Estimating Ideal Points at the Supreme Court Using Agenda-Setting Votes

Ben Johnson*
Department of Politics
Princeton University

January 6, 2018

Abstract

This article develops a new method for estimating the ideological preferences of Supreme Court justices based on their agenda-setting votes. While careful attention to agenda-setting is important in its own right, ideological measures from these votes have much to recommend them in general. Compared to the more familiar votes on case dispositions, there are far more agenda-setting votes on a far more stable set of issues. The estimated model takes into account case importance, justices’ individual preferences over docket size, and their ideological leanings. The method successfully predicts votes and returns new useful measures for scholars of the Court.

*Department of Politics, Princeton University 130 Corwin Hall, Princeton, NJ 08544-1012 E-mail: bbj@princeton.edu. Thanks to Chris Achen, Deborah Beim, Charles Cameron, Brandice Canes-Wrone, John Kastelic, John Londregan, Nolan McCarty, Kelly Rader, Marc Ratkovic, Sepehr Shahshahani, and Keith Whittington for helpful comments and sound advice. I thank the Center for the Study of Democratic Politics at Princeton for financial support and Julie Nelson and Rebecca Voth for research assistance.
Introduction

Models of spatial voting are at the heart of modern theories and empirical studies of electoral, legislative, and judicial politics. These models assume that actors have preferences that may be represented as a point in some policy space. The point that represents an actor’s most preferred policy is usually referred to as an “ideal point.” Spatial models typically make the subsequent assumption that policymakers support proposals that move policy closer to their ideal policies.

Evaluations of theories of judicial behavior have relied heavily on measures of justices’ ideal points. The most common of these measures are based on justices’ votes to reverse or affirm a lower court, but there are significant drawbacks to focusing attention solely on the disposition of cases. While parties appeal thousands of cases to the Court each year, the Court decides only a select few. Studying only those few cases means that scholars often ignore the bulk of the Court’s work. What is more, since those few cases are not randomly selected from the appealed cases, any measures derived from that subset will almost certainly be biased.

This article measures Supreme Court ideology at the agenda-setting—or certiorari—stage. The Court rules on thousands of certiorari petitions every term while it only decides a few dozen cases. Further, since the Court does not have control over whether a petition is filed—whereas it has total control over whether a case is decided—ideological measures recovered from the certiorari stage are almost certainly less biased. Thus a focus on agenda-setting provides more and better data from which scholars can study judicial ideology.

Certiorari is also independently important. While it rarely receives headlines, sophisticated observers and the justices themselves recognize the paramount importance of the Court’s agenda-setting decisions. Justice Brennan called it “second to none in importance,” (Brennan Jr., 1972) while Justice Thurgood Marshall remarked that the power to not decide a case is “among the most important things done by the Supreme Court,” (Marshall and Tushnet, 2001) and a leading legal scholar of Supreme Court practice has suggested the Court’s discretionary power over its docket “may be more important than what the Court decides on the merits” (Hartnett, 2000, p. 1737). So, understanding how ideology affects certiorari is valuable in itself.

Scholars have not been idle in this effort. Previous empirical studies have shown that specific factors—among them circuit splits, amicus briefs, and recommendations from the Solicitor General—signal that a case is important and are likely candidates for a grant of certiorari (Caldeira and Wright, 1988; Ulmer, 1983, 1984). Alternatively, studies have looked at the relationship of cert votes to subsequent votes on the merits, finding a correlation between voting to grant cert—the shorthand for certiorari—and voting to reverse on the merits (Palmer, 1990). Consistent with this, qualitative evidence suggests justices care about both the intrinsic legal importance of the case as well as the ideological effects of the likely case outcome (Perry, 1991).
This article builds on this earlier work and develops a structural model of voting at the certiorari stage that accounts for both case importance and potential ideological gains. The model builds on existing ideal point models but differs in several ways from the more familiar Martin-Quinn scores (Martin and Quinn, 2002). The main difference is that the ideal points estimated here come from a different set of votes. The model itself contains three key modifications to adapt the standard model to the context of certiorari. First, the model takes seriously the larger set of voting options at certiorari. As will be described more fully later, justices have three options instead of two at the certiorari stage. Second, the model accounts for justices’ preferences for clarifying law: for replacing an unclear precedent with a clear rule. Third, justices differ in their willingness to take on cases in general. For instance, some may place a higher value on their leisure time than others. Once developed, the model is estimated using original data hand-coded from the Blackmun Archive. The estimation returns several different values of substantive interest, including: ideal point estimates that are more stable than those recovered from merits votes, information about justices’ preferences over the size of the docket, and whether the case—if taken—is expected to move the law to the left or to the right. The model also could easily be extended to a multidimensional or dynamic context or combined with a model of voting at disposition to provide a richer understanding of ideology at the Court over time and throughout the decision-making process.

**Existing Approaches**

Given its substantive importance, the process through which the Court grants the writ of certiorari—the mechanism through which the Court agrees to decide a case—scholars in both law and political science have examined the process. The best source of information on certiorari in general remains Perry (1991), which makes use of countless interviews with justices and clerks to provide a qualitative look at the process. Empirical scholars have narrowed in on specific questions over time. For instance Caldeira and Wright (1988), McGuire and Caldeira (1993), and Black and Owens (2009) look at how case-specific factors affect certiorari decisions. Caldeira, Wright, and Zorn (1999), Ulmer (1983), and Ulmer (1984) all find that the Court is more likely to take a case when there is a disagreement between lower courts or when a lower court splits from clear Supreme Court precedent. Similarly, Palmer (1982), Boucher and Segal (1995), Caldeira, Wright, and Zorn (1999), and Benesh, Brenner, and Spaeth (2002) consider whether cert votes are affected by the anticipated vote on the merits. Relatedly, Palmer (1990) finds that there is a strong tendency for justices to vote to grant petitions in cases where they will subsequently vote to reverse. While Boucher and Segal (1995), Caldeira and Wright (1988), Caldeira, Wright, and Zorn (1999), and Cameron, Se-

---

1 Also, Martin-Quinn scores are based on a dynamic model, whereas my model is static. As such, it is more like the model from Clinton, Jackman, and Rivers (2004).
gal, and Songer (2000) examine certiorari as a method the Court can use to induce compliance by lower courts, other scholars, such as Harvey and Friedman (2009) and Owens (2010) have debated the existence and extent of Congress’ influence on the Court’s agenda.

Surprisingly, despite the wide-ranging interest in certiorari, scholars have yet to apply the tools of the scaling revolution to the cert process. This is surprising since the use of item response theory (IRT) models to estimate measures of ideology has become ubiquitous in the discipline. The project began with Poole and Rosenthal’s efforts to scale congressional rollcall votes (Poole and Rosenthal, 1985, 1991). The technique has since been expanded to state legislatures, (Shor and McCarty, 2011), and the British House of Commons, (Kellermann, 2012), among other political institutions. Similarly, judicial politics scholars have used these tools to recover ideal point measures from the justices’ votes to reverse or affirm lower courts on the merits (Martin and Quinn, 2002; Bailey, 2007).

These models rely on a spatial theory of voting that has roots that go back over three-hundred years. During the time of the French Revolution, supporters of the Ancien Regime sat on the right of the General Assembly while the revolutionaries sat on the left. Over the decades, supporters of monarchy continued to sit on the right side of the chamber while more revolutionary types sat on the left, while moderates were in the middle. For a century, “left” and “right” referred to the seating arrangements, but by the 1870s, French political parties adopted these words to describe their own policies. Modern spatial voting theory entered the mainstream of political science with the development of the median voter theorem (Black, 1948; Downs, 1957).

Spatial voting theories begin with the assumption that policies may be represented as points in a space. For example, if politics is one-dimensional — say liberal and conservative — then policies may be placed on a line with more liberal policies to the left of more conservative policies. As with policies, individuals may also be identified as points within this space, with conservative legislators to the right of progressives. The point on the line that represents the legislator is her ideal point. The policy that occupies the same point as the legislator is that legislator’s ideal policy.

Building from this heuristic, scaling models assert that individuals prefer policies “closer” to policies that are farther away. That is, when faced with a choice between two policies, A and B, a legislator prefers whichever policy is closer to her ideal policy. In a one-dimensional legislative context, the status quo policy is a point on the line, and the bill up for a vote represents a different policy that also sits on the line. The point exactly half-way between the status quo and the proposed policy is the cutpoint that divides legislators. All else equal, those on the proposal side of the

---

2In contrast, Segal and Cover (1989) develops measures of judicial ideology from newspaper editorials arising between the date of nomination and confirmation. By construction, these scores do not depend on any decision a justice makes as a justice. One advantage measures from cert votes has over newspaper based scores is that the model can be re-estimated for each justice in every term, whereas Segal-Cover scores provide a single number for every justice. How important this distinction is will vary by need.
cutpoint are closer to the proposed bill than the status quo and should vote Aye. Those on the other side of the cutpoint are closer to—and therefore prefer—the status quo. They should vote Nay.

A particularly prominent and user-friendly scaling model built on this spatial framework was provided by Clinton, Jackman, and Rivers (2004). Their model operationalized the spatial voting model by making individual legislator’s utility from a given policy a function of the squared distance between her ideal point and that policy plus a random shock. By assuming the random shock has a standard normal distribution, it is straightforward to use a Bayesian probit model to jointly estimate the ideal points of the legislators and the cutpoints of the various rollcalls. The challenge is to ensure that the structural model is theoretically grounded in the institutional realities of the data generating process (Clinton and Meirowitz, 2003).

A Brief Overview of the Cert Process

Among the most important, nebulous, and political decisions the Court makes in any case is whether to take it in the first place. The importance of agenda-setting will be unsurprising to social scientists, but it is also quite apparent to legal scholars who have made careful studies of the process and the justices themselves (Hartnett, 2000; Brennan Jr., 1972; Marshall and Tushnet, 2001).

The rules governing agenda-setting at the Court have changed dramatically over time. For more than a century, the Supreme Court’s docket was largely mandatory. In 1925, however, Chief Justice Taft and the other justices pressed for, and Congress passed, the Judges’ Bill, which gave the Court greater control over the cases it took. While the bill did not eliminate the Court’s mandatory jurisdiction entirely, Congress continued to whittle away at the remaining parts of it until by the late 1980s, the Court had almost complete discretion over its caseload. This discretion operates through the writ of certiorari.

The Court receives between 6,000-8,000 petitions for the writ every year. From these, the Court will grant certiorari only about 1% of the time. Granting the writ is a matter of pure discretion. The justices rarely provide reasons for their votes, and there is no clear formula for what makes a petition certworthy. The guidelines, such as they are, appear in Rule 10 of the Supreme Court’s Rules. The Court professes an interest in resolving circuit splits, clarifying federal law, rebuking lower courts that misstate precedent, and deciding important questions. Still the guidelines provide little useful guidance in practice. Indeed, Perry (1991) calls the criteria “tautological.”

The formal certiorari process begins when a party to a case files a petition. Given the number of petitions, the Court streamlines consideration of the petitions by way of the “cert pool.” Not all justices participate in the cert pool. For the years considered in this paper, every justice save
Justice Stevens participated in the pool. Those justices who do participate assign their clerks to contribute memos to the pool. The process works as follows. When a petition arrives, the Clerk of the Court assigns the petition to one of the clerks. That clerk reviews the case file and writes a memo to be circulated to all participating justices. The memo describes the facts of case, the arguments of the parties, the legal issues in question, and makes a recommendation as to whether or not to grant the petition. The memos are often the only information the justices have about the petitions, as even the most certworthy petitions get minimal review at the preliminary stage.

Once justices receive the memo, they create the “Discuss List.” The Chief Justice gets things started by circulating a preliminary list of cases to discuss at the weekly conference. Any justice may add a case to the list on his own initiative by sending a request to the Chief. The list circulates throughout the week with cases being added or removed, and once finalized, all petitions not on the list are summarily denied. Interestingly, the Court maintains a strong norm to not discuss petitions prior to conference. Those discussions at conference, though, are quite limited.

At the conference, the justices vote on each petition included on the Discuss List. A justice may vote to deny the petition, grant it, or Join-3. The last of these options is something of a “timid grant” (Perry, 1991). It has utility because the Court follows a minority voting rule at the agenda-setting stage. The Rule of Four requires that it only takes four justices to support a petition for it to be granted. The Join-3 option simply says that a justice is willing to be the fourth vote. That is, it counts as a grant if three other justices vote to grant the petition, but it would count as a denial if two or fewer justices vote to grant. In the extreme case where two justices voted to grant and seven voted Join-3, the petition would be denied. If the Court grants the petition, the Court will take the case and subsequently decide it on the merits under majority rule.

The Structural Model

This section develops a structural model of voting during the Court’s certiorari process. Justices are \( i \in \{1, 2, ..., I\} \) with ideal points \( x_i \in \mathbb{R} \), and cases are \( j \in \{1, 2, ..., J\} \). Let \( z_{ij} \) be justice \( i \)'s vote at cert in case \( j \). When presented with a case, each justice chooses an action \( z_{ij} \in \{\text{Deny, Join-3, Grant}\} \). If the Rule of Four is satisfied—that is if at least four justices vote to grant or if three vote to grant and one more casts a Join-3—the Court takes the case, decides it,
and issues a new policy.

In contrast to the usual assumption that policies are points on the real line, in this model, policies are random variables. This reflects the inconsistency and uncertainty inherent in the judicial hierarchy and the Court’s decision-making process. Consider first the policy that would result if the Court takes the case. From the perspective of the justices as they decide whether or not to accept the case at all, they almost certainly lack the information—merits briefs, amicus briefs, time to think through the issues and circulate opinions—to be certain of the eventual outcome. That said, they may have an idea about the range and probabilities of different possible outcomes. This uncertainty is straightforwardly represented by a random variable, $P_j$, with finite mean and standard deviation, $\mu_{p_j}$ and $\sigma_{p_j}$.

When thinking of the status quo policy, it is important to keep in mind that the Court professes a desire to take cases to resolve splits in lower courts. The existence of a split is a clear indication that there is not a single status quo policy; rather, there is a range of policies that are applied differently in lower courts. Even when there is not a clear split, it may be that a statute or precedent is sufficiently unclear that the same rule, though stated identically in different jurisdictions, is applied unevenly. Under either interpretation, a better way to model current policy is as a random variable rather than a point. Therefore, denote the status quo as $Q_j$, with finite mean and standard deviation, $\mu_{q_j}$ and $\sigma_{q_j}$, respectively.\(^7\)

Combining these two permits the model to account for the Court’s desire to “bring clarity” to the law. When the Court can resolve circuit splits and other status quo policies with high variances, it reduces the variance. Similarly, if the current law is relatively clear (low variance), then even if taking the case would likely mean that the new policy would be closer to a justice in expectation, if the new policy would introduce uncertainty (higher variance), the justice may be unwilling to make that trade.

To build out the model, begin with the policy concerns from which measures of ideology will emerge. The policy utility for justice $i$ if the Court rejects case $j$ is

$$V_{i,j}^{\text{policy (deny)}} = \int_{Q_j} \left[-(x_i - q_j)^2 f(q_j) dq_j\right]$$

$$= \int_{Q_j} \left[-(x_i^2 + q_j^2 - 2x_i q_j) f(q_j) dq_j\right]$$

$$= -\left(x_i^2 + \mu_{q_j}^2 + \sigma_{q_j}^2\right) - 2x_i \mu_{q_j},$$

where $f(q_j)$ is the pdf of $Q_j$. Similarly, if $g(p_j)$ is the pdf of $P_j$ and the Court takes the case, the

\(^7\)Allowing $Q$ and $P$ to be random variables is a less restrictive assumption than that they are fixed points. The model is unchanged if $Q$ and $P$ are fixed.
justice receives a policy payoff of:

\[
V_{ij}^{\text{policy}}(\text{grant}) = \int_{P_j} - \left(x_i - p_j\right)^2 g(p_j) dp_j \\
= \int_{P_j} - \left(x_i^2 - 2x_i p_j - p_j^2\right) g(q_j) dp_j \\
= -\left(x_i^2 + \mu_{p_j}^2 + \sigma_{p_j}^2\right) - 2x_i\mu_{p_j} 
\]

(2)

Define the net policy payoff from taking the case to be

\[
\theta_{ij}^{\text{policy}} = V_{ij}^{\text{policy}}(\text{grant}) - V_{ij}^{\text{policy}}(\text{deny}) \\
= x_i^2 + \mu_{q_j}^2 + \sigma_{q_j}^2 + 2x_i\mu_{q_j} - x_i^2 - \mu_{p_j}^2 - \sigma_{p_j}^2 - 2x_i\mu_{p_j} \\
= \left(\mu_{q_j}^2 - \mu_{p_j}^2\right) + \left(\sigma_{q_j}^2 - \sigma_{p_j}^2\right) - 2x_i\left(\mu_{p_j} - \mu_{q_j}\right) 
\]

(3)

Two parts of the final line in equation 3 are substantively interesting. The second term compares the variances of the status quo and proposed policies. As previously suggested, if the current policy has a high variance and the proposed policy will create a more consistent rule (low variance), then the net payoff from this piece is high. This would make every justice more likely to grant the petition. In contrast, if the current policy has a low variance and changing the law would introduce uncertainty, this term becomes negative, and every justice is more likely to deny the petition.

Finally, the third quantity interacts a justice’s ideal point with the signed distance between the expected values of the two policies.\(^8\) Larger distances between policy options mean the magnitude of this term will increase because the difference within the parentheses will be larger. Note that if the difference inside the parentheses is positive, then the expected value of future policy is larger than the status quo policy. That is, the policy the Court is expected to announce is to the right of the status quo. Importantly, while this difference is not fully identified, once we orient the space to put conservatives to the right of liberals, the sign is identified. Similarly, we can be confident that if the difference in expected means in one case is six and in another it is two, we can be confident that while both cases would likely move the law to the right, the former will move the law further.

Cases also vary in importance. For example, there is strong evidence that when the Solicitor General wants the Court to take a case, the Court does so. Similarly, when there is a lot of money at stake, when a federal statute is called into question, or when lower courts flout clear precedent, the Court takes notice. Most cases, however, are not intrinsically certworthy. Denote the intrinsic certworthiness of a case as \(\kappa_j\). This is a case-specific payoff given to each justice for taking the case. Higher values of \(\kappa_j\) signal important cases, and when \(\kappa_j\) is large, it is more likely that every justice will vote to grant the case. On the other hand, some cases are such a waste of time, they

\(^8\)The part in parentheses is the item discrimination parameter in the traditional IRT framework.
could provide negative utility.

Finally, just as cases vary, justices differ in their preferences for taking cases. There are several possible reasons for this. First, justices differ in their views of the proper role of the Court. For instance, Byron White thought it was the Court’s job to decide lots of cases, and in particular, to settle virtually all circuit splits. Other justices think the Court should save review for important issues of national concern where the Court can lead the way. Second, some justices may have more energy or be more efficient at writing opinions. If a justice is relatively young and eager, she may want to take lots of cases. If a justice is getting older and slowing down, spending the time to read all of the briefs and write the opinions may be more taxing. Similarly, if one justice places a premium on spending time at a vacation property in Florida while another lives in D.C. and enjoys working, the latter will likely want to take more cases than the former. The model accounts for these differences with a justice-specific constant, $\alpha_i$. This term is effectively justice $i$’s opportunity cost for taking a case.

Define these non-policy benefits and costs for maintaining the status quo or taking the case as follows:

\[
V_{ij}^{non-pol}(deny) = 0
\]

\[
V_{ij}^{non-pol}(grant) = \kappa_j - \alpha_i
\]  

(4)

\[
\theta_{ij}^{non-pol} = V_{ij}^{non-pol}(grant) - V_{ij}^{non-pol}(deny) = \kappa_j - \alpha_i
\]  

(5)

Now, define a latent variable

\[
z_{ij}^* = \theta_{ij}^{policy} + \theta_{ij}^{non-pol} + \epsilon_{ij}
\]

\[
= \Lambda_j - \alpha_i - x_i G_{j}^{cert} + \epsilon_{ij}
\]  

(6)

where $\Lambda_j = (\mu_{qj}^2 - \mu_{pj}^2) + (\sigma_{qj}^2 - \sigma_{pj}^2) + \kappa_j$, $G_{j}^{cert} = 2 (\mu_{pj} - \mu_{qj})$, and $\epsilon_{ij} \sim \mathcal{N}(0, 1)$ is a random shock to a particular justice’s utility in a particular case.

To complete the model, recall that the qualitative evidence suggests justices treat the Join-3 vote as a “timid” vote to grant. This suggests that to cast such a vote, the justice must expect positive utility from taking the case, but not enough to cast a “strong” vote to grant. This interpretation suggests a level of sincerity on the part of the justices that seems strange given the curious nature of the Join-3 and the obvious strategic opportunities it would provide.

However, the assumption of sincerity is consistent both with the justices’ qualitative accounts and the data. Given the way the Join-3 operates, there is no strategic reason for the justice who votes last to ever cast a Join-3.\footnote{The justices vote in order of seniority, so the junior justice always votes last.} If three justices have already voted to grant, it will function as...
a grant. If less than three have voted to grant, a Join-3 is functionally a deny. If justices use the
Join-3 strategically, then there is little reason for justices at the end of the line to make use of the
vote. And yet, in the data Justice Souter cast eighteen Join-3 votes as the junior justice in 1990.
Justice Thomas cast seven as the junior justice in 1992, and Justice Ginsburg cast six as the junior
justice in 1993. This behavior is strong evidence that the justices’ spoke truthfully to H.W. Perry
when they called the Join-3 a “timid” vote to grant, and it strengthens the argument in favor of
treating join-3 votes as an intermediate category.

Treating the Join-3 as a weak grant suggests using an ordered probit specification instead of
the usual probit. But one cannot assume that all justices think of Join-3s in the same way. As with
preferences over taking cases at all, Justices differ in how much they like to use the Join-3. For
instance, in 1989, Justice Blackmun cast 45 Join-3 votes while Justice Stevens did not cast a single
Join-3. This means the threshold that separates the Join-3 from the grant votes must be allowed to
vary across justices and estimated accordingly. Denote the threshold level of utility required for
justice $i$ to move from a Join-3 to a solid Grant as $\lambda_i$.

Once this is done, the voting at the cert stage can be modeled as follows.

**Model 1.**

$$
z_{ij}^{\text{cert}} = \begin{cases} 
0 (\text{Deny}), & \text{if } z_{ij}^* < 0 \\
1 (\text{Join-3}), & \text{if } 0 \leq z_{ij}^* < \lambda_i \\
2 (\text{Grant}), & \text{if } z_{ij}^* \geq \lambda_i 
\end{cases} \quad (7)
$$

Statistically, the model has many advantages over current measures of ideology at the Court.
First, there are far more non-unanimous cert votes in a term than votes on disposition. Second, by
introducing the Join-3 as an intermediate category, the cert model has more granular information
about justices’ preferences, which allows for more precise estimates. Third, insofar as certiorari is
the selection stage, running statistical models on the subsequent dispositional stage may introduce
selection bias. The subsequent votes at disposition are not random draws from the set of possible
cases; they are selected by the justices themselves. To be sure, even the cases considered for
certiorari are unrepresentative of the cases filed in lower courts across the country, and it is quite
possible that even the results in this paper suffer from selection bias (Kastellec and Lax, 2008).
Still, there is plainly less selection at cert than at disposition.

**Formal Comparison with the Martin-Quinn and Bailey Models**

The most obvious difference between the above model and that from Martin and Quinn (2002) is
that their model is dynamic, while the model above examines only a single period. Accordingly,
I compare the cert model to the static version of the Martin-Quinn model, which reduces to that of Clinton et al. (2004). Bailey (2007) uses this same model, though it expands the set of voters to include legislators and the Executive. After setting aside these obvious differences, several important distinctions remain.

The first is that their model describes the disposition of a case, while the present model examines the cert stage. Therefore, the set of possible actions is different. At disposition, the justices can either vote to reverse or affirm on the merits. Thus, in contrast to the cert model, the options at disposition are binary. Their model also differs in that the policy options facing the justices are points, \( r_j, a_j \in \mathbb{R} \), for the policies that will result from reversing or affirming, respectively. In contrast, the certiorari model assumes that the policy options are random variables rather than points. That said, there are clear similarities. In both models, justice \( i \) has ideal point \( x_i \in \mathbb{R} \), each justice chooses between policies based on a quadratic loss function, and both presume random, normal shocks to the justices utility functions.

The Martin-Quinn model asserts that justice \( i \) will vote to reverse in case \( j \) if \( u_{ij}^{\text{rev}} > u_{ij}^{\text{aff}} \). Define a latent variable, \( y_{ij}^* \), as follows:

\[
y_{ij}^* = u_{ij}^{\text{rev}} - u_{ij}^{\text{aff}} \\
= (a_j - x_i)^2 - (r_j - x_i)^2 + \epsilon_{ij} \\
= (a_j^2 - r_j^2) + x_i (r_j - a_j) + \epsilon_{ij} \\
= b_j + x_i G_{ij}^{\text{disp}} + \epsilon_{ij}
\] (8)

where \( b_j = a_j^2 - r_j^2 \) and \( G_{ij}^{\text{disp}} = 2(r_j - a_j) \).

The actual votes of the justices are then described by the following model:

**Model 2.**

\[
y_{ij} = \begin{cases} 
0 \text{ (affirm),} & \text{if } y_{ij}^* < 0 \\
1 \text{ (reverse),} & \text{if } y_{ij}^* \geq 0 
\end{cases}
\] (9)

So the probability that a justice votes to affirm is \( Pr(y^* < 0) = Pr(b_j + x_i G_{ij}^{\text{disp}} + \epsilon_{ij} < 0) = Pr(\epsilon_{ij} < -b_j - x_i G_{ij}^{\text{disp}}) = \Phi(-\tilde{z}_{ij}), \) where \( \tilde{z}_{ij} = b_j + x_i G_{ij}^{\text{disp}} \).

The comparison between equation 6, describing the latent variable for certiorari, and equation 8, describing disposition, is instructive. In the former, there is a justice fixed-effect that measures how costly it is for a justice to take a case. In effect, it controls for different justices’ preferences over docket size. Since there is no similar, theoretically compelling reason to think that justices have a preference to reverse rather than affirm, equation 8 does not include such a variable.
The case-specific constant, or difficulty parameter, is estimated in both models. The interpretation, however, differs between the two models. In the cert model, this parameter includes the difference between the variances of two different random variables and a certworthiness constant in addition to the difference of the squared expected values. In the disposition model, since there is no theoretical reason to prefer reversing and policies are points, the difficulty parameter is simply the difference of squares.

These differences in interpretation matter if the researcher is interested in locating the cutpoint for or midpoints between policies. Under the standard model, the midpoint between the current status quo policy (from affirming) and the proposed policy (following reversal) is easily recoverable as

\[ m_j = \frac{a_j}{G_{\text{disp}}} = \frac{a_j + r_j}{2}. \]

The analogous quotient from the cert model is

\[ \Lambda_j \frac{G_{\text{cert}}}{G_{\text{disp}}} = \mu_{q_j} + \mu_{p_j} + \frac{\sigma_{q_j}^2 - \sigma_{p_j}^2}{G_{\text{cert}}} + \kappa_j, \]

which will not be at the midpoint because it is affected by the variance and certworthiness components of the fixed-effect. Further, this quantity is not a cutpoint at all, since individual justices have different reserve utilities. If every justice had the same utility threshold to move from a deny to a Join-3 and then from Join-3 to grant, then it would be straightforward to estimate a common cutpoint. But if justice A requires a larger payoff to vote to grant than justice B, then the expectation of a common cut-point vanishes.

**Data and Estimation**

The certiorari process unfolds behind closed doors, and the votes are not made public by the Court. Fortunately, over the years, justices have kept the records of these votes, which become accessible to scholars with the release of the justices’ papers. The data for this study come from the papers of Justice Blackmun as maintained in the Blackmun Archive (Epstein, Segal, and Spaeth, 2007). The Archive contains a docket sheet for almost every case from 1986 through 1993. On these docket sheets, Justice Blackmun recorded the votes of each justice. The docket sheets are available in pdf formats, so to turn the docket sheets into usable data, I had a research assistant hand-code the cert votes into a database. To check for quality, the resulting coded votes were checked against similar coding efforts made by other scholars.\(^{11}\)

\(^{11}\)For this, I am thankful to Justice Stras and Margaret M. Cordray for sharing their data.
Figure 1: There are Far More Non-Unanimous Votes at Certiorari than Disposition

One of the advantages of these data is that there is so much of it, at least relative to the justices’ votes on the merits. While merits votes are easily retrievable and the subject of much study, the Court’s light caseload in recent years means that there are less data to work with on the merits than at certiorari. Figure 1 compares the number of non-unanimous certiorari petitions to the number of cases decided non-unanimously on the merits.\textsuperscript{12} The figure shows that in every period there are significantly more agenda-setting votes available for analysis than merits votes. In total, there are 2289 cert petitions in the dataset where the voting was not unanimous. This yields well over twenty-thousand votes across thirteen justices.

To estimate the model, I follow the standard assumption that the error terms in equation 6 are independent. Under that assumption, the likelihood is

\[
L(\hat{z}_{ij}|z) \propto I \prod_{i=1}^{I} \prod_{j=1}^{J} \{ \Phi (-\hat{z}_{ij}) \}^{1(z_{ij}=\text{deny})} \times \\
\{ \Phi (\lambda_i - \hat{z}_{ij}) - \Phi (-\hat{z}_{ij}) \}^{1(z_{ij}=\text{Join}-3)} \times \\
\{ 1 - \Phi (\lambda_i - \hat{z}_{ij}) \}^{1(z_{ij}=\text{grant})}
\]  

\textsuperscript{12}Unanimous votes are removed from the analysis because they do not offer any way to differentiate among the justices.
where \( z_{ij} = \Lambda_j - \alpha_i - x_iG_j \). As is the case with these kinds of ideal point models, the likelihood is not identified unless the researcher restricts parameters. But if estimated in a Bayesian context, such identification is not absolutely necessary. Instead, careful selection of priors guarantees a well-behaved posterior.

I place relatively uninformative priors on all parameters. Specifically, I assume normal, mean 0 priors with a standard deviation of 5 for all justice-specific and case-specific parameters. These priors are quite conservative. They incorporate no external information about the justices’ ideologies or preferences over taking cases. Similarly, the priors build in no information about the relative importance of cases or their ideological valences.

Posterior distributions for individual parameters are recovered as in Clinton, Jackman, and Rivers (2004) with two modifications. First, the block of justice-specific parameters includes a constant term in the cert model that is absent from their model. Second, I incorporate an additional step in the estimation algorithm to draw justices’ individual thresholds between the Join-3 vote and a grant. The estimation process and diagnostics are detailed further in the accompanying appendix.

Recovering Quantities of Interest

The cert model yields three particular measures of interest. First, it provides a new set of ideal points for each justice. Second, the model measures each justices’ relative willingness to take cases, or equivalently, preferences over taking more or fewer cases during a term. Finally, it predicts the anticipated direction of the decision, i.e. whether the law is anticipated to move to the left or to the right as a result of the subsequent case outcome as well as providing a measure of case importance.

Figure 2 plots the posterior distributions of the ideal points for each justice in 1988 based on their cert votes. Consistent with Martin and Quinn’s estimates of ideology, Byron White is the median justice in 1988, while O’Connor, Scalia, and Kennedy are virtually indistinguishable ideologically. One difference between their scores and the cert data appears in the rank ordering of Justices Stevens and Blackmun. At cert, Blackmun appears more centrist than Stevens, while their scores show Stevens to the right of Blackmun.13 Similarly, certiorari data suggests that Justice Brennan was the most liberal justice, while measures drawn from dispositional votes places Justice Marshall in the left-most slot.

The cert model also estimates justices’ relative willingness to take cases or use the Join-3. Recall that the model assumes that justices derive a common amount of utility from the importance

---

13 Similarly, a static model of the dispositional votes in 1988 places Blackmun to the left of Stevens. In both the dynamic Martin-Quinn scores and the static model of dispositional votes, the two converge in 1989, and Stevens then moves clearly to the left of Blackmun in 1990.
of a case but a justice-specific ideological payout. But while justices not only differ in their total utilities, they also differ in how much utility they require in order to cast a particular vote. Some place a low bar and others use a higher threshold. Similarly, justices differ in their willingness to deploy the Join-3.

Figure 3 below shows both of these measures for 1988. The left figure shows each justice’s threshold for denying a cert petition. That is, any petition that does not provide a justice with at least the amount of utility shown in the left figure should be denied. The measures suggest that justices differ in their willingness to deny cases. Some use a higher threshold—deny more cases—than others. This signals a desire to take fewer cases.

These results are consistent with qualitative accounts. Scholars and observers have noted Justice White’s preference to take lots of cases. Specifically, he believed the Court should almost always get involved if there was a clear circuit split. This should imply that White had a very low reserve utility when compared to his colleagues. As expected, Justice White had the lowest reserve utility indicating he favored taking more cases than other justices.

Similarly, it appears that conservative justices have lower thresholds than more liberal justices. This makes sense, since the Rehnquist Court was famously conservative, and the more liberal justices would prefer to take fewer cases in such an environment. Fewer cases means their ideological adversaries have fewer opportunities to work mischief. The way to keep cases from coming onto the docket is to vote to deny, which should suggest that liberal justices should have relatively higher reserve utilities. This, too, is borne out in the results.

While it is important to note that the scale provides no meaningful information about the absolute magnitude of these effects, it is important that there are significant differences between justices. Statistically, this means estimates from a model that does not control for differences in justices’ preferences for taking cases would be inconsistent. Substantively, it indicates that the justices have meaningfully different preferences over the size of the docket.

The justices also differ in their willingness to deploy the Join-3 as a device. Figure 3b shows the mean and 95% credible intervals of the amount of marginal utility a petition must offer a justice
(a) Justices have different thresholds to deny cases.  
(b) Justices differ in willingness to use Join-3.

Figure 3: Means and 95% credible intervals of justices’ thresholds to deny or grant a cert petition in 1988.

to move from a Join-3 to a full grant. Some justices almost never use the Join-3. In 1988, Justice Kennedy cast only one such vote, Justices Scalia and Stevens cast two, while Brennan and Marshall cast three and four, respectively. Accordingly, these justices have small Join-3 windows. On the other hand, Justice Blackmun cast 31 Join-3 votes and Rehnquist 23. The model captures their greater willingness to vote Join-3 and estimates a larger Join-3 window for these justices.

Finally, the model returns some useful—albeit imprecise—information about the individual petitions. In particular, estimating the model returns information about the importance of each cert petition and the anticipated ideological shift in the law that would result from taking the case. Case importance is roughly captured by the case-specific constant, $\Lambda_j$, while information about the ideological valence of petitions is found in $G_{j}^{cert}$. Figure 4 plots these values for each cert petition from 1988.

Unsurprisingly, the petitions the Court granted tend to be more important. This is clear in figure 4 where case importance is placed on the vertical dimension. The blue dots representing granted petitions are generally more important than (located above) the orange dots representing the denied petitions.\textsuperscript{14} Interestingly, there is no similar bias to the right or left of the ideological spectrum. The Court grants petitions that are favored by both the ideological left and right.

**Validation**

As an initial check, I look to see if the model is doing a reasonable job classifying the votes. As a baseline, I compare the cert model’s performance to a simplistic model that assumes every justice plays a pure strategy—always grant or always deny. For example, Chief Justice Rehnquist voted

\textsuperscript{14}The cluster of granted petitions at the top of the figure are petitions where
to deny cert in just over half of the nonunanimous cert petitions included in the dataset for 1989.\footnote{While the model estimates Join-3 votes separately, for the analysis presented in this and the following sections, I treat Join-3 votes as votes to grant.} A naive model that predicted Rehnquist always voted to grant would correctly classify just over half of his cert votes correctly. In contrast, the cert model correctly predicts well over 80\% of the Chief’s votes. Figure 5 below compares the classification success rate for the model presented above compared each justice’s modal vote. In the figure, the purple range shows the mean number of votes correctly classified across all 400 runs of the Gibbs sampler along with the 95\% credible intervals. The black dots show the proportion of votes predicted by assuming the justice cast his...
most common vote on every petition. The figure clearly demonstrates that the model predicts cert votes better than a pure strategy of either “deny” or “grant.”

Figure 5: Model Predicts Better than the Mode

But validation requires more than this, since as with all ideal point models, the cert model, at its core, looks for the best way to summarize data. So far, it is only clear that the summary is better at classifying votes than a naive model. The challenge now is to show that the model actually capture something of substantive value. It is always a challenge to determine whether the estimates returned in item response models represent the preferences of the actors or if they are picking up something else. Since statistics cannot validate the substance of the results, the arguments in favor of the returned measures must rely instead on the conformity of the results with qualitative accounts and consistency with alternative measures. Figure 3 above provides a first step in this direction. As argued above, the measures of the justices’ thresholds for denying certiorari comport with qualitative accounts of Justice White’s behavior and strategic accounts of liberal justices displaying a reticence to add cases to the docket of a conservative Court. Similarly, the Join-3 windows comport with the qualitative and quantitative records that suggest justices vary widely on their willingness to use the Join-3 as a tool during the cert process.

Whereas those measures are entirely new, measures of judicial ideology are presently available and widely accepted as at least approximate measures of justices’ ideologies. Validation of ideal points, then, depends in large part on whether they seem plausible, which means a clear relationship with existing measures.
Figure 6a plots the correlation between the ideal points recovered by the cert model and those recovered from dispositional votes in 1989. Ideal points recovered from dispositional votes are widely considered to capture ideological positions along a liberal-conservative dimension. The correlation between the ideal points is .91 for 1988. Figure 6b shows that 1988 is not an outlier. In every term between 1986 and 1992, the ideal points correlate at above .9. The strong correlation between the ideal points recovered from the cert model and the ideology scores from the merits votes suggests that the justices’ votes in each period reflect similar ideological impulses. The figure also shows that 1993 is something of an outlier—though examination of the deviation is deferred to a later section.

(a) Comparison of ideal points recovered from votes at cert and disposition in 1988. (b) Annual correlation of ideal points drawn from cert and disposition.

Figure 6: Recover Similar Ideal Points from Certiorari and Disposition

Having built, estimated, and validated the cert model, one may be left with the question of whether all of this machinery was necessary. This objection could take two forms. First, one might ask why a new model is necessary. One may object that the most obvious and direct path would be to recode Join-3 votes as grants and then run a standard scaling package. If this procedure recovered useful measures, then there is no reason to invent a new model. But such measures are often implausible. For instance, in 1991, the correlation between the ideal points recovered in this way and Martin-Quinn scores is 0.001. That said, in other years, this method works well. The correlation for 1987 is 0.93. Essentially, ideal points recovered using the standard Clinton et al. model are unstable.

The reason for this instability is that the traditional model does not account for justices’ different thresholds for denying cert. What the traditional model does is find the single dimension that best summarizes the data. In some years, for example 1987, that dimension is ideology. But other times, notably 1991, the justices’ threshold values explain the most. The blue bars in figure 7 show the correlation between the ideology scores recovered from the cert model and the single

---

16These points are computed using the ideal package from the pscl library in R.
dimension recovered from running the standard model—that lacks a justice-specific constant—on the data. Similarly, the orange bars show the correlation between the justices’ threshold values from the cert model and the single dimension recovered from the traditional model. The figure shows that there is a great deal of instability in decision-making at certiorari: at least in the first dimension. In some years—1986, 1989, and 1993—variation in cert votes are largely due to differences in justices’ preferences over the size of the docket. In contrast, 1987 and 1992 are years where traditional ideology seems to explain more of the cert votes.

This figure shows that the bias from mis-specifying the ideal point model is more than a theoretical possibility. Estimating “ideal points” from cert votes without accounting for the justices’ differing preferences for docket size leads to absurd outcomes. For instance, these scores show Justice White as the most extreme justice in all but one term from 1986 until he left the Court after the 1992 term. In contrast, Justice White is an ideological moderate in both the traditional Martin-Quinn scores and in the ideal points recovered from the cert model. White’s extreme willingness to take lots of cases makes him stand out at certiorari, and if that characteristic is not captured by a justice-specific constant, it works its way through the ideal point measure making White look like
an extremist.

The results so far suggest that the cert model does a good job of classifying cert votes and of returning substantively interesting measures and that certiorari is sufficiently different from the dispositional phase that it requires a bespoke model to understand it. Still, some may raise a second objection to all of this statistical machinery based on the question of whether the results are substantively interesting or represent a meaningful advance beyond existing measures of judicial ideology. There are several reasons to be optimistic about information recovered from cert votes as compared to votes on the merits. Two are particularly important: First, since scholars can only observe merits votes for cases the Court selects from the cert round, ideal points based on dispositional votes come from a non-random sample of the possible cases. This is prima facie evidence that the ideal point measures will not be consistent. Examining the cert votes reduces the level of selection bias. Second, there are many more votes at cert than at disposition. In some years, there are more than three-times as many non-unanimous cert votes, which provides more information from which to derive ideal points.

Together this suggests that cert provides more and better data from which to derive measures of ideology. Since the dispositional docket is small and subject to the preferences of a Court with a changing membership, the rolcall votes at the disposition stage may vary from year-to-year. Thus even if justices’ ideologies do not change, measures of ideology may shift. Such changes would result from a change in the composition of the docket, but would appear as a change in ideology. Put differently, measures will be unstable over time if the characteristics of the cases over time are unstable (Ho and Quinn, 2010). Scholars have attempted to address these issues either by bridging (Bailey, 2007) or by statistical modeling (Martin and Quinn, 2002).

In contrast, certiorari provides a large pool of cases that emerge from a process largely outside of the justices’ direct control. This suggests that the composition of the certiorari agenda will be relatively stable over time. If so, this should result in more stable estimates of ideology without resorting to modeling or bridging.17 Figure 8 below shows the year-to-year correlation between three different measures over consecutive years. The black circles show how the ideal points recovered from dispositional votes in one year correlate with ideal points from those estimated from the previous year’s dispositional votes. Similarly, the blue diamonds show the year-to-year correlation of ideal point measures derived from cert votes. The orange squares show the same correlations for the justices’ threshold values for denying certiorari.

Ideal points derived from merits votes in individual years are somewhat consistent, though in 1990 and 1993, the year-to-year consistency falls to about .8.18 This suggests that when Souter—

17 To be clear, the model as presented does not yield scores that are directly comparable over time. Still, the changes between recovered ideal points are small relative to those recovered from dispositional votes. Those changes are in the process of creating dynamic scores through use of a dynamic linear model or bridging.

18 Of course, Martin-Quinn scores are remarkably stable, but that is a function of their estimation strategy. Their
a Republican appointee—replaced the leader of the liberal bloc (Justice Brennan) in 1990, the ideal points of the remaining justices—as revealed through the merits votes—shifted. They then remained relatively stable until 1993 when Ginsburg replaced White. In contrast, the ideal points for the remaining justices recovered from cert votes remained far more stable in 1990 and 1993.

This discussion points toward a possible explanation of what makes 1993 such an outlier. Recall from figure 6 that until 1993, ideal points recovered from cert votes correlated at above .9 with those recovered from merits votes. Similarly, figure 8 suggests that measures of each justice’s reserve utilities shifted sharply in 1993, and the ideal points recovered from the merits votes also noticeably changed. The most likely cause of these deviations is a change in Court personnel.

Byron White left the Court in 1993 and was replaced by Justice Ginsburg. For present purposes, White and Ginsburg differ in two important ways. First, White was a centrist conservative. He was often the median justice, and he dissented in *Roe v. Wade*. In contrast, Ginsburg—the co-founder of the Women’s Rights Project at the ACLU—was noticeably more liberal. This likely signaled to the rest of the Court that some case outcomes would be more liberal going forward and presumably made conservatives less likely to vote for certiorari on the margin and liberals more likely to do so. Second, and almost certainly more importantly, Justice White was the most prolific advocate...
for taking more cases. Since Justice White was almost always in favor of granting cert, the other justices faced the problem of forming a coalition of three from the remaining eight. Upon White’s departure and his replacement by Ginsburg—who had more typical appetites for taking cases—the game changed and justices’ strategies had to adjust. It became harder to assemble a four-vote coalition. Cases that would have made the docket if White was still present to cast the fourth vote were instead denied.

This changed the composition of the Court’s docket, and it almost certainly influenced the ideal points recovered from the merits votes. In contrast, the petitions considered at cert were likely quite similar across the 1992 and 1993 terms. This explains why the year-over-year correlation of ideal points from the votes on the merits is only about .8, while the correlation for ideal points recovered using the cert model is .99. If we assume that the ideology should be roughly stable from year to year, this is evidence that the cert votes are better data from which to recover ideal points.

**Conclusion**

Modern scaling methods have revolutionized the study of political institutions. This is the first article to bring these tools to bear on agenda-setting at the Supreme Court. The certiorari process—through which the Court decides which cases to decide—is an incredibly important but often overlooked institution. By developing a structural model of voting at certiorari and deploying a novel dataset of certiorari votes, I provide improved estimates of judicial ideology and entirely new measures of justices’ willingness to take cases or use the Join-3 vote. The ideal point estimates come from more and better data than alternative measures and are intrinsically more stable. Finally, the model provides an empirical foundation for studies of the Court’s agenda-setting process in general.

January 6, 2018

**References**


Appendix

The results presented assume a probit specification, so $\epsilon_{ij} \sim \mathcal{N}(0, 1)$ in equation 6. Estimating the remaining parameters from that equation requires assuming a prior distribution for each of the terms. I use uninformative normal priors across the board; thus $\Lambda_j$, $\alpha_i$, $x_i$, and $G_j^{cert}$ are all given priors of $\mathcal{N}(0,5^2)$. The Gibbs sampler used to recover the posterior distribution of each parameter runs as described in Clinton, Jackman, and Rivers (2004). The primary difference between their estimation procedure and the cert model is the ordinal specification. To recover the individual justices’ threshold cutoffs between Join-3 and grant votes—$\lambda_i$ from model 1—I draw from $U(0, z_i)$, where $z_i$ is the minimum value of $z_{ij}^*$ across all $j \in J$ holding $i$ fixed.

The algorithm is coded in R and after a burn-in of 50,000 cycles, the computer runs through the sampler another 400,000 times storing the output for every 800th observation. This leaves 500 draws from the posterior of each parameter. For the ideal point estimates, these thinned observations have AR(1) parameters averaging about 0.08 (max 0.11 Marshall).