# Estimating Preferences of Circuit Judges: A Model of Consensus Voting

Joshua B. Fischman *University of Virginia*

**Abstract**

This paper develops a consensus voting model for estimating preferences of federal circuit court judges. Unlike standard ideal point models, which assume that judges vote sincerely for their preferred outcomes, the consensus model accounts for the norm of consensus in the courts of appeals by including a cost of dissent in the judicial utility function. A test of the consensus voting model on a data set of asylum appeals demonstrates that it provides a substantially better fit than a comparable sincere voting model and also generates more accurate predictions of voting probabilities. The model generates credible estimates of the impact of panel composition on case outcomes, which is surprisingly large in the asylum cases. Even though 95 percent of these decisions were unanimous, roughly half of the cases could have been decided differently if assigned to different panels.

## 1. Introduction

In recent years, methods for analyzing roll-call votes in legislatures have been applied to study judicial decisions in multimember courts. These methods, which estimate judges' ideal points from the voting alignments in their decisions, have generated measures of ideology for judges in courts such as the U.S. Supreme Court (Martin and Quinn 2002; Bailey 2007), the European Court of Human

This paper is adapted from a working paper previously circulated under the title "Decision-Making under a Norm of Consensus: A Structural Analysis of Three-Judge Panels." I would especially like to thank my dissertation advisers, Glenn Ellison, Stephen Ryan, and Jim Snyder, as well as Michael Abramowicz, Frank Cross, Rajeev Dehejia, Michael Gilbert, Dan Ho, Yannis Ioannides, Lewis Kornhauser, Jackie Chou Lem, Anup Malani, David Matsa, Greg Mitchell, Susan Rose-Ackerman, Matthew Stephenson, an anonymous referee, and seminar participants at the Massachusetts Institute of Technology, Tufts University, Harvard Law School, Yale University, the University of Toronto, the conference on Empirical Legal Studies, the American Law and Economics Association, the National Bureau of Economic Research, Northwestern University School of Law, the University of Southern California School of Law, the University of Virginia School of Law, and the University of Texas School of Law. I am grateful to David Law for sharing his data and for many helpful conversations.

[*Journal of Law and Economics,* vol. 54 (November 2011)]

Rights (Voeten 2007), and the Canadian Supreme Court (Alarie and Green 2007). These methods have proved far less useful, however, for analyzing decision making in federal circuit courts. Instead, empirical studies of circuit judges typically explain their behavior in terms of observable characteristics such as political party, gender, or judicial common-space scores derived from ideology estimates of judges' appointing presidents and home-state senators (Epstein et al. 2007; Giles, Hettinger, and Peppers 2001).

An important limitation of conventional ideal point models is that they assume that actors vote sincerely for their preferred outcomes. Numerous studies of circuit courts, however, have demonstrated that judges' votes are strongly influenced by their panel colleagues (for example, Revesz 1997; Cross and Tiller 1998; Sunstein et al. 2006; Peresie 2005; Boyd, Epstein, and Martin 2010). I have argued that these panel effects are due primarily to a norm of consensus on the courts of appeals (Fischman 2011). Others have attributed them to dissent aversion (Posner 2008, pp. 32–34), collegial deliberation (Edwards 2003), strategic motivations (Cross and Tiller 1998), and group polarization (Sunstein et al. 2006). Notwithstanding disagreement about the causes of panel effects, these studies provide compelling evidence that the assumption of sincere voting does not hold in circuit court panels.

A second limitation of most ideal point models is that they only make inferences from nonunanimous votes. In circuit courts, however, most cases are decided by consensus, with unanimity rates exceeding 95 percent in some circuits (Cohen 2002, p. 102). In addition, circuit judges typically decide cases in rotating three-judge panels, so that many pairs of judges are never observed disagreeing with each other or disagree in only a small number of cases. Once unanimous decisions are removed, disagreements among the judges may be too sparse to analyze using conventional ideal point models.

As an alternative to these models, this article develops a model of consensus voting that incorporates institutional features of circuit court panels. Like conventional ideal point models, the consensus voting model employs a random utility framework, but it differs in two key respects. First, it accounts for the influence of panel colleagues by incorporating a cost of dissent into the judicial utility function. Second, it exploits the random composition of panels and the random assignment of cases in the courts of appeals in order to isolate the impact of judicial preferences on case outcomes.

A test of the consensus voting model using a data set of asylum cases in the Ninth Circuit Court of Appeals shows that it has superior explanatory power compared to a sincere voting model. The consensus voting model also significantly outperforms a regression model that predicts voting behavior using a judge's political party. Most important, the consensus voting model can generate credible estimates of the probability of a liberal decision from any hypothetical panel and can also estimate the impact of panel composition on case dispositions. Models that assume sincere voting, by contrast, generate estimates of panel voting probabilities that exhibit severe bias toward the mean and therefore understate the true differences in judges' voting behavior.

The article proceeds as follows. Section 2 provides a brief overview of ideal point models and discusses why the standard formulation is not appropriate for analyzing circuit court decisions. It then constructs two models for analyzing panel votes: a sincere voting model that exploits unanimous opinions but maintains the assumption of sincere voting and a consensus voting model that incorporates the cost of dissent. Section 3 describes the Ninth Circuit asylum data analyzed in this paper. Section 4 presents the results, including a detailed comparison of the models. Section 5 discusses some applications of the consensus voting model and demonstrates how it can be used to estimate the impact of panel composition on case outcomes. Section 6 concludes. All technical details are provided in the Appendix.

## 2. Theoretical Framework

### 2.1. Standard Ideal Point Models

Ideal point models, such as the NOMINATE models of Poole and Rosenthal (1997) and those based on item response theory (for example, Clinton, Jackman, and Rivers 2004), simultaneously estimate an ideal point for each legislator as well as an item difficulty and discrimination parameter characterizing each bill. The bill-specific parameters are estimated from the voting alignments and can be estimated with reasonable precision when there are sufficiently many legislators and nonunanimous votes.

To use the standard terminology of the ideal point literature, let $x_i$, $x_j$, and $x_k$ be the ideal points of the judges who participate in case $t$, and let $y_{it} = 1$ if judge $i$ votes to reverse the lower court and $y_{it} = 0$ if judge $i$ votes to affirm. Assuming a quadratic utility function[1] in a Euclidean policy space, a one-dimensional spatial voting model has a reduced form equivalent to the standard item response model used for scoring standardized tests:

$$y_{it} = \begin{cases} 1 & \text{if } b_t(x_i - m_t) + \varepsilon_{it} \geq 0 \\ 0 & \text{if } b_t(x_i - m_t) + \varepsilon_{it} < 0, \end{cases} \tag{1}$$

where the location parameter $m_t$ is the cut point in the policy space representing the boundary between votes to affirm and votes to reverse, $b_t$ is the discrimination parameter, and $\varepsilon_{it}$ is an error term typically assumed to be normal, logistic, or uniform. The bill parameters and the voters' ideal points are then jointly estimated. Implicit in equation (1) is a critical assumption: that the error terms $\varepsilon_{it}$ are independent of each other and of the other judges' ideal points.

In the legislative setting, the parameters $m_t$ and $b_t$ are determined by the spatial

[1] The NOMINATE models of Poole and Rosenthal (1997) assume a Gaussian utility function and have a reduced form that is different from the item response theoretic models. Nevertheless, the argument in the following discussion applies as well to the NOMINATE models.

locations of the policy alternative and status quo, which are fixed before a roll call is taken. Although this interpretation is often given in ideal point models of courts, the meaning of these parameters is less clear, since a court is typically not restricted to a dichotomous choice between a status quo and a unique alternative. There may be a variety of remedies available to an appellate court, and numerous legal grounds for each, giving rise to multiple possible discrimination parameters.

One immediate consequence of the item response formulation in equation (1) is that $\Pr(y_{it} = 1)$ should be statistically independent of $x_j$ and $x_k$; the ideal points of judge *i*'s panel colleagues should have no impact on judge *i*'s vote. When item response models are used in the educational testing context, this assumption is reasonable: one student's answer should be independent of other students' ability parameters. In the legislative setting, however, the independence assumption may be inappropriate due to party whipping (Snyder and Groseclose 2000), proxy voting (Rosenthal and Voeten 2004), or strategic behavior[2] (Spirling and McLean 2007). The independence assumption is most clearly inapplicable to circuit court panels, as numerous studies have shown that judges are significantly influenced by their panel colleagues.

A second consequence of the standard ideal point formulation is that, in the absence of distributional assumptions about $m_t$ and $b_t$, unanimous opinions are not informative about judges' ideal points. Because the ideological directions of the voting outcomes in each case (the signs of the discrimination parameters $b_t$) are estimated from the voting alignment of the judges in that case, it is impossible to estimate the direction of the outcome when a decision is unanimous. Thus, in order to draw inferences about judges' preferences from unanimous votes, it is necessary to impose restrictions on the sign of $b_t$, and the data must be coded in a manner that determines which direction is the liberal outcome in each case. It is also necessary to make assumptions about the distribution of the cut points $m_t$; otherwise, it is impossible to determine the extent to which unanimous decisions are the result of easy cases or panels being ideologically homogeneous. Models that avoid making parametric assumptions about the distribution of the cut points, such as the NOMINATE models of Poole and Rosenthal (1997) or the conditional fixed-effects logit used in Fischman and Law (2009), automatically treat unanimous decisions as easy cases and hence uninformative about the preferences of the judges.

The exclusion of unanimous votes has little impact in many applications of ideal point estimation. In studies of Congress, each term has hundreds of nonunanimous roll calls, and subjectively coding the ideological direction of each bill would be controversial. In studies of the U.S. Supreme Court, such as Martin and Quinn (2002) and Bailey (2007), the models make weak distributional assumptions about the case parameters but similarly avoid making judgments about the ideological direction of case outcomes. This is feasible because there

[2] Poole and Rosenthal (1997, pp. 15–17) show that ideal point models are robust to certain types of strategic voting, such as logrolling and agenda setting.

are enough nonunanimous decisions by the Supreme Court to generate useful ideal point estimates without directional coding of case outcomes.

In the circuit courts, however, distributional assumptions about case parameters are both necessary and justified. When the data are limited to cases involving a single legal issue, the direction of each decision can usually be coded in a straightforward manner. Furthermore, the random assignment of cases to panels means that the case cut points may be modeled as following a distribution that is independent of the judges' ideal points. Finally, with only three votes per case, the cut points cannot be precisely estimated from the voting alignments. Because of the problem of incidental parameters (Neyman and Scott 1948), the model would also be unable to generate consistent estimates of the judges' ideal points (Londregan 1999). Peress (2009) shows that this is also true in Bayesian ideal point models.

Because circuit court decisions are predominantly unanimous, the cases with dissents may be too sparse to generate useful estimates of judicial preferences. In the asylum data analyzed in this paper, for example, most pairs of judges are never observed disagreeing with each other. Those that do disagree follow nearly perfect spatial voting, which makes it difficult to estimate judicial preferences using a probabilistic voting model (Rosenthal and Voeten 2004). Whether this pattern holds for other data sets of circuit court voting remains to be seen, but this example clearly raises issues about the applicability of the standard ideal point framework to circuit court voting.

In all subsequent discussions, I make the following assumptions in order to allow inferences to be made from unanimous decisions. First, I assume that judges' votes can be coded in a manner that distinguishes liberal from conservative outcomes. Second, I assume a constant discrimination parameter $b_t = 1$, which simplifies the exposition and estimation.[3] Finally, I assume that the cut points $m_t$ follow a normal distribution. I use the term "sincere voting model" to denote an item response model that incorporates these assumptions. The consensus voting model uses the same distributional assumptions but relaxes the assumption of sincere voting.

### *2.2. Sincere Voting Model*

In traditional ideal point models, the cut point and discrimination parameters are derived endogenously from the hypothesized locations of the status quo and alternative policies. In a judicial decision, however, there is not necessarily a fixed alternative or status quo. Instead, I parametrize judicial utility in terms of a judge's indifference point, so that a judge's preferred outcome depends on the

[3] In large roll calls, the discrimination parameter can provide useful information regarding how sharply a particular question distinguishes the voters along ideological lines. In agnostic voting models, such as that of Martin and Quinn (2002), the discrimination parameter also serves the purpose of estimating which voting outcome represents the liberal direction. With only three votes per case, ideologically coded case dispositions, and data comprising homogeneous cases, the discrimination parameter is less essential. Including it would be possible but would complicate the estimation, requiring the likelihood function to be estimated using two-dimensional numerical integration.

position of the case cut point relative to this indifference point. The intensity of a judge's preference is determined by the distance between these two points.

Let $x_i$, $x_j$, and $x_k$ be the indifference points for the judges who participate in case $t$, and let the case be represented by a cut point $m_t$. The space is oriented so that a larger $x_i$ represents a more liberal judge, while a larger $m_t$ corresponds to a weaker claim. A judge's preferred disposition is determined by the position of the cut point relative to the judge's indifference point: a judge will prefer to grant relief to the claimant if $x_i + \varepsilon_{it} > m_t$, where $\varepsilon_{it} \sim N(0, 1)$ is an error term representing an idiosyncratic shock to judge $i$'s preferences in case $t$. For simplicity, let $x_{it} = x_i + \varepsilon_{it}$, which may be interpreted as judge $i$'s location in case $t$.

Denote the voting outcome as $y_{it} = 1$ if judge $i$ votes in the liberal direction and $y_{it} = 0$ if judge $i$ votes in the conservative direction. The utility of judge $i$ in case $t$ is defined as follows:

$$U_i(x_i, m_t, \varepsilon_{it}) = \begin{cases} \min\{0, x_{it} - m_t\} & \text{if } y_i = 1 \\ \min\{0, m_t - x_{it}\} & \text{if } y_i = 0. \end{cases}$$

Thus, a judge has a utility of zero when she votes in favor of her preferred outcome. When the judge votes against her preferred outcome, she incurs a disutility that is increasing in the distance between her indifference point and the cut point.

In the sincere voting model, for a given cut point $m_t$,

$$\Pr(y_{it} = 1 | m_t) = \Phi(x_i - m_t).$$

Exploiting the random assignment of cases, let $m_t = \mathbf{z}_t\boldsymbol{\gamma} + \eta_t$, where $\mathbf{z}_t$ is a (possibly null) vector of characteristics of case $t$, $\boldsymbol{\gamma}$ is a vector of parameters to be estimated, and $\eta_t \sim N(0, \sigma^2)$ is a random effect that is independent of $\mathbf{z}_t$ and the $x_i$ terms. The parameter $\sigma$, which is estimated in the model, is the standard deviation of the unobserved component of the cut points. When $\sigma$ is large relative to the $x_i$ terms, this typically means that there are more easy cases, since there are more cut points outside the range of most of the judges' indifference points.

The probability of a liberal vote by judge $i$ in case $t$ may be expressed in terms of the observable variables by integrating over the random effect:

$$\Pr(y_{it} = 1 | \mathbf{z}_t) = \int \Phi(x_i - \mathbf{z}_t\boldsymbol{\gamma} - \sigma u)\phi(u)du. \quad (2)$$

As in traditional ideal point models, this probability is independent of $x_j$ and $x_k$. The sincere voting model does not account for the possibility that judge $i$'s vote may be influenced by judges $j$ and $k$, contrary to much of the empirical literature on panel effects in circuit courts. However, the judges' votes may still be correlated because of the unobserved random effect.

### 2.3. Consensus Voting Model

In the consensus voting model, dissenting is costly, so judges' votes are no longer determined solely by their preferences for case outcomes. The utility function of judge $i$ now takes the form

$$U_i(x_i,\ m_t,\ \varepsilon_{it}) = \begin{cases} \min\{0,\ x_{it} - m_t\} - cI_d & \text{if } y_i = 1 \\ \min\{0,\ m_t - x_{it}\} - cI_d & \text{if } y_i = 0, \end{cases}$$

where $c$ represents the cost of dissent and $I_d$ is a dummy variable indicating whether judge $i$ dissented. Thus, a judge incurs a disutility $c$ whenever her vote differs from the votes of the other judges on the panel. Although the indifference points $x_i$ vary by judge, the model assumes a uniform cost of dissent for all judges. All of the judges are assumed to have perfect information about each other's payoffs.

The cost of dissent may reflect a variety of factors. Epstein, Landes, and Posner (2011) characterize this cost in terms of effort and collegiality. Ginsburg (1990) emphasizes that too many separate opinions may harm the institutional legitimacy of the court and weaken the signaling value of dissent. In addition, the cost of dissent may include the lost opportunity to negotiate with the majority for a more limited holding.

The equilibrium concept employed here is the strong Nash equilibrium[4] (Aumann 1959), which requires that the voting profile be immune from deviations by any coalition of judges. Note that the standard Nash equilibrium concept could result in multiple equilibria if the cost of dissent were sufficiently large. For example, if a panel of judges agreed on the disposition of a case, it would be a Nash equilibrium for them to vote unanimously for their preferred outcome or unanimously against it. In either case, no judge would deviate unilaterally from the equilibrium. The latter Nash equilibrium, which is implausible, would not be a strong Nash equilibrium, since a coalition consisting of the entire panel would deviate.

For any combination of $x_{it}$, $x_{jt}$, $x_{kt}$, and $m_t$, there will be a unique strong Nash equilibrium. Note that the maximum utility is zero, which is achieved whenever a judge votes for his preferred outcome and is in the panel majority. Since there will always be at least two judges who agree on the disposition, they can vote sincerely and achieve the maximum utility, which will not be affected by the third judge's vote. If the third judge also agrees on the disposition, the vote will be unanimous, and clearly no coalition could benefit by deviating. If the third judge prefers the opposite outcome, she will dissent if her distance $|x_{it} - m_t|$ from the cut point is greater than $c$. If her distance from the cut point is less than $c$, then she will join the majority.

The above discussion is formalized in the following proposition.

**Proposition 1.** Consider a panel in case $t$ consisting of judges $i$, $j$, $k$, where $x_{it} < x_{jt} < x_{kt}$, and let the case cut point be $m_t$. Then there will be a unique[5]

[4] The coalition-proof Nash equilibrium concept of Bernheim, Peleg, and Whinston (1987) would suffice here as well.

[5] There will be multiple equilibria in the trivial case in which equality is strict, which occurs with zero probability.

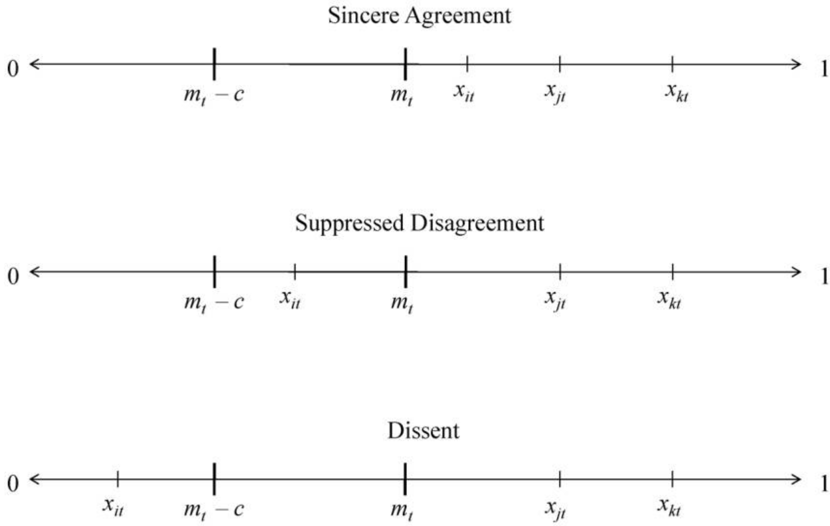


**Figure 1.** Spatial representation of consensus voting model

strong Nash equilibrium, which takes the following form:

$$(y_{it}, y_{jt}, y_{kt}) = \begin{cases} (1,\ 1,\ 1) & \text{if } x_{it} \geq m_t - c \quad \text{and} \quad x_{jt} \geq m_t \\ (0,\ 1,\ 1) & \text{if } x_{it} < m_t - c \quad \text{and} \quad x_{jt} \geq m_t \\ (0,\ 0,\ 1) & \text{if } x_{jt} < m_t \quad \text{and} \quad x_{kt} \geq m_t + c \\ (0,\ 0,\ 0) & \text{if } x_{jt} < m_t \quad \text{and} \quad x_{kt} < m_t + c. \end{cases}$$

Figure 1 provides a spatial representation of the consensus voting model. In the examples depicted in Figure 1, judges $j$ and $k$ always prefer the liberal outcome, since they are to the right of the cut point, while the position of judge $i$ varies. In the top example, judge $i$ is also to the right of the cut point, so there is sincere agreement, and the case will be decided unanimously in the liberal direction.

The middle example illustrates suppressed disagreement. Here, judge $i$ is to the left of the cut point and would prefer the conservative outcome. However, judge $i$'s disutility from dissenting outweighs the disutility she would incur from voting against her preferred outcome, as measured by her distance $|x_{it} - m_t|$ from the cut point. Hence, judge $i$ will join the majority to form a unanimous opinion.

In the bottom example, judge $i$ prefers the conservative outcome and is sufficiently far from the cut point that her disutility from joining the majority would outweigh the cost of dissent. In this example, judge $i$ will issue a dissenting opinion.

In the top and bottom examples, both the sincere voting model and the consensus voting model predict the same outcome. It is only in the middle example that the predictions of the two models differ: the sincere voting model predicts a dissenting opinion, while the consensus voting model predicts a unanimous decision. How often this situation arises depends on the magnitude of the cost of dissent relative to the other parameters.

Since the predictions of the model are probabilistic, it is impossible to know whether a particular unanimous vote reflects sincere agreement. Nevertheless, by recognizing the possibility of suppressed disagreement, the consensus voting model accounts for the influence of panel colleagues in interpreting judges' votes. To illustrate, consider two examples in the data involving Judge Thomas Nelson, who has a moderate voting record in asylum cases. In a series of 11 cases, Judge Nelson was empaneled with Judge Stephen Reinhardt, who has an extremely liberal voting record in asylum cases, and Judge Phyllis Kravitch, a Carter appointee.[6] The panel decided all of the cases unanimously, granting relief to asylum petitioners in 10 out of 11 of the cases, far higher than Judge Nelson's typical rate. In another series of cases, Judge Nelson was empaneled with Judges Diarmuid O'Scannlain and Ferdinand Fernandez, both of whom have extremely conservative voting records in asylum cases. The latter panel unanimously denied relief in all of the 12 cases it heard. Aside from the panel compositions, there do not appear to be any other common characteristics among these cases that would explain the disparate outcomes. Although these examples are anecdotal, they suggest that Judge Nelson's votes were strongly influenced by his panel colleagues in those cases. Treating all of his votes as sincere would result in a distorted estimate of his indifference point.

In another example, Judge William Fletcher was empaneled with Judges O'Scannlain and Reinhardt, who occupied opposite ends of the ideological spectrum. The panel granted relief in four out of 14 cases, with two conservative dissents by Judge O'Scannlain and seven liberal dissents by Judge Reinhardt. In these cases, Judge Fletcher's votes are likely to be accurate indicators of his preferences, because he was presumably the pivotal judge.

Under the assumptions of the consensus voting model, the probability that judge $i$ will cast a liberal vote in case $t$ is

$$\begin{aligned}\Pr(y_{it} = 1|m_t) = {} & \Phi(x_i - m_t) \\ & + \Phi(x_j - m_t)\Phi(x_k - m_t)[\Phi(x_i - m_t + c) - \Phi(x_i - m_t)] \\ & - \Phi(m_t - x_j)\Phi(m_t - x_k)[\Phi(x_i - m_t) - \Phi(x_i - m_t - c)].\end{aligned} \tag{3}$$

The first term is the voting probability under the sincere voting model, and the second and third terms adjust for the possibility of suppressed disagreement. Note that this probability is strictly increasing with $x_j$ and $x_k$, the indifference points of the panel colleagues.

[6] Judge Kravitch was sitting by designation from the Eleventh Circuit and did not hear any other asylum cases in the Ninth Circuit during the period of study.

As in the sincere voting model, the cut point can be expressed as $m_t = z_t\gamma + \eta_t$, and the unconditional probability may be calculated by integrating over the random effect, as in equation (2).

The following proposition summarizes some key properties of the consensus voting model.

**Proposition 2.** The consensus voting model satisfies the following properties:

*a*) If there is no cost of dissent ($c = 0$), then the consensus voting model yields the same predictions as the sincere voting model.

*b*) If the cost of dissent is positive ($c > 0$), then the probability that a judge votes in the liberal direction is strictly increasing in the indifference points of the judge's panel colleagues.

*c*) The greater the cost of dissent, the greater the impact of the panel colleagues' indifference points on the judge's vote.

*d*) The expected proportion of unanimous opinions is strictly increasing in the cost of dissent, keeping all other parameters fixed.

*e*) For a given $x_i$, $x_j$, $x_k$, and $m_t$, the sincere voting model and the consensus voting model will both predict the same probability of a liberal (or conservative) case disposition.

See the Appendix for the proof. A consequence of part *a* is that the sincere voting model can be viewed as being nested within the consensus voting model and may thus be tested by restricting $c = 0$. Parts *b* and *c* explain the impact of the cost of dissent on judges' votes: it increases the degree to which a judge's vote is influenced by the indifference points of her panel colleagues. Thus, as stated in part *d*, a higher cost of dissent corresponds to a higher rate of unanimity. Although the proportion of unanimous opinions is not strictly increasing in $\sigma$, a very large $\sigma$ would mean that cut points fluctuate substantially, suggesting more easy cases. Thus, a high unanimity rate may stem from a large $c$, a large $\sigma$, or both: a large $c$ would represent a strong norm of consensus, while a large $\sigma$ would represent a preponderance of easy cases. Part *e* follows from the fact that the disposition of a case is determined by the median judge in both models.

### 2.4. Identification

Spatial voting models require identifying restrictions in order to be invariant to translation, rescaling, and reflection. Both the sincere voting model and the consensus voting model are invariant to translation because of the assumption that $E[\eta_t] = 0$. If $\mathbf{z}$ is a null vector, this is equivalent to restricting the cut points to have mean zero; otherwise, the vector $\mathbf{z}$ must have full rank and may not include a constant term. The scale of the space is normalized by the variance of the idiosyncratic error term $\varepsilon_{it}$. Because there is no discrimination parameter in this model, there is no need to normalize the judges' indifference points. Finally, there is no reflection invariance because the voting outcomes are directionally coded; a higher $x_i$ always means a stronger propensity to vote in the liberal direction. The scale and reflection invariance can also be seen by noting that the likelihood function (provided in the Appendix) is not invariant to a scale transformation.[7]

The cost of dissent is identified by the degree to which judges' votes are influenced by their colleagues. The prior example involving Judge Nelson suggests a high cost of dissent. When empaneled with liberal colleagues, he consistently joined them in providing relief to asylum petitioners, but he consistently voted to deny relief when empaneled with conservative colleagues.[8] Formally, the identification of the cost of dissent comes from equation (3) and proposition 2*c*. If the cost of dissent is zero, then a judge's vote in a particular case should be uncorrelated with the voting record of his panel colleagues in other cases. This correlation increases strictly with the cost of dissent.

Another requirement for identification of the cost of dissent is that there be variation, yet some overlap, in the composition of panels. If there were only three judges who decided all cases together, the cost of dissent could not be identified. The same would be true if there were multiple panels with fixed composition. Finally, the random assignment of cases ensures that differences in voting behavior among judges are due to differences in judicial preferences, and not to case selection. However, it is impossible to identify the impact of any case characteristics (components of $z$) that are perfectly collinear with the presence of a particular judge on a panel.

To understand the identification of $\sigma$, note that a larger $\sigma$ increases the correlation of the three judges' votes in a particular case and, hence, the overall unanimity rate, as in proposition 2*e*. Thus, $\sigma$ is identified by the excess unanimity above the level that would occur in either model without the random effect. In practice, the identification of $\sigma$ can be extremely weak in the sincere voting model. This occurs because the scale of the space is normalized by the magnitude of the idiosyncratic errors. When the judges follow nearly perfect spatial voting, the scale will be weakly identified. To avoid computational difficulties, I impose the constraint $\sigma \leq 5$, which is sufficient to enable estimation of the model. The fitted voting probabilities are not sensitive to the choice of constraint.

The identification of the model was also verified using Monte Carlo simulations. Simulated data sets were generated using a variety of parameter values, and the model was then used to estimate the parameters from the simulated data. In all cases, the model was able to derive maximum likelihood estimates that were close to the parameters used to generate the data.

[7] The identification of the sincere voting model can also be understood by recognizing that it is functionally equivalent to a random-effects probit model.

[8] The model assumes that judges are influenced by the decisions of their panel colleagues, rather than their characteristics; in Manski's (1993) terminology, the panel interactions are "endogenous" rather than "contextual" effects. Although this assumption is consistent with most theories of panel effects, much of the current literature interprets these effects in terms of colleagues' characteristics, such as party, gender, or race. Fischman (2011) shows that the influence of panel colleagues appears to be highly stable across a wide range of studies when interpreted as an endogenous effect.

### 2.5. Estimation

The maximum likelihood model is estimated using Newton's method, with neutral starting values. The model estimates indifference points for individual judges with at least 10 votes in the data. Judges with fewer votes are grouped by party of appointment, with an additional error term to account for within-group heterogeneity. Judges with perfect liberal or perfect conservative voting rates are constrained to have $x_i = \pm\infty$, as appropriate. These judges are not reported in the results.

The vector $\mathbf{z}$ of case characteristics includes only dummy variables representing the year in which the case was decided, with 1992 as the omitted year. Cases are randomly assigned to panels, but only within the same time period, so time dummies are necessary to control for possible changes in the composition of cases over time. Dummies for major countries of origin were included in alternative specifications, but these did not substantively affect the results and are not reported. Because panels are sometimes assigned multiple cases with common characteristics, all standard errors are clustered to account for the potential similarity of cases decided by the same panel.

## 3. Data

The data examined in this paper are taken from Law (2005) and consist of 1,892 asylum cases decided by the Ninth Circuit Court of Appeals between 1992 and 2001. A foreign national seeking asylum in the United States may petition the Department of Justice and, if denied, may file a claim before an administrative immigration judge. Asylum claims may also be raised in deportation hearings. The legal standard for asylum cases is that the petitioner must demonstrate a "well-founded fear of persecution on account of race, religion, nationality, membership in a particular social group, or political opinion" (8 U.S.C. sec. 1101[a][42][A] [2006]). Denials of asylum may be appealed to an administrative appellate court, the Board of Immigration Appeals (BIA), and then to the federal circuit court corresponding to the region in which the claimant resides. All circuit court appeals are initiated by asylum petitioners; the government does not appeal asylum cases to federal court, since the attorney general may unilaterally overturn adverse BIA decisions (Legomsky 2010). Many of the cases involve review of factual determinations made by the immigration judge, and 92 percent of the decisions are unpublished.

An important feature of the data, consistent with a central assumption of the model, is that cases are randomly assigned to panels. In the Ninth Circuit, appeals may be grouped together on the court calendar to be heard by the same panel. The stated goals of the assignment process are to "balance judges' workloads" and to allow a single panel to hear "unrelated appeals involving similar legal issues" (9th Cir. R., intro., E[1]). Judges are randomly assigned to panels using a computer algorithm that requires that each panel consist of at least two active judges and "aim[s] . . . to enable each active judge to sit with every other active and senior judge approximately the same number of times over a two-year period" (sec. E[5]). When panels are composed, "the clerk does not know which cases ultimately will be allocated to each of the panels" (sec. E[2]). The only exception to random assignment is that "a case heard by the court on a prior appeal may be set before the same panel upon a later appeal" (sec. E[4]).

As a consequence of these rules, the characteristics of the cases assigned to a particular panel will be independent of the indifference points of the judges on that panel, which will also be independent of each other. However, when a panel is assigned multiple asylum appeals, or hears multiple appeals arising out of the same case, these cases may have characteristics in common.[9]

Although the circuit rules ensure randomization, some deviation from random assignment may occur if cases are settled on the basis of the judges assigned to them. In fact, studies have shown in other contexts that judges may influence both case outcomes and settlement decisions (for example, Waldfogel 1995; de Figueiredo 2005). However, settlement rates in asylum cases are quite low in the Ninth Circuit,[10] and the panel of judges is announced "on the Monday of the week preceding argument" (9th Cir. R., intro., E[3]), after most of the litigation costs have been sunk. Finally, a simulation-based statistical test found that the data were consistent with random assignment. Details are provided in the Appendix.

Each judge's vote is coded on the basis of whether it provides some sort of relief to the asylum petitioner. If the judge votes to provide any relief—a remand, a grant of asylum, or an order withholding deportation—that vote is coded as liberal; otherwise the vote is labeled conservative. The overall liberal voting rate in the data is 18 percent. Democratic appointees support asylum claimants 25 percent of the time, compared to 12 percent for Republican appointees. Ninety-five percent of the decisions were unanimous.

As in many prior studies of voting in appellate panels, there are strong panel effects in asylum voting. Judges appointed by Democratic presidents vote to provide relief to asylum claimants 35 percent of the time when on a panel with two other Democratic appointees, but only 15 percent of the time when the panel colleagues are both Republican appointees. Republican appointees favor the asylum claimant 20 percent of the time when empaneled with two Democratic appointees, but only 6 percent of the time when on an all-Republican panel. These statistics reveal a strong impact of panel colleagues on a judge's vote: a Republican judge with two Democratic colleagues votes in the liberal direction more often than a Democratic judge with two Republican colleagues.

Table 1 provides summary statistics for the 27 Ninth Circuit judges who

[9] Most commonly, this will occur when asylum petitioners from the same country of origin are assigned to the same panel. Occasionally, petitioners raising a common procedural claim may also be grouped together.

[10] The settlement rate for asylum cases ranged between 6 and 8 percent from 1994 to 1999 and increased to 14 percent in 2000 (Palmer, Yale-Loehr, and Cronin 2005).

**Table 1**
**Voting Statistics in Ninth Circuit Asylum Cases, 1992–2001, by Judge**

| Judge | % Liberal Votes: Unanimous Decisions | % Liberal Votes: Nonunanimous Decisions | Number of Votes: Unanimous Decisions | Number of Votes: Nonunanimous Decisions | Party of Appointment |
|---|---|---|---|---|---|
| O'Scannlain | 3 | 8 | 184 | 24 | Republican |
| Wallace | 4 | 0 | 134 | 9 | Republican |
| Farris | 4 | 0 | 126 | 5 | Democrat |
| Leavy | 5 | 0 | 144 | 2 | Republican |
| Rymer | 5 | 0 | 184 | 6 | Republican |
| Fernandez | 6 | 17 | 207 | 6 | Republican |
| Sneed | 8 | 0 | 110 | 1 | Republican |
| Kozinski | 10 | 0 | 132 | 8 | Republican |
| Brunetti | 11 | 0 | 205 | 3 | Republican |
| Tashima | 11 | 20 | 136 | 5 | Democrat |
| Silverman | 11 | 0 | 114 | 2 | Democrat |
| Trott | 12 | 8 | 173 | 12 | Republican |
| Thompson | 12 | 0 | 140 | 5 | Republican |
| Hall | 13 | 0 | 125 | 5 | Republican |
| Kleinfeld | 14 | 40 | 128 | 5 | Republican |
| Canby | 15 | . . . | 109 | 0 | Democrat |
| Beezer | 15 | 17 | 142 | 6 | Republican |
| Hug | 15 | 75 | 139 | 4 | Democrat |
| Thomas | 16 | 93 | 105 | 14 | Democrat |
| Nelson, T. | 16 | 0 | 195 | 3 | Republican |
| Schroeder | 16 | 100 | 268 | 3 | Democrat |
| Goodwin | 22 | 50 | 156 | 4 | Republican |
| Hawkins | 22 | 57 | 177 | 14 | Democrat |
| Browning | 23 | 100 | 136 | 3 | Democrat |
| Pregerson | 37 | 100 | 167 | 29 | Democrat |
| Fletcher, B. | 46 | 100 | 116 | 9 | Democrat |
| Reinhardt | 51 | 100 | 87 | 28 | Democrat |
| Averages: | | | | | |
| All Democrats | 21 | 72 | 2,541 | 173 | |
| All Republicans | 12 | 17 | 2,835 | 127 | |
| Overall | 16 | 48 | 5,376 | 300 | |

participated in at least 100 cases, with voting data displayed separately for unanimous and nonunanimous decisions. These statistics show why the inclusion of unanimous decisions is necessary for estimating the judges' preferences. The first data column reveals that there is substantial variation in the judges' voting rates, even when restricted to the unanimous decisions: Judge O'Scannlain favors the asylum claimant 3 percent of the time in unanimous cases, compared to 51 percent for Judge Reinhardt. The nonunanimous decisions, on the other hand, are quite sparse. Even among the judges who participated in at least 100 cases, a majority have five or fewer votes in nonunanimous cases. Eleven of these judges always vote in the conservative direction in nonunanimous cases, and five always vote in the liberal direction. The votes in nonunanimous cases are also highly concentrated among judges at the extreme ends of the spectrum. Of these 100 cases, 72 involve at least one of five judges—Reinhardt, Harry Pregerson, Betty Fletcher, O'Scannlain, and J. Clifford Wallace—and 27 include two of these judges.

Another complication is that the nonunanimous cases follow nearly perfect spatial voting: it is possible to select ideal points and case cut points that correctly classify all but one of the votes in the data set.[11] When voting behavior is nearly deterministic, probabilistic voting models that rely solely on observed disagreements can be difficult to estimate (Rosenthal and Voeten 2004). Finally, because the cases are decided by rotating three-judge panels, most pairs of judges are never compared with each other in nonunanimous cases, either directly or through indirect comparisons with other judges.[12]

The final column of Table 1 provides each judge's party of appointment. Although asylum voting rates are roughly correlated with party, there are some exceptions, most notably Judge Joseph Farris, who was appointed by President Carter but has one of the most anti-asylum voting records in the Ninth Circuit. The heterogeneity among the judges' voting rates within each party demonstrates that there is substantial measurement error associated with relying on political party to explain judicial behavior.

## 4. Results

Table 2 presents parameter estimates for both the sincere and consensus voting models, including indifference point estimates for all judges who participated in at least 100 cases. In the consensus voting model, the cost of dissent is positive and significant at the 1 percent level. It is also large relative to the dispersion of the judges' indifference points, exceeding the distance between Judge Pamela Rymer, one of the most conservative judges, and Judge Michael Hawkins, one of the most liberal. The standard deviation $\sigma$ of the random effect is small relative to the cost of dissent and the ideological distances among the judges, suggesting that the high rate of unanimous opinions is not merely driven by a preponderance of easy cases.

The magnitudes of the judges' indifference points appear to be much larger in the sincere voting model than in the consensus voting model. However, $\sigma$ is also much larger, meaning that there is a greater variation in the cut points. Thus, the larger magnitudes of the indifference points in this sincere voting model do not translate into greater differences in voting behavior. The rank ordering of the judges is similar in both models, but many of the differences between pairs of judges are not statistically significant.

[11] The only vote that cannot be perfectly classified arises from a pair of cases in which Judges Andrew Kleinfeld and M. Margaret McKeown alternate taking the liberal and conservative sides.

[12] Poole (2005, p. 48) advocates using a nonparametric optimal classification algorithm to estimate voter preferences when voting is nearly deterministic. Because there is no basis for comparing most pairs of judges in the asylum data, however, the optimal classification algorithm cannot generate useful bounds on the judges' indifference points.

Table 2
Results

| | Consensus Voting Model | | Sincere Voting Model | |
|---|---|---|---|---|
| | Estimate | SE | Estimate | SE |
| Judge: | | | | |
| Wallace | −3.46 | 1.06 | −8.52 | 1.26 |
| Farris | −3.08 | 1.23 | −8.84 | .81 |
| O'Scannlain | −2.92 | .90 | −8.25 | .73 |
| Rymer | −2.40 | .68 | −8.24 | .65 |
| Leavy | −2.22 | .66 | −7.97 | .96 |
| Kozinski | −2.22 | .69 | −8.12 | .66 |
| Fernandez | −2.06 | .72 | −8.05 | .57 |
| Brunetti | −1.89 | .87 | −7.38 | .78 |
| Trott | −1.91 | .63 | −7.35 | 1.23 |
| Thompson | −1.78 | .64 | −6.86 | .51 |
| Sneed | −1.71 | .75 | −7.23 | .45 |
| Hall | −1.68 | .64 | −7.63 | .68 |
| Nelson, T. | −1.39 | .57 | −6.62 | 4.42 |
| Canby | −1.35 | .60 | −6.89 | .78 |
| Silverman | −1.32 | .78 | −7.57 | 1.10 |
| Tashima | −1.28 | .66 | −7.72 | .64 |
| Beezer | −1.28 | .61 | −6.50 | .66 |
| Goodwin | −1.19 | .86 | −7.07 | .90 |
| Kleinfeld | −1.09 | .67 | −6.92 | .74 |
| Schroeder | −.80 | .59 | −6.61 | .71 |
| Hug | −.76 | .67 | −5.59 | .62 |
| Browning | −.54 | .53 | −6.18 | .64 |
| Hawkins | −.39 | .51 | −5.73 | .63 |
| Thomas | .24 | .53 | −5.50 | .67 |
| Fletcher, B. | .52 | .44 | −5.07 | 2.71 |
| Pregerson | .77 | .46 | −4.48 | .92 |
| Reinhardt | 1.65 | .52 | −1.51 | 1.11 |
| Model parameters: | | | | |
| Cost of dissent | 2.16 | .17 | .00 | . . . |
| SD of random effect ($\sigma$) | .71 | .32 | 5.00 | . . . |
| log Likelihood | −973.19 | | −1,159.75 | |

**Note.** Reported are maximum likelihood estimates from the consensus voting model and the sincere voting model. Standard errors are clustered for panels who decided multiple cases. Provided are indifference point estimates for all judges with at least 100 votes. Year dummies are not reported. The cost of dissent and $\sigma$ are fixed at 0 and 5, respectively, in the sincere voting model.

Since proposition 2*a* shows that the sincere voting model coincides with the consensus voting model when $c = 0$, the fit of the models can be compared using a likelihood ratio test on this restriction. The difference between the log likelihoods for the two models, shown in Table 2, is highly significant for a single-parameter restriction.[13] Thus, the hypothesis that judges vote independently can be rejected. This result is consistent with the literature on panel effects,

[13] A cluster-adjusted likelihood ratio test (Rotnitzky and Jewell 1990) yields a $\chi^2_1$-test statistic of 128.9.

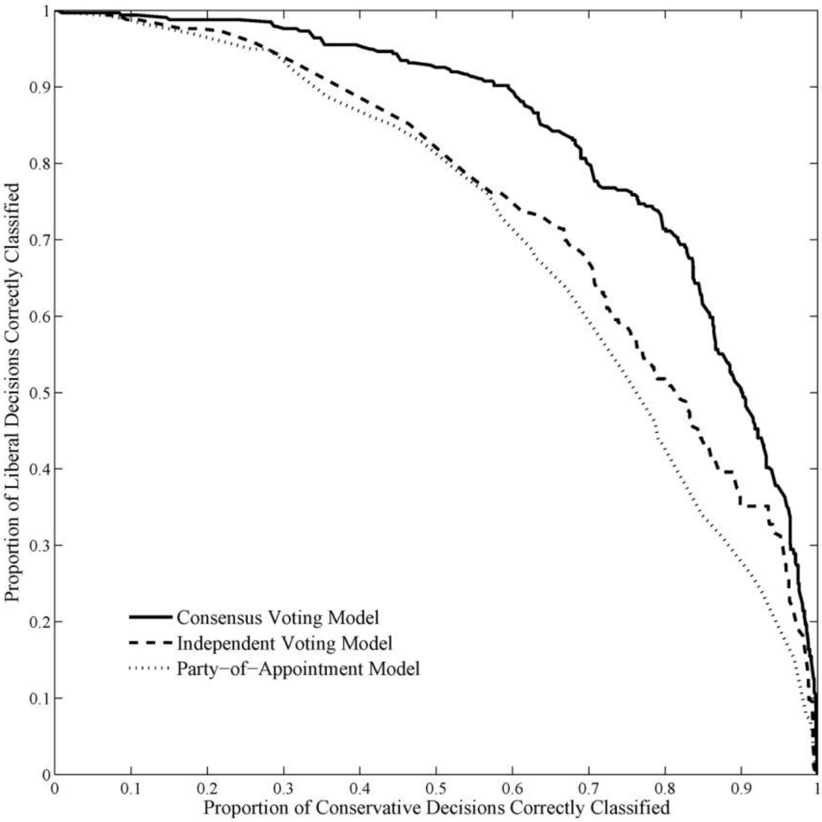


**Figure 2.** Receiver operating characteristic curves depicting proportion of panel decisions correctly classified for consensus voting model, sincere voting model, and party-of-appointment model.

although the test employed here does not require the availability of proxy variables, such as party of appointment, that are correlated with judicial preferences.

## *4.1. Classification Rates*

To compare how well the models classify panel decisions, Figure 2 provides receiver operating characteristic (ROC) curves for the consensus voting model, the sincere voting model, and a party-of-appointment model[14] similar to those

[14] The party-of-appointment model predicts votes on the basis of whether a judge and the judge's colleagues were appointed by Republican or Democratic presidents. To generate the fitted probabilities, I estimated a random-effects probit model using a dummy variable indicating whether the judge was appointed by a Democrat, a variable for the number of Democrat-appointed panel colleagues, and year dummies. The panel probability of a liberal decision represents the probability that

commonly used in studies of circuit courts. The party-of-appointment model accounts for interdependence in panel voting but does not analyze circuit judges individually; the sincere voting model analyzes judges individually but assumes strictly independent voting. These curves plot the proportion of correctly predicted liberal and conservative panel decisions as the classification threshold varies from 0 to 1. The ROC curve for the consensus voting model dominates the ROC curves for both the sincere voting model and the party-of-appointment model: for every ratio of liberal and conservative decisions correctly classified, the consensus voting model correctly classifies a greater proportion of decisions overall. The sincere voting model also dominates the party-of-appointment model. At the thresholds that equalize classification error among liberal and conservative decisions, the consensus voting model correctly classifies 76 percent of votes, the sincere voting model correctly classifies 67 percent, and the party-of-appointment model correctly classifies 65 percent.

### *4.2. Bias in Predicted Probabilities*

Although the indifference point estimates can be used to make ordinal comparisons, they do not have a cardinal interpretation (Ho and Quinn 2010). One advantage of the consensus voting model is that these indifference point estimates can be used to calculate accurate voting probabilities for any real or hypothetical panel. Figure 3 provides a calibration plot of panel voting probabilities from both the consensus and sincere voting models. For both models, I sorted the cases according to the fitted probability of a liberal panel decision,[15] grouped the cases into quintiles, and plotted the average fitted probability of a liberal decision in each quintile against the actual proportion of liberal decisions in that quintile. All five quintiles of the consensus voting model are close to the 45-degree line, revealing a lack of visible bias. The sincere voting model, on the other hand, generates voting probabilities that are clearly biased toward the mean, overpredicting liberal voting rates for the lowest quintile and underpredicting liberal voting rates for the three highest quintiles.

The explanation for this bias is that a norm of consensus pushes judges' voting rates toward the mean. A judge on the conservative end of the spectrum will often be empaneled with colleagues who are more liberal. These panel colleagues will, on average, influence the judge to vote more liberally than if the judge were always voting sincerely. Similarly, consensus voting will typically push a liberal judge in the conservative direction. Because the sincere voting model fails to account for this, it estimates all judges to be more moderate than they actually are.

The consensus voting model also generates accurate voting probabilities for

at least two judges favor the liberal outcome. A model that used judicial common-space scores provided a fit that was nearly identical to that of the party-of-appointment model. For a more detailed comparison of these measures, see Fischman and Law (2009).

[15] The fitted probability of a liberal panel vote is calculated by summing the probability of all voting profiles in which at least two judges support the liberal outcome.

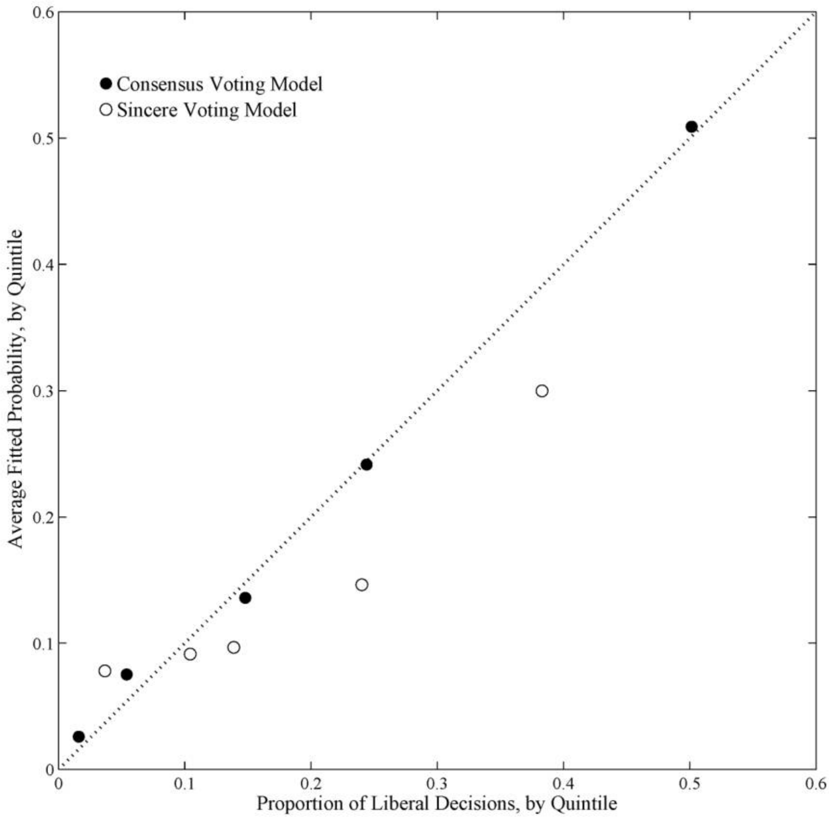


**Figure 3.** Calibration plot comparing fitted probabilities versus actual panel decisions, averaged by quintile.

individual judges. Figure 4 provides a calibration plot comparing the average fitted probability of a liberal vote[16] and the proportion of liberal votes for each judge who participated in at least 30 cases. The fit is reasonably close, and there is no perceptible bias. The most prominent outlier is Judge Warren Ferguson, who wrote five dissents in 61 cases, resulting in a rate of dissent four times higher than the average. The fit in Figure 4 is remarkable in light of the fact that the model assumes a uniform cost of dissent for all judges. If there were substantial heterogeneity in the cost of dissent, this would lead the model to provide a poor fit for some of the judges. The close fit in Figure 4 suggests that

[16] The fitted probability of a liberal vote for a judge in a particular case is calculated by summing the probability of all voting profiles in which the judge casts a liberal vote. This probability is then averaged over all cases in which the judge participates. The fitted probabilities therefore incorporate the influence of the judges' panel colleagues in various cases.

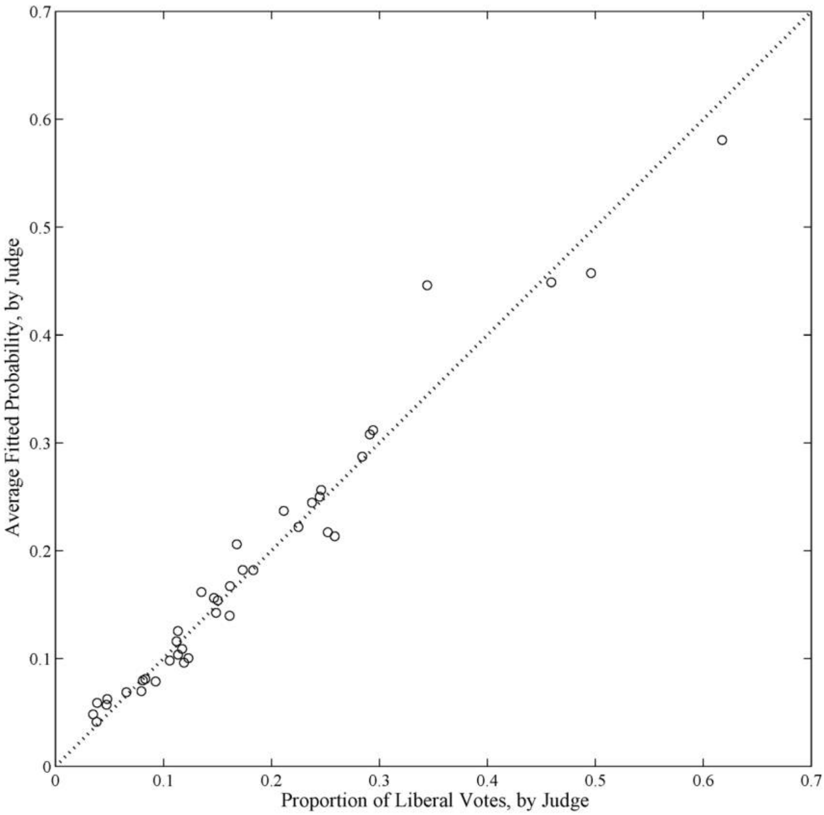


**Figure 4.** Calibration plot comparing fitted probabilities versus actual votes, averaged by judge, for consensus voting model.

the assumption of a uniform cost of dissent explains the voting behavior in the asylum cases reasonably well.

## 5. Applications of the Consensus Voting Model

### *5.1. Autonomous Voting Rates*

The indifference point estimates generated by the consensus voting model locate the judges in an abstract policy space. To make these estimates more tangible, the model can be used to estimate a hypothetical autonomous voting rate for each judge. This represents the proportion of cases in which the judge would support the liberal outcome if she were voting autonomously in every case. Since a judge's observed voting rate incorporates the accumulated influence of panel colleagues in various cases, the autonomous voting rate better represents her intrinsic preferences. It more accurately predicts how the judge would vote if she were empaneled with like-minded colleagues or if she were the pivotal member of a panel.

A judge's autonomous voting rate is determined by calculating the expected proportion of cases in which the judge's indifference point is to the right of the cut point:

$$\hat{r}_i = \frac{1}{T}\sum_{t=1}^{T}\int \Phi(\hat{x}_i - z_t\gamma - \sigma u)\phi(u)du, \tag{4}$$

where $\hat{x}_i$ is the estimate of judge $i$'s indifference point and $T$ is the number of cases. Note that the quantity inside the summation takes the same form as in the sincere voting model in equation (2), but the estimate $\hat{x}_i$ is derived from the consensus voting model.

Table 3 provides estimated autonomous voting rates for all Ninth Circuit judges who decided at least 30 cases in the data. Note that the autonomous voting rates are more polarized than the judges' observed voting rates. For liberal judges, the observed voting rates are lower than the autonomous voting rates because the former incorporate the influence of panel colleagues, who are typically more conservative. Judge Reinhardt, for instance, has an observed voting rate of 62 percent but an estimated autonomous voting rate of 92 percent. Conservative judges, on the other hand, have autonomous voting rates that are lower than their observed voting rates, since the latter reflect the influence of more liberal colleagues. Judge Wallace voted in the liberal direction 3 percent of the time, but his autonomous voting rate is estimated to be only .5 percent.

For any hypothetical panel, the median autonomous voting rate among the three judges provides a simple estimate of the probability that the panel would reach a liberal decision. For example, consider a panel consisting of Judges Dorothy Nelson, Hawkins, and Sidney Thomas, who have observed voting rates of 24, 25, and 25 percent, respectively, but autonomous voting rates of 47, 42, and 61 percent. A naive observer might speculate, on the basis of the observed voting rates, that the panel would have a 25 percent probability of reaching a liberal decision. However, the judges' observed voting rates do not accurately reflect how they would behave when empaneled with like-minded colleagues. In fact, the consensus voting model predicts a 44 percent probability of a liberal decision, which is roughly in line with the median autonomous voting rate of these three judges.

### *5.2. The Impact of Panel Composition*

How much the outcomes of cases depend on the judges selected to decide them is a central inquiry in the study of circuit court decision making. The estimates of panel voting probabilities generated by the consensus voting model can be used to measure the impact of panel composition on case outcomes. To

**Table 3**
**Observed and Autonomous Voting Rates (%)**

| Judge | Observed Voting Rate | Autonomous Voting Rate Estimate | Autonomous Voting Rate SE |
|---|---|---|---|
| Wallace | 3 | .5 | 1 |
| Farris | 4 | 1.1 | 2 |
| O'Scannlain | 4 | 1.5 | 2 |
| Rymer | 5 | 4 | 2 |
| Leavy | 5 | 5 | 3 |
| Fernandez | 7 | 7 | 5 |
| Wiggins | 8 | 5 | 5 |
| Sneed | 8 | 11 | 7 |
| Graber | 8 | 7 | 5 |
| Kozinski | 9 | 5 | 3 |
| Brunetti | 11 | 8 | 8 |
| Silverman | 11 | 18 | 13 |
| Trott | 11 | 8 | 4 |
| Thompson | 12 | 10 | 5 |
| Alarcon | 12 | 14 | 10 |
| Hall | 12 | 11 | 6 |
| Skopil | 14 | 18 | 12 |
| Canby | 15 | 17 | 8 |
| Norris | 16 | 17 | 17 |
| Nelson, T. | 16 | 16 | 7 |
| Tashima | 11 | 18 | 9 |
| Beezer | 15 | 18 | 8 |
| Kleinfeld | 15 | 23 | 12 |
| Hug | 17 | 31 | 15 |
| Schroeder | 17 | 30 | 12 |
| McKeown | 18 | 25 | 11 |
| Fletcher, W. | 21 | 32 | 20 |
| Fletcher, B. | 50 | 70 | 6 |
| Reinhardt | 62 | 92 | 5 |

**Note.** Provided are observed voting rates and estimated autonomous voting rates under the consensus voting model for all judges who participated in at least 30 cases. The autonomous voting rate represents the rate at which the judge would support the liberal outcome, if the judge were voting autonomously in every case. Standard errors for the autonomous voting rates are calculated using the delta method.

generate a distribution of panel decision probabilities, I used the consensus voting model to estimate the probability of a liberal decision for 13,000 simulated panels, chosen from the judges (shown in Table 3) who participated in at least 30 decisions.[17] The distribution of predicted decision rates is shown in Table 4.

It is immediately evident that panel composition has a large impact on the decisions in asylum cases. Going from the fifth percentile to the ninety-fifth percentile of the spectrum increases the probability of a liberal decision from 3

[17] Thirteen panels can be selected without replacement from among the 40 judges. This selection process was repeated 1,000 times. The probability is calculated assuming a randomly chosen case from the data.

Table 4
**Simulated Probability of Liberal Decision**

| Panel | Probability (%) |
|---|---|
| Most conservative | .2 |
| 5th percentile | 3.3 |
| 25th percentile | 8.9 |
| Median | 16.1 |
| 75th percentile | 26.8 |
| 95th percentile | 49.4 |
| Most liberal | 85.8 |

to 49 percent. Thus, in at least 46 percent of the asylum cases, the outcome would be different if the case were assigned to a panel from the liberal end of the spectrum rather than a panel from the conservative end.

Asylum law may well allow for greater judicial discretion because of the ambiguity of the legal standards and the low salience of the cases. Determining whether panel composition has as much of an impact in other areas of case law will require further study.

### *5.3. Sincere Agreement Rates*

The consensus voting model can also be used to estimate how often unanimous opinions stem from sincere agreement as opposed to suppressed disagreement. For each unanimous case $t$, I simulated random draws of $\varepsilon_{it}$, $\varepsilon_{jt}$, $\varepsilon_{kt}$, and $\eta_t$ and used the parameter estimates to determine whether each random draw would result in sincere agreement, suppressed disagreement, or a dissenting opinion. Let $C_t$ denote the proportion of draws resulting in suppressed disagreement and $S_t$ the proportion resulting in sincere agreement. By Bayes's rule, the probability that unanimity in case $t$ resulted from sincere agreement is represented by the fraction $S_t/(S_t + C_t)$.

Although 95 percent of cases in the data were decided unanimously, the simulation results suggest that the judges sincerely agreed only 55 percent of the time. In another 40 percent of the cases, panel unanimity was due to suppressed disagreement. The high rate of unanimity on the courts of appeals is sometimes cited as evidence that most cases are easy and that circuit judges do not exercise much discretion (see, for example, Edwards 1991, 1998; Tamanaha 2007). These estimates, however, suggest that internal disagreement is far more widespread than dissent rates indicate.

## 6. Conclusion

As Judge Harry Edwards (2003, p. 1656) noted, "A model that takes each appellate judge as an atomized individual casting a purely individual vote in any given case will not produce a good explanation of how judges decide cases." The consensus voting model provides a more realistic account of voting behavior in circuit courts and explains the voting data far better than a model of sincere voting. The estimates from the consensus voting model suggest that roughly half the asylum cases could have been decided differently if they had been assigned to different panels. This model can also be applied to other data sets to determine to what extent this result generalizes to other circuits and to other areas of case law. Coding the cases in greater detail would make it possible to compare the impact of case characteristics and panel composition on outcomes. The estimates of judicial indifference points and autonomous voting rates can also be used as ideology scores in other studies, but care must be taken to avoid circularity when such measures are used in subsequent regressions.

The distortions that arise from the sincere voting assumption may also be relevant in other consensus-oriented voting bodies, such as the U.S. Supreme Court in earlier eras (Dorff and Brenner 1992; Epstein, Segal, and Spaeth 2001; Post 2001), the International Court of Justice (Voeten 2007), many state supreme courts (Brace and Butler 2001), and the Federal Open Market Committee (Meade 2005). The model constructed here for circuit court panels exploits random assignment and rotating panels to identify the cost of dissent. Whether comparable consensus voting models can be constructed for other voting bodies will depend on the institutional characteristics of those bodies or the availability of documents that shed light on internal deliberations.

## Appendix

### *A1. Likelihood Function for the Sincere Voting Model*

Let $q_{it} = 1$ if $y_{it} = 1$ and $q_{it} = -1$ if $y_{it} = 0$. For a given cut point $m_t$, the probability of any voting alignment in case $t$, conditional on $m_t$, can be expressed as follows:

$$\Pr(y_{1t}, y_{2t}, y_{3t}|m_t) = \Phi[q_{1t}(x_1 - m_t)]\Phi[q_{2t}(x_2 - m_t)]\Phi[q_{3t}(x_3 - m_t)].$$

Since $m_t = z_t\gamma + \eta_t$, where $\eta_t \sim N(0, \sigma^2)$, the likelihood is obtained by integrating over the distribution of cut points:

$$L_t = \int \Pr(y_{1t}, y_{2t}, y_{3t}|z_t\gamma + \sigma u)\phi(u)du, \tag{A1}$$

which can be approximated using Gauss-Hermite quadrature. A more detailed treatment of the random-effects probit model, which is functionally equivalent, can be found in Butler and Moffitt (1982).

### *A2. Likelihood Function for the Consensus Voting Model*

The probability of reaching a unanimous liberal outcome through sincere agreement is the same as in the sincere voting model, $\Phi(x_1 - m_t)\Phi(x_2 - m_t)\Phi(x_3 - m_t)$. The probability of reaching the same outcome as a result of suppressed disagreement on the part of judge 3, for example, would be

$\Phi(x_1 - m_t)\Phi(x_2 - m_t)[\Phi(x_3 - m_t + c) - \Phi(x_3 - m_t)]$. Summing over all possibilities for suppressed disagreement yields

$$\begin{aligned}\Pr[(y_{1t}, y_{2t}, y_{3t}) = (1, 1, 1)|m_t] = {} & \Phi(x_1 - m_t + c)\Phi(x_2 - m_t)\Phi(x_3 - m_t) \\ & + \Phi(x_1 - m_t)\Phi(x_2 - m_t + c)\Phi(x_3 - m_t) \\ & + \Phi(x_1 - m_t)\Phi(x_\tau - m_t)\Phi(x_3 - m_t + c) \\ & - 2\Phi(x_1 - m_t)\Phi(x_2 - m_t)\Phi(x_3 - m_t).\end{aligned} \tag{A2}$$

The probability of a unanimous conservative outcome takes a similar form:

$$\begin{aligned}\Pr[(y_{1t}, y_{2t}, y_{3t}) = (0, 0, 0)|m_t] = {} & \Phi(m_t - x_1 + c)\Phi(m_t - x_2)\Phi(m_t - x_3) \\ & + \Phi(m_t - x_1)\Phi(m_t - x_2 + c)\Phi(m_t - x_3) \\ & + \Phi(m_t - x_1)\Phi(m_t - x_2)\Phi(m_t - x_3 + c) \\ & - 2\Phi(m_t - x_1)\Phi(m_t - x_2)\Phi(m_t - x_3).\end{aligned} \tag{A3}$$

For a nonunanimous outcome—say, $(y_{1t}, y_{2t}, y_{3t}) = (1, 1, 0)$—the probability conditional on the cut point takes the following form:

$$\Pr[(y_{1t}, y_{2t}, y_{3t}) = (1, 1, 0)|m_t] = \Phi(x_1 - m_t)\Phi(x_2 - m_t)\Phi(m_t - x_3 - c),$$

and the other nonunanimous outcomes can be derived in similar fashion. The likelihood for case $t$ can then be computed by integrating over the distribution of cut points, as in equation (A1).

When the number of observations for a judge is below a certain threshold, the judge is assigned to a group based on observable characteristics. I group these judges by party of appointment, so that judges in the Democratic group have indifference points $x_i \sim N(x_{\text{Dem}}, \tau^2)$ and judges in the Republican group have indifference points $x_i \sim N(x_{\text{Rep}}, \tau^2)$, where $x_{\mathbf{Dem}}$, $x_{\mathbf{Rep}}$, and $\tau$ are additional parameters to be estimated. Then $(x_{\text{Dem}} - m_t)/(1 + \tau^2)$ or $(x_{\text{Rep}} - m_t)/(1 + \tau^2)$ may be substituted for $x_i - m_t$ in the likelihood function.

### *A3. Proof of Proposition 2*

Part *a* is an immediate consequence of the fact that the likelihood functions for the two models are equal when $c = 0$. For part *b*, note that differentiating equation (3) with respect to $x_j$ yields

$$\begin{aligned}\frac{d}{dx_j}\Pr(y_{it} = 1|m_t) = {} & \phi(x_j - m_t)\Phi(x_k - m_t)[\Phi(x_i - m_t + c) - \Phi(x_i - m_t)] \\ & + \phi(m_t - x_j)\Phi(m_t - x_k)[\Phi(x_i - m_t) - \Phi(x_i - m_t - c)],\end{aligned}$$

which is strictly positive when $c > 0$. Since this is true for all $m_t$, it follows that

$$\frac{d}{dx_j}\Pr(y_{it} = 1) > 0.$$

For part *c*, note that

$$\frac{d^2}{dx_j dc}\Pr(y_{it} = 1|m_t) = \phi(x_j - m_t)\Phi(x_k - m_t)\phi(x_i - m_t + c)$$

$$+ \phi(m_t - x_j)\Phi(m_t - x_k)\phi(x_i - m_t - c) > 0,$$

hence

$$\frac{d^2}{dx_j dc}\Pr(y_{it} = 1) > 0.$$

To show part *d*, note that equations (A2) and (A3) are both strictly increasing in *c*.

Part *e* follows immediately from the fact that the consensus voting model only affects the issuance of a dissenting opinion. Since the judges in the majority always vote sincerely, the case disposition will be the same in either model.

### *A4. Statistical Test for Random Assignment*

A key identifying assumption in this paper is that judges are randomly assigned to panels and to cases.[18] Without random assignment, the model would not be identified, since it would be impossible to distinguish between selection of judges into cases and the effect of judicial preferences and consensus voting.

The test of random assignment examines the number of Democratic judges assigned to claimants for each country of origin. Because of the way cases are clustered—a single panel may hear up to 50 asylum cases together, many of which may be from the same country—the hypothesis is tested using Monte Carlo simulation.[19] Under the assumption of random assignment, the proportion of Democratic judges assigned to claimants from each country should be representative of the proportion of Democratic appointees in the entire pool. The assignment of cases is simulated by randomly matching clusters of cases with three-judge panels that decided cases in the same year and calculating the percentage of Democratic appointees associated with claimants from each country in each simulated matching. Repeating this random-assignment simulation 10,000 times generates a simulated distribution of the number of Democratic judges assigned to claimants from each country.

[18] Note that when multiple cases are simultaneously assigned to a single panel, the characteristics of these cases may be correlated with each other but should be independent of the judges' indifference points. Formally, the identifying assumption is that if a panel of judges with indifference points $x_1$, $x_2$, $x_3$ is assigned cases with cut points $m_1, \ldots, m_t$, then the $x_i$ terms will be independent of each other and each $x_i$ will be independent of each $m_j$. There are many instances in the data in which a panel is assigned cases that clearly have common characteristics, such as the same country of origin. However, identification of the model does not require that the $m_j$ terms be independent of each other.

[19] The simulation method here closely follows the method developed in Abrams, Bertrand, and Mullainathan (forthcoming).

Random assignment is tested by comparing the actual percentage of Democratic appointees for each country with the simulated distribution. For country $k$, let $v_k$ be the actual percentage of Democratic appointees and $F_k$ be the empirical cumulative distribution function derived from the simulations. Then, under the assumption of random assignment, the test statistic $F_k(v_k)$ should be uniformly distributed.

For example, there were 318 claimants from Nicaragua in the data, and 40 percent of the judges assigned to those cases were Democrats. In the simulated matchings, Nicaraguan claimants were assigned a lower percentage of Democrats 29.5 percent of the time and a higher percentage 70.5 percent of the time. Thus, the test statistic for Nicaragua would be .295.

This process yields test statistics associated with the 25 most frequently occurring countries of origin (those countries for which there are at least 10 cases) and an additional test statistic for all remaining countries grouped together. A Kolmogorov-Smirnov test fails to reject the assumption that these test statistics are distributed uniformly at the significance level $p < .10$, suggesting that the data are consistent with the assumption that cases are randomly assigned.

## References

Abrams, David, Marianne Bertrand, and Sendhil Mullainathan. Forthcoming. Do Judges Vary in Their Treatment of Race? *Journal of Legal Studies*.

Alarie, Benjamin R. D., and Andrew Green. 2007. The Reasonable Justice: An Empirical Analysis of Frank Iacobucci's Career on the Supreme Court of Canada. *University of Toronto Law Journal* 57:195–226.

Aumann, Robert. 1959. Acceptable Points in General Cooperative *n*-Person Games. Pp. 287–324 in *Contributions to the Theory of Games IV*, edited by Albert W. Tucker and Robert D. Luce. Princeton, N.J.: Princeton University Press.

Bailey, Michael A. 2007. Comparable Preference Estimates across Time and Institutions for the Court, Congress, and Presidency. *American Journal of Political Science* 51:433–48.

Bernheim, B. Douglas, Bezalel Peleg, and Michael D. Whinston. 1987. Coalition-Proof Nash Equilibria I: Concepts. *Journal of Economic Theory* 42:1–12.

Boyd, Christina L., Lee Epstein, and Andrew D. Martin. 2010. Untangling the Causal Effects of Sex on Judging. *American Journal of Political Science* 54:389–411.

Brace, Paul, and Kellie Sims Butler. 2001. New Perspectives for the Comparative Study of the Judiciary: The State Supreme Court Project. *Justice System Journal* 22:243–62.

Butler, J. S., and Robert Moffitt. 1982. A Computationally Efficient Quadrature Procedure for the One-Factor Multinomial Probit Model. *Econometrica* 50:761–64.

Clinton, Joshua, Simon Jackman, and Douglas Rivers. 2004. The Statistical Analysis of Roll Call Data. *American Political Science Review* 98:355–70.

Cohen, Jonathan M. 2002. *Inside Appellate Courts: The Impact of Court Organization on Judicial Decision Making in the United States Courts of Appeals*. Ann Arbor: University of Michigan Press.

de Figueiredo, John M. 2005. Strategic Plaintiffs and Ideological Judges in Telecommunications Litigation. *Journal of Law, Economics, and Organization* 21:501–23.

Dorff, Robert H., and Saul Brenner. 1992. Conformity Voting on the United States Supreme Court. *Journal of Politics* 54:762–75.

Edwards, Harry T. 1991. The Judicial Function and the Elusive Goal of Principled Decisionmaking. *Wisconsin Law Review* 1991:837–65.

———. 1998. Collegiality and Decision Making on the D.C. Circuit. *Virginia Law Review* 84:1335–70.

———. 2003. The Effects of Collegiality on Judicial Decision Making. *University of Pennsylvania Law Review* 151:1639–90.

Epstein, Lee, William M. Landes, and Richard A. Posner. 2011. Why (and When) Judges Dissent: A Theoretical and Empirical Analysis. *Journal of Legal Analysis* 3:101–37.

Epstein, Lee, Andrew D. Martin, Jeffrey A. Segal, and Chad Westerland. 2007. The Judicial Common Space. *Journal of Law, Economics, and Organization* 23:303–25.

Epstein, Lee, Jeffrey A. Segal, and Harold J. Spaeth. 2001. The Norm of Consensus on the U.S. Supreme Court. *American Journal of Political Science* 45:362–77.

Fischman, Joshua B. 2011. Interpreting Circuit Court Voting Patterns: A Social Interactions Framework. Working Paper. University of Virginia School of Law, Charlottesville.

Fischman, Joshua B., and David S. Law. 2009. What Is Judicial Ideology, and How Should We Measure It? *Washington University Journal of Law and Policy* 29:133–214.

Giles, Micheal W., Virginia A. Hettinger, and Todd Peppers. 2001. Picking Federal Judges: A Note on Policy and Partisan Selection Agendas. *Political Research Quarterly* 54:623–41.

Ginsburg, Ruth Bader. 1990. Remarks on Writing Separately. *Washington Law Review* 65:133–50.

Ho, Daniel E., and Kevin M. Quinn. 2010. How Not to Lie with Judicial Votes: Misconceptions, Measurement, and Models. *California Law Review* 98:813–76.

Law, David S. 2005. Strategic Judicial Lawmaking: Ideology, Publication, and Asylum Law in the Ninth Circuit. *University of Cincinnati Law Review* 73:817–66.

Legomsky, Stephen H. 2010. Restructuring Immigration Adjudication. *Duke Law Journal* 59:1635–721.

Londregan, John. 1999. Estimating Legislators' Preferred Points. *Political Analysis* 8:35–56.

Manski, Charles F. 1993. Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies* 60:531–42.

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* 10:134–53.

Meade, Ellen E. 2005. The FOMC: Preferences, Voting, and Consensus. *Federal Reserve Bank of St. Louis Review* 87:93–101.

Neyman, J., and Elizabeth L. Scott. 1948. Consistent Estimates Based on Partially Consistent Observations. *Econometrica* 16:1–32.

Palmer, John R. B., Stephen W. Yale-Loehr, and Elizabeth Cronin. 2005. Why Are So Many People Challenging Board of Immigration Appeals Decisions in Federal Court? An Empirical Analysis of the Recent Surge in Petitions for Review. *Georgetown Immigration Law Journal* 20:1–100.

Peresie, Jennifer L. 2005. Female Judges Matter: Gender and Collegial Decisionmaking in the Federal Appellate Courts. *Yale Law Journal* 114:1759–90.

Peress, Michael. 2009. Small Chamber Ideal Point Estimation. *Political Analysis* 17:276–90.

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

Poole, Keith T., and Howard Rosenthal. 1997. *Congress: A Political-Economic History of Roll Call Voting.* New York: Oxford University Press.

Posner, Richard A. 2008. *How Judges Think.* Cambridge, Mass.: Harvard University Press.

Post, Robert. 2001. The Supreme Court Opinion as Institutional Practice: Dissent, Legal Scholarship, and Decisionmaking in the Taft Court. *Minnesota Law Review* 85:1267–1390.

Revesz, Richard L. 1997. Environmental Regulation, Ideology, and the D.C. Circuit. *Virginia Law Review* 83:1717–72.

Rosenthal, Howard, and Erik Voeten. 2004. Analyzing Roll Calls with Perfect Spatial Voting: France, 1946–1958. *American Journal of Political Science* 48:620–32.

Rotnitzky, Andrea, and Nicholas P. Jewell. 1990. Hypothesis Testing of Regression Parameters in Semiparametric Generalized Linear Models for Cluster Correlated Data. *Biometrika* 77:485–97.

Snyder, James M., Jr., and Tim Groseclose. 2000. Estimating Party Influence in Congressional Roll-Call Voting. *American Journal of Political Science* 44:193–211.

Spirling, Arthur, and Iain McLean. 2007. UK OC OK? Interpreting Optimal Classification Scores for the U.K. House of Commons. *Political Analysis* 15:85–96.

Sunstein, Cass R., David Schkade, Lisa M. Ellman, and Andres Sawicki. 2006. *Are Judges Political? An Empirical Analysis of the Federal Judiciary.* Washington, D.C.: Brookings Institution Press.

Tamanaha, Brian Z. 2007. How an Instrumental View of Law Corrodes the Rule of Law. *DePaul Law Review* 56:469–505.

Tiller, Emerson H., and Frank B. Cross. 1998. Judicial Partisanship and Obedience to Legal Doctrine: Whistleblowing on the Federal Courts of Appeal. *Yale Law Journal* 107: 2155–76.

Voeten, Erik. 2007. The Politics of International Judicial Appointments: Evidence from the European Court of Human Rights. *International Organization* 61:669–701.

Waldfogel, Joel. 1995. The Selection Hypothesis and the Relationship between Trial and Plaintiff Victory. *Journal of Political Economy* 103:229–60.