variable, we see a broad range of effect sizes, though the range is greater for Actual Conflict. There are justices who

[5]Black and Douglas appear in both the 1939 and the 1968 samples and Brennan, Marshall, and White appear both in 1968 and 1982, of course. Thus, we have 22 unique justices. However, in the interest of avoiding needless linguistic and modeling complications, we treat Douglas–1939 as a different justice than Douglas–1968, and so forth.

[6]There are several reasons to prefer this approach to a multilevel random coefficients model. The two most important are that our fixed-effects interaction approach allows us to estimate our quantities of interest directly, whereas a random coefficients model would allow only indirect (Empirical Bayes) prediction of justice-specific coefficients—which, in turn, are only *related* to our quantity of interest, the differences in predicted probabilities. Secondly, with only 27 justices, it is doubtful that we can rely on the asymptotic properties of a multilevel, random-coefficient model, which are valid as the number of clusters (i.e., justices) approaches infinity.

[7]See Caldeira and Lempert (2020) for a detailed discussion of this statistical approach. The regressions on which these estimates are based use robust standard errors, clustered by docket number, and also include fixed effects for Court term.

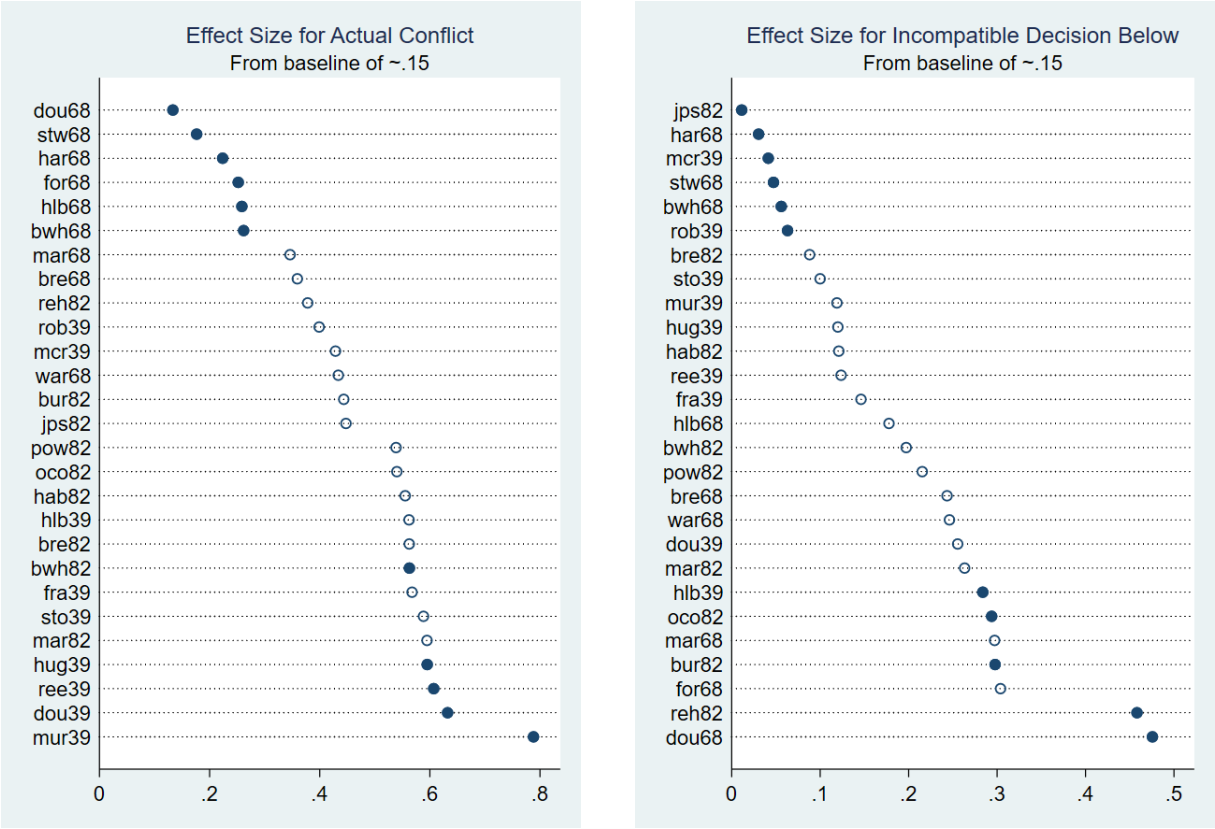

Figure 1. Individual justice effect sizes for effects summarized in Table 1. Solid circle indicates that the effect for a given justice is statistically significantly different from the average effect for all other justices; hollow circle indicates otherwise.

clearly stand out from their colleagues: Murphy (OT 1939) for Actual Conflict and Douglas (1968) and Rehnquist (1982) for Incompatible Decision Below, but even disregarding those three we see great variance in the weight that justices place on these covariates. Interestingly, though the weight that late-period (1982) Justice White places on Actual Conflict is significantly greater than that of the average justice in the sample, he is not among the most extreme justices in this respect.

The amount of justice-level heterogeneity for these two variables is not atypical. Table 1 summarizes the range of effect sizes for each covariate. Clearly, there is significant heterogeneity for each of the variables—the spread is particularly broad for U.S. Petitioner, Actual Conflict, Incompatible Decision Below, and Amicus Present. In the Appendix, we graphically present the effect estimates, for each covariate, for each justice in our sample. These graphs

illustrate that the wide range of effect sizes in Table 1 is not due to rare outliers; rather, justices tend to vary substantially in how much weight they put on each covariate—especially the covariates that have large effects on the Court's collective decision.

Covariate	Mean Effect	Min Effect	Max Effect
U.S. Petitioner	.47	−.02	.77
Intermediate Reversal	.10	.02	.20
Alleged Conflict	.08	−.03	.22
Actual Conflict	.45	.13	.79
Civil Liberties	.03	−.08	.24
Incompatible Decision Below	.19	.01	.48
Amicus Present	.33	.20	.49
Dissent Below	.10	−.01	.26

Table 1. Mean, Minimum, and Maximum Effect Sizes, Over Justices, on Cert Votes in OT 1939, 1968, and 1982.

Having demonstrated significant justice-level heterogeneity, we now ask whether this heterogeneity can be systematically explained by justice-level characteristics. In particular, after a review of relevant literature, we focus on justice experience and ideological extremism as potential explanators of effect heterogeneity.

Explaining Justice-Level Heterogeneity

As discussed above, the literature on justice-level heterogeneity at cert is very sparse, so we turn to studies that assess justice behavior at the merits stage to derive our hypotheses. One body of literature that addresses questions analogous to our own includes works that seek to measure whether justice attributes affect the extent to which a justice's vote (or other behavior on the merits) is influenced by *legal factors* (as opposed—explicitly or implicitly—to *ideological factors*). Since several of our covariates can be fit into the "legal" or the "ideo-

logical" categories, we might ask whether results from the merits stage also hold at the cert stage.

Before delving into that literature, it is worth spending some time categorizing our covariates as legal or ideological. Probably the most straightforwardly legal covariate is Actual Conflict, which is explicitly referred to in the Court Rules as a non-ideological reason for granting cert (see further Black and Owens 2009; Lindquist and Klein 2006; Perry 1991). A strong argument can also be made for U.S. Petitioner as a legal factor. The most direct evidence is from Johnson, Wahlbeck and Spriggs (2006, 107), which shows that legal arguments from the Solicitor General and from attorneys representing the federal government are higher-quality than those of other lawyers; for evidence that this holds not only at the merits stage, but also on cert, see Budziak and Lempert (2018, 47). The categorization of amicus briefs is more contested. Still, there is a line of theoretical (see Black and Owens 2009, 1070) and empirical (see Collins 2008*a*) work that argues for interpreting amicus briefs as indicators that a case is *legally* salient.

The status of Intermediate Reversal and Dissent Below is ambiguous. Though these types of disagreements in the lower courts may signal that the case is not legally trivial, it may also signal ideological divisions in a given case. Thus, we do not categorize these variables as either legal or ideological. Nor do we think that issue area (Civil Liberties) is a legal factor per se. And finally, Incompatible Decision Below is defined in terms of ideology, so this covariate is the most obviously ideological (and non-legal). In sum, then, we categorize Actual Conflict and U.S. Petitioner as clearly legal, Amicus Present as arguably so, and Incompatible Decision Below as clearly ideological.

One area of focus in the relevant merits-stage literature has been justice experience. Systematic analyses dating back at least to Snyder (1958) have suggested that shorter-tenured justices decide cases in a less ideologically extreme manner. Considering the relative weight of ideological and legal factors as a function of tenure, Hurwitz and Stefko (2004) hypothesize that justices who are newer to the Court will weight precedent relatively more heavily than

ideological considerations, compared to their more experienced colleagues. The authors justify this expectation on two grounds:

> First, as Snyder suggested, newcomer justices are generally less comfortable pursuing ideologically extreme agendas from the bench than veteran justices, such that newcomers would be more likely to make decisions that conform to existing precedent. Second, appointees to the Supreme Court are typically drawn from lower level courts, particularly in the post-World War II era. We can logically assume that their experience on inferior courts, where institutional constraints on personal agendas are more prevalent, will have conditioned and socialized new justices to make decisions in ways their new institution does not necessarily command.

Their empirical results are consistent with these expectations: the longer a justice serves, the less likely is her merits vote to follow applicable Supreme Court precedent, and more likely to be consistent with her ideology (as measured by Segal and Cover (1989) scores). Subsequently, in a wider-ranging analysis, Wedeking (2012) confirms Hurwitz and Stefko (2004). Drawing on this line of literature, we hypothesize that the "legal" variables that affect the cert vote—Actual Conflict, U.S. Petitioner, Amicus Present—will affect less experienced justices to a greater degree than they affect more experienced justices.

Ideological extremism has received less attention in the literature. Still, there is some evidence that ideological moderates weight legal factors more heavily, relative to ideological factors, than their more extreme colleagues. Enns and Wohlfarth (2013, 1090) theorizes that the "swing justice" in a given case, who is typically an ideological moderate, "will be the least responsive to attitudinal considerations and most responsive to legal and strategic factors" in casting his merits vote. Enns and Wohlfarth (2013) finds evidence for this proposition, although testing it involves serious measurement challenges. Segal and Spaeth (1996), which shows that of fifteen justices considered, only Powell and Stewart—two moderates—are appreciably affected by stare decisis when voting on the merits, is also consistent with the idea that legal factors affect moderates more strongly. Accordingly, we hypothesize that justices closer

to the Court median will be more affected by the "legal variables" than will those further from the median.

To test these hypotheses, we estimate a logit regression predicting whether a justice voted to grant cert (=1) or not(=0) with the standard covariates—Actual Conflict, U.S. Petitioner, Amicus Present, Dissent Below, Intermediate Reversal, Alleged Conflict, Incompatible Decision Below, and Civil Liberties—each interacted, in turn, with two variables: (1) the justice's Martin-Quinn (2002) *Distance to the Median* justice and (2) his *Tenure* on the Court, in years, at the time of the vote. We relegate the regression table to the Appendix; here, we focus on our quantity of interest, the difference in the effect of each covariate as we go from the 10th to the 90th percentile in Tenure and Distance to the Median, respectively.[8] In other words, we calculate and test the statistical significance of a set of second differences. As an example, the second difference for U.S. Petitioner as a function of Tenure is formally defined as, $\Delta\Delta[\Pr(Y)] = [\Pr(Y|\text{Tenure} = 1, \text{U.S. Petitioner} = 1) - \Pr(Y|\text{Tenure} = 1, \text{U.S. Petitioner} = 0)] - [\Pr(Y|\text{Tenure} = 26, \text{U.S. Petitioner} = 1) - \Pr(Y|\text{Tenure} = 26, \text{U.S. Petitioner} = 0)]$(see Berry, DeMeritt and Esarey 2010). For comparability of effect sizes, we again calculate them from a baseline value of approximately .15.

We find initial support for our hypotheses related to both justice experience and extremism. In particular, we find that more experienced and more ideologically extreme justices weight Actual Conflict and U.S. Petitioner—the two largest, and the most clearly legal, factors affecting the cert vote—significantly less heavily than do more junior and moderate justices. For example, the presence of the U.S. as petitioner increases the probability of a grant by a justice with Tenure at the 10th percentile by .50. This increase for a senior justice, with Tenure at the 90th percentile, is just .30. The second difference of .20 is significant at the .05 level.

Additionally, there is a bit of evidence that whether there is Amicus Present, a factor we classified as "arguably legal," is weighted less significantly by longer-tenured and more

[8]For Tenure, these values are 1 and 26 years, for Distance to Median, the values are 0 and 4.4 Martin-Quinn units.

Covariate	Effect as F(Tenure) Junior	Senior	Second Diff	Effect as F(Distance to Median) Moderate	Extreme	Second Diff
U.S. Petitioner	.50*	.30*	−.20*	.59*	.24*	−.35*
Intermediate Reversal	.11*	.08*	−.03	.11*	.11*	−.00
Alleged Conflict	.04*	.09*	.05*	.04*	.05*	.01
Actual Conflict	.46*	.34*	−.12*	.50*	.37*	−.13*
Civil Liberties	.03	.05*	.03	.04	.05*	.02
Incompatible Decision Below	.18*	.18*	−.00	.18*	.24*	.06†
Amicus Present	.35*	.29*	−.07	.34*	.24*	−.11†
Dissent Below	.11*	.06*	−.05	.08*	.11*	.04

Table 2. Effect sizes and second differences for eight covariates on the probability of a cert grant vote as a function of, respectively, justice Tenure and Distance to Median. A Junior and Moderate justice is one who is at the 10th percentile of Tenure and Extremism, respectively. A Senior and Extreme justice is at the 90th percentile of the respective measure. See text for details, and Table 4 in the Appendix for the regression on which these estimates are based. (*: $p < 0.05$; †: $p < 0.1$.)

extreme justices—though the second differences are large in magnitude and in the expected direction, they are not significant at the .05 level.

Two other results are worth noting. First, we find that experience does not affect the extent to which justices take into account the ideological direction of the decision below—the second difference for Incompatible Decision Below as a function of Tenure is ≈ 0. This is in contrast to some (but not all) of the literature on extremism in merits voting as a function of experience. Second, although—as it stands to reason—more extreme justices weight the ideological direction of the decision below more heavily than do relative moderates, the difference is not huge, and only marginally statistically significant. Thus, it appears that even relative moderates put a fair amount of weight on the ideological implications of the cert vote.[9]

[9]There is one unexpected result: more experienced justices appear to put greater weight on *Alleged* Conflict than do newer justices. We speculate that more experienced justices may be less rigid in construing a conflict—perhaps feeling less bound by formal rules and definitions.

Next, we evaluate the robustness of these results. Specifically, we address the concern we raised in the introduction: in short, whether the significant second differences are driven by a single justice. Recall, this is of consequence because if some results *are* driven by a single justice, we can have little confidence in attributing the significant second differences to our observed justice attributes (i.e., experience and extremism) as opposed to unmeasurable justice-level idiosynracies. To use the formal terms, we seek to address whether any justice is unusually influential, as an outlier with high leverage. Though the broader temporal scope of our analysis reduces the risk that our results are not robust in this sense, it does not eliminate it.

We proceed by re-estimating the model in Table 4 27 times, jackknifing out (omitting) from the sample a different justice each time. Then, we use the estimates from each model to estimate 27 sets of second differences, and their associated significance levels, as we did in Table 2. We consider significant second differences robust if, for every one of the 27 jackknife replications, the p value remains below .05. Our approach is analogous to the DFBETAS method of detecting influential observations (e.g., Kennedy 2003, 379), adapted to our quantity of interest (second differences as opposed to OLS coefficients) and considering influential justices (clusters) rather than influential observations.

	Effect as F(Tenure)		Effect as F(Distance to Median)	
Covariate	Minimum Second Diff	Maximum p-value	Minimum Second Diff	Maximum p-value
U.S. Petitioner	−.17	.02	−.30	.00
Actual Conflict	−.05	.34	−.03	.64

Table 3. Minimum second differences in probability of a cert grant vote, and maximum associated p-values, calculated across 27 replications jackknifing out a different justice each time, for two covariates as functions of justice Tenure and Distance to Median. See text for details.

As Table 3 shows, the results for U.S. Petitioner are robust, but not the results for Actual Conflict. No matter which justice is jackknifed out, the second differences remain large in magnitude and statistically significant at the .05 level for U.S. Petitioner—we can confidently

conclude that more junior and more moderate justices give more weight to a petition when the U.S. seeks review. However, for actual conflict, when Douglas (1968) is jackknifed out, the magnitude of the second differences decrease precipitously, and become statistically non-significant. To put it somewhat informally, Douglas is unusually influential here, due to the relatively low weight he puts on U.S. Petitioner, and his relatively extreme values for Distance to Median and Tenure. To be clear, the results in Table 2 should not be considered invalid: by all rights, Douglas *should* be included in the sample—it is not that there was some error in recording his data, or that he should be thought of as drawn from a different population than the other justices. However, we cannot say with confidence that those results for Actual Conflict are in fact driven by experience and extremism, as opposed to some justice-specific idiosyncrasy.[10] By contrast, we can be confident that our results for U.S. Petitioner are not driven by such idiosyncrasies.

Conclusion

We have explored heterogeneity in the certiorari vote at the Supreme Court, in the widest-spanning analysis of individual justices' votes in the literature. Two central findings emerge.

First, there is substantial justice-level heterogeneity in the weight that justices place on the standard factors shaping the cert vote (U.S. Petitioner, Actual Conflict, Alleged Conflict, Intermediate Reversal, Dissent Below, Amicus Present, Incompatible Decision Below). Thus, court-level analyses obscure substantial and interesting variation justices' individual agenda-setting calculations.

Second, some of this heterogeneity is associated with justice experience and ideological extremism, partially in theoretically predicted ways. The two unambiguously legal factors shaping the cert vote, Actual Conflict and U.S. Petitioner are weighted more heavily by newer

[10] Alternative approaches to assessing the robustness of our results, including two-way clustering by docket and justice (Cameron, Gelbach and Miller 2011; Gu and Yoo 2019) and substituting rank-based measures for Tenure and Distance to the Median (Iman and Conover 1979) lead to the same conclusions as the robustness checks we present here. The alternative robustness checks are included with the replication code.

justices and less extreme justices. This is consistent with expectations that more junior justices will emphasize formal, legal criteria when voting to grant cert, and that relative moderates will give more weight to legal factors over ideological ones. However, we cannot rule out the possibility that these results for Actual Conflict are driven by justice-specific idiosyncrasies, as opposed to experience and extremism per se.

This leads to a note of caution, applicable broadly to justice-level studies where there are relatively few justices, and covariates that are constant within justices, or nearly so, are of interest. To reduce the likelihood that results are driven by unmeasured justice-specific confounders (what we have called *idiosyncracies*), it is worth re-estimating results with individual justices, in turn, jackknifed out of the sample. Given that we have seen that this danger is present even in our relatively large sample with 27 justices included, it is a threat worth considering in studies that are narrower in temporal scope.

As this work is only an initial exploration of justice-level heterogeneity in agenda setting, several promising avenues for future research are apparent. For example, researchers could consider adding more terms and justices to analyses such as those presented above (though data limitations will make this difficult). Researchers might also test whether heterogeneity on the cert vote can be explained by systematic justice-level factors other than experience and extremism (or perhaps by case-level factors).

References

Bartels, Brandon L. 2011. "Choices in Context: How Case-Level Factors Influence the Magnitude of Ideological Voting on the U.S. Supreme Court." *American Politics Research* 39(1): 142–175.

Benesh, Sara C., David A. Armstrong II and Zachary Wallander. 2020. "Advisors to Elites: Untangling their Effect." *Journal of Law and Courts* 8(0): 0–00.

Berry, William D., Jacquline H.R. DeMeritt and Justin Esarey. 2010. "Testing for Interactions in Binary Logit and Probit Models: Is a Product Term Essential?" *American Journal of Political Science* 54(1): 248–266.

Black, Ryan C. and Christina L. Boyd. 2012*a*. "The Role of Law Clerks in the U.S. Supreme Courts Agenda-Setting Process." *American Politics Research* 40(1): 147–173.

Black, Ryan C. and Christina L. Boyd. 2012*b*. "U.S. Supreme Court Agenda Setting and the Role of Litigant Status." *Journal of Law, Economics, and Organization* 28(2): 286–312.

Black, Ryan C. and Ryan J. Owens. 2009. "Agenda Setting in the Supreme Court: The Collision of Policy and Jurisprudence." *Journal of Politics* 71(3): 1062–1075.

Black, Ryan C. and Ryan Owens. 2012. "Consider the Source (and the Message) Supreme Court Justices and Strategic Audits of Lower Court Decisions." *Political Research Quarterly* 65(2): 385–395.

Broyde, Michael J. 1987. "The Intercircuit Tribunal and Perceived Conflicts: An Analysis of Justice White's Dissents from Denial of Certiorari During the 1985 Term." *NYU Law Review* 62(3): 610–651.

Budziak, Jeffrey and Daniel Lempert. 2018. "Assessing Threats to Inference with Simultaneous Sensitivity Analysis: The Case of U.S. Supreme Court Oral Arguments." *Political Science Research and Methods* 6(1): 33–56.

Caldeira, Gregory A. and Daniel Lempert. 2017. "Agenda Control in the Hughes Court, OT 1939." Presented at 2017 Annual Meeting of the Midwest Political Science Association.

Caldeira, Gregory A. and Daniel Lempert. 2020. "Selection of Cases for Discussion: The U.S. Supreme Court, OT 1939, 1968, and 1982." *Journal of Law and Courts* 8(2): 381–395.

Caldeira, Gregory A. and John R. Wright. 1990. "The Discuss List: Agenda Building in the Supreme Court." *Law and Society Review* 24(3): 807–836.

Caldeira, Gregory A., John R. Wright and Christopher J.W. Zorn. 1999. "Sophisticated Voting and Gate-keeping in the Supreme Court." *Journal of Law, Economics, and Organization* 15(3): 549–572.

Caldeira, Gregory and John Wright. 1988. "Organized Interests and Agenda Setting at the U.S. Supreme Court." *American Political Science Review* 82(4): 1109–1127.

Cameron, A. Colin, Jonah B. Gelbach and Douglas L. Miller. 2011. "Robust Inference with Multiway Clustering." *Journal of Economic and Business Statistics* 29(2): 238–249.

Collins, Paul M. 2008*a*. "Amici Curiae and Dissensus on the U.S. Supreme Court." *Journal of Empirical Legal Studies* 5(1): 143–170.

Collins, Paul M. 2008*b*. "The Consistency of Judicial Choice." *Journal of Politics* 70(3): 861–873.

Cordray, Richard and Margaret Meriwether Cordray. 2004. "The Philosophy of Certiorari: Jurisprudential Considerations in Supreme Court Case Selection." *Washington University Law Quarterly* 82(2): 389–452.

Enns, Peter K. and Patrick C. Wohlfarth. 2013. "The Swing Justice." *Journal of Politics* 75(4): 1089–1107.

Epstein, Lee, Jeffrey A. Segal and Harold J. Spaeth. 2007. "The Digital Archive of the Papers of Justice Harry A. Blackmun." Available at: http://epstein.wustl.edu/blackmun.php.

Feller, Ari and Chris C. Holmes. 2009. "Beyond Toplines: Heterogeneous Treatment Effects in Randomized Experiments.". Downloaded from: http://www.stat.columbia.edu/ gelman/stuff_for_blog/feller.pdf.

Gu, Ariel and Hong Il Yoo. 2019. "vcemway: A one-stop solution for robust inference with multiway clustering." *The Stata Journal* 19(4): 900–912.

Hurwitz, Mark S. and Joseph V. Stefko. 2004. "Acclimation and Attitudes: NewcomerJustices and Precedent Conformance on the Supreme Court." *Political Research Quarterly* 57(1): 121–129.

Iman, Ronald L. and W.J. Conover. 1979. "The Use of the Rank Transform in Regression." *Technometrics* 21(4): 499–509.

Johnson, Benjamin. 2018. "The Supreme Court's Political Docket: How Ideology and the Chief Justice Control the Court's Agenda and Shape Law." *Connecticut Law Review* 50(3): 581–640.

Johnson, Timothy, Paul J. Wahlbeck and James F. Spriggs. 2006. "The Influence of Oral Arguments on the U.S. Supreme Court." *American Political Science Review* 100(1): 99–113.

Kennedy, Peter. 2003. *A Guide to Econometrics*. 5th ed. Cambridge, MA: MIT Press.

Lam, Patrick. 2013. "A Framework for Estimating Individual Causal Effects.". Downloaded from: http://patricklam.org/publication/dissertation/intro.pdf.

Lane, Elizabeth A. and Ryan C. Black. 2017. "Agenda Setting and Case Selection on the U.S. Supreme Court." In Oxford Research Encyclopedia Of Politics (Online). Doi: 10.1093/acrefore/9780190228637.013.91.

Lindquist, Stefanie A. and David E. Klein. 2006. "The Influence of Jurisprudential Considerations on Supreme Court Decisionmaking: A Study of Conflict Cases." *Law & Society Review* 40(1): 135–162.

Martin, Andrew D. and Kevin M. Quinn. 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the US Supreme Court, 1953–1999." *Political Analysis* 10(2): 134–153.

McGuire, Kevin T. and Gregory A. Caldeira. 1993. "Lawyers, Organized Interests, and the Law of Obscenity: Agenda Setting in the Supreme Court." *American Political Science Review* 87(3): 717–726.

NYU Law Review Supreme Court Project. 1984*a*. "Appendices." *New York University Law Review* 59(6): 1406–1929.

NYU Law Review Supreme Court Project. 1984*b*. "Appendix B." *New York University Law Review* 59(5): 823–1004.

Perry, H.W. 1991. *Deciding to Decide: Agenda Setting in the U.S. Supreme Court*. Cambridge, MA: Harvard University Press.

Rainey, Carlisle. 2016. "Compression and Conditional Effects: A Product Term is Essential When Using Logistic Regression to Test for Interaction." *Political Science Research and Methods* 4(3): 621–639.

Robertson, Reynolds and Francis Kirkham. 1936. *Jurisdiction of the Supreme Court of the United States*. St Paul, Minnesota: West Publishing.

Schoenherr, Jessica A. and Ryan C. Black. 2019. "Friends with Benefits: Case Significance, Amicus Curiae, and Agenda Setting on the U.S. Supreme Court." *International Review of Law and Economics* 58(1): 43–53.

Segal, Jeffrey A. and Harold J. Spaeth. 1996. "The Influence of Stare Decisis on the Votes of United States Supreme Court Justices." *American Journal of Political Science* 40(4): 971–1003.

Segal, Jeffrey and Albert Cover. 1989. "Ideological Values and the Votes of Supreme Court Justices." *American Political Science Review* 83(2): 557–565.

Shapiro, Stephen M., Kenneth S. Geller, Timothy S. Bishop, Edward A. Hartnett and Dan Himmelfarb. 2019. *Supreme Court Practice*. 11th ed. New York: Bloomberg BNA.

Snyder, Eloise C. 1958. "The Supreme Court as a Small Group." *Social Forces* 36(3): 232–238.

Spaeth, Harold J. 2001. "The Vinson-Warren Supreme Court Judicial Database Documentation." Available at: http://artsandsciences.sc.edu/poli/juri/vinwar_codebook.pdf.

Stern, Robert L. and Eugene Gressman. 1950. *Supreme Court Practice*. Washington, D.C.: Bureau of National Affairs.

Sullivan, J. Thomas. 2002. "Justice White's Principled Passion for Consistency." *Journal of Appellate Practice and Process* 4(1): 79–87.

Wedeking, Justin. 2012. "Why Do Policy-Motivated Justices Conform to Unfavorable Precedents? The Role of Social-Legal Backgrounds and Precedential Characteristics." *Justice System Journal* 33(1): 69–95.

Wolfman, Bernard, Jonathan L. Silver and Marjorie A. Silver. 1973. "Behavior of Justice Douglas in Federal Tax Cases." *Pennsylvania Law Review* 122(2): 235–365.

Appendix

Figures

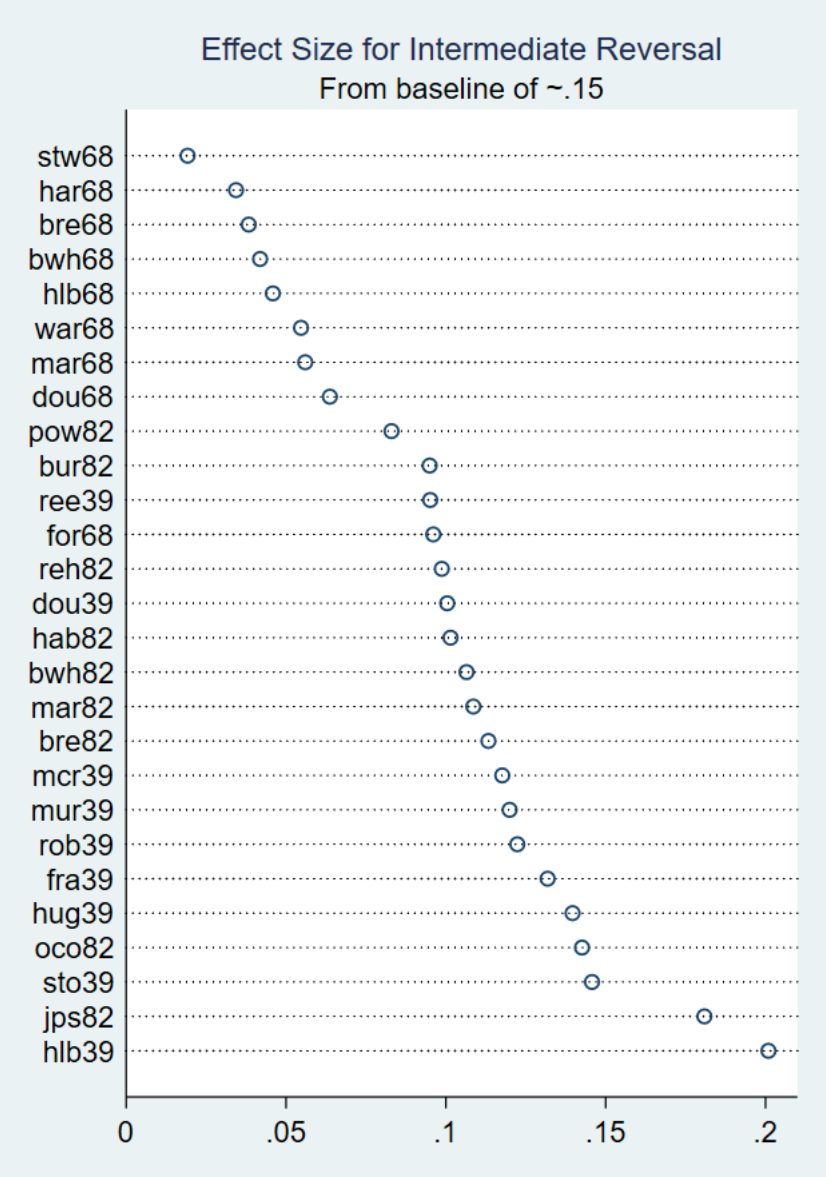

Figure 2. Individual justice effect sizes for effects summarized in Table 1. Solid circle indicates that the effect for a given justice is statistically significantly different from the average effect for all other justices; hollow circle indicates otherwise.

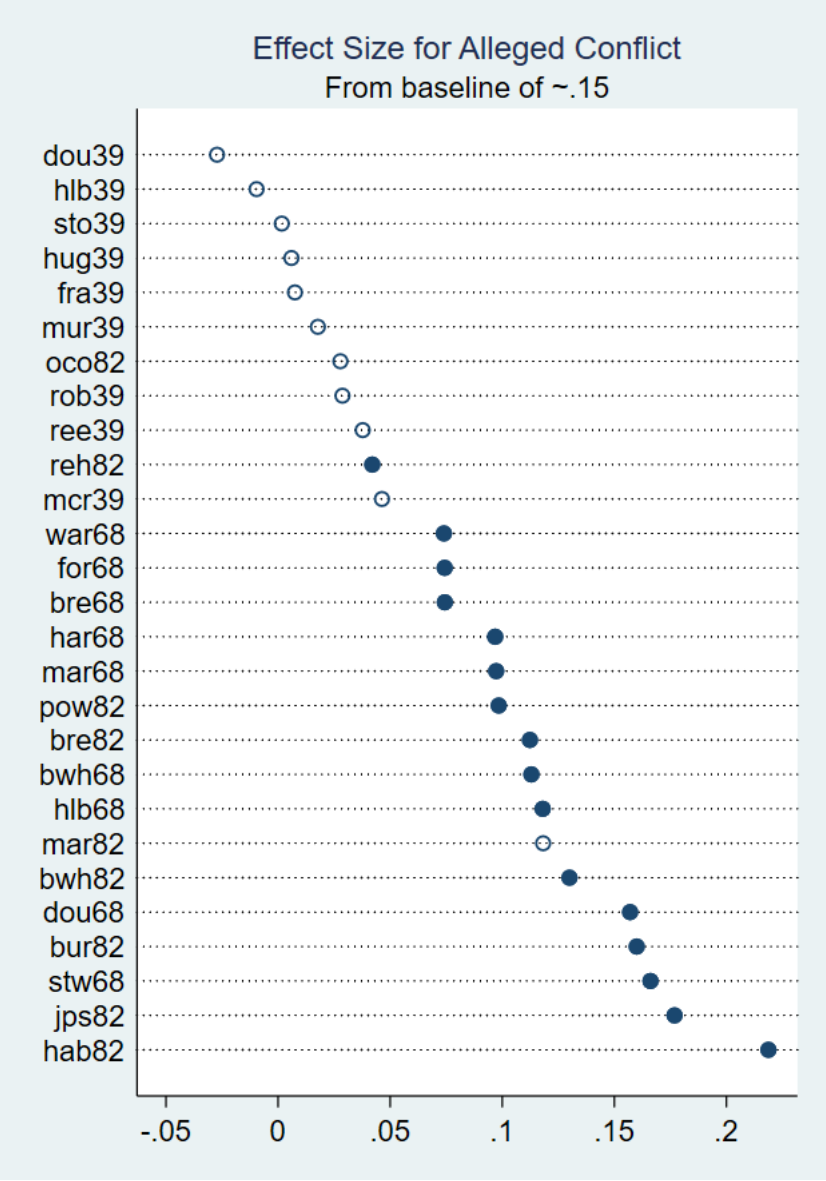

Figure 3. Individual justice effect sizes for effects summarized in Table 1. Solid circle indicates that the effect for a given justice is statistically significantly different from the average effect for all other justices; hollow circle indicates otherwise.

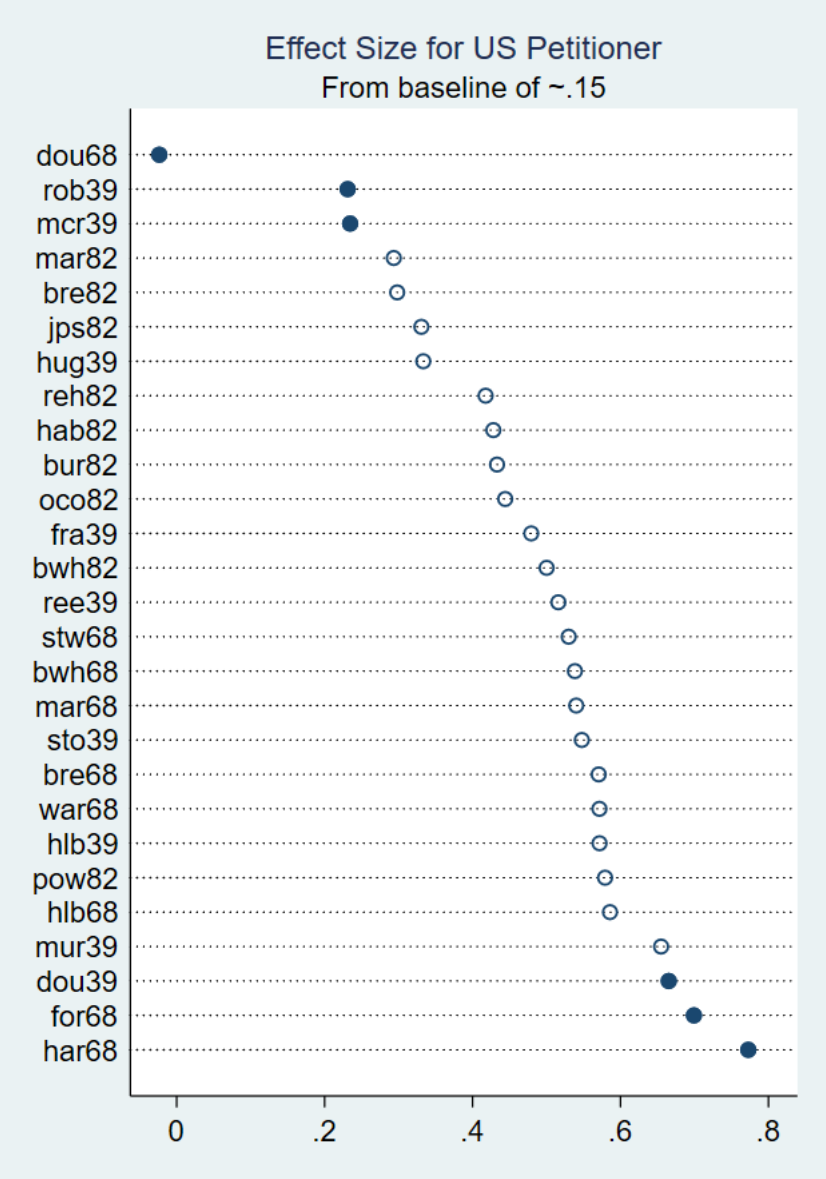

Figure 4. Individual justice effect sizes for effects summarized in Table 1. Solid circle indicates that the effect for a given justice is statistically significantly different from the average effect for all other justices; hollow circle indicates otherwise.

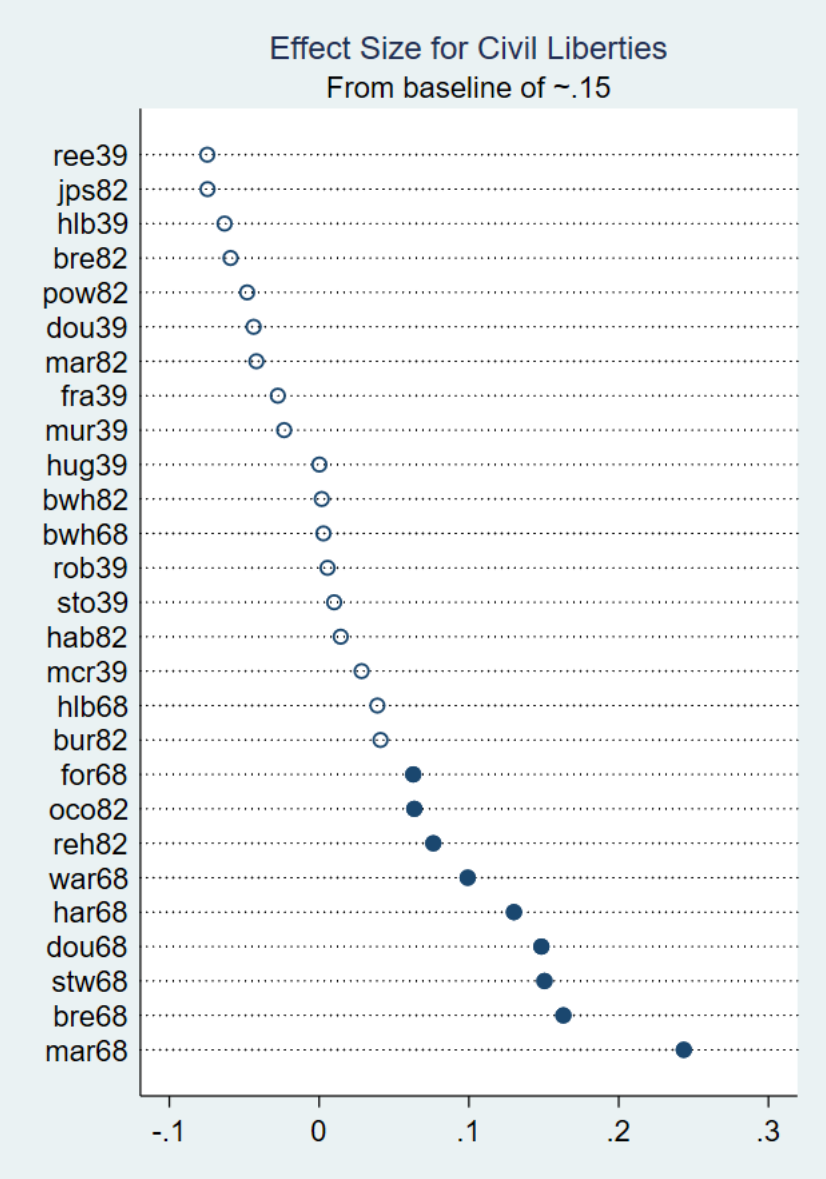

Figure 5. Individual justice effect sizes for effects summarized in Table 1. Solid circle indicates that the effect for a given justice is statistically significantly different from the average effect for all other justices; hollow circle indicates otherwise.

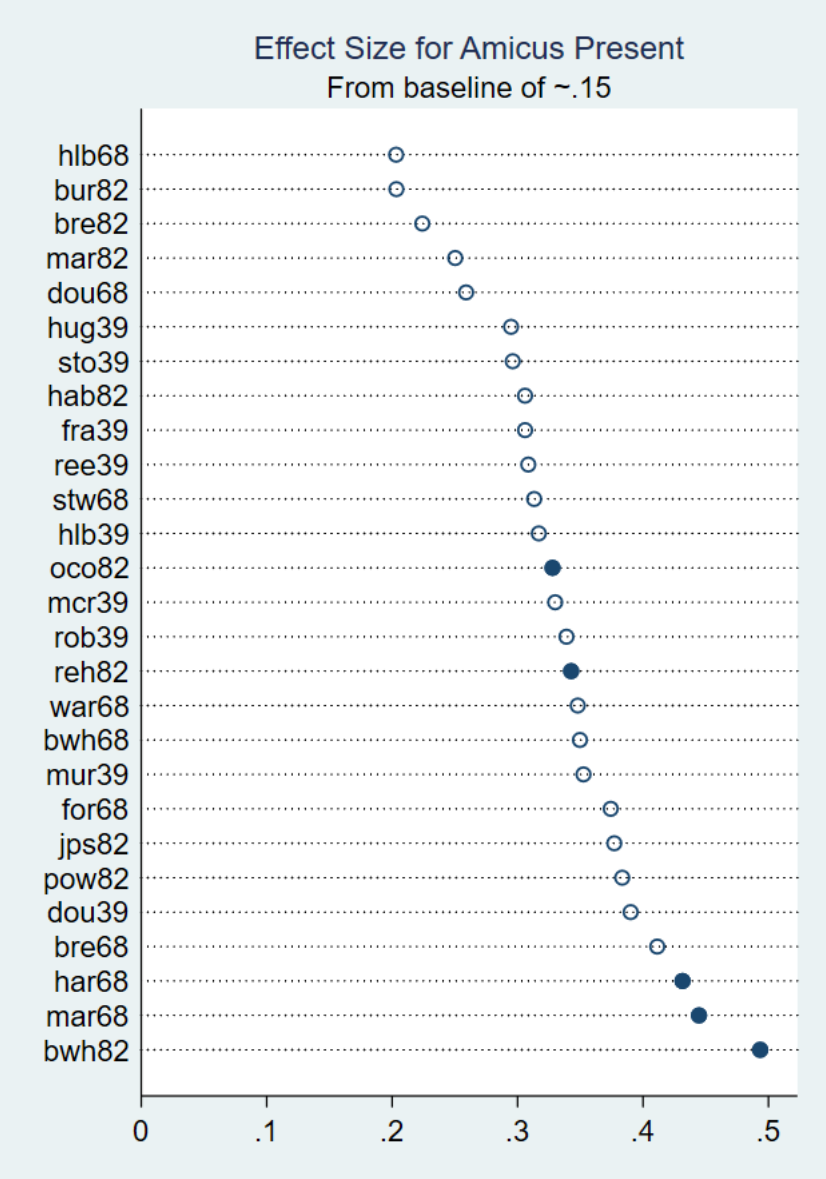

Figure 6. Individual justice effect sizes for effects summarized in Table 1. Solid circle indicates that the effect for a given justice is statistically significantly different from the average effect for all other justices; hollow circle indicates otherwise.

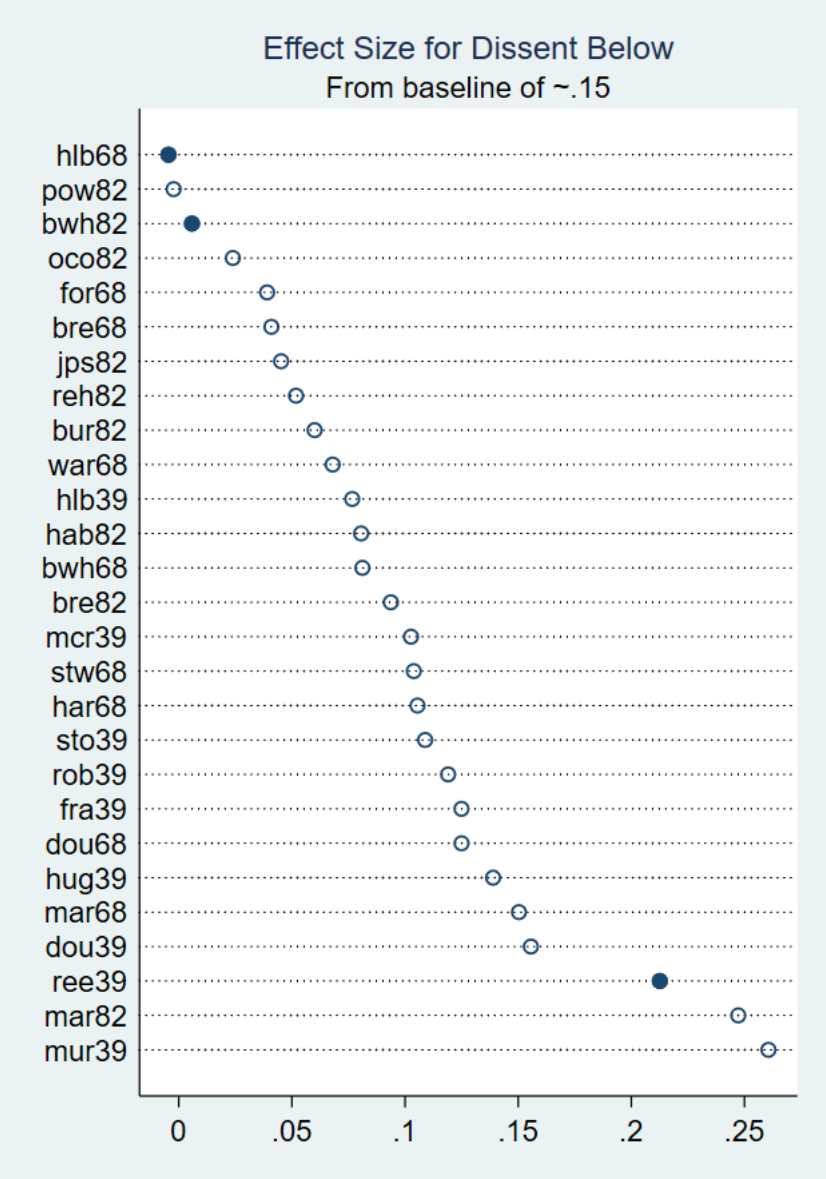

Figure 7. Individual justice effect sizes for effects summarized in Table 1. Solid circle indicates that the effect for a given justice is statistically significantly different from the average effect for all other justices; hollow circle indicates otherwise.

Table

Covariate	β	s.e.
US Petitioner	2.731*	(0.238)
Actual Conflict	2.298*	(0.161)
Alleged Conflict	0.309*	(0.136)
Dissent Below	0.567*	(0.159)
Intermediate Reversal	0.752*	(0.130)
Incompatible Decision Below	0.968*	(0.129)
Civil Liberties	0.221†	(0.135)
Amicus Present	1.920*	(0.220)
Tenure	−0.001	(0.006)
Distance to Median	−0.014	(0.030)
OT 1968	−1.592*	(0.136)
OT 1982	−2.216*	(0.139)
Constant	−2.652*	(0.133)
× with Tenure:		
US Petitioner	−0.008	(0.011)
Actual Conflict	−0.012	(0.008)
Alleged Conflict	0.019*	(0.007)
Dissent Below	−0.019*	(0.008)
Intermediate Reversal	−0.012†	(0.007)
Incompatible Decision Below	−0.007	(0.006)
Civil Liberties	0.009	(0.006)
Amicus Present	−0.004	(0.010)
× with Distance to Median:		
US Petitioner	−0.320*	(0.065)
Actual Conflict	−0.085*	(0.035)
Alleged Conflict	−0.049†	(0.030)
Dissent Below	0.114*	(0.034)
Intermediate Reversal	0.033	(0.029)
Incompatible Decision Below	0.079*	(0.032)
Civil Liberties	−0.005	(0.028)
Amicus Present	−0.090*	(0.046)

Table 4. Dependent Variable: Did a justice vote to grant cert (=1) or not (=0), all paid dockets, OTs 1939, 1968, 1982. $N = 32,867$. Logit regression; robust standard errors clustered on docket number. (*: $p < 0.05$; †: $p < 0.1$.)