Who Controls Opinion Content? Testing Theories of Authorship using Case-Specific Preference Estimates for the US Supreme Court*

Benjamin E. Lauderdale
Associate Professor
Department of Methodology
London School of Economics
b.e.lauderdale@lse.ac.uk

Tom S. Clark
Asa Griggs Candler Professor
Department of Political Science
Emory University
tom.clark@emory.edu

November 19, 2013

ABSTRACT

Recent research has demonstrated that the preferences of US Supreme Court justices are not simply unidimensional. We demonstrate a new approach to Bayesian preference estimation that estimates case-specific preferences for justices, using a conditional autoregressive model with citation counts determining the correlation between justices' preferences across cases. By using citations between cases to identify the most relevant precedent cases, we are able to describe variation in revealed preferences across areas of the law. In applications that test theories of bargaining on the Court, these estimates enable stronger identification of variation in preferences while holding the composition of the Court constant. We show that Chief Justices from 1946 to 2005 strategically assign authorship to their colleagues in cases where those colleagues are more closely in agreement with the Chief Justice, and patterns of other justices joining those opinions are consistent with the idea that authors shape the content of opinions.

*We thank Andrew Martin and Maya Sen for helpful comments and suggestions. Earlier versions of this paper were presented at the 2013 meetings of the European Political Science Association and the American Political Science Association, the Princeton University Public Law Colloquium, and at Emory University.

1. INTRODUCTION

In *Kyllo v. United States (2001)*, the US Supreme Court held that the use of an infrared camera to reveal evidence of illegal drug production in a home was a search for the purposes of the Fourth Amendment, and therefore required a warrant. This decision created an important legal distinction from previous decisions which held that a visual inspection of a home was not a Fourth Amendment search (e.g. *California v. Ciraolo*, 1986). More recently, in *Florida v. Jardines* (2013), the court held that purposefully bringing a trained dog past a home where police suspected drugs were being grown also constituted a Fourth Amendment search. *Kyllo* is mentioned 7 times in the *Jardines* majority opinion, and 6 times in the dissent.[1]

Both *Kyllo* and *Jardines* were decided by 5-4 majorities with seemingly unusual alignments of justices. In *Kyllo*, the Court divided: Scalia, Thomas, Ginsburg, Souter, Breyer versus Stevens, Rehnquist, O'Connor, Kennedy. In *Jardines*, the Court divided: Scalia, Thomas, Ginsburg, Sotomayor, Kagan versus Breyer, Alito, Roberts, Kennedy. Of the five justices from 2001 who remained on the Court in 2013, only one of these justices (Breyer) switched sides. The unusual alliance of Scalia, Thomas and Ginsburg held that both infrared cameras and dogs noses constitute searches, while Kennedy held that neither does. If Justice Breyer was the pivotal justice in *Kyllo*,[2] the votes in these cases are all spatially consistent, with the only difference being that Breyer thinks walking a trained dog past a house is not a search, while pointing an infrared camera at the house is.However, while the votes in

[1]We did not need to wait for the decision to know that *Jardines* was closely related to *Kyllo*. The word "Kyllo" appears 28 times in the brief from the respondent in *Jardines*, and 38 times in the brief from the petitioner.

[2]This is not the only way to reconcile this pattern of votes, if Souter was pivotal but Breyer was next most weakly in favor of holding for a search, that would also yield the requisite pattern of preferences.

these two cases are spatially consistent with one another with respect to how broadly Fourth Amendment protections apply, they are highly inconsistent with decisions in other areas of the law, where Scalia and Thomas find themselves on the opposite side of the Court from Ginsburg and the other justices that joined them in these decisions.

Consistent but unusual voting patterns among the justices are not unique to Fourth Amendment cases; scholars of the Court know well that different areas of the law are associated with different cleavages on the bench. These regularities pose a challenge for the study of judicial behavior. On one hand, conceiving of justices' relative preferences as stable across areas of the law provides a great deal of analytic leverage and has the great benefit of parsimony (see, for example, the myriad applications of Martin and Quinn 2002). At the same time, the qualitatively-documented variation of preferences across areas of the law has implications for the politics of judicial decision-making.

Importantly, concern with such variation is not simply motivated by a desire to empirically measure relatively slight fluctuations around more important general patterns. Instead, systematic variation in preferences across substantive areas of the law implicates some of the most significant areas of research in judicial politics. In particular, judicial coalition-building and opinion-writing are directly implicated by systematic variation in judicial alignments. Scholars have long been interested in how voting coalitions shape the bargaining that produces Supreme Court opinions. If the justices do systematically vary in their relative alignment across areas of the law, then the Court is composed not simply of one dominant voting coalition but instead of predictable, but varying, voting coalitions. This, in turn, gives rise to empirical questions about patterns of opinion assignment, authorship, and content (e.g., Maltzman, Spriggs and Wahlbeck 2000; Carrubba, Friedman, Martin and Vanberg 2012).

In this spirit several recent studies have developed estimates of judicial preferences that account for variation across substantive issues. Lauderdale and Clark (2012) describe a kernel weighted optimal classification approach to preference estimation, which produces non-

probabilistic estimates using majority opinion citations and expert coding of cases. Lauderdale and Clark (Forthcoming) describe a many dimensional Bayesian ideal point estimator, which produces posterior distributions of case-specific preferences as mixtures of a large number of issue-specific dimensions, estimating these mixtures of dimensions using a topic model on the opinion texts. In this paper, we consider both a new approach to case-specific preference estimation and a new kind of data about case similarity. Instead of specifying our case-specific preference estimates as convex combinations of preferences along dimensions defined by a topic model (Lauderdale and Clark Forthcoming), we use the technique of conditional autoregressive (CAR) models from spatial statistics (Besag 1974) to find "local" averages of judicial behavior. Put differently, we aim to smooth justices' binary case votes into continuous measures of latent case preferences, by borrowing information from legally similar cases. To identify the set of cases that are most doctrinally relevant—adjacent in the spatial statistics sense—we use a weighted adjacency measure based on the counts of citations between opinions on different cases.

This approach dispenses with the idea of estimating a small number of preference dimensions that predict case-specific preferences, as is done in standard ideal point estimation techniques (Poole and Rosenthal 1985; Jackman 2001; Martin and Quinn 2002), and instead attempts to directly construct case-specific preference dimensions. While we apply this model to the U.S. Supreme Court, it is applicable to any other preference estimation problem where it is possible to produce a suitable adjacency matrix for the votes. In some sense, what we are doing is no longer "ideal point" estimation, as that terminology is based on estimating voter locations under a low-dimensional spatial model, typically using a random utility derivation of a choice between two points in the policy space to motivate an estimator (Poole 2005). Here, we are directly estimating latent preferences on individual cases (votes): an approach that is more akin to a nonparametric smoother that attempts to infer a mean function from points scattered across a space. In nonparametric regression

problems with continuous explanatory variables, there is usually just a single observation at any point in a space. Thus, the only way to estimate a mean function is by averaging across nearby points in the space. Here, we only observe a binary vote for each justice in each case,[3] so if we want to know which justices were likely to be closer to making the opposite decision and which were farther, we must employ some source of information to identify which cases are substantively proximate.

Even though our estimates of relative positions on individual cases will necessarily have significant uncertainty, they enable more credible identification strategies in applications to studying how bargaining on the Court depends on justices' relative preferences. Previously, such analyses have often relied on weakly-identified temporal variation in the preferences of justices (Ho and Quinn 2010). The preference variation we recover is more robustly identified, as it derives from contemporaneous variation in behavior of justices across areas of the law. Using our estimates, we revisit a central substantive question in the literature on the Court: how much influence do individual members of the Court exert over the majority opinion? In particular, we consider two decisions made in every case: to whom is majority opinion authorship assigned, and which justices join the resulting majority opinion? With respect to the first question, we find evidence that chief justices strategically assign authorship to associate justices in cases where their preferences are more proximate, within the set of cases where the chief and a given associate justice are both in the majority. With respect to the second question, we find evidence that justices who are in the majority are less likely to join the majority opinion when their preferences are further from the author and when they are further from the median of the majority coalition. Together, these results suggest that both the composition of the majority and the identity of the author have some influence on the content of the opinion, and that chief justices are strategically responsive to the latter fact

[3]There is additional information in the pattern of joining opinions, which we use separately in one of our applications.

in making assignment decisions.

2. OPINION-WRITING AND THE STATISTICAL STUDY OF JUDICIAL VOTING

When the Court decides a case, the Chief Justice, if he is in the majority, has the power to select the majority opinion-writer. (If the Chief is not in the majority, the the senior member of the majority has the power.) Next, the majority opinion writer drafts an opinion and has an incentive to attain the assent of at least four other justices (because only an opinion signed by a majority of justices has the weight of binding precedent). These two stages of Supreme Court decision-making have given rise to several theoretical questions. Can a Chief Justice strategically shape the outcome of cases with his power to assign majority opinions? Who controls the content of the opinion, given the complexities of unstructured bargaining among a group of justices? These two broad questions have motivated myriad quantitative studies of judicial behavior.

However, the example with which we began highlights a tension inherent in these studies. On one hand, social science is concerned with the study of systematic behavioral patterns and the institutional forces that shape those patterns. On the other hand, the study of law and courts often begins with the claim that each case is different and presents unique questions and issues to the judges who decide it, directly implicating those fundamental questions about judicial bargaining and behavior. For much of its history, the political science subfield of judicial politics has resolved this tension by assuming that while cases may be unique, what is systematic is judges' underlying preferences; cases may all be different, but judges respond systematically and consistently across all cases. Indeed, the "attitudinal model" (e.g., Segal and Spaeth 2002) holds as its primary assumption the idea that Supreme Court justices can be ordered consistently according to unidimensional ideologies, and that what cases do is to divide them into voting blocs. Supreme Court justices are ideologically-driven

actors, and the legal nuances of a given case do not affect how they align relative to one another with respect to the disposition of the case.

However, the past decades have witnessed an important development in the study of judicial voting. Scholars have increasingly asked how judges' preferences might vary across legal dimensions and questions (e.g., Kornhauser 1992; Lax 2007; Bailey and Maltzman 2011; Lauderdale and Clark 2012; Lauderdale and Clark Forthcoming). What it means for the "law to matter" in these studies is not that law is a constraint on judges' capacity to do what they want, but rather that the law defines the dimensions along which judges evaluate cases and affects how judges balance competing interests. In many respects, this development has brought the political science of law and courts together with the account of judicial choice that is embodied in qualitative, biography-based accounts of judicial careers, which often emphasize the extent to which judges' views of the world are nuanced and variable across issue domains. The implication of this perspective seems to challenge the theoretical justification for ideal point estimation that the systematic component of judicial preferences is *sufficiently* characterized with one or two dimensions of disagreement (Grofman and Brazill 2002; Martin and Quinn 2002). This assumption is at odds with new developments in the study of the Court.

But the statistical study of judicial voting has reflected a unidimensional attitudinal model of preferences as much for practical statistical reasons as substantive ones. The estimation of high-dimensional models becomes difficult, or even impossible, with small numbers of voters and simple binary votes (e.g., Londregan 2000; Lauderdale and Clark Forthcoming). Nevertheless, considerable insights have been gleaned from these traditional approaches to judicial preferences. Perhaps the most well-known approach has been to use item-response theory (IRT) models (Jackman 2001) to study judicial voting in terms of a single preference dimension. While there exist many possible theoretical justifications for this model, the existing research has rested the IRT model on a random utility model of

judicial voting. In this model, each justice is assumed to (1) derive some utility from voting for each of the parties (sides) in a case and (2) choose to vote for whichever of those two choices yields higher utility, subject to some random error. In a landmark paper, Martin and Quinn (2002) introduce a dynamic IRT model to estimate judicial preferences that can vary from term-to-term. Bailey (2007) proposes a model that includes additional bridging data across time to identify preference dynamics as well as bridging data across different voting bodies, to identify the relative preferences of Congress and the Supreme Court.

More recently, scholars have begun to focus on using more than just votes to estimate judicial preferences, enabling characterizations of preferences that vary by legal issue as well as time. Lauderdale and Clark (2012) use kernel-weighted optimal classification to estimate case-specific rank orderings of Supreme Court justices. Bailey and Maltzman (2011) use information about the subject matter of cases to evaluate how justices systematically vary across areas of the law; relatedly, Lauderdale and Clark (Forthcoming) combine text analysis with voting data to estimate a high-dimensional ideal point model.[4] In what follows, we combine the goal of estimating preferences that vary across areas of law, the logic of using auxiliary data contained in this intellectual trajectory of ideal point estimation, and a new kind of data about case similarity. This enables us to develop an estimator that recovers estimates of judicial preferences in which justices' preferences can vary flexibly, but nonetheless systematically, from case-to-case. The key to our innovation lies in pushing to its logical conclusion the idea that auxiliary information about the similarity among votes can be used to identify a statistical model in which preferences vary as a function of the legal issues raised by each case. As we show in Section 5, our method for preference estimation can then be used to revisit, with more powerful analytic leverage, the two fundamental questions

[4]Indeed, the use of auxiliary information to develop richer ideal point estimators is becoming increasingly common in a variety of institutional contexts (Clinton and Meirowitz 2003; Jessee 2009; Zucco, Jr and Lauderdale 2011).

driving studies of judicial bargaining and opinion-writing.

3. MODEL

Our estimation approach diverges from previous Bayesian ideal point estimators in several respects. We do not base our model on a random utility model that describes a choice between binary policy alternatives under spatial preferences (Poole 2005). For our purposes it is enough to assume that our voters (justices) have latent preferences for each side of each case relative to one another, and that these are correlated across cases. To this end, we do not aim to estimate a small (Jackman 2001), or even a large (Lauderdale and Clark Forthcoming) number of dimensions to summarize behavior. We aim to estimate latent preferences on every vote. That is, we want to know which justices were close to the cutpoint in a case, and in what order of preferences they were likely to have been arranged. To do this, instead of modeling the latent preferences on each vote in terms of a small number of latent dimensions, we estimate them conditional on each other, subject to an assumption that latent preferences are more similar on substantively similar cases than on substantively distant cases. From a more mechanical perspective, we aim to *smooth* the binary measures of justice votes into continuous measures of justice preferences, with the smoothing occurring across substantively similar cases.

To achieve this smoothing, our model for each justice's case specific preferences is a conditional autoregressive (CAR, Besag 1974) prior for the covariance of a justice's preferences across cases. Most often CAR models are used in geographic contexts, where physical adjacency determines the covariance structure. Here, extending previous research, we use citation data to construct a measure of legal adjacency between cases (Lauderdale and Clark 2012; Clark and Lauderdale 2012). Data on citation counts provide a similarity weight dictating the strength of correlation between the latent preferences for each pair of cases.

The latent preferences of justice $i \in \{1, 2, \ldots, n\}$ for case $j \in \{1, 2, \ldots, m\}$ have an expected value of the citation count weighted mean of her preferences in all other cases, with a total precision determined by the number of citations and an estimated parameter λ_ψ. The following formulation is based on the notation and Gibbs sampler described by Besag, Green, Higdon and Mengersen (1995).

Let Y be an $n \times m$ matrix of votes, where $y_{ij} = 1$ if justice i is in the majority in case j, $y_{ij} = 0$ if justice i is in the minority in case j, and is missing otherwise. Let ψ_{ij} be the latent preferences for justice i on case j, let α_j be the cutpoint for case j, and let $\beta_j \in \{-1, 1\}$ be the polarity of case j. The observed decision is

$$\begin{aligned} y_{ij} = 1 \quad &\Longleftrightarrow \quad \psi_{ij} > \alpha_j \cdot \beta_j \\ y_{ij} = -1 \quad &\Longleftrightarrow \quad \psi_{ij} < \alpha_j \cdot \beta_j \end{aligned}$$

That is, we assume there exists a threshold value for each case that divides the justices into voting coalitions depending on whether their each justice's latent utility is greater or lesser than that threshold. The β parameter captures the polarity of the case majority—if a case majority is voting in a "liberal" direction, then having latent preferences greater than the case location (α_j) implies that a justice did *not* vote with the majority; when the majority is voting in a "conservative" direction, the opposite relationship will exist. Hence, a conservative outcome will be associated with $\beta = 1$, and a liberal outcome will be associated with $\beta = -1$. While in principle one could consider using expert codings (such as from the Supreme Court Database (Spaeth, Epstein, Ruger, Whittington, Segal and Martin N.d.)) to fix β for each case, we follow the practice of previous ideal point estimation approaches of allowing the voting data to inform the case polarity and estimate β.

We assume the justices' latent case-specific utilities, ψ_{ij}, have a joint prior distribution that depends on the strength of the relationship between pairs of cases j and j' (including

cases that justice i did and did not vote on). While the strength of the relationship between case pairs could depend on many factors, in this paper we model it purely as a function of the citations between case j and j'. Let the $m \times m$ matrix $A = \{a_{jj'}\}$ measure the relationship between all pairs of cases. We assume $a_{jj'}$ is the fraction of citations in case j that are to case j', plus the fraction of citations in case j' that are to case j. Usually there will only be citations in one direction; however, cases that are decided concurrently may cite each other reciprocally. Thus, the A matrix is symmetric, with $a_{jj} \equiv 0$. There are many ways one could define this matrix in terms of citations and/or other data, and so exploration of alternative specifications here is an important area for further development.

Our model of the justice's latent preferences is given by the following:

$$\psi_{ij} \quad \propto \quad \lambda_{\psi}^{m/2} exp \left\{ -\frac{1}{2} \lambda_{\psi} \sum_{j} \sum_{j'} a_{jj'} \left(\psi_{ij} - \psi_{ij'} \right)^2 \right\} \tag{1}$$

This distribution implies that the justices' latent preferences in each case (ψ_{ij}) are, in expectation, a weighted average of their latent preferences across all other cases ($\psi_{ij'}$), weighted by the strength of the relationship to each other case ($a_{jj'}$).

This *intrinsic conditional autoregressive* prior is not a proper multivariate normal distribution, and requires us to impose a constraint to identify a particular posterior distribution. The prior only identifies the latent preferences relative to one another: the mean of the ψ_{ij} is therefore not identified (Besag et al. 1995). To see why, observe that each element of ψ_{ij} has an expected value that is a precision-weighted average of the other elements $\psi_{ij'}$ (where $j' \neq j$); however this leaves the overall mean $\sum_j \psi_{ij}$ unidentified. To identify the scale, we put independent standard normal priors over the case locations/cutpoints α_j:

$$\alpha_j \quad \sim \quad \mathcal{N}(0, 1) \tag{2}$$

By assuming the case cutpoints fall in in a range dictated by a standard normal prior, we

guarantee that the justices' case-specific latent preferences ψ_{ij} will lie in a similar range. We also standardize the α_j at each iteration of the Gibbs Sampler (see below), with proportionate rescaling of all other parameters.

The Court takes many cases that are decided unanimously: in such cases the cutpoint α_j is either greater than or less than all of the voting justices' latent preferences. We include these unanimous cases, as they are part of the citation network and help to strengthen the estimation of the correlation structure of preferences across areas of the law. However, we cannot, from the voting data alone, determine whether a case is unanimous because the cutpoint was to the left or to the right of all the justices, and so in such cases we will have a bimodal posterior distribution for the cutpoint. This posterior is unimodal when considered conditional on the polarity of the case β_j. This indeterminacy is important in the design of the Gibbs sampler described in the appendix, as the sampler needs to be able to jump between posterior density associated with left decisions and right decisions.[5]

To summarize, this model represents a minimalist approach to modeling the justices' ideal points. We assume that the justices' preferences in any given case are a weighted average of their preferences in related cases, with weights determined by how closely related those cases are. The observed votes result from whether the latent preferences of a justice are above or below a cutpoint that is specific to that case, and whether the case has a left majority or a right majority is estimated from the data at the same time.

The model generates several quantities of interest. First, the posterior distributions of the ψ_{ij} provide our beliefs about the relative positions of justices on a given case. These allow us to construct probabilistic estimates of which justices were likely to have been pivotal

[5]This issue is most serious in unanimous cases, but will also arise in non-unanimous votes where the justice alignment is unusual and therefore the polarity of the case is ambiguous. One could solve this problem if one used exogenous coding of the case polarities based on substantive criteria; however these would be likely to be least reliable in precisely the cases where voting coalitions were most unusual.

in a given case, and they can also be compared to the estimate of the cutpoint α_j for that case, in order to yield a visual representation of preferences in that case. Second, we can define the conditionally expected decision of each justice as $\mu_{ij} = \sum_{j'} a_{jj'} \psi_{ij'}$. This weighted average of preferences in related cases is our expectation for the justice's position on a given case, conditional on all his or her other decisions. Compared across justices, this is a useful summary of how the justices are aligned at a given "location" in the law, as opposed to on a given case. Given a set of citation weights and combined with λ_ψ, we could use this quantity to construct predictive distributions of judicial preferences for cases that have not yet been decided. Third, we can define a justice's mean preferences $\theta_i = \left(\sum_j \psi_{ij}\right)/m$, which capture the justice's general position and which are functionally very similar to unidimensional ideal point estimates of judicial preferences.

To estimate the model, we specify a Gibbs sampler for MCMC simulation of this model in an appendix to this paper. For this model, MCMC is expensive both in terms of computation and storage. Where there are M votes and N voters, each step of the Gibbs sampler involves doing computations comparable to fitting a regression by OLS on M observations, $M \times N$ times. We estimate the model using a function written in C++ and linked to R using Rcpp (Eddelbuettel and François 2011). Storing the resulting $M \times N$ latent preferences ψ_{ij} for each iteration of the sample consumes storage at a high rate. The analyses in this paper are based on a 5000 iteration sample (after 500 iteration burn in), thinned to every tenth iteration.

4. RESULTS

4.1. *Median Justices*

Lauderdale and Clark (2012) employ kernel-weighted optimal classification (KWOC) to show

that the justices' rank orderings vary from case-to-case, giving rise to substantial variation in who occupies the critical "median justice" position in any given case. Our approach allows us to estimate, similarly, who is the median justice in any given case, with the additional benefit of being able to incorporate uncertainty in our estimates. Specifically, we compute the probability of any justice being the median or pivotal justice in every case that the justice heard, given the data and our model. This is simply a matter of identifying the median justice at each iteration of the simulation in every case, and then taking a mean across iterations to compute the posterior probabilities of being pivotal for each justice for each case. In Figure 1, we average these estimates across terms, generating estimates of the fraction of cases in each term for which each justice was the pivotal justice.

A number of striking patterns emerge from this analysis. First, as Lauderdale and Clark (2012) show, the relative degree to which the role of median is concentrated on particular justices varies over time. For example, during the 1970s, the pivotal role was played disproportionately by Justices Stewart, White, Blackmun and Powell. By contrast, throughout the late-1990s and early-2000s, the pivotal role was shared disproportionately by two justices—O'Connor and Kennedy. Indeed, the central lesson we can draw from Figure 1 is that no single justice ever dominates all cases heard by the Court.[6] Put differently, knowing who is pivotal depends on knowing something about the substantive nature of what the Court is deciding. Compared to common public perceptions of a Court dominated by a small number of pivotal justices, these estimates reveal that the variation in the rate at which justices are pivotal is not nearly as unequal as some presume. In terms of the average rate over their careers, the least frequently pivotal justice in the data set is Douglas, the most frequently pivotal is White, but they only differ by about a factor of four in the fraction of cases where

[6]Our data sources do not enable us to consider the Court since O'Connor's retirement. While we expect that this period may exhibit the greatest concentration of the median role, it is still unlikely that justice Kennedy is median in more than a third of cases.

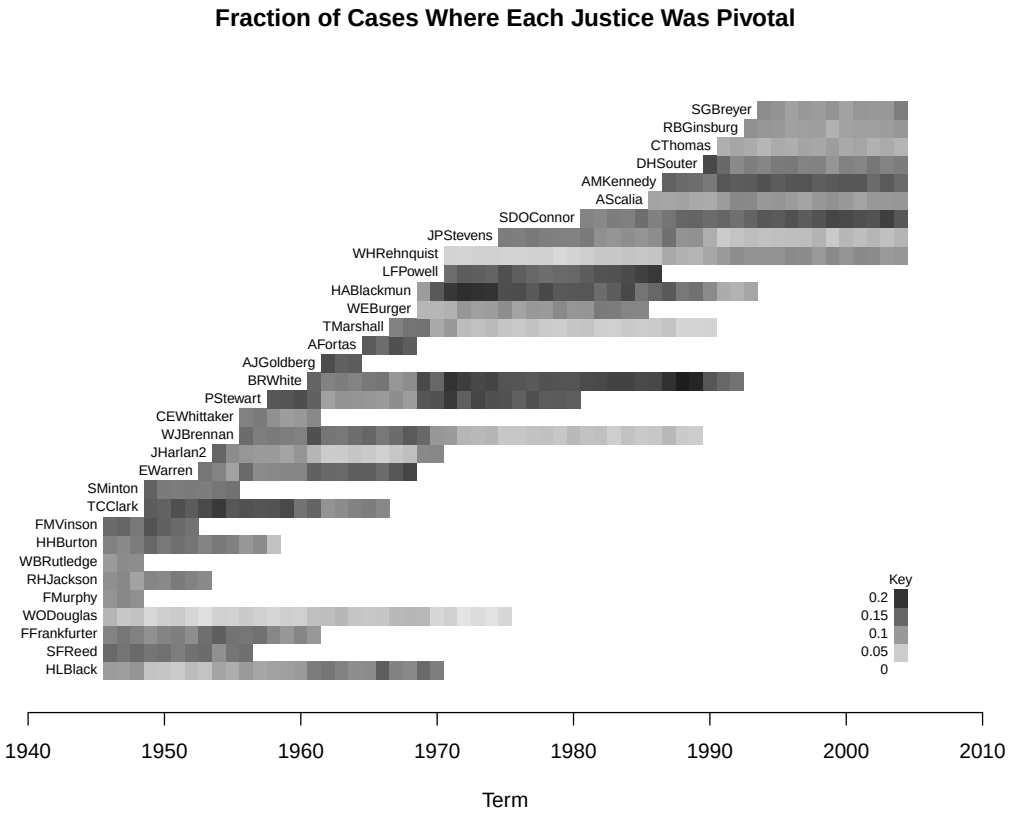

Figure 1: Estimated fraction of cases in which each justice was the pivotal justice, for each term.

we estimate them to have been pivotal. In some sense this should not be a surprise. There are always justice coalitions that cannot be rationalized by a constant unidimensional ordering of justices, and these occur in a substantial number of cases. These indicate that there must be case-to-case variation in the relative preferences of justices, and that the pivotal justices in such cases may be individuals who one thinks of as inhabiting the extremes of the Court more generally.

Figure 1 also demonstrates a second pattern. Probability of being pivotal is not something that often shifts sharply and dramatically. Rather, we often see a justice's pivotality waxing or waning over the course of his or her career. Justice White, for example, became increasingly pivotal during the late-1960s and early-1970s. Justice Blackmun, initially a frequent pivotal voter, became less so over the course of his career. Occasionally, though, we see sharp shifts involving many justices. In a short period of time during the late 1960s and early-1970s, Justices Black, Warren, Harlan and Fortas left the Court, while Burger, Blackmun, Powell and Rehnquist joined the Court. This had the consequence of making Justices Marshall and Brennan much less frequently pivotal, as justices who were often to their left were replaced by justices who were rarely to their left. However, the relatively smooth pattern of transitions, accompanied by an even smoother evolution in the issues and questions the Court addresses, leads to a picture of Supreme Court voting coalitions that are constantly evolving and does not change sharply in response to single justice replacements.

4.2. *Case-specific Preferences*

Figure 2 shows the relative positions of the justices voting on six high profile cases. These plots show preferences relative to the cutpoint, and so they slightly overstate the uncertainty of justices' positions relative to one another. The top two cases show *Roe v. Wade* and *Planned Parenthood v. Casey*. The alignments of justices within the voting coalitions in

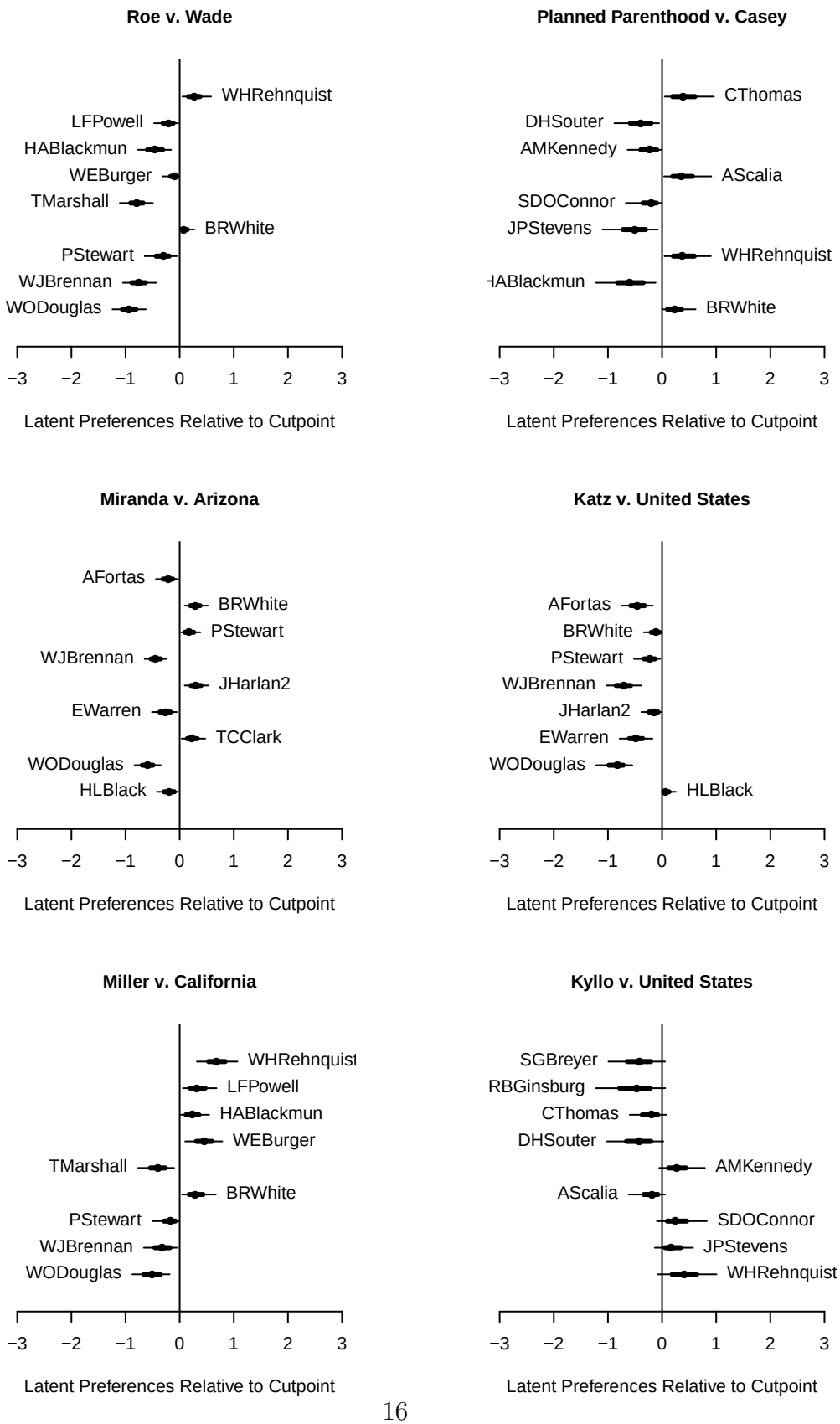

Figure 2: Estimated fraction of cases in which each justice was the pivotal justice, for each term.

these cases are roughly as we would expect them to be. In the 7-2 decision in *Roe*, the justice most tenuously in the majority is Burger, who rarely voted for the pro-choice side of subsequent abortion cases. While the majority opinion author Blackmun was fourth from the left, near the middle of the Court in *Roe*, our estimates in *Casey* place him farthest to the pro-choice end of any justice. Blackmun's concurrence/dissent in the complicated disposition of *Casey* was notably apocalyptic in reference to the four justices who voted to strike down *Roe*, reinforcing the idea that Blackmun was likely furthest of any justice on the left of Casey from voting the other way.

In *Miranda*, we see the expected result that the more ardent civil libertarians Brennan and Douglas were to the left of the remaining justices in the majority: Black, Fortas and Warren. Black, while typically on the left of the Court during his career, shifted right late in his career and particularly so in criminal justice cases. *Katz v. US* is another example of this: we are able to estimate that Black's dissent was from the right, not the left, because of Black's behavior in related cases.

Kyllo, the final panel, is an interesting case because it involves the unusual alliances of justices discussed at the beginning of this paper. It also is a good case for demonstrating the fact that our model is not always confident of the polarity of the decision. Unlike the other 5 cases plotted in Figure 2, in *Kyllo* the 95% intervals of some justices cover the estimated cut point. This happens because our model estimates the probability that the majority is on the left (i.e., $\beta_j = -1$) to be substantially less than 100%. Substantively, we would argue that the expansive definition of what constitutes a 4th Amendment search is a left decision, and so placing Thomas and Scalia to the left of Stevens, O'Connor, Kennedy and Rehnquist is in fact the correct alignment, as opposed to placing Ginsburg, Souter and Breyer to their right. However, given only the voting and citation data, there is some remaining ambiguity in this case, precisely because the alignment of the justices is unusual.

As we can see from these plots, our estimates of justices' relative locations are not precise

for any particular case, but do reflect meaningful variation in justices' relative preferences across different areas of the law. Our ability to pin down the relative locations of justices varies by case, as a function of the number of citations connecting that case to other cases. For example, *Roe* is a heavily cited case, with a large number of subsequent cases elaborating the Court's doctrine on abortion rights. As a result, we can infer a great deal about the relative positions of the various justices on this case. In contrast, for a case like *Kyllo*, there are fewer related cases with more heterogenous rulings, and so our uncertainty about the relative positions of justices is greater.

5. APPLICATIONS

Returning now to the two overarching questions motivating studies of judicial bargaining, we consider the Chief Justice's decision to assign a majority opinion and also the individual justices' decisions whether to join the majority opinion.

5.1. *Modeling Opinion Assignment*

A class of formal (as well as informal) models of opinion-writing predicts that who writes the opinion could matter greatly for the content of the opinion (Lax and Cameron 2007; Bonneau, Hammond, Maltzman and Wahlbeck 2007; Cameron and Clark Forthcoming; Maltzman, Spriggs and Wahlbeck 2000). We refer to these models of bargaining and opinion assignment as "author influence models" models and contrast them with "monopoly models," which predict that opinion authorship does not matter. The key distinction between these two classes of models is that the former predicts a set of strategic incentives for opinion assignment, whereas the latter predicts no strategic incentives concerning opinion assignment.

The empirical literature has examined opinion assignment, in part to help adjudicate

among these competing models of bargaining. Some of this evidence suggests that the power to assign the majority opinion is used strategically to influence case outcomes (Lax and Rader N.d.; Maltzman, Spriggs and Wahlbeck 2000; Cameron and Clark Forthcoming). One of the most powerful research designs in the literature leverages "vote fluidity" to evaluate strategic opinion assignment. Vote fluidity refers to the idea that justices who are "marginal", in the sense that they are ideologically proximate to the minority coalition, have the potential to "defect" and switch their votes from the majority coalition to the minority coalition. Past empirical studies have argued that justices who are closest to being on the fence between the two coalitions are assigned the majority opinion at a disproportionately high rate (Brenner 1982; Brenner and Spaeth 1988). Lax and Rader (N.d.) push this research design further by connecting it to four specific models of opinion-writing: two monopoly models and two author influence models. They find evidence consistent with the claim that the Chief Justice uses opinion assignment to strategically maintain coalitions and induce opinions that are aligned with his preferences.

However, it has been widely documented that there exists a norm by which opinion assignments are distributed evenly among the justices (e.g., Brenner and Palmer 1988; Maltzman and Wahlbeck 1996; Maltzman, Spriggs and Wahlbeck 2000). Assuming this norm is binding, if the justices' relative preferences do not vary from case-to-case, then it is not possible for the Chief Justice to assign cases disproportionately to his ideological allies. However, if relative preferences do vary case-to-case, then the Chief could use that variation to strategically assign cases to justices *when* they are most aligned with him, possibly relieving some of the constraint imposed by the norm of balanced workloads. Relying on measures of preferences that do not allow for case-to-case variation in preferences has prevented previous analyses from comparing two competing accounts of strategic opinion assignment. According to the vote fluidity logic, the Chief Justice assigns opinions to marginal justices in order to hold coalitions together. According to author influence models, the Chief Justice has an incentive

to assign opinions to his ideological allies.

Our estimation strategy allows us to revisit this debate with analytic leverage that was not previously available. Our estimates recover the arrangement of justices relative to each other in a way that varies from case-to-case. As a consequence, we can exploit potentially consequential variation in which justices are pivotal or in key institutional positions. Thus we specify an empirical model in which opinion assignment is estimated as a function of either a justice's ideological distance to the Chief Justice or ideological distance to the case-specific voting threshold (the point of indifference between the two coalitions). An alternative possible operationalization of the vote fluidity model is that distance to the median, rather than the coalition division, is what really matters. Thus, we also consider each justice's distance to the Court median. Finally, we consider each justice's distance to the coalition median, which, though not predicted by the median of the majority coalition model, potentially captures the influence of the center of the majority coalition. In order to test these varying predictions about authorship assignment, we construct case-specific measures of the median justice, the ideal point of the majority coalition median, the voting cut-point, and the Chief Justice. We then calculate the absolute distance for each justice to each of these points, for each case.

We specify a hierarchical conditional logit model, which we customize for these data. Let $Y_{ij} = 1$ if justice i is in the majority in case j. Where $A_{ij} \in \{0, 1\}$ indicates whether justice j was the author in case i, we fit a conditional logit model as a function of the position of that justice relative to the four positions. X_{1ij} is the estimated distance from justice i to the cutpoint in case j; X_{2ij} is the estimated distance to the median of the court; X_{3ij} is the estimated distance to the median of the majority coalition; X_{4ij} is the estimated distance to the chief justice. We limit our attention to the cases where the chief justice is in the majority and is therefore assigning the author of the opinion. Where $t \in \{1, 2, \ldots, 27\}$ is the natural court for case j, and $r \in \{1, 2, 3, 4\}$ is the chief justice for case j, our model has the following

form:

$$A^*_{ij} = \gamma^{court}_{it} + \sum_{k=1}^{4} \delta^{court}_{kt} X_{kij} \text{ if } Y_{ij} = 1 \quad (3)$$

$$A^*_{ij} = -\infty \text{ if } Y_{ij} \neq 1 \quad (4)$$

$$p(A_{ij} = 1) = \frac{exp(A^*_{ij})}{\sum_{i=1}^{n} exp(A^*_{ij})} \quad (5)$$

Notice that we assume in our specification that if a justice is not on the court or not in the majority, her probability of authoring is zero ($A^*_{ij} = -\infty$). Among those who are on the court and in the majority (including the chief justice), the probability of authoring is generated by a conditional logistic model where the latent utility of authorship for each justice is based on a justice and natural court specific intercept γ_{it}, plus the effects of the distance measures, which depend on a justice and natural court specific coefficient δ_{kt}. We generally expect the δ_{kt} to be zero or negative for most of the distance measures we consider, as the theories of authorship assignment we consider yield predictions where authorship probability declines as a justice's position gets further from the cut point, the court median, the majority median and/or the chief justice.

The slope and intercepts from this model are indexed by the natural court; however we specify the model as a hierarchical model, in which natural courts are nested within Chief Justice regimes. Such a specification allows us to make statements about the average relationships among all courts, among natural courts during a particular Chief Justice's tenure, etc. Thus, we include the following parameters and prior distribution assumptions in our model: $\gamma^{court}_{it} \sim \mathcal{N}\left(\gamma^{chief}_{ir}, \sigma^2_{\gamma^{chief}}\right)$, $\gamma^{chief}_{ir} \sim \mathcal{N}\left(\gamma_i, \sigma^2_{\gamma}\right)$, $\delta^{court}_{kt} \sim \mathcal{N}\left(\delta^{chief}_{kr}, \sigma^2_{\delta^{chief}}\right)$, $\delta^{chief}_{kr} \sim \mathcal{N}\left(\delta_k, \sigma^2_{\delta}\right)$, and $\sigma^2_{\gamma}, \sigma^2_{\delta}, \sigma^2_{\gamma^{chief}}, \sigma^2_{\delta^{chief}} \sim \mathcal{E}(1)$.

The hierarchical model for γ and δ reflects the fact that baseline authorship probabilities for a given justice will depend on the other justices on the court, and particularly on the Chief Justice. For example, if we set the $\beta_{kt} = 0$, our null model below, we would be assuming

that the relative positions of justices on a given case versus other cases do not matter, but the γ_{it} would still allow for the possibility that some justices are more likely to author than others, given that they are both in the majority. It is important to allow this to vary not only by natural court, but also by chief justice, so that this intercept can absorb the effect of average distance to a given chief justice.

For purposes of interpreting coefficient magnitude, it is useful to note the variation in each of the distance measures. The standard deviations of these measures, across all justices voting with the majority, are similar in magnitude: distance to cutpoint 0.34; distance to court median 0.20; distance to majority median 0.19; distance to chief 0.32. We program and simulate our model in `JAGS` (Plummer 2008) via `R` (R Development Core Team 2008), with reported results based on two parallel chains of 2000 iterations recorded after a burn-in of 500 iterations. We summarize the core findings in Figure 3, which shows the posterior estimates of the slope coefficients for each natural court (posterior means and central 95% credible intervals).

A number of findings stand out. There are weak and inconsistent associations between the probability of being assigned the majority opinion and a justice's distance to the (1) case-specific cut-point, (2) the case-specific Court median, and (3) the median of the majority coalition. All have substantial posterior density on both sides of zero in nearly every natural court. We estimate a much larger and more consistent relationship between authorship and distance from the Chief Justice under every Chief except Vinson. Thus, by far the strongest association, both in magnitude and statistical evidence, is that increasing distance to the Chief is associated with declining probability of receiving authorship assignment.[7]

[7]When we compare model fit to simpler models excluding subsets of these measures, using the DIC fit statistic (Spiegelhalter, Best, Carlin and der Linde 2002), we find that the presented model with all four distance measures fits about as well as a model with only the distance to the Chief Justice, and much better than any model using only the other variables.

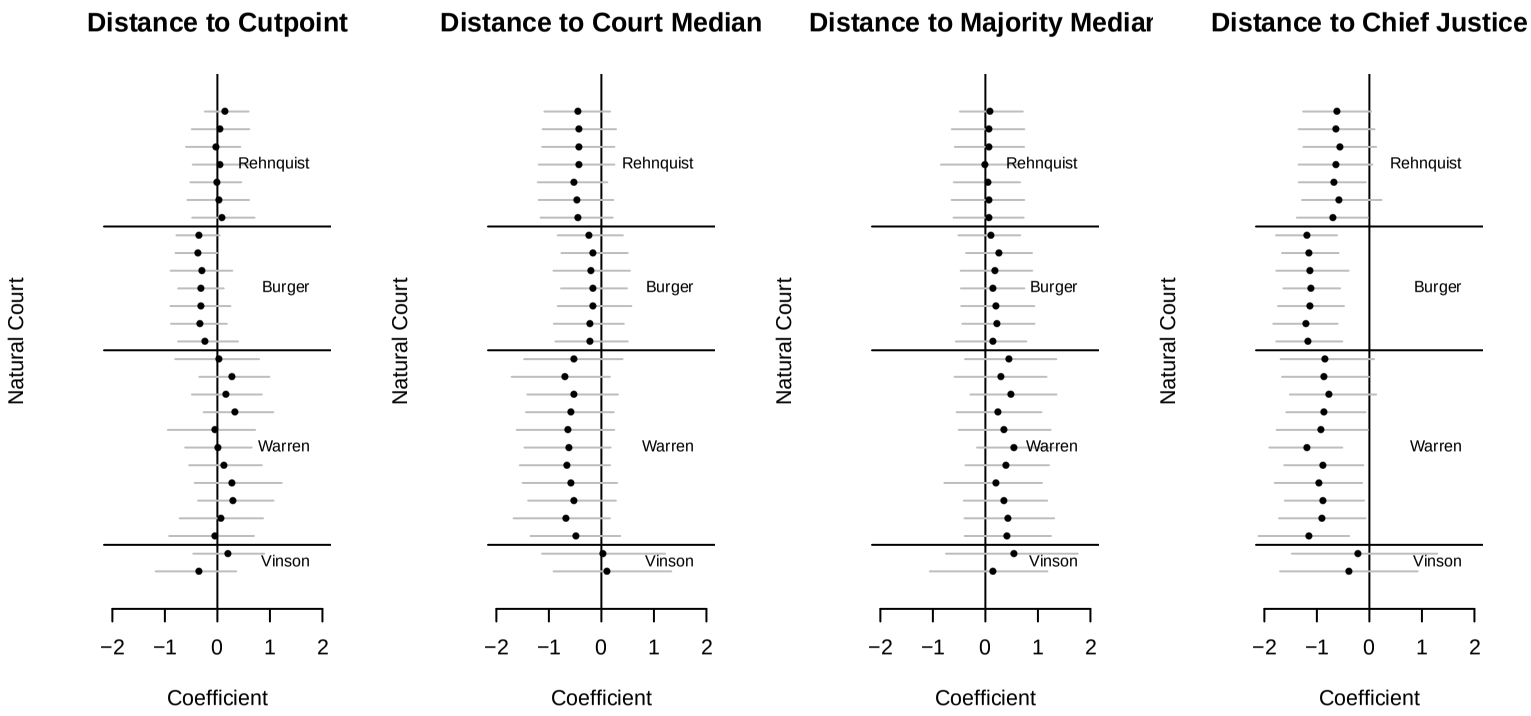

Figure 3: Coefficients of hierarchical conditional logit model for authorship, for each natural court in chronological order, with central 95% posterior intervals.

To help interpret these results, it is worth recalling the derivation of our distance measures. The model shows us that Chief Justices are more likely to assign authorship to an associate justice when the associate justice's voting in similar cases is closer to the Chief's voting in those similar cases. This evidence points to an insight that could only be recovered by contemplating case-specific preferences. Due to an empirical focus on static (Segal and Cover 1989) preferences, or preferences that can only vary term-by-term (Martin and Quinn 2002), scholars have missed a potential implication of the norm of proportionate assignment of opinions among the justices. Rather than simply being required to assign opinions to all justices and therefore strategically selecting "marginal" justices when there is a risk that a coalition will fall apart, it may instead be that the Chief Justice follows the norm of proportionate assignment by leveraging case-by-case variation to assign opinions to justices when they are most closely aligned. The Chief Justice is able to utilize *variation in which justices most strongly share his views* to help sustain a pattern of relatively equal assignment across cases. The fact that being close to the median of the Court or the majority has a much smaller effect, if any at all, indicates that justices are at most rarely able to leverage their relative positions on individual cases to gain the right to author opinions.

5.2. *Modeling the Decision to Join an Opinion*

Part of the motivation behind the studies seeking to understand the choice of opinion author is an interest in the influence individual members of the Court have over the content of the Supreme Court's opinions. Various theories and empirical tests have appeared in the literature (for a review, see Clark and Lauderdale 2010). The key issue at hand concerns how the Court's institutional arrangements shape the way in which the collective views of nine justices will be distilled into a single statement of law in the form of the Court's opinion. Among the many competing theories are claims that the median justice will control

the Court's opinion (this argument is a direct application of Black 1948). Bonneau et al. (2007) argue that if the Supreme Court were to operate as a closed-rule institution, then the logic of Romer and Rosenthal (1978) would apply, and the opinion author would have some degree of control over the opinion. Other arguments are more oriented around the Court's actual institutional rules and constraints and make arguments about author influence (e.g., Lax and Cameron 2007) or the difference in bargaining leverage between members of the majority coalition and members of the minority coalition (Carrubba et al. 2012; Cameron and Kornhauser N.d.). Related arguments contend that justices likely to "switch" sides and are therefore more marginal members of the majority coalition have special influence (e.g., Lax and Rader N.d.).

These (sometimes) competing theories yield a set of predictions on which our estimates provide new empirical leverage. Carrubba et al. (2012) argue that the pattern of join decisions—whether individual justices sign onto a majority opinion—reveals information about where an opinion is located in the underling latent ideology dimension. Justices who are more proximate to the opinion should be more likely to join the opinion. Therefore, we can evaluate whether how close a justice is to any theoretically-predicted point is predictive of her decision to join the majority opinion. From the existing literature, we identify five such points: (1) the cut point dividing justices into voting coalitions (more marginal justices have influence); (2) the Court median (median justice theory); (3) the Chief Justice (assignment power); (4) the majority coalition median; and (5) the opinion author.

We model the decision to join a majority opinion by each member i in the majority voting coalition of case j. Let $O_{ij} = 1$ if justice i joins the majority opinion in case j and $O_{ij} = 0$ if she does not join the majority opinion. We specify the decision as a function of each justice's distance to each of the theoretically-implicated points in the voting space. For this model, we also add in dummy variables Z_{lj} for the total number of judges l in the majority coalition on the decision in case j, to capture the fact that the incentives to join change with

the number of justices in the majority, not just their relative positions. Coefficients for each of these variables are estimated for each natural court, hierarchically modeled within Chief Justices, hierarchically modeled within the entire period.

$$O^*_{ij} = \gamma^{court}_{it} + \sum_{k=1}^{5} \delta^{court}_{kt} X_{kij} + \sum_{l=1}^{9} \nu_{lt} Z_{lj} \quad (6)$$

$$p(O_{ij} = 1) = \frac{\exp(O^*_{ij})}{1 + \exp(O^*_{ij})} \quad (7)$$

As in the previous empirical model, the slope and intercepts from this model are indexed by the natural courts, which we estimate as nested within Chief Justice regimes. Thus, we include the parameters and prior distribution assumptions in our model given above. We also model the dummy variables for size of the majority coalition in exactly the same way. Simulation details are also as in the previous application. The estimates for the coefficients on each of the five distance variables are reported by natural court in Figure 4.

The relationships between the distance measures and the patterns of joining opinions are more complex than we found under the authorship model, where there were weak relationships for all variables other than distance to the Chief Justice. Starting with the left column of Figure 4, and moving through the columns in order, we see that at least since the end of the Warren Court, there is little association between distance to the cutpoint and joining the majority opinion, conditional on the other distance measures. Before that, there was a small positive association, where justices further from the cut point were more likely to join an opinion, other distance measures equal. In the second column, we see a consistently positive relationship between distance to the median justice and joining the majority opinion. In the third column, and consistent with Carrubba et al. (2012), the further a justice is from the median of the majority coalition, the less likely she is to join the majority opinion. There is little evidence that distance to the Chief Justice matters where justices farther from the Chief were less likely to join (other distances equal). Finally, the most powerful relationship

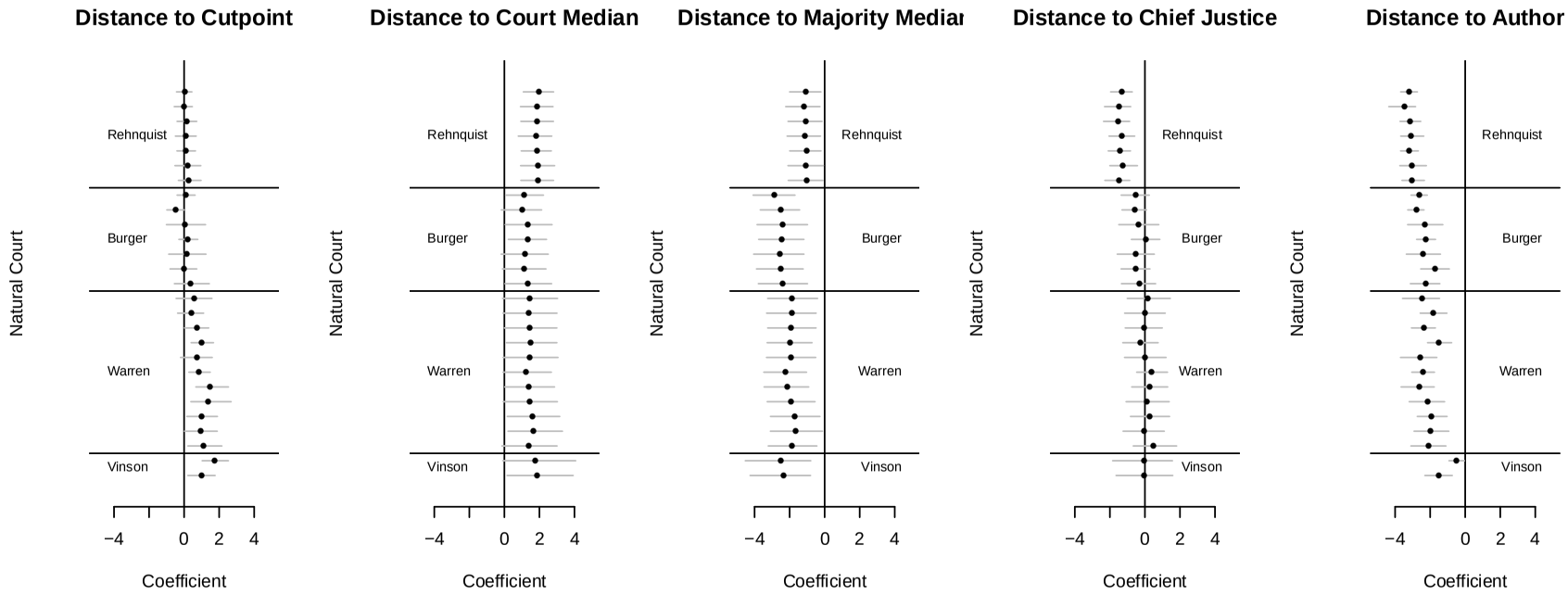

Figure 4: Coefficients of hierarchical logit model for joining the majority opinion, for each natural court in chronological order, with central 95% posterior intervals.

in the data, and one that has grown more powerful over time, is the negative association between joining the majority opinion and distance from the opinion author. The other notable time variation in the estimates is that under Rehnquist, distance to the majority median was somewhat supplanted by distance to the Chief Justice in shaping joining behavior, perhaps suggesting that Rehnquist was better able to control the content of majority opinions relative to previous Chiefs.

If we synthesize these findings, focusing on the most robust and consistent relationships, we find a suggestion that the opinion itself is influenced by the coalition median as well as the author of the opinion (see also Cameron and Kornhauser N.d.). At the same time, we find evidence that the further a justice is from the median justice, the more likely she is to join the majority opinion. In order to interpret this finding, it is important to recognize what it means to hold all the other measures constant. To imagine what it means to hold constant distance to the author and distance to the majority median, while increasing distance to the median, it is helpful to contemplate the most extreme justice on the Court, who is part of a majority and where the author is in the block of three justices on her side of the median. The further those three justices and the extreme justice get from the median, the more likely the extreme justice is to join the opinion. In some sense, these are increasingly precarious coalitions with the median, and given the location of the author within the increasingly extreme group of four justices, it makes sense that the extreme justice would be more inclined to support the opinion.

The major implication of these results is that distance to the opinion author is the most powerful predictor of the decision to join a majority opinion, and this has become more true over time. The simultaneous predictive power of the opinion author and the majority median is consistent with existing formal models of majority median influence (e.g., Carrubba et al. 2012; Cameron and Kornhauser N.d.). Connecting these results back to those in the previous application, and recognizing that the Chief can only select the author, not the

majority coalition, there is a strong suggestion that a strategic Chief Justice ought to use his assignment power to select opinion authors who share his own views in any given case, just as we saw in the previous analysis. Taken together, then, these two analyses suggest a strategic dynamic by which the Chief Justice knows that, given a particular majority coalition, the opinion author can influence the opinion content and therefore strategically selects different opinion authors as they vary form case-to-case in the extent to which they are ideologically aligned with the Chief.

6. DISCUSSION AND CONCLUSION

Scholars have long struggled with how law and politics interact in the course of judicial decision-making. We contend that the legal issues presented in cases are consequential for understanding systematic variation in the cleavages that characterize judicial politics. Rather than law acting as a constraining force on judicial policy-making, law can be interpreted as the context in which judges make their decisions. Further, empirically capturing the extent to which Supreme Court justices vary in their relative preferences across areas of the law has direct implications for studying decision-making. This is so most directly because decisions on the US Supreme Court are made collectively. When deciding the disposition of a case, the justices vote and a majority determines who wins and who loses. When writing an opinion that states the rationale for their decision, the justices similarly interact, bargaining among themselves over the opinion's content. Usually, a single opinion is joined by a majority of the justices, constituting the Opinion of the Court and binding precedent, and the authority to decide who writes the opinion of the Court is given to a single justice: the Chief Justice if he is in the majority. The outcomes of this group decision-making and opinion-writing process are a function of justices' relative alignment with each other, which varies across substantive questions presented to the Court. The potential for voting coalitions to vary systematically

across areas of the law opens up a host of theoretical and empirical puzzles (e.g., Lauderdale and Clark 2012).

Thus, to study the politics of bargaining in this context, we propose a method for estimating *case-specific* preferences. The underlying statistical logic of our approach is that of smoothing. Justices' votes are binary, and we aim to give a smooth, continuous summary of the preferences that generate those votes. To generate this, we have to take account of the varying polarity and cutpoints of individual cases. The model we present in this paper is not the only way to generate such measures, but it is a particularly simple way of doing so. Citations are an especially attractive basis for generating case-specific preferences because they closely track the Court's own sense of which cases are doctrinally relevant. Moreover, they are fairly predictable based on the legal briefs presented to the Court, so they open up the possibility of predicting the relative positions of justices in advance of a decision. Indeed, to the extent that the distribution of citations in opinions is potentially endogenous to bargaining over the opinion content, employing citations from other sources could provide additional predictive power. While we have not explored this possibility directly in this paper, the model we present could be straightforwardly applied using the citations in legal briefs rather than opinions, the only obstacle is collecting the data. Such an analysis could not predict the cutpoint in a pending case, but it could generate probabilities that each justice would be pivotal.

Armed with our estimates of preferences, an analysis of opinion-writing and opinion content reveals new evidence of the strategic incentives facing the justices. For example, while scholars have frequently observed the constraining effect of the norm of equal assignment of opinion-writing responsibilities across the justices and proposed norm-based justifications for this practice, our analysis suggests a norm of balanced workloads may be supported by the Chief Justice's strategic assignment of opinions to justices who are closest to him *in a given case* (cf. Lax and Cameron 2007). At the same time, we find evidence in support of a

recently-developed class of models in which the center of the majority exercises a gravitational force on the content of the opinion (as distinguished from the center of the Court as a whole), while the author still has substantial discretion as well (e.g., Lax and Cameron 2007; Carrubba et al. 2012; Cameron and Kornhauser N.d.). Taken together, these findings suggest a complex interaction among assignment power and bargaining among the justices that could not necessarily be documented in the absence of case-specific estimates of judicial preferences. While this is far from the final word in the study of Supreme Court bargaining, we anticipate the estimation strategy and empirical applications reported here will open the door to future empirical investigations of many theoretical predictions from the literature.

Finally, the estimates we develop here have potential applications beyond those motivating our study. By smoothing the justices' voting behavior across legally-similar cases, our model can provide insight for litigation strategy aimed at identifying the types of arguments most likely to succeed. Similarly, one could use our estimates to evaluate which cleavages on the Court are most likely to spur interbranch conflict. Do Courts with a common cleavage across substantive issues spur more backlash from Congress than Courts with myriad cleavages running across different issues? Does a justice's contemporary pivotality affect the contentiousness of confirmation hearings for her successor? We anticipate scholars will find these and other potential applications fruitful avenues for research.

REFERENCES

Bailey, Michael A. 2007. "Comparable Preference Estimates across Time and Institutions for the Court, Congress, and Presidency." *American Journal of Political Science* 51(3):433–448.

Bailey, Michael A. and Forrest Maltzman. 2011. *The Constrained Court: Law, Politics, and the Decisions Justices Make*. Princeton University Press.

Besag, Julian. 1974. "Spatial Interaction and the Statistical Analysis of Lattice Systems." *Journal of the Royal Statistical Society (Series B)* 36(2):192–236.

Besag, Julian, Peter Green, David Higdon and Kerrie Mengersen. 1995. "Bayesian Computation and Stochastic Systems." *Statistical Science* 10(1):3–41.

Black, Duncan. 1948. "On the Rationale of Group Decision-making." *Journal of Political Economy* 50(1):23–34.

Bonneau, Chris W., Thomas H. Hammond, Forrest Maltzman and Paul J. Wahlbeck. 2007. "Agenda Control, the Median Justice, and the Majority Opinion on the U.S. Supreme Court." *American Journal of Political Science* 51(4):890–905. Forthcoming.

Brenner, Saul. 1982. "Strategic Choice and Opinion Assignment on the U.S. Supreme Court: A Reexamination." *Western Political Quarterly* 35(2):204–211.

Brenner, Saul and Harold J. Spaeth. 1988. "Majority Opinion Assignments and the Maintenance of the Original Coalition on the Warren Court." *American Journal of Political Science* 32(1):72–81.

Brenner, Saul and Jan Palmer. 1988. "The Time Taken to Write Opinions as a Determinant of Opinion Assignments." *Judicature* 72:179–84.

Cameron, Charles M. and Lewis A. Kornhauser. N.d. "Bargaining on Appellate Courts." Princeton University working paper.

Cameron, Charles M. and Tom S. Clark. Forthcoming. The Chief Justice and Procedural Power. In *The Chief Justice: Appointment and Influence*. University of Michigan Press.

Carrubba, Clifford J., Barry Friedman, Andrew D. Martin and Georg Vanberg. 2012. "Who Controls the Content of Supreme Court Opinions?" *American Journal of Political Science* 56(2):400–412.

Clark, Tom S. and Benjamin Lauderdale. 2010. "Locating Supreme Court Opinions in Doctrine Space." *American Journal of Political Science* 54(4):871–890.

Clark, Tom S. and Benjamin Lauderdale. 2012. "The Genealogy of Law." *Political Analysis* 20(3):329–350.

Clinton, Joshua D. and Adam Meirowitz. 2003. "Integrating Roll Call Analysis and Voting Theory: A Framework." *Political Analysis* 11:381–396.

Eddelbuettel, Dirk and Romain François. 2011. "Rcpp: Seamless R and C++ Integration." *Journal of Statistical Software* 40(8):1–18.

Grofman, Bernard and Timothy J. Brazill. 2002. "Identifying the Median Justice on the Supreme Court through Multidimensional Scaling: Analysis of "Natural Courts" 1953-1991." *Public Choice* 112(1/2):55–79.

Ho, Daniel E. and Kevin M. Quinn. 2010. "How Not to Lie with Judicial Votes: Misconceptions, Measurement, and Models." *California Law Review* 98(3):813–876.

Jackman, Simon. 2001. "Multidimensional Analysis of Roll Call Data via Bayesian Simulation: Identification, Estimation, Inference and Model Checking." *Political Analysis* 9(3):227–241.

Jessee, Stephen. 2009. "Spatial Voting in the 2004 Presidential Election." *American Political Science Review* 103:59–81.

Kornhauser, Lewis A. 1992. "Modeling Collegial Courts II: Legal Doctrine." *Journal of Law Economics & Organization* 8:441–70.

Lauderdale, Benjamin E. and Tom S. Clark. 2012. "The Supreme Court's Many Median Justices." *American Political Science Review* 106(4):847–866.

Lauderdale, Benjamin E. and Tom S. Clark. Forthcoming. "Scaling Politically Meaningful Dimensions Using Texts and Votes." *American Journal of Political Science* p. Available online at http://userwww.service.emory.edu/ tclark7/ldairt.pdf.

Lax, Jeffrey R. 2007. "Constructing Legal Rules on Appellate Courts." *American Political Science Review.* 101(3):591–604.

Lax, Jeffrey R. and Charles M. Cameron. 2007. "Bargaining and Opinion Assignment on the U.S. Supreme Court." *Journal of Law, Economics, and Organization* 23(2):276–302.

Lax, Jeffrey R. and Kelly T. Rader. N.d. "Bargaining Power in the Supreme Court." Columbia University working paper.

Londregan, John. 2000. "Estimating Legislators' Preferred Points." *Political Analysis* 8(1):35–56.

Maltzman, Forrest, James F. Spriggs, II and Paul J. Wahlbeck. 2000. *Crafting Law on the Supreme Court: the Collegial Game.* Cambridge University Press.

Maltzman, Forrest and Paul J. Wahlbeck. 1996. "Strategic Policy Considerations and Voting Fluidity on the Burger Court." *American Political Science Review* 90-3 (Sept.):581–92.

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

Plummer, Martyn. 2008. *Just Another Gibbs Sampler.* http://www-ice.iarc.fr/ martyn/software/jags/: .

Poole, Keith T. 2005. *Spatial Models of Parliamentary Voting.* Cambridge University Press.

Poole, Keith T. and Howard Rosenthal. 1985. "A Spatial Model for Legislative Roll Call Analysis." *American Journal of Political Science* 29(2):357–384.

R Development Core Team. 2008. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Romer, Thomas and Howard Rosenthal. 1978. "Political Resource Allocation, Controlled Agendas, and the Status Quo." *Public Choice* 33:27–43.

Segal, Jeffrey A. and Albert D. Cover. 1989. "Ideological Values and the Votes of U.S. Supreme Court Justices." *American Political Science Review* 83(3):557–565.

Segal, Jeffrey A. and Harold J. Spaeth. 2002. *The Supreme Court and the Attitudinal Model Revisited*. Cambridge University Press.

Spaeth, Harold, Lee Epstein, Ted Ruger, Keith Whittington, Jeffrey Segal and Andrew D. Martin. N.d. "The Supreme Court Database.".

Spiegelhalter, D. J., N. G. Best, B. P. Carlin and A. Van der Linde. 2002. "Bayesian Measures of Model Complexity and Fit (with Discussion)." *Journal of the Royal Statistical Society (Series B)* 64(4):583–616.

Zucco, Jr, Cesar and Benjamin E. Lauderdale. 2011. "Distinguishing Between Influences on Brazilian Legislative Behavior." *Legislative Studies Quarterly* .

APPENDIX: GIBBS SAMPLER

We draw a posterior sample from this model by MCMC, using a Gibbs sampler with one Metropolis-Hastings step. This step is needed because of the deterministic relationship between latent votes, the cutpoint, and the case polarity; without a step that proposes to flip several quantities at once, the sampler could never explore both the possibility of a left majority and the possibility of a right majority for the same case. Let $a_{j+} = \sum_{j'} a_{jj'}$ be the sum of the bidirectional citation fractions linking case j to all other cases, and let W be an $m \times m$ matrix with diagonal elements $w_{jj} = a_{j+}$ and off-diagonal elements $w_{jj'} = -a_{jj'}$.

- For each case j,
  - Propose flip of case polarity $\beta_j^{\dagger} = -\beta_j^{(t-1)}$ and a reflection $\alpha_j^{\dagger}$ of the current cutpoint $\alpha_j^{(t-1)}$ across the mean of ψ_j, where $\beta_j^{(t-1)}$ and $\alpha_j^{(t-1)}$ are the values selected in the previous iteration of the sampler. Compute the Metropolis-Hastings acceptance ratio:

$$r = \prod_i \frac{\Phi\left(\sqrt{\tau_{ij}} \cdot (\alpha_j^{\dagger} - \mu_{ij}) \cdot (-1)^{\frac{y_{ij}\beta_j^{\dagger}+1}{2}}\right)}{\Phi\left(\sqrt{\tau_{ij}} \cdot (\alpha_j^{(t-1)} - \mu_{ij}) \cdot (-1)^{\frac{y_{ij}\beta_j^{(t-1)}+1}{2}}\right)} \tag{8}$$

$$\tau_{ij} = \lambda_{\psi} a_{j+} \tag{9}$$

$$\mu_{ij} = \frac{\sum_{j'} a_{jj'} \psi_{ij'}}{a_{j+}} \tag{10}$$

    With probability $min(r, 1)$, we accept the flipped case polarity $\beta_j^{(t)} = \beta_j^{\dagger} = -\beta_j^{(t-1)}$ and the reflected cutpoint $\alpha_j = \alpha_j^{\dagger}$; otherwise $\beta_j^{(t)} = \beta_j^{(t-1)}$ and $\alpha_j^{\dagger} = \alpha_j^{(t-1)}$.

  - For each justice i: draw ψ_{ij} from a singly truncated normal distribution consistent

with α_j and β_j:

$$\psi_{ij}^{*\dagger} \sim t\mathcal{N}\left(\mu_{ij}, 1/\tau_{ij}, \alpha_j^{\dagger}, \infty\right) \quad \text{if } y_{ij} \cdot \beta_j^{\dagger} = 1 \tag{11}$$

$$\psi_{ij}^{*\dagger} \sim t\mathcal{N}\left(\mu_{ij}, 1/\tau_{ij}, -\infty, \alpha_j^{\dagger}\right) \quad \text{if } y_{ij} \cdot \beta_j^{\dagger} = -1 \tag{12}$$

– Draw α_j from the singly or doubly truncated normal distribution of values consistent with ψ_{ij} and β_j. For case j, let $\psi_{ij}^{(max,-1)}$ be the maximum value of ψ_{ij} among dissenting justices i ($y_{ij} = -1$), let $\psi_{ij}^{(min,-1)}$ be the minimum value of ψ_{ij} among dissenting justices i, and let $\psi_{ij}^{(max(1))}$ and $\psi_{ij}^{(min,1)}$ be the corresponding quantities for justices in the majority. Let $u_j = 1$ if $y_{ij} = 1 \forall i$, and $u_j = 0$ otherwise.

$$\alpha_j \sim t\mathcal{N}\left(0, 1, \psi_{ij}^{(max,-1)}, \psi_{ij}^{(min,1)}\right) \quad \text{if } \beta_j = 1 \;\&\; u_j = 0 \tag{13}$$

$$\alpha_j \sim t\mathcal{N}\left(0, 1, \psi_{ij}^{(max,1)}, \psi_{ij}^{(min,-1)}\right) \quad \text{if } \beta_j = -1 \;\&\; u_j = 0 \tag{14}$$

$$\alpha_j \sim t\mathcal{N}\left(0, 1, -\infty, \psi_{ij}^{(min,1)}\right) \quad \text{if } \beta_j = 1 \;\&\; u_j = 1 \tag{15}$$

$$\alpha_j \sim t\mathcal{N}\left(0, 1, \psi_{ij}^{(max,1)}, \infty\right) \quad \text{if } \beta_j = -1 \;\&\; u_j = 1 \tag{16}$$

- Draw λ_ψ:

$$\lambda_\psi \sim \mathcal{G}\left(a_\psi + \frac{m \cdot n}{2}, b_\psi + \frac{1}{2}\sum_{i=1}^{n} \psi_i^{\top} W \psi_i\right) \tag{17}$$

Each iteration of the above steps provides a single draw from the posterior distribution of our model. We use an improper prior $a_\psi = b_\psi = 0$, as the scale identification is provided by the prior on and renormalization of α_j.