Selection of Cases for Discussion

THE US SUPREME COURT, OCTOBER TERM 1939, 1968, AND 1982

GREGORY A. CALDEIRA, Ohio State University
DANIEL LEMPERT, SUNY Potsdam

ABSTRACT

The first, hidden stage of the Supreme Court's agenda-setting process is the formation of the "discuss list," the small set of cases actually considered in conference. Yet few have systematically considered the operation of and the influences on this critical initial phase of decision making. No systematic, empirical work makes comparisons over time of how features of cases shape the makeup of the chief justice's discuss list. Here, we examine the composition of the discuss list through a multivariate analysis of all paid petitions for certiorari filed in October Term 1939, 1968, and 1982. We are thereby able to compare the tendencies and efficacy of three long-serving chief justices—Hughes, Warren, and Burger—in making up the discuss list. And, methodologically, we present an alternative to the "observed-value" and the "representative-case" methods of calculating effect sizes for second differences, with software to implement our proposal.

I. INTRODUCTION

Agenda setting in the US Supreme Court is best conceptualized as occurring in two stages. The first, hidden stage of the Supreme Court's agenda-setting process is the formation of the "discuss list," the small set of cases actually considered in conference; those not selected are denied without a discussion or vote. In the second stage, the conference discusses and votes on the small set of remaining petitions. The latter stage has been the subject of much research. Even though the first stage, the creation of the discuss list, is a critical juncture, few have systematically considered the operation of and the influences on this initial phase of decision making (but see Provine 1980, 26–30; Caldeira and Wright 1990, 809–15;

We appreciate helpful discussions with Larry Baum, Jeff Budziak, Rob Hinckley, Matt Hitt, Kyle Kopko, and Steven Nawara and useful comments from Kevin McGuire and two anonymous reviewers. A version of this article was presented at the 2018 American Political Science Association Annual Meeting, where Doug Rice and other panelists and audience members provided valuable feedback. We thank Audra J. Smith for excellent research assistance. For use of data for OT 1968 and OT 1982, we thank John Wright and Christopher Zorn. Contact the corresponding author, Gregory A. Caldeira, at caldeira.1@osu.edu.

Journal of Law and Courts (Fall 2020) 2164-6570/2020/0802-00XX$10.00. Electronically published August 20, 2020.

Black and Boyd 2013). And no systematic empirical work makes comparisons over time of how features of cases shape the makeup of the chief justice's discuss list.

In the modern Supreme Court, the chief justice has many powers and duties (Cross and Lindquist 2006; Ruger 2006; Johnson 2018). Composing and distributing the discuss list is one of the most burdensome and important. The power to make up this list arguably gives the chief justice a first-mover advantage in setting the Court's agenda and at a minimum builds credit with colleagues for performing an arduous and unwelcome task. Managing the discuss list, as such, poses a challenge to and opportunity for the chief justice to exercise whatever skills or other advantages he can muster.

Here, we compare how three chief justices selected the cases to be discussed at conference during the October Terms (OTs) 1939, 1968, and 1982. Our statistical analysis spans a substantial swath of modern political history and major changes in the formal and informal operations of the Court. This lengthy span is important and valuable because of the limited temporal scope of previous work (Caldeira and Wright [1990], OT 1982; Black and Boyd [2013], samples from OTs 1986, 1987, 1991, and 1992).

Unlike previous scholars, we focus on the chief justice's selections, not the ultimate discuss list after associates have added cases. This allows us to assess specifically the tendencies and efficacy of the chief justices and thus present systematic evidence heretofore lacking on the relative merits of Hughes, Warren, and Burger as Court "task leaders" at the initial stage of agenda setting (Danelski 2016). This is not to say that the chiefs' choices were inconsequential: we can show that the chiefs' choices have a clear impact on the Court's ultimate agenda, especially on the Hughes and Warren Courts.[1]

Specifically, of the cases in our samples, we find that the associate justices added only nine cases for discussion to the set of 347 Hughes selected in 1939—and none of these cases were ultimately granted cert. Warren was similarly influential in 1968: associates added 24 cases to the set of 315 cases Warren chose to discuss, but only two of the associate-added cases were granted.

We can thus substantiate theorizing by Cameron and Clark (2016, 204) that the choices the chief justices make at this stage have appreciable downstream effects, since indeed there is evidence that "the chief is able to 'seed' the discuss list with cases . . . he favors [and] hold back cases he wishes to avoid" and thereby exert "special influence" on the cases the Court decides on the merits. True, Burger was challenged much more often, and the associates he served with were also more successful in having the cases they

1. In addition to the substantive reasons we point out immediately after this, there are also statistical reasons to focus on the chief's choice: modeling the associates' additions to the discuss list presents serious challenges. The central problem is that when deciding which cases to discuss, associates do not confront the same set of cases that the chief does. Specifically, they cannot add (or remove) cases that have already been added by the chief. Separate logit models for the chief's and the associates' choices would not be technically inappropriate, but comparisons between the chief's choice and the associates' choices would be uninformative at best and misleading at worst given the drastically different samples of cases. Alternatively, the associates' choices could be modeled with a selection model (with the chief's choice as the first stage), but the required identifying assumptions are implausible in this context.

added granted.[2] In the results section, we consider how our main empirical results could shed light on this discrepancy.

Of course, there were also changes in the Court's internal operations and in the environment within which it operated over the years we cover, as we discuss in what follows. Accordingly, we also test hypotheses that are derived from theoretical expectations of how institutional and environmental changes may have influenced the selection of cases for discussion. Although definitively disentangling the effects of the individual chiefs from the effects of the institutional context is difficult, our hypotheses, whether about the chiefs or about the institution, are largely supported.

II. FROM FULL CONSIDERATION TO THE DISCUSS LIST

During the hearings for the Judges' Bill of 1925, Justice Van Devanter stated that, as a matter of course, each justice read and prepared a memorandum on each petition for certiorari, and then the full Court considered each case in conference. And for a few years after the passage of the Judges' Bill, the Court actually did give all petitions at least perfunctory discussion at conference (see generally Hartnett 2000). But by the mid-1930s, Chief Justice Hughes had established a procedure for "special listing" of cases that the conference would not consider and that would be denied without discussion. The exact date of the change is not known, but references to cases being systematically denied without discussion exist as early as 1931, and it is clear that special listing was systematically implemented by 1935 (Provine 1980, 26–30; Caldeira and Wright 1990, 810; Ward and Weiden 2007, 111–12).[3] Presumably, as the justices accumulated more experience under the new regime of the Judges' Bill and as writs of certiorari became the dominant part of the Court's workload, the impetus for the change was Hughes and his colleagues' recognition that a significant proportion of petitions coming to the Court were nonmeritorious.

From the mid-1930s forward, then, the chief justice and his clerks would make up and circulate this special list of cases, also known as the "dead list," soon after the briefs and records for cases circulated to the associate justices; then, as the individual chambers reviewed cases, any of the associate justices could remove a case from the list.[4] Since the 1950s (the exact year is unclear), the Court has reversed the formal procedure, and now the chief justice circulates a "discuss list," the cases to be taken up, instead of a "special list" or "dead list," the cases to be denied (Stevens 1983). Cases not on the discuss list are denied without discussion. In what follows, we use *discuss list* to refer to both the special and discuss lists when the distinction is not important.[5]

---

2. Burger added 215 cases to the discuss list; his colleagues added 205, of which 42 were granted.

3. Charles Evans Hughes to Harlan Fiske Stone, September 30, 1935, box 75, Stone Papers, Library of Congress.

4. Interestingly, Justice Douglas in his certiorari memoranda denotes these cases as "black listed." Evidence from the papers indicates that associate justices almost never removed cases from Hughes's initial list.

5. As we note later on, the percentage of cases the chief places on the discuss list has decreased over time in the terms we cover—although we cannot attribute the decrease to the change in nomenclature and institutional practice.

III. THEORY AND HYPOTHESES

Our initial theoretical expectations derive from Caldeira and Wright (1990), who divide the set of covariates expected to influence the makeup of the discuss list into those of high quality (i.e., providing valuable, reliable information about the merits of a petition) and those of low quality (i.e., providing only weak or noisy information about the merits of a petition). High-quality indicators include the presence of actual or "square" conflict (i.e., conflict among the circuits or any other conflict described in Supreme Court Rule 10), the United States as petitioner, the ideological direction of the lower-court decision, and the presence of amicus briefs. Low-quality indicators include the presence of a dissent in the court below (usually a three-judge panel), an intermediate reversal (whether the court below the Supreme Court reversed the court below it), issue area, and alleged conflict.

The classification of indicators into these two categories is relatively straightforward, but the role played by the ideological direction of the lower-court decision at the discuss-list stage deserves some comment. On the one hand, no political scientist will be surprised that ideology plays a significant role at the agenda-setting stage (e.g., Caldeira, Wright, and Zorn 1999).[6] On the other hand, ideological "manipulation" of the discuss list by the chief justice is normatively disfavored. The discuss list's stated purpose is to eliminate frivolous petitions—a "housekeeping function," as it were. Justice Douglas is typical in calling it an "efficiency device" that eliminates "patently frivolous cases that [do not] present any substantial question, that [are] not even worthy of conference discussion."[7] That the justices and clerks interviewed in Perry (1991, 85–91) for the most part denied that Burger engaged in ideological manipulation of the discuss list (cf. Caldeira and Wright 1990, 826) suggests the hold of this normative ideal.

Normative claims aside, prior research suggests that, empirically, at the discuss-list stage, each covariate, whether high quality or not, should have some effect (Caldeira and Wright 1990).[8] Here, however, we are primarily interested in a different question: Do Hughes, Warren, and Burger put differential weight on indicators in selecting cases for discussion?

We offer several theoretical expectations for differences among chief justices. Two hypotheses are quite straightforward. First, it is well understood that civil liberties issues (including criminal procedure and civil rights) concerned the Court little before Vinson's

6. As Caldeira and Wright (1990, 815) put it: "In our view, the formation of the discuss list is, first and foremost, a political process, driven by the ideological stakes so often at issue in the great matters brought before the justices. It is the initial skirmish in the battle for public policy, and the content of the discuss list holds enormous implications for the eventual shape of decisions on the merits."

7. William O. Douglas and Walter F. Murphy, William O. Douglas Oral History Interviews, cassette 5, December 27, 1961, Public Policy Papers, Department of Rare Books and Special Collections, Princeton University Library, https://findingaids.princeton.edu/collections/MC015/c05.

8. In Black and Boyd (2013), the regression coefficients attending alleged conflict, intermediate reversal, and issue area are not significant, but the sample size is much smaller than that in Caldeira and Wright (1990) (727 vs. 1,771), which cautions against putting too much weight on differences in significance.

Table 1. Relative Frequencies of Each Covariate, for OT 1939, 1968, and 1982

Covariate	OT 1939	OT 1968	OT 1982
US petitioner	.09	.03	.03
Intermediate reversal	.25	.20	.25
Alleged conflict	.29	.53	.69
Actual conflict	.09	.11	.06
Civil liberties	.07	.48	.51
Compatible ideology with decision below	.58	.33	.62
Amicus present	.02	.03	.08
Dissent below	.12	.14	.12
N	651	1,067	1,759

Note.—For example, 9% of non-IFP cert petitions in OT 1939 were filed by the United States.

chief justiceship and that they were a particular emphasis for the Warren Court (Schwartz 1983, 215–61; Urofsky 1999; Belknap 2007, 150–86). Insofar as we can assume that the chief justice's agenda setting played a part in these differences, we hypothesize that the presence of civil liberties as the primary issue in a case will have a greater effect on Warren than on Hughes and, to a lesser extent, on Burger. Second, as we note in table 1 and as Caldeira and Wright (1988) observe, the number of amicus briefs on certiorari has increased over time; this leads us to hypothesize that the signaling value of amicus briefs will decrease over the course of the period we cover. The idea is that when amicus briefs are rare, they are "news" and thus informative, but as they become more numerous and as participants realize the impact they have on the Court, they become ordinary and thus less informative.[9] Therefore, we hypothesize that the presence of amicus briefs will have the greatest effect on Hughes's discuss list, when amicus briefs were rare, and the smallest effect on Burger's list, when these briefs were more plentiful.

Historical and anecdotal accounts of Hughes, Warren, and Burger as chief justices inform other expectations. Chief Justice Hughes had a reputation for excellent administrative skills, mastery of the Court's caseload, and skills as a leader across the board, and he also had the advantage of his service as an associate justice and as a lawyer practicing before the Court (McElwain 1949; Ross 2007; Danelski 2016). His legal acumen was held in unanimously high regard: typically, he is described as presenting cases "comprehensively, precisely, and impartially" (Goldstein 2011, 724). Of course, during his tenure as chief justice, the Court's caseload was small in comparison to the post–World War II period, and there were relatively few *in forma pauperis* (IFP) petitions; also, until his final two terms, he continued to preside over a relatively consensual group. Yet he and his associates had small staffs, did most of their own work, and had to deal with briefs of tremendously variable quality, arguments from lawyers of heterogeneous skills, and a carryover of diversity cases even after *Erie v. Tompkins* (304 U.S. 64 [1938]).

9. For evidence of an analogous pattern, see Caldeira, Wright, and Zorn (1997) on the declining impact of briefs amicus curiae on certiorari as filings proliferated.

In contrast to Hughes, Warren came to the Court as an administrator. His days as a practicing lawyer were in the 1930s, and, accordingly, he could not rely on his intellectual ability or experience to establish mastery of the cases and issues coming to the Court (Schwartz 1983; Belknap 2007). He also faced a larger caseload, a rising tide of IFPs (see, e.g., Boskey 1946), a stable of strong and contentious associate justices, a bar of mixed quality, and a more controversial mix of cases. To his advantage, he was well organized, had a larger staff than did Hughes (see, e.g., Peppers 2006, 145–205), had the advantage of a well-established routine of handling the caseload inherited from Hughes and Vinson, and, like Hughes, could draw on his moral authority and reputation for fairness.

In contrast to and unlike Hughes and Warren, Burger came to the Court with long experience as a federal appellate judge and with an interest in, and clear ideas about, how to administer the Court. He also had the advantage by the mid-1970s of a much larger staff than his predecessors. To his disadvantage, he faced an exploding caseload, of both paid and unpaid cases; a set of strong-minded, confident, and outspoken associates; and more and more politically controversial and contentious issues. And he followed a beloved chief. He quickly established a reputation for pushing his own agenda and at times for incompetence; generally, he was not regarded as trustworthy. Specifically, Burger's normatively dubious ideological machinations in conference (strategic passing, misrecording of votes, manipulation of opinion assignment) have been widely reported (e.g., Woodward and Armstrong 1979; Schwartz 1990; Johnson, Spriggs, and Wahlbeck 2005; Goldstein 2011). Indeed, there is at least an oblique suggestion by a law clerk that Burger engaged in similar behavior when composing the discuss list (Perry 1991, 86). Accordingly, we have grounds to hypothesize that ideological compatibility with the decision below will affect Burger's decision to slate a case for discussion to a greater degree than it will affect Hughes's decision or perhaps Warren's decision.

Because qualitative and historical scholarship holds Hughes's skills as a lawyer in particularly high regard but assesses the abilities of Warren and (particularly) Burger less positively, two further hypotheses are suggested. We hypothesize that the presence of actual conflict will affect Hughes's selection of cases the most and Burger's the least—not only because Hughes had a sharper eye but also because he likely valued the presence of conflict more than did the others. Relatedly, we hypothesize that alleged conflict will affect Burger the most and Hughes the least.

To derive expectations about dissent below and intermediate reversal (i.e., reversal in the court below), we consider the relative prevalence of these indicators among petitions to the Court, across the three terms (see table 1). Both of these indicators are considered "low cost" (i.e., easily discerned) and "low value" (Caldeira and Wright 1990). Still, their value may vary across terms as a function of their relative prevalence—a signal that is relatively rare may have greater probative value. We see, however, that the relative prevalence of these indicators is very similar across terms. Accordingly, we propose that the effect of dissent below and intermediate reversal will be similar across terms.

Finally, we turn to the United States as petitioner. Each of the three chiefs were—somewhat roughly speaking in the case of Hughes—ideologically compatible with the

contemporaneous presidential administration in the term we consider.[10] Existing studies have consistently indicated that the petitions by the United States are much more likely to be granted and—in the few terms for which we have evidence—discussed. Thus, we hypothesize no significant differences in effects for a US petitioner on Hughes's, Warren's, or Burger's discuss list.[11]

IV. DATA AND MEASURES

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, Marshall's certiorari memoranda. The data for the OT 1982 are from, and as described in, Caldeira and Wright (1990).

We concisely describe the operationalization of our variables (see Caldeira and Wright [1990] and Caldeira and Lempert [2017] for additional details). Our outcome variable, discuss, is whether the chief justice did (=1) or did not (=0) set a case for discussion. US 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 two or more cases occurring in different circuits or 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 1984a, 1984b) 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.

---

10. By OT 1939, Hughes consistently recognized the constitutionality of the Roosevelt administration's program.

11. Epstein and Posner (2018, 846–47) present evidence that there has been some decline in the Court's deference to the executive, at least on merits votes, but that decline does not set in until about 1990, after our data end.

Civil liberties is coded 1 if the primary substantive issue area of the case involves civil liberties (including criminal procedure and civil rights). Compatible ideology with decision below equals 1 if the decision below is liberal (following generally the criteria in Spaeth [2007]) in 1939 and 1968 and if it is conservative in 1982; in other words, we consider Hughes and Warren attitudinally liberal and Burger conservative. Obviously, we recognize that this categorization of Hughes is contestable, but in OT 1939 Hughes voted on the merits to reverse 39 of 65 conservative lower-court decisions and to affirm 44 of 61 liberal lower-court decisions (Spaeth et al. 2017). And as we describe in what follows, our interpretation of relevant results is robust to categorization of Hughes as conservative. We omit observations in which the lower-court decision has no discernible ideological implication. If amicus curiae briefs in support of or opposed to certiorari appear in a case, we code amicus present as 1, and 0 otherwise. Finally, dissent below is coded 1 if, in the court immediately below the Supreme Court, one or more judges dissented.

V. METHOD

Our primary question is how the effect of each (potential) influence on the makeup of the discuss list varies across the terms 1939, 1968, and 1982. For example, does the presence of the United States as a petitioner influence Hughes more than Warren or Burger? Thus, generically, we are interested in comparing effect sizes across groups. With binary outcomes, the calculation of effect sizes to be used in such a comparison involves subtle and more difficult methodological choices than commonly recognized. The two prominent approaches to calculating effect sizes after modeling binary outcomes are the "representative-case" method and the "observed-value" method (Hanmer and Kalkan 2013). The former calculates effect sizes by holding "control" covariates constant at some representative or interesting set of values; the latter sets "control" covariates at their in-sample observed values and calculates an effect size by averaging over each observation in the sample (see, e.g., Hanmer and Kalkan 2013; Long and Mustillo 2019). We argue in what follows for a different approach to calculating effect size differences across groups that bases such calculations on the comparison of units with similar "baseline probabilities" of a positive outcome. An informal sketch of our argument follows; specifics and the information about the Stata program that implements our recommendation are in the appendix (available online).

To compare effect sizes across groups of some variable Z on a binary outcome Y, it is typical to calculate the change in $\Pr(y = 1)$ as z goes from some baseline value to some reference value for each group and then calculate the size and significance of the difference between these two effects. Because models for binary outcomes (in our case, a logit) are nonlinear, the effect size depends not only on the estimated coefficients (as in ordinary least squares) but also on the values that the other independent variables $\mathbf{X}$ take on. Thus, an important, but perhaps undertheorized, question is how to set $\mathbf{x}$ for a given substantive inquiry. Define $\Pr(y = 1|z = \text{baseline value}, \mathbf{x}, \text{group} = g)$ as the baseline probability for group g. Observe that both common approaches to setting $\mathbf{x}$ for calculation of effect sizes—the

representative-case and the observed-value approaches—may entail calculations that are based on very different baseline probabilities across groups, if **x** is fixed across groups (as it is in all applications of which we are aware). The central issue we address is that effect size comparisons based on substantially different baseline probabilities may be strongly influenced by the nonlinear functional form of the logit model (i.e., where on the logit curve the baseline probability for each group is located). We argue that for many substantive inquiries, a more appropriate comparison is to set the values of the control covariates **x** so that baseline probabilities are as similar as possible across groups.

Consider our application. It is unsurprising that (even conditional on covariates) the baseline probability that the chief selects a case for discussion is greater in 1939 than in 1968, which is greater than in 1982. The increasing number of petitions facing a chief, combined with a relatively small number of petitions granted and heard, ensures that cases with comparable attributes will have different probabilities of being granted (Caldeira and Lempert 2017) and discussed. Whether one considers a representative case or averages over observations via the observed-value approach, baseline probabilities of a positive outcome will be significantly different across terms. If one is interested in comparing the effects of a covariate across terms, we argue, the appropriate comparison is based not on cases with identical attributes and very different baseline probabilities of being discussed but rather on cases with very similar baseline probabilities of being discussed, even if their attributes differ. Thus, as we do in what follows, one might compare how the presence of actual conflict changes the probability of making the discuss list for a case with an approximately 50% chance of being discussed ceteris paribus, for Hughes, Warren, and Burger.

VI. RESULTS

We present our results in table 2. For each covariate and term, we give the effect sizes (i.e., first differences); in the rightmost column, we list whether the second differences for each pair of terms, for each covariate, are significantly different from 0. We calculate the effect sizes and second differences as follows. For each term-pair, we estimate a logit model in which each covariate is interacted with a group (i.e., term) indicator.[12] With these estimates, we use sdcasepick (see the appendix) to select the set of observations that (1) have a value 0 for the covariate of interest (e.g., in row 1, US petitioner) and (2) have a predicted probability of being discussed between .45 and .55. Then, from that set, sdcasepick selects the pair of observations, consisting of one in each term, whose members have the smallest absolute difference between their predicted probabilities of being discussed. (The mean of the 24 [eight covariates, three term-pairs] smallest distances

12. For the term pairs 1939/1968 and 1939/1982, we estimate a Firth logit model (Firth 1993; Coveney 2015) since US petitioner and amicus briefs perfectly predict Hughes's decision to place a case on the discuss list. Since the sign and significance of the coefficients in interactive models are not particularly informative when one is interested in the outcome Pr(Y) (e.g., Rainey 2016, 625), we do not present the regressions themselves here.

Table 2. Effect Sizes for Each Covariate on the Probability That the Chief Justice, in OT 1939, 1968, and 1982, Selects a Case for Conference Discussion

Covariate	Effect Size			Significant Difference between Terms?
	OT 1939	OT 1968	OT 1982	
US petitioner	.50	.49/.48	.32	39/82 (<.1)
Intermediate reversal	.17	.14	.11	
Alleged conflict	.03 (NS)	.21	.08 (NS)	39/68; 68/82 (<.1)
Actual conflict	.43/.42	.38	.35	
Civil liberties	−.07 (NS)	.17/.16	.06 (NS)	39/68; 68/82 (<.1)
Compatible ideology with decision below	−.11	−.10	−.29	39/82; 68/82
Amicus present	.47	.30/.31	.17	39/82
Dissent below	.19	.16/.15	.16/.15	

Note.—See text for details of effect size calculation. The final column notes whether the difference in effects between pairs of terms is significant at the .05 or (if explicitly noted) the .1 level. For example, the effect of alleged conflict is significantly greater in OT 1968 than in 1939 at the .05 level, and it is greater in OT 1968 than in 1982 at the .1 level. All effects are significant at the .05 level, except, in OT 1939 and OT 1982, alleged conflict and civil liberties. NS = not significant.

between observations is .0036, with a range of [.0004, .0095].) More colloquially, we select the pair of observations that are most similar in baseline probability, while having an approximately 50:50 chance of being discussed. Thus, we ask, how much does the presence of a covariate change the probability that a case is discussed, from a baseline of about .5?

Although our hypotheses are about differences between effects across terms (i.e., second differences), we note here that all effects (i.e., first differences) are significantly different from 0 at the .05 level, except for, in OT 1939 and OT 1982, alleged conflict and civil liberties. Thus, we see that most covariates, whether "low-quality" indicators or not, affect the chief's selection of cases to discuss; this is consistent with existing research.

Turning first to US petitioner, we see that this variable exerts a huge effect on the chief's decision to discuss, although the effect size is appreciably smaller for Burger than for Hughes and Warren. Still, even in OT 1982, US petitioner is surpassed only by actual conflict in terms of effect magnitude. And in OT 1939 and OT 1968, an otherwise "marginal" case (i.e., with approximately a 50:50 chance of being placed on the discuss list) becomes, in effect, certain to be set for discussion if the United States is the petitioning party. Thus, our hypotheses are, in the main, supported, although we did not predict the (marginally significant) decrease in effect size between OT 1939 and OT 1982 (about which more in the conclusion).

Turning next to intermediate reversal and dissent below, we find that the effects for both of these variables are modest and remain relatively constant across all three of our terms. This is consistent with our expectations, which were based on the observation that the relative frequency of petitions involving intermediate reversals and dissents below was quite similar across OTs 1939, 1968, and 1982. In addition, the moderate effect size—between .1 and .2 in each term—is consistent with both variables' status as low quality.

The results for alleged conflict are unexpected. Although, in line with our expectations, a mere allegation of conflict has very little effect on Hughes's decision to discuss a case, we find that the effect of alleged conflict is significantly greater for Warren than Hughes and appreciably greater ($p < .1$) for Warren than Burger. We consider possible explanations for this result in Section VII.

The effects for actual conflict are more in line with our expectations. We find that, in all three terms, this high-quality indicator has large effects. A petition that is marginal in the absence of conflict becomes very likely to be discussed if a square conflict is present. We do see some decrease in effect from Hughes to Warren to Burger, as hypothesized, although the differences are not statistically significant.

Turning to civil liberties, we find, as expected, that the effect is greatest for Warren's discuss list, and the difference between OT 1968 and OT 1939 is statistically significant at the .05 level, while the difference between OT 1968 and OT 1982 is significant at the .1 level. The results here line up quite well with historical accounts of the Court's agenda in these periods and demonstrate that even at the initial stage of agenda setting, cases involving civil liberties were favored in the Warren Court substantially more so than in the Hughes and Burger Courts.

The patterns for the presence of amicus briefs are also consistent with expectations, as we see a sizable and consistent drop-off in effect size over terms. This is consistent with our hypothesis that the effect of amicus briefs would decrease as they become more common. Indeed, we find that although a case that is otherwise marginal would be practically certain to be discussed in OT 1939 if an amicus brief is present, the presence of an amicus brief in an otherwise marginal case in OT 1982 makes only about as much difference as the presence of a dissent in the court below. This difference between OT 1939 and OT 1982 is statistically significant as well.

Finally, consider the effect of ideology. Here, we find rather dramatic confirmation of our hypothesis. Although both Hughes and Warren were less likely to set for discussion cases in which they favored the decision below ideologically, the magnitude of the effect for Burger is nearly three times larger.[13] A case that has about a 50:50 chance of being discussed when ideologically discordant with Hughes or Warren drops to a probability of about .4 if, all other attributes equal, it is ideologically concordant with Hughes or Warren. But a liberal case that is marginally likely (i.e., has about a one-in-two chance) to make the discuss list under Burger has only about a one-in-five chance of being added to the list if that same case involves a conservative decision below instead. Indeed, we find that this difference is significant not only substantively but also statistically. This result may well shed light on why Burger's discuss list was challenged so much more often than that of

13. Recall that compatible ideology is based on the classification of Hughes as a liberal; if we categorize him instead as conservative, the finding would be that he set cases for discussion counterattitudinally (perhaps discarding his ideological inclinations in favor of the Court's preferences), which would distinguish him from Burger even more.

Hughes or Warren. Certainly by 1982, if not much earlier, it is likely that Burger's ideological approach to setting cases for discussion had become apparent to his colleagues, naturally leading the associates to "supplement" his list with "rapid fire," as Perry's (1991, 87) informant puts it.

VII. DISCUSSION AND CONCLUSION

Although we have found support for several of our hypotheses, a few things have surprised us. Here, we speculate about potential explanations for these results. First, we found that although the presence of the United States as a petitioner still had a very substantial effect on Burger's decision to list a case, it carried more weight under Hughes and Warren. One possible explanation is a decrease in the standing of the solicitor general during the early Reagan years; as Caplan (1988, 264–66) reports, the justices, across the ideological spectrum, perceived a decline in the commitment of the solicitor general's office to its traditional strategy of litigation, of reliance on the small set of criteria on which the Court itself relies. Burger, in particular, is described objecting to Solicitor General Rex Lee's and his office's willingness "to be a spokesman for the 'reactionary' Reagan Administration" (265). More generally, we know empirically that the solicitor general has pursued somewhat different strategies of petitioning over time—filing 56 petitions in OT 1939, 30 in OT 1968, and 48 in OT 1982—with varying degrees of success (see, e.g., Cordray and Cordray 2010).

Second, compared to Hughes and Burger, Warren was more willing to set cases for discussion on the mere allegation of conflict, perhaps as a result of his more general openness to considering important policy issues in the absence of traditional criteria such as square conflict. For one, the Warren Court devoted a large share of its plenary agenda to cases of state and federal criminal procedures, many of which came to the Court as IFPs. We know, of course, that the justices vary considerably in their propensity to vote to grant certiorari, from Byron R. White, who at least toward the end of his career made a point of dissenting in dozens of cases he thought the Court should hear, to Thurgood Marshall, who in his later years seldom voted for cert, believing that the conservative majority would simply engage in "mischief" if permitted to decide a case on the merits.

Finally, we reported some decrease in the effect of actual conflict from Hughes to Warren to Burger, although none of these differences was statistically significant. The decline in the "punch" of actual conflict may well be a function of the capabilities of the three chief justices, as we hypothesized. Yet, in the hearings on the Judges' Bill of 1925, the justices testified that the Court would grant certiorari on "conflicts" as a matter of course, and Robertson and Kirkham (1936) and Stern and Gressman (1950) repeated this claim. It remains an open question whether the Court ever granted certiorari on all conflicts, as the justices had warranted, but we know from our data that as early as OT 1939, the Court denied certiorari on a substantial number of square conflicts. Stern (1953), of Stern and Gressman's *Supreme Court Practice* fame, offers an alternative explanation. He suggests that the Supreme Court, as time passed after the passage of the Judges' Bill, has become less

likely to grant certiorari in the presence of conflicts that (*a*) are no longer live (e.g., because resolved by statutory amendment), (*b*) involve a case that would be decided by the Court on grounds that would not resolve the conflict, and (*c*) involve a case that is at an interlocutory stage. And it is possible that the Court—and thus the chief justice making up the discuss list—has placed additional emphasis on such exceptions and perhaps added others.

A visitor from the Supreme Court of the 1930s would scarcely recognize the Court of the 1980s, let alone the Court of the 2010s, so radically has it changed. An institution in which justices employed a solitary law clerk; wrote their own opinions; discussed most of the cases; issued short, unadorned opinions; seldom issued dissents; and decided numerous cases—a diverse mixture of common-law, statutory, and constitutional—has morphed into an institution in which the justices act as partners in individual law offices of four clerks; publish long, scholarly opinions; and decide on a small, rarefied set of cases. And yet the decision-making calculus of the chief justice in setting the discuss list evinces marked stability across time, with square conflicts, the United States as a petitioner, and amicus briefs as high-impact and relatively consistent forces. Nonetheless, as we hypothesized, Hughes, Warren, and Burger differed in the weights they placed on features of cases and in the extent to which political ideology shaped their decisions. As such, we conclude that both institutions and individuals play an important role at the initial agenda-setting stage. We have been able to show this through the first systematic comparative account of how chief justices select the set of cases for discussion. And we have presented a method of calculating second differences that is broadly relevant for research involving comparison of effect sizes across groups.

REFERENCES

Belknap, Michael R. 2007. *The Supreme Court under Earl Warren, 1953–1969*. Columbia: University of South Carolina Press.

Black, Ryan C., and Christina L. Boyd. 2013. "Selecting the Select Few: The Discuss List and the US Supreme Court's Agenda-Setting Process." *Social Science Quarterly* 94 (4): 1124–44.

Boskey, Bennett. 1946. "The Supreme Court's 'Miscellaneous' Docket." *Harvard Law Review* 59 (4): 604–6.

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

Caldeira, Gregory A., and John R. Wright. 1988. "Organized Interests and Agenda Setting at the U.S. Supreme Court." *American Political Science Review* 82 (4): 1109–27.

———. 1990. "The Discuss List: Agenda Building in the Supreme Court." *Law and Society Review* 24 (3): 807–36.

Caldeira, Gregory A., John R. Wright, and Christopher J. Zorn. 1997. "Sophisticated Judicial Behavior: Agenda-Setting via the Discuss List." Paper presented at the 1997 Annual Meeting of the American Political Science Association.

———. 1999. "Sophisticated Voting and Gate-Keeping in the Supreme Court." *Journal of Law, Economics, and Organization* 15 (3): 549–72.

Cameron, Charles M., and Tom Clark. 2016. "The Chief Justice and Procedural Power." In *The Chief Justice: Appointment and Influence*, ed. Artemus Ward and David J. Danelski, 202–34. Ann Arbor: University of Michigan Press.

Caplan, Lincoln. 1988. *The Tenth Justice: The Solicitor General and the Rule of Law*. New York: Vintage.

Cordray, Margaret Meriwether, and Richard Cordray. 2010. "The Solicitor General's Changing Role in Supreme Court Litigation." *Boston College Law Review* 51 (5): 1323–82.

Coveney, Joseph. 2015. "FIRTHLOGIT: Stata Module to Calculate Bias Reduction in Logistic Regression." Statistical Software Components S456948, Boston College Department of Economics, revised July 25.

Cross, Frank B., and Stefanie Lindquist. 2006. "Doctrinal and Strategic Influences of the Chief Justice." *University of Pennsylvania Law Review* 154 (6): 1665–707.

Danelski, David J. 2016. "The Influence of the Chief Justice in the Decisional Process of the Supreme Court." In *The Chief Justice: Appointment and Influence*, ed. Artemus Ward and David J. Danelski, 19–46. Ann Arbor: University of Michigan Press.

Epstein, Lee, and Eric A. Posner. 2018. "The Decline of Supreme Court Deference to the President." *University of Pennsylvania Law Review* 166 (4): 829–60.

Firth, David. 1993. "Bias Reduction of Maximum Likelihood Estimates." *Biometrika* 80 (1): 27–38.

Goldstein, Joel K. 2011. "Leading the Court: Studies in Influence as Chief Justice." *Stetson Law Review* 40 (3): 717–62.

Hanmer, Michael J., and Kerem Ozan Kalkan. 2013. "Behind the Curve: Clarifying the Best Approach to Calculating Predicted Probabilities and Marginal Effects from Limited Dependent Variable Models." *American Journal of Political Science* 57 (1): 263–77.

Hartnett, Edward A. 2000. "Questioning Certiorari: Some Reflections Seventy-Five Years after the Judges' Bill." *Columbia Law Review* 100 (7): 1643–738.

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 (1): 581–640.

Johnson, Timothy, James F. Spriggs, and Paul J. Wahlbeck. 2005. "Passing and Strategic Voting on the U.S. Supreme Court." *Law and Society Review* 39 (2): 349–74.

Long, J. Scott, and Sarah A. Mustillo. 2019. "Using Predictions and Marginal Effects to Compare Groups in Regression Models for Binary Outcomes." *Sociological Methods and Research*. https://doi.org/10.1177/0049124118799374.

McElwain, Edwin. 1949. "The Business of the Supreme Court as Conducted by Chief Justice Hughes." *Harvard Law Review* 63 (1): 5–26.

NYU Law Review Supreme Court Project. 1984a. Appendices. *New York University Law Review* 59 (6): 1406–929.

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

Peppers, Todd. 2006. *Courtiers of the Marble Palace*. Stanford, CA: Stanford Law and Politics.

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

Provine, Doris Marie. 1980. *Case Selection in the United States Supreme Court*. Chicago: University of Chicago 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): 621–39.

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

Ross, William G. 2007. *The Chief Justiceship of Charles Evans Hughes, 1930–1941*. Columbia: University of South Carolina Press.

Ruger, Theodore W. 2006. "The Chief Justice's Special Authority and the Norms of Judicial Power." *University of Pennsylvania Law Review* 154 (6): 1551–74.

Schwartz, Bernard G. 1983. *Super Chief: Earl Warren and His Court*. New York: New York University Press.

———. 1990. *The Ascent of Pragmatism: The Burger Court in Action*. Boston: Addison-Wesley.

Spaeth, Harold J. 2007. "The Judicial Research Initiative—United States Supreme Court Judicial Databases." University of South Carolina, Department of Political Science. http://artsandsciences.sc.edu/poli/juri/burger_codebook.pdf.

Spaeth, Harold J., Lee Epstein, Jeffrey A. Segal, Ted Ruger, Andrew D. Martin, and Sarah Benesh. 2017. The Supreme Court Database. http://scdb.wustl.edu.

Stern, Robert L. 1953. "Denial of Certiorari despite a Conflict." *Harvard Law Review* 66 (3): 465–72.

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

Stevens, John Paul. 1983. "The Life Span of a Judge-Made Rule." *New York University Law Review* 58 (1): 1–21.

Urofsky, Melvin J. 1999. *Division and Discord: The Supreme Court under Stone and Vinson, 1941–1953*. Columbia: University of South Carolina Press.

Ward, Artemus, and David L. Weiden. 2007. *Sorcerers' Apprentices: 100 Years of Law Clerks at the United States Supreme Court*. New York: New York University Press.

Woodward, Bob, and Scott Armstrong. 1979. *The Brethren*. New York: Avon.