

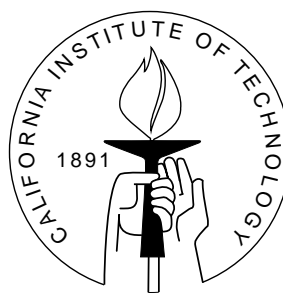
DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

# **CALIFORNIA INSTITUTE OF TECHNOLOGY**

PASADENA, CALIFORNIA 91125

THE VALUE OF INFORMATION IN THE COURT.  
GET IT RIGHT, KEEP IT TIGHT.

Matias Iaryczower and Matthew Shum



## **SOCIAL SCIENCE WORKING PAPER 1313**

September 2009  
revised July, 2010

# **The Value of Information in the Court. Get it Right, Keep it Tight.**

Matias Iaryczower and Matthew Shum

## **Abstract**

We estimate an equilibrium model of decision-making in the US Supreme Court which takes into account both private information and ideological differences between justices. We present a measure of the value of information in the court: the probability that a justice votes differently than what she would have voted for in the absence of case-specific information. Our results suggest a sizable value of information: in roughly 44% of cases, justices' initial leanings are changed by their personal assessments of the case. Our results also confirm the increased politicization of the Supreme Court in the last quarter century. We perform counterfactual simulations to draw implications for institutional design.

JEL classification numbers: D72, D78, C13

Key words: Supreme Court, common values, strategic voting, mixture model

# The Value of Information in the Court. Get it Right, Keep it Tight.\*

Matias Iaryczower and Matthew Shum<sup>†</sup>

“A judge sworn to decide impartially can offer no forecasts, no hints, for that would show not only disregard for the specifics of the particular case, it would display disdain for the entire judicial process.” *From the statement submitted to the Senate Committee on the Judiciary by Justice Ruth Ginsburg.*

“One should always subordinate his own personal views, whether they be economic, social, political, or whatever they may be, because when you are talking about your own views you are only one of millions of individuals in the country. When you are interpreting the law, perhaps you have a special skill and special training that does give you the right to pass on these questions. I have to confess that in this open area, sometimes inevitably, a man is the product of his own background and he may be somewhat influenced. But I will do my very best to subordinate those considerations because I think that is the duty of any judge.” *Justice John Stevens, in his Senate confirmatory hearing*

## 1 Introduction

It is a commonplace, in the press and popular discussions, to characterize the Supreme Court in terms of the ideological divisions among its members. In the Roberts’ 06-09 court, for example, Chief Justice Roberts, and Justices Scalia, Thomas, and Alito are typically thought of as the Court’s conservative wing, while Justices Kennedy and O’Connor are depicted as moderates, and Justices Stevens, Ginsburg, and Breyer as part of the Court’s liberal wing.

This ideological characterization of the court can be a useful starting point in analyzing the behavior of Supreme Court justices. But a purely ideological account is

---

\*The title follows words from Justice Ruth Ginsburg, “If confirmed, I will take the counsel to heart and strive to write opinions that both ‘get it right’ and ‘keep it tight’ ” (statement submitted to the Senate Committee on the Judiciary).

<sup>†</sup>Division of Humanities and Social Sciences, California Institute of Technology, Pasadena, California 91125, USA, emails: miaryc@hss.caltech.edu, mshum@caltech.edu

incomplete at best. Judging entails deciding *what was decided*; it requires understanding the case under consideration, understanding the body of the law, and interpreting the meaning of the law as it applies to the case. As Justice Ruth Ginsburg put it, “[E]ach case is based on particular facts and its decision should turn on those facts and the governing law, stated and explained in light of the particular arguments the parties or their representatives choose to present.”<sup>1</sup> This particular decision-making process is the reason why the qualifications of candidates to the Supreme Court receive close scrutiny in the press, and why the competence of candidates is a significant factor explaining whether senators vote to confirm a nominee (Segal, Cameron, and Cover (1990, 1992), Epstein, Lindstadt, Segal, and Westerland (2006)). Qualifications matter because the decision-making process in the Court is not *only* ideological.<sup>2</sup>

In this paper, we build on the existing literature to incorporate the value of information into the purely ideological framework of the spatial model. We provide an analysis of decision-making in the court taking into account not only the possible bias or ideology of justices, but also the information available to the justices in each case, as well as their ability – or skill, as Justice Stevens puts it – to map the law and the specifics of the case to an outcome. In this context, we ask: does the case information have enough power to overturn the prior biases and ideological considerations of the justices? Our analysis allows us to quantify precisely the degree to which justices “subordinated” their personal views and “interpreted” the law.

To tackle this question, we consider a model in which ideology interacts with common values in an incomplete information context.<sup>3</sup> We then estimate the parameters of the model from Supreme Court voting outcomes. In order to do this, we introduce a new estimation approach that allows us to handle our model of voting with common values and strategic agents.

In the model, we assume that the goal of any justice  $i$  in any given case  $t$  is that she (in the *expressive* voting model) or the Court (in the *strategic* voting model) rules according to  $i$ ’s own best understanding of how the law applies to the particulars of the case. We maintain that it is the residual uncertainty in the meaning of the law which allows justices to differ in their opinions about a case. With anything less than complete certainty, opinions can differ among justices because of idiosyncratic thresholds of proof brought by ideological differences, because of differences in the information that is effectively available to each justice, or because of differences in the ability to evaluate the available information in different contexts.

In particular, we assume that before ruling in each case  $t$ , each justice observes a private signal, which reflects her understanding of the particulars of the case. The precision ( $\theta_{it}$ ) of each justice’s signal measures her ability to map the specifics of the case

---

<sup>1</sup>From the statement submitted to the Senate Committee on the Judiciary by Justice Ruth Ginsburg.

<sup>2</sup>See also Cameron and Kornhauser (2008) and Lax and Cameron (2007).

<sup>3</sup>Such models of voting have been proposed in the information aggregation literature on voting. See for example Austen-Smith and Banks (1996), Feddersen and Pesendorfer (1997, 1998). Our model is closest to Duggan and Martinelli (2001).

to the meaning of the law. The imprecision in the information leaves room for interpretation, which in turn allows ideological biases to come into play. These biases could reflect variation across a liberal/conservative dimension, theoretical arguments about the law, or other determinants for a non-neutral approach to this case. In the model, this bias or ideology boils down to a threshold  $\pi_{it}$  such that the justice prefers to rule for the Plaintiff in this case if and only if the probability that the law favors the Plaintiff is at least  $\pi_{it}$ .<sup>4</sup> Information precision and bias then interact to produce outcomes. Higher precision means that it is typically more clear for the justice whether the ruling should favor the Plaintiff or the Defendant according to the body of law. A larger bias means that despite her case information, a justice persists in going with her preconception of how to rule in a case like this. In the extreme, with  $\pi_{it} \approx 0$  (or  $\pi_{it} \approx 1$ ), justice  $i$  will vote almost completely in line with her ideology. On the other hand, when  $\pi_{it} = 1/2$  for all  $i$ , the setting boils down to an unbiased, pure common values model.

In the estimation, we recover the values of  $(\theta_{it}, \pi_{it})|X_t$  for each justice  $i$  conditioning on observable covariates  $X_t$  of the cases and the justices. Thus justices' bias and information precision vary with the type of issue, characteristics of the Plaintiff and the Defendant, and other case characteristics, as well with their own judicial experience, court experience, and other individual characteristics of each justice. We estimate the model in two steps. In the first step, using the observed votes, we estimate a "reduced-form" model of justices' probabilities of voting in favor of the Plaintiff when the law favors the Plaintiff and when the law favors the Defendant. In the second step, we recover the structural parameters characterizing justices' preferences and information  $\{(\theta_{it}, \pi_{it})\}_{i=1}^n|X_t$ , using the equilibrium conditions in the voting model. We do this for both the expressive voting model (where justices care about getting their decision right) and the strategic voting model, where justices are concerned about getting the court's decision right, and therefore "learn" from their peers in equilibrium.

Our approach allows us to disentangle the effects of ideology and information for each justice, and therefore to quantify the tradeoffs between ideology and information in the court. The main result of the paper is a measure of the value of information in the court. Our measure, FLEX, is the probability that a justice  $i$  votes differently than what he or she would have voted for in the absence of case specific information.

The results suggest a sizable value of information in the court: in roughly 44% of cases, justices' initial leanings – which reflect their priors or their ideological biases – are changed by the case-specific private information of the justices. Moreover, the temporal evolution of FLEX scores suggest that the increased politicization of the Supreme Court appointment process pointed out by the literature has become uniformly marked only in the last quarter-century, following the failed nomination of justice Bork (i.e., within the Rehnquist and Roberts courts). Indeed, for Economics and Federalism cases, FLEX scores have fallen about 40% from the Warren to the Roberts courts, but for basic rights and criminal issues we see a marked decrease in FLEX scores (of about 20% and 16%

---

<sup>4</sup>We write  $\pi_{it}$  and  $\theta_{it}$  instead of simply  $\pi_i, \theta_i$  because in the estimation we allow these parameters to change with the characteristics of each case  $X_t$ .

respectively) only during the Rehnquist and Roberts courts.

In other results, we compare the value of information across individual justices and issues. We show that justices with the lowest FLEX scores tend to have a liberal judicial philosophy, while justices with the highest FLEX scores are not necessarily conservative, and include a more than proportional share of chief justices. We also show that the value of information in the court is lower in issue areas in which ideological considerations tend to weigh more heavily – such as Basic Rights. We also consider whether justices are differently predisposed when the government is involved in a case (they are), and when the court considers the constitutionality of a law enacted by Congress (they are, in Basic Rights cases).

Finally, we use counterfactual simulations to draw implications for institutional design. In particular, we compare the performance of the court with a counterfactual scenario in which ruling against the Defendant requires the unanimous consent of the justices. In a small committee composed of a heterogeneous group of individuals in terms of both preferences and abilities, such as the Court, majority rule does not generically dominate unanimity rule. The results for the Court however are in line with the asymptotic results first pointed out by Feddersen and Pesendorfer (1998) (see also Dugan and Martinelli (2001) and Meirowitz (2002)), that unanimity rule leads to a larger probability of error than simple majority rule.

The rest of the paper is organized as follows. Section 2 presents the relation with the literature. Section 3 introduces the theoretical model. Section 4 describes the data, and Section 5 describes the estimation procedure. Section 6 discusses the results. Section 7 concludes.

## 2 Related Literature

This paper builds on the significant contributions of a large literature. A first group of papers uses justices’ voting data - the proportion of votes in favor and against the Defendant, or the proportion of liberal and conservative votes - to test various hypothesis in reduced form models. Segal and Cover (1989) show that the ideology of each justice - as measured by the proportion of liberal and conservative statements in newspaper editorials - is highly correlated with the votes of justices in civil liberties cases. Segal, Epstein, Cameron, and Spaeth (1995) expands the coverage of the original SC scores, and shows that the correlation is lower for other justices and other issues (economic regulation). Epstein and Mershon (1996) further argue that newspaper editorials are tilted towards a few “splashy” civil liberties issues, and show that the scores have little explanatory power for most non-civil liberty areas. Epstein, Hoekstra, Segal, and Spaeth (1998) argue that the preferences of justices - as measured by the proportion of liberal votes on civil liberties cases - changes through time.<sup>5</sup> Finally, Landes and Posner (2008)

---

<sup>5</sup>See also Segal and Spaeth (1993) and Epstein and Knight (1997).

argues that members of a liberal or conservative minority do not tend to vote more often with the majority the larger the majority is. They also show that justices appointed by Democratic presidents (but not those appointed by Republican presidents) vote more liberally the fewer of them there are.

The first group of papers measures the ideological preferences of justices with the proportion of liberal statements in newspaper editorials or directly with the proportion of liberal votes by each justice. A second group of papers employs a radically different approach to recover the ideology of the justices from the data. [Martin and Quinn \(2002, 2007\)](#) build on the influential literature analyzing voting records in legislatures ([Poole and Rosenthal \(1985, 1991\)](#), [Heckman and Snyder \(1997\)](#), [Clinton, Jackman, and Rivers \(2004\)](#)). The main idea here is to assume that the voting data is generated by a precise model of behavior - the sincere spatial voting model (SVM) commonly employed in political science - and then estimate the parameters of the model from the voting data (i.e., structural estimation). Building on the findings of [Epstein, Hoekstra, Segal, and Spaeth \(1998\)](#), Martin and Quinn also allow ideal policies to change flexibly through time, but the underlying theoretical model is otherwise the same as in the above papers.<sup>6,7</sup>

For all its attractive properties, the sincere spatial voting model (SVM) has one severe limitation when applied to the analysis of voting in the Court: it is a pure private values model in which ideology is the only determinant of voting behavior. This precludes the possibility of common values and dispersed information which, as we argued above, seem central to the nature of decision-making in the court. In this paper, we therefore structurally estimate a model that allows both ideology and precision of private information to come into play. With common values and dispersed information, strategic considerations - which are absent in the sincere voting spatial model - come into play.<sup>8,9</sup>

In this paper we introduce a new estimation procedure to deal with ideology and common values in the context of equilibrium behavior. The closest effort is that of [Iaryczower, Katz, and Saiegh \(2009\)](#), who model strategic voting and common values in

---

<sup>6</sup>More recently, [Degan and Merlo \(2008\)](#) and [de Paula and Merlo \(2009\)](#) consider the nonparametric identification and estimation of the ideological voting model. [Coate and Conlin \(2004\)](#), [Coate, Conlin, and Moro \(2008\)](#), and [Kawai and Watanabe \(2009\)](#) also perform structural estimation of strategic voting (ie. “pivotal voting”) models, with ideological voters.

<sup>7</sup>See [Lim \(2008\)](#) for structural estimation of a model that incorporates career concerns into judges’ behavior.

<sup>8</sup>See however [Londregan \(1999\)](#), [Clinton and Meirowitz \(2003\)](#) and [Clinton and Meirowitz \(2004\)](#), who analyze the spatial voting model without assuming sincere voting, paying attention to agendas and sequence.

<sup>9</sup>It should be emphasized that we are referring here to strategic considerations that are internal to the Court. Justices may also be strategic in response to the behavior of political actors outside of the Court (the president, Congress). Whether Justices indeed respond or not to these outside pressures is a matter of debate in the literature, captured in the ‘attitudinal’ vs ‘rational choice’ camps (see [Segal and Spaeth \(1993\)](#), [Gely and Spiller \(1990\)](#); ?). Clearly, however, for all the rationality in our model, this paper is not any more in the ‘rational choice’ than in the ‘attitudinal’ camps. Also see [Cameron, Segal, and Songer \(2000\)](#) and [Daughety and Reinganum \(2006\)](#) for strategic models of incomplete information explaining how appellate judges may influence the cases that the SC chooses to hear.

Congress.<sup>10</sup> The underlying theoretical model in that paper, however, is designed to deal with the bicameral aspect of Congress, and is otherwise less flexible than the model we consider here.

### 3 The Model

The court is composed of  $n$  justices,  $i = 1, \dots, n$ , who consider  $T$  independent cases,  $t = 1, \dots, T$ . In each case  $t$ , justice  $i$  can rule in favor or against the Defendant. We denote this ruling by  $v_i^t \in \{0, 1\}$ , with  $v_i^t = 0$  indicating a ruling in favor of the Defendant and  $v_i^t = 1$  a ruling in favor of the Plaintiff. The court aggregates the decisions of the individual justices by simple majority rule; i.e. rules in favor of the Plaintiff ( $v_t = p$ ) if  $\sum_i v_i^t \geq R_s \equiv \frac{n+1}{2}$  and in favor of the Defendant ( $v_t = d$ ) otherwise.

We consider two related models of individual behavior. In the *expressive* or sincere voting model, we assume that in deciding their vote, justices care only about their individual vote. In the *strategic* or outcome-oriented voting model, we assume instead that justices care about the ruling of the court. We assume that the goal of any justice  $i$  in any given case  $t$  is that she (in the expressive voting model) or the court (in the strategic voting model) rules according to  $i$ 's own best understanding of how the law applies to the particulars of the case.

Specifically, before ruling in each case  $t$ , each justice  $i$  observes a private signal  $s_{it} = \omega_t + \sigma_{it}\varepsilon_t$ , where  $\varepsilon_t \sim \mathcal{N}(0, 1)$ . Here  $\omega_t \in \{0, 1\}$  is an unobservable variable – for both the econometrician and the justices – indicating whether the meaning of the law favors the Plaintiff ( $\omega_t = 1$ ) or the Defendant ( $\omega_t = 0$ ), and  $\theta_{it} = 1/\sigma_{it}$  is a scale parameter that parametrizes the informativeness of  $i$ 's signals. This parameterization of the information structure satisfies the Monotone Likelihood Ratio Property (MLRP), which is important in what follows.

Justices care about this information because their payoffs are state dependent. In particular, we assume that given  $\pi_{it} \in (0, 1)$ , justice  $i$  has a payoff of  $-\pi_{it}$  when the law favors the Defendant but she/the court rules in favor of the Plaintiff ( $v_t = 1$  when  $\omega_t = 0$ ) and of  $-(1 - \pi_{it})$  when the law favors the Plaintiff but instead she/the court rules in favor of the Defendant ( $v_t = 0$  when  $\omega_t = 1$ ). The payoffs of  $v_t = \omega_t = 0$  and  $v_t = \omega_t = 1$  are normalized to zero. Thus given information  $E$ , Justice  $i$  votes to rule against the Defendant in  $t$  if and only if  $\Pr^i(\omega_t = 1|E) \geq \pi_{it}$ . Equivalently, justice  $i$  votes to rule against the Defendant in case  $t$  given  $E$  if and only if the likelihood ratio  $\Pr^i(E|\omega_t = 1)/\Pr^i(E|\omega_t = 0)$  is larger than  $\frac{\pi_{it}}{1-\pi_{it}} \frac{1-\rho_t}{\rho_t}$ , where  $\rho_t \equiv \Pr(\omega_t = 1)$  denotes justices' common prior probability of the unobserved state  $\omega_t$ . Note that since  $\omega_t$  is assumed to be unobservable, there is always information that would make any two justices disagree about a case. Moreover, if sufficiently biased, two justices can disagree almost always.

---

<sup>10</sup> In [Chiang and Knight \(2008\)](#), common values enter a non-strategic model in which voters gain information about candidates from newspaper endorsements.



In particular, with  $\pi_{it} \approx 0$  (or  $\pi_{it} \approx 1$ ), justice  $i$  is almost always ideological. On the other hand, when  $\pi_{it} = 1/2$  for all  $i$ , the setting boils down to an unbiased, pure common values model.<sup>11</sup>

The two alternative models of behavior differ in how much information each justice has *in equilibrium*. In the expressive voting model, justices care about their own ruling, and therefore vote based on their own information  $s_{it}$ , i.e., rule against the Defendant whenever  $\Pr^i(\omega_t = P|s_{it}) \geq \pi_{it}$ . Then  $E$  consists only of  $s_{it}$ , and  $i$  votes to rule against the Defendant if

$$\frac{\Pr(s_{it}|\omega_t = 1)}{\Pr(s_{it}|\omega_t = 0)} = \frac{\phi(\theta_{it}[s_{it} - 1])}{\phi(\theta_{it}s_{it})} \geq \frac{\pi_{it}}{1 - \pi_{it}} \frac{1 - \rho_t}{\rho_t} \quad (1)$$

Let  $s_{it}^{exp}$  denote the value of  $s_{it}$  that solves (1) with equality. By the MLRP the ratio  $L(s) \equiv \Pr(s|\omega_t = 1)/\Pr(s|\omega_t = 0)$  is increasing in  $s$ , so that  $i$  rules against the Defendant whenever  $s_{it} \geq s_{it}^{exp}$  and in favor of the Defendant otherwise. This cutoff point  $s_{it}^{exp}$  completely characterizes behavior in the expressive voting case. Therefore we can write the likelihood of the justices' votes in case  $t$  in the expressive voting model as

$$\Pr(v_t) \equiv \sum_{\omega_t} \Pr(\omega_t) \prod_{i=1}^n [1 - \Phi(\theta_{it}[s_{it}^{exp} - \omega_t])]^{v_{it}} \Phi(\theta_{it}[s_{it}^{exp} - \omega_t])^{1-v_{it}} \quad (2)$$

In the strategic voting model, justices care about the ruling of the court. As a result, any justice  $i$  then considers the implications of her vote assuming that she is pivotal for the decision. (This supposition is not correct when the justice is not in fact pivotal, but for the same reason these mistakes have no cost for the outcome-oriented justice.) Here, the relevant information for justice  $i$  in case  $t$  is not only her private information  $s_{it}$ , but also the equilibrium information contained in the event that  $i$  is pivotal for the court's decision, given the equilibrium strategy profile followed by the remaining justices.

Let  $\mu_{jt} : \mathbf{R} \rightarrow [0, 1]$  denote the strategy of justice  $j$ , where  $\mu_{jt}(s_{jt}) \equiv \Pr(v_{jt} = 1|s_{jt})$ . Then (1) becomes

$$\frac{P_{\mu_{-i,t}}(piv_{it}|\omega_t = 1)}{P_{\mu_{-i,t}}(piv_{it}|\omega_t = 0)} \frac{\phi(\theta_{it}[s_{it} - 1])}{\phi(\theta_{it}s_{it})} \geq \frac{\pi_{it}}{1 - \pi_{it}} \frac{1 - \rho_t}{\rho_t} \quad (3)$$

As before, the MLRP implies that the best response to any strategy  $\mu_{-i}$  of the remaining justices is a cutoff strategy, such that  $i$  rules against the Defendant ( $\mu_{it}(s_{it}) = 1$ )

---

<sup>11</sup>In our setting, justices share common priors, but their ideological biases are captured by the  $\pi_i$  parameters. See [Froeb and Kobayashi \(1996\)](#) for a model where justices' biases are manifested in their priors. We also attempted to estimate a model where priors  $\rho$ , as well as bias  $\pi$ , differed across justices  $i$  and cases  $t$ . The resulting model was not well-identified – particularly, we could not disentangle the effect of a variable on a justice's priors vs his probability of voting for the plaintiff. As a result, we could not obtain reliable estimates, and did not pursue this.

if  $s_{it}$  implies (3), and in favor of the Defendant ( $\mu_{it}(s_{it}) = 0$ ) otherwise.<sup>12</sup> This in turn implies that all responsive equilibria are cutoff equilibria; i.e., that any equilibrium is characterized by cutpoints  $s_{it}^*$  for each justice  $i = 1, \dots, n$  such that justice  $i$  votes against the Defendant if and only if  $s_{it} \geq s_{it}^*$ . Now, given cutoff strategies,  $\Pr(v_{it} = 1 | \omega_t) = \int \mu_{it}(s) \phi(\theta_{it}[s - \omega_t]) ds = [1 - \Phi(\theta_{it}[s_{it}^* - \omega_t])]$ . Therefore from (3), and letting  $\mathcal{C}_{R-1}^i$  denote the set of coalitions  $C \subset N \setminus i$  with  $R - 1$  members,  $\{s_{it}^*\}_{i=1}^n$  is given by the  $n$  equations

$$\frac{\sum_{C \in \mathcal{C}_{R-1}} \left( \prod_{j \in C} [1 - \Phi(\theta_j[s_j^* - 1])] \right) \left( \prod_{j \neq i, j \notin C} \Phi(\theta_j[s_j^* - 1]) \right) \frac{\phi(\theta_{it}[s_{it}^* - 1])}{\phi(\theta_{it}s_{it}^*)}}{\sum_{C \in \mathcal{C}_{R-1}} \left( \prod_{j \in C} [1 - \Phi(\theta_j s_j^*)] \right) \left( \prod_{j \neq i, j \notin C} \Phi(\theta_j s_j^*) \right) \frac{\phi(\theta_{it}s_{it}^*)}{\phi(\theta_{it}s_{it}^*)}} = \frac{\pi_{it}}{1 - \pi_{it}} \frac{1 - \rho_t}{\rho_t} \quad (4)$$

The cutpoints  $\{s_{it}^*\}$  completely characterize equilibrium behavior. Therefore we can write the likelihood of the justices' votes in case  $t$  in the strategic voting case as

$$\Pr(v_t) \equiv \sum_{\omega_t} \Pr(\omega_t) \prod_{i=1}^n [1 - \Phi(\theta_{it}[s_{it}^* - \omega_t])]^{v_{it}} \Phi(\theta_{it}[s_{it}^* - \omega_t])^{1-v_{it}} \quad (5)$$

The likelihood functions for the expressive and the strategic models (Eqs. 2,5) are identical, except for the cutoff points:  $s_t^{exp}$  for the expressive model, and  $s_t^*$  for the strategic model. In principle, uniqueness of equilibrium is not guaranteed here, as any set of cutpoints satisfying Eq. (4) for the strategic model (Eq. (1) for the expressive model) constitute an equilibrium. The validity of our empirical results below do not require a unique equilibrium; only that the *same equilibrium* is being played within the data sample, so that the data can be pooled, as we have done, to estimate the model parameters.<sup>13</sup>

## 4 Data

Our data derives from two sources. The first is a database of votes and case-specific information from the *Original United States Supreme Court Judicial Database* (Spaeth

<sup>12</sup>The proof of this result follows Duggan and Martinelli (2001), and is included in the online appendix for convenience.

<sup>13</sup> This is the same assumption which is required in the recent empirical literature on estimating dynamic games (eg. Bajari, Benkard, and Levin (2007), Aguirregabiria and Mira (2007)). Another set of papers have explicitly allowed for different equilibria to be manifested in different observations of the dataset, resulting in only partial identification of model parameters (eg. Ciliberto and Tamer's 2009 analysis of airline entry games). It would be challenging to implement such an approach in our voting model, because there is already one case-specific unobservable (the state  $\omega_t$ ). If in addition, the equilibrium played were also assumed to vary across cases, there would be *two* case-specific unobservables, which would raise some tricky identification issues.

(2008)). The second is a database of justice specific information comes from the *United States Supreme Court Justice Database* (Epstein, Walker, Staudt, Hendrickson, and Roberts (2008)).

The first database begins with the first term of the Warren Court (1953), continues through the Burger and Rehnquist Courts, and finishes with the 2008 term of the Roberts Court. For our purposes, it will be useful to distinguish periods in which the composition of the court remains unchanged (this is called a *natural court* in the literature). Given changes in the composition of the court, this creates a number of natural courts per chief justice. As we will explain later, we will focus on decisions in which nine justices vote.<sup>14</sup> This restricts the list of natural courts in our sample to those with nine members. The upper panel of Table 1 (in the Appendix) presents the lists of all such natural courts, together with the number of cases per issue and the percentage of decisions favoring the Plaintiff.

We distinguish between four classes of issues: *Criminal* (includes Criminal Procedure), *Basic Rights* (includes Civil Rights, First Amendment, Due Process, Privacy, and Judicial Power), *Economic* (includes Economic Activity, Unions and Attorneys) and *Federal* (includes Federalism, Interstate Relations, and Federal Taxation).<sup>15</sup> We code the type of Plaintiff and Defendant as one of three classes: *U.S. Government* (includes the U.S. Government itself, Federal Agencies and Congress), *Local Government* (state governments, local governments, boards of education, and state courts) and *Private Party* (individuals, employees, businesses, nonprofit organizations, politicians, aliens and sovereigns). We code the type of law under consideration as one of three classes: *Judicial Review* (judicial review at the national level), *Statutory Interpretation* (statutory construction at the national level), and *Others* (includes judicial review at the state level, supreme court supervision of lower federal courts, interpretation of administrative regulation or rule or executive order, interpretation of state laws, and federal common law). We also include information about whether lower courts agreed on a decision or not.

The second database provides us with information about each justice. We include their political party affiliation at time of nomination, their prior judicial experience, and the years of experience in the court at the time of the decision. We also include the Segal-Cover (SC) score of the nominees ideology and qualifications.<sup>16</sup> The lower panel of Table 1 (in the Appendix) summarizes this information for each justice in our data.

---

<sup>14</sup>We include all such cases except memorandum cases and decrees, as well as those in which the court has original jurisdiction. To avoid repetitions we set  $anal = 0$  (this conforms to standard practice).

<sup>15</sup>A fifth residual category groups Miscellaneous cases (Spaeth's issues 980-99).

<sup>16</sup>These scores were derived through content analyses of newspaper editorials written between the date of the Presidents nomination and the date of the Senates final action over the nomination (see Segal and Cover (1989), Segal, Cameron, and Cover (1990, 1992), and Epstein, Lindstadt, Segal, and Westerland (2006)).

## 5 Estimation: description and identification

In this section, we describe the estimation and identification of our model. Clearly, identification of our voting model from vote data alone is challenging: as Londregan (1999), among others, have noted, from binary data on votes it is difficult to recover estimates of the continuous preference distributions of voters, without additional modeling and/or parametric restrictions. In our model, these additional restrictions come in the form of the parametric restrictions on voters’ preferences and the information structure, detailed in Section 3 above. Here, we present an argument as to the identification of these model parameters from the observed vote data. We start with a more intuitive discussion of identification, by relating the structural parameters to be estimated to observed quantities in the data. We then proceed to a more formal statistical discussion of identification.

At an intuitive level, the key for identification is that the common value induces a correlation of votes in equilibrium: all justices tend to receive larger signals when the law favors the Plaintiff, and all justices tend to receive smaller signals when the law favors the Defendant. Suppose first that cases are homogeneous, so that justices’ types and prior are invariant across cases. If justices’ quality of information were large relative to their bias, and the prior relatively uninformative (say  $\pi_i \approx 1/2$  for all  $i$  and  $\rho \approx 1/2$ ), the court would “flip-flop” evenly between unanimous pro-defendant and pro-plaintiff decisions. Now suppose that instead  $\rho \approx 2/3$ . Since in this case the law favors the Plaintiff more frequently, justices will tend to receive large signals more frequently (moreover, to compensate for the larger prior, justices will also use strategies that are more favorable for the Plaintiff). As a result, the majority of the court would rule for the Plaintiff more often than before. Therefore, the frequency in which the majority decision favors the Plaintiff tracks  $\rho$ : a larger frequency corresponds to a larger estimated prior  $\rho$ .<sup>17</sup> Suppose next that we change the bias of one justice  $i$  in our previous example so that her bias is large relative to the quality of her information. Then while all other justices will alternate between sometimes finding for the Plaintiff and sometimes for the Defendant,  $i$  will stay put in one decision. This illustrates that low variability in individual decisions corresponds to large bias. Finally, return to the previous example in which all justices are moderate. As we pointed out before, if the quality of information is sufficiently high for all justices, then we would expect these to be unanimous votes. But as the quality of information of some justices is lower, these justices would disagree with the majority more often. Thus justices with variable voting records who tend to be in the minority are associated with a low quality of information.

Now, as it is, this identification scheme may seem unreasonable, because it appears to penalize “maverick” justices who go against the grain by assigning them a low precision parameter. For this reason, in the empirical work, we control for many case-specific covariates, and take into account inherent differences among justices due to ideology, judicial experience, etc. Therefore, justices with low  $\theta$ ’s are those who have attributes that characterize justices who vote erratically, *even after* taking characteristics of the case into account: these are not maverick justices, but rather *inconsistent* ones.

---

<sup>17</sup>Below, we discuss in more detail how to disentangle priors from case selection.

**A more formal argument.** Our formal identification argument has two parts, which will later be mimicked for estimation: first, we show that the justices’ priors and their “reduced-form” voting probabilities are identified from the vote data; second, we show that the parameters of voters’ preferences and the information structure are identified from the reduced-form vote probabilities. For simplicity, we will assume here that all cases are homogeneous, in the sense that all the parameters of the model,  $\{(\theta_i, \pi_i)\}_{i=1}^n$ , as well as  $\rho$ , are assumed to be identical across all cases. However, these identification arguments continue to hold if all the parameters, as well as  $Pr(v_t)$ , depend on characteristics of case  $t$ , which we denote  $X_t$ . In our empirical work below, we will control for case heterogeneity using a rich set of covariates, so that the justice-specific parameters  $\theta_{it}$  and  $\pi_{it}$  vary across both justices  $i$  and cases  $t$ .

**First step.** We introduce the following notation:

$$\begin{array}{ll} \text{Priors:} & \rho \equiv \Pr(\omega_t = 1) \\ & 1 - \rho = \Pr(\omega_t = 0) \end{array} \quad \begin{array}{ll} \text{Voting Probs.:} & \gamma_{i,1} \equiv \Pr(v_{it} = 1 | \omega_t = 1) \\ & \gamma_{i,0} \equiv \Pr(v_{it} = 1 | \omega_t = 0) \end{array}$$

Given this notation, the first-step estimation problem (in both the expressive as well as the strategic case) is to maximize the following reduced-form likelihood function for the votes:

$$\begin{aligned} \max_{\{\gamma_{i,1}, \gamma_{i,0}\}_{i=1}^n, \rho} \Pr(v_t) &= \rho \prod_{i=1}^n [\gamma_{i,1}^{v_{it}} (1 - \gamma_{i,1})^{1-v_{it}}] + (1 - \rho) \prod_{i=1}^n [\gamma_{i,0}^{v_{it}} (1 - \gamma_{i,0})^{1-v_{it}}] \\ \text{s.t. } &\gamma_{i,1} \geq \gamma_{i,0}, \quad \forall i. \end{aligned} \quad (6)$$

Conditional on the state  $\omega_t$ , the individual votes  $v_{it}$  are independent across the justices  $i$ . Thus, the vector of votes  $v_t$  follows a multivariate mixture distribution, with mixing probability  $\rho$ . Identification of the state-specific voting probabilities  $\{\gamma_{i,1}, \gamma_{i,0}\}_{i=1}^n$  and the mixing probability  $\rho$  are available in, eg., [Hall and Zhou \(2003\)](#), [Hu \(2008\)](#), and [Kasahara and Shimotsu \(2007\)](#).

Intuitively, the unconditional correlation among justices’ votes is crucial to identification. If there were only one decision maker, for example, it would not be possible to disentangle the independent effects of ideology and information.<sup>18</sup> Moreover, the inequality  $\gamma_{i,1} > \gamma_{i,0}$ , which is implied by the monotone likelihood ratio property, is also important: without this assumption, the voting probabilities would only be identified up to an arbitrary classification of  $\omega_t$ . This inequality resolves this classification problem by setting  $\gamma_{i,1}$  ( $\gamma_{i,0}$ ) equal to the maximum (minimum) of the two identified voting probabilities.

---

<sup>18</sup>Specifically, for the  $n = 9$  justices on the US Supreme Court, the vote vector  $v_t$  can take  $2^9$  values; with a large enough dataset, it is possible to estimate the probability that  $v_t$  takes each of these values by the empirical frequency. On the other hand, there are only 19 parameters (18 vote probabilities, and  $\rho$ ) to estimate. Since  $2^9 \gg 19$ , the relevant necessary condition for identification is satisfied.

The above-cited papers contain constructive identification proofs, which can be directly mimicked for estimation. For our purposes, we found it more convenient to maximize the likelihood function (6) directly. This constituted the first step of our estimation procedure.

**Second step.** Using the estimates of the two justice-specific vote probabilities  $\hat{\gamma}_{i,1}$  and  $\hat{\gamma}_{i,0}$ , from the first step, we recover the two structural parameters,  $\pi_i$  and  $\theta_i$ , for each justice  $i$ . Recall our earlier assumptions that justice  $i$ 's private information is  $s_{it} = \omega_t + \frac{1}{\theta_i} \varepsilon_{it}$ , with  $\varepsilon_{it} \sim \mathcal{N}(0, 1)$ . Then  $\gamma_{i,1} \equiv 1 - \Phi(\theta_i[s_i^* - 1])$  and  $\gamma_{i,0} \equiv (1 - \Phi(\theta_i s_i^*))$ . Solving these equations for  $\theta_i$  and  $s_i^*$  given  $\hat{\gamma}_{i,1}$  and  $\hat{\gamma}_{i,0}$  (and substituting  $\Phi^{-1}(\gamma_{i,1}) = -\Phi^{-1}(1 - \gamma_{i,1})$ ) gives <sup>19</sup>

$$\hat{\theta}_i = \Phi^{-1}(1 - \hat{\gamma}_{i,0}) - \Phi^{-1}(1 - \hat{\gamma}_{i,1}); \quad \hat{s}_i = \frac{\Phi^{-1}(1 - \hat{\gamma}_{i,0})}{\Phi^{-1}(1 - \hat{\gamma}_{i,0}) + \Phi^{-1}(\hat{\gamma}_{i,1})} \quad (7)$$

Note that the estimate of  $\hat{\theta}_i$ , the precision of  $i$ 's information, is given by the difference between the conditional probabilities of voting in favor of the Plaintiff when the law favors the Plaintiff ( $\omega = 1$ ) and when the law favors the Defendant ( $\omega = 0$ ). This implies that precision is increasing in the probability of correctly ruling in favor of the Plaintiff ( $\gamma_{i,1}$ ), and decreasing in  $\gamma_{i,0}$ , which is the probability of incorrectly ruling against the defendant. This is very intuitive in light of the theoretical model.

The estimate of the equilibrium cutpoint, instead, is a decreasing function of the ratio between  $\Phi^{-1}(\hat{\gamma}_1)$  and  $\Phi^{-1}(1 - \hat{\gamma}_0)$ . Thus  $\hat{s}$  is decreasing in the ratio of the probability of voting correctly in favor of the Plaintiff ( $\gamma_1$ ) *relative* to the probability of correctly voting in favor of the Defendant ( $1 - \gamma_0$ ). When this ratio is large, for instance – indicating a bias towards the plaintiff – the cutpoint  $\hat{s}$  will be small, implying that the justice requires a low informational threshold to vote in favor of the plaintiff.

In order to recover the bias parameter  $\pi_i$ , we use the equilibrium voting condition, which differs between the expressive and strategic models. In the case of the expressive voting model, this is given by

$$\frac{\phi(\theta_i[\hat{s}_i - 1])}{\phi(\theta_i \hat{s}_i)} = \frac{\hat{\pi}_i^{exp}}{1 - \hat{\pi}_i^{exp}} \frac{1 - \hat{\rho}}{\hat{\rho}}, \quad (8)$$

while in the strategic voting model (with majority rule  $R$ ) this is given by

$$\left[ \frac{1 - \Phi(\theta_i[\hat{s}_i - 1])}{1 - \Phi(\theta_i \hat{s}_i)} \right]^{R-1} \left[ \frac{\Phi(\theta_i[\hat{s}_i - 1])}{\Phi(\theta_i \hat{s}_i)} \right]^{n-R} \frac{\phi(\theta_i[\hat{s}_i - 1])}{\phi(\theta_i \hat{s}_i)} = \frac{\hat{\pi}_i^*}{1 - \hat{\pi}_i^*} \frac{1 - \hat{\rho}}{\hat{\rho}} \quad (9)$$

---

<sup>19</sup>Note that for each justice, we use the estimates of  $\gamma_{i,0}, \gamma_{i,1}$  to recover the two quantities  $\theta_i$  and  $s_i$ . For this reason, we consider a one-parameter specification of the information structure; with additional parameters, we might not have identification.

For both models, plugging in our estimates of  $\theta_i$  and  $\hat{s}_i$  into the appropriate equilibrium condition allows us to recover estimates of  $\hat{\pi}_i^{exp}$  and  $\hat{\pi}_i^*$  for the expressive and strategic models, respectively.

Note that, in recovering  $\theta_i$ , it was not necessary to specify whether justices vote expressively or strategically. An assumption regarding strategic or expressive voting is required only for recovering  $\pi_i$ . This distinction between  $\theta_i$  and  $\pi_i$  is a remarkable property of this problem. It implies that the precision estimate is independent of whether justices care about the court ruling or about their own vote being correct, and therefore of whether justices use the information contained in the event of them being pivotal or simply best respond to their own private information.

## 5.1 Estimation: Accommodating case and justice heterogeneity

While our foregoing discussion of identification assumed that all cases are homogeneous, this was mainly for convenience, and our empirical model accommodates case-level heterogeneity by allowing the reduced-form parameters of the model, which are recovered in the first step of the estimation procedure, to depend quite flexibly on observable characteristics  $X_t$ . Specifically, we parameterize justices' priors in case  $t$ ,  $\rho_t \equiv \Pr(\omega_t = 1)$ , as a logit probability which depends on the characteristics  $X_t$ :

$$\rho(X_t; \beta) \equiv \frac{\exp(X_t' \beta)}{1 + \exp(X_t' \beta)}, \quad \in [0, 1].$$

Once the prior probability  $\rho_t$  varies across cases, so will the equilibrium strategies  $s_{it}^*$ , and hence so will the justice-specific conditional probabilities of ruling against the Defendant  $\gamma_{it,1}$  and  $\gamma_{it,0}$ . Because of this, the model becomes more difficult. One possibility is to undertake “direct” estimation where, for each value of the parameters, we need to solve for the equilibrium cutpoints using equations (4) for each justice and each case. Obviously, this is computationally quite cumbersome. Therefore, we propose a procedure that builds on our previous methodology. Since we know that the heterogeneous cases will cause the justice-specific probabilities of ruling in favor of the Plaintiff to vary across cases, we parameterize these in the following way, which also restricts  $\gamma_{i,t,1} \geq \gamma_{i,t,0}$ , for all  $X_t$ :

$$\begin{aligned} \gamma_{i,0}(\zeta, \eta) &= \frac{\exp(Z_i' \zeta + X_t' \eta)}{1 + \exp(Z_i' \zeta + X_t' \eta)}, \quad \in [0, 1]; \\ \gamma_{i,1}(\zeta, \eta, \alpha, \delta) &= \frac{\gamma_{i,0} + \exp(Z_i' \alpha + X_t' \delta)}{1 + \exp(Z_i' \alpha + X_t' \delta)}, \quad \in [\gamma_{i,0}(\zeta, \eta), 1]. \end{aligned} \tag{10}$$

In the first stage, we estimate the parameters  $(\beta, \delta, \eta)$  as well as the justice-specific variables  $(\alpha_i, \zeta_i)$  for  $i = 1, \dots, n$ . For this, we maximize the following likelihood function

$$\begin{aligned} \max_{\alpha, \beta, \zeta, \eta, \delta} \sum_t \log & \left[ \rho(X_t; \beta) \cdot \prod_{i=1}^n \{ \gamma_{i,1}(\zeta, \eta, \alpha, \delta)^{v_{it}} (1 - \gamma_{i,1}(\zeta, \eta, \alpha, \delta))^{1-v_{it}} \} \right. \\ & \left. + (1 - \rho(X_t; \beta)) \cdot \prod_{i=1}^n \{ \gamma_{i,0}(\zeta, \eta)^{v_{it}} (1 - \gamma_{i,0}(\zeta, \eta))^{1-v_{it}} \} \right]. \end{aligned} \tag{11}$$

For the second stage, we use the predicted values of  $\gamma_{i,t,1}$  and  $\gamma_{i,t,0}$  to recover case and justice specific values of  $\theta_{it}$  and  $s_{it}^*$ , using the equations in (7). We can then compute the bias estimates solving the  $n$  equations (4) for the strategic voting model, or (1) (with equality) for the expressive voting model. Note that, when the voting probabilities  $\gamma_{i,0}$  and  $\gamma_{i,1}$  are case-specific and depend on the covariates  $X$  and  $Z$ , then so will the model parameters  $\theta_{it}$  and  $\pi_{it}$ .

**Disentangling Priors and Endogenous Case Selection.** Up to now, we have implicitly assumed that all the cases heard by the Court are exogenously chosen; i.e., we have not explicitly modeled an agenda-setting stage. However, it is well-known that case selection can be endogenous, both because the Supreme Court must decide (via a vote) whether or not to “grant cert” (that is, to hear) a case that has been brought to its attention, and also because petitioners and lower courts may selectively recommend cases to the Supreme Court for which, given the ideological leanings of the justices, the plaintiff has a high probability of winning.<sup>20</sup>

In our empirical model, this endogenous case selection is not explicitly accommodated, and will thus be captured in the parameter  $\rho$  describing justices’ common prior beliefs about the “right” judgment in the cases. To see this intuitively, consider the likelihood problem (11). Note that – for given covariates  $X_t$  – the parameter estimates  $\beta$  in  $\rho(X_t, \beta)$  should be set so that  $\rho(X_t, \beta)$  is high (resp. low) when justices vote more often in favor of (resp. against) the plaintiff.

This suggests that in general it will be difficult to distinguish a shift in justices’ prior beliefs (about randomly assigned cases) from case selection, because both will lead, all else equal, to a higher probability of voting in favor of the plaintiff. This difficulty in disentangling beliefs and case selection implies that the estimates of  $\rho|X_t$  should not be taken out of the context of those courts for which it was computed. On the other hand, letting  $\rho$  vary in response to the voting data allows us to “control” for case selection. To capture this we include in  $X_t$  both variables that capture justices’ priors and case selection. In particular, to capture agenda-setting considerations we include chief justice dummies in  $X_t$ . Later, we will gauge the importance of agenda-setting by examining the coefficients on these variables.

## 6 Results

In this section, we describe our results for the heterogeneous model described above. As before, we restrict attention to cases in which all nine justices voted.<sup>21</sup> The covari-

<sup>20</sup>See Cameron, Segal, and Songer (2000) and Daughety and Reinganum (2006) for more information on the Supreme Court case selection process.

<sup>21</sup>Note that the equilibrium cutpoint of each justice will be different for each different composition of the *voting members* of the court, implying different conditional probabilities of ruling in favor of the Plaintiff in each state for each configuration of *voting members*, even fixing the covariates  $X_t$ . Including



ates are those that were described in Section 4. As case-specific covariates, we include characteristics of the Plaintiff and the Defendant (whether Plaintiff and Defendant are a Local Government, the Federal Government, or private parties), the authority for decision (whether this is a case that involves the interpretation of a Federal Law, a challenge that a Federal Law is unconstitutional, or others), and the disposition of the case by lower courts (whether the lower courts agree or not). To further control for endogenous case selection, we also include the identity of the chief justice at the time of consideration of the case (Warren, Burger, Rehnquist, or Roberts). In order to allow maximal flexibility in the order of justices' bias along different issues, we estimate the model separately for cases involving Basic Rights, Economic, Criminal, and Federal issues.<sup>22</sup> As justice-specific covariates we include the number of years of prior judicial experience, the political party of the President that nominated the justice (Democratic or Republican Nominee), and the Segal-Cover measures of ideology and quality. We also include three variables that vary per case and justice. These are each justice's years of experience at the Court at the time of the ruling, and, for each justice  $i$ , the average Segal-Cover scores of justices other than  $i$  sitting in the Court that ruled in the case.

Table 2 presents the "first stage" MLE estimates of the coefficients of the common prior function  $\rho(X_t)$ , and of the state-contingent probabilities of ruling against the Defendant  $\gamma_0(X_t, Z_{it})$  and  $\gamma_1(X_t, Z_{it})$ . Note that for all issues other than Federal (for which the small sample size leads to uniformly larger standard errors), all the coefficients of the case-specific and justice-specific variables are statistically significant in the specification of at least one of our first-stage parameters. Either the justices' common prior that the Plaintiff should win the case, or the individual probabilities of ruling correctly and incorrectly in favor of the Plaintiff are significantly different depending on whether the Plaintiff or the Defendant are themselves the Federal Government, a Local Government, or a private party, on whether previous courts agreed on a ruling or not, etc.

The coefficients on the average SC quality and ideology measures for the other justices (shaded in Table 2), merit additional discussion. Under simple expressive voting, a justice's vote is not affected by her colleagues, so that the coefficients on these covariates should be zero. Under strategic voting, however, the justices' votes are interdependent, and these coefficients should be significantly nonzero. Including the covariates for the average Segal-Cover scores for the other justices therefore allows us to informally test the strategic vs. the expressive voting model. We see that for the basic rights and criminal cases, these variables are significant, but not in the other cases. This suggests that for the two largest subsets of the cases, the strategic voting model is appropriate.

[Table 2 about here]

---

only the votes in which all justices vote therefore dramatically reduces the number of parameters to be estimated. This still leaves a significant number of cases in the sample (see Table 1).

<sup>22</sup>The results of carrying out our estimation pooling all votes and introducing "issue" as an additional covariate are otherwise similar to the issue-by-issue estimation (results are available from the authors upon request).

## 6.1 The Value of Information in the Court

Given the first stage coefficients we can compute, for any case  $t$  with characteristics  $X_t$ , the common prior  $\rho_t = \rho(X_t)$ , as well as the conditional probabilities  $\gamma_{i,t,0} = \gamma_0(X_t, Z_{it})$  and  $\gamma_{i,t,1} = \gamma_1(X_t, Z_{it})$  that a justice with characteristics  $Z_{it}$  in case  $t$  rules against the Defendant in each state of nature. We can then use the predicted values of  $\gamma_{i,t,1}$  and  $\gamma_{i,t,0}$  to recover case and justice specific values of  $s_{it}^*$ , and the “deep parameters”  $\theta_{it}$  and  $\pi_{it}$  (for both the strategic and the expressive voting models). In particular, we can do this for cases with characteristics  $X_t = x$  and the actual justices and courts observed in the realized history. To describe the main results we will focus for the most part on cases of statutory interpretation in which both the Plaintiff and the Defendant are private parties, and in which lower courts have agreed on a ruling. We then consider comparative statics from this initial exercise.

To reinforce the logic of the model, we begin by presenting the complete set of estimates for a single court. Table 3 presents the estimates for each issue for the longest-lasting natural court in our data: REHN7, with chief justice Rehnquist, between 1994 and 2004. For each parameter, standard errors (computed using the Delta method) are given in parentheses.

[Table 3 about here]

For each issue, we indicate the MLE estimate of the common prior probability that the law favors the Plaintiff. Thus in Criminal Procedure the prior is lower than in all other issues, and moreover favors the Defendant ( $\rho = 0.333$ ).<sup>23</sup> In all other issues the prior favors the Plaintiff, and is most favorable to the Plaintiff in Economics ( $\rho = 0.610$ ).

The first two columns present the MLE estimates of the probability that justice  $i$  rules in favor of the Plaintiff when the law favors the Defendant ( $\gamma_{it0}$ ) and when the law favors the Plaintiff ( $\gamma_{it1}$ ). Thus, taking Economics for example, justice Breyer had a probability of  $\gamma_{it1} = 0.94$  of correctly ruling for the Plaintiff, and a probability of  $1 - \gamma_{it0} = 1 - 0.14 = 0.86$  of correctly ruling in favor of the Defendant. Column 3 presents the estimate of the informativeness or precision of each justice’s signal. As we pointed out earlier, this is an increasing function of the difference between the probability that justice  $i$  rules in favor of the Plaintiff when the law favors the Plaintiff and when the law favors the Defendant. The larger precision for Breyer relative to Stevens in Economics, for example, reflects both a higher probability of ruling for the Plaintiff when the law favors the Plaintiff (0.94 vs 0.82) and a smaller probability of ruling for the Plaintiff when the law favors the Defendant (0.14 vs 0.16).

---

<sup>23</sup>However it should be noted that most frequently, Criminal Procedure cases have the Government as Plaintiff or Defendant (as opposed to here, where we consider both Plaintiff and Defendant to be private parties). When we condition for the US Government as Plaintiff, the prior belief that the Plaintiff is right increases to  $\rho = 0.915$  for the REHN7 court. We return to this later in the paper.

The fourth column presents the equilibrium cutpoint. Thus, taking again Economics for example, Justice Breyer would vote for the Plaintiff after observing a signal  $s_i \geq 0.42$ , but it would take more evidence (a signal above 0.48) for Justice Souter to rule in favor of the Plaintiff, and even more for Justice Stevens to vote in the same way (a signal above 0.52). This results from three factors. The first is the precision of private information. According to the estimates, Breyer has a more precise signal than Souter, who in turn has a more precise signal than Stevens. Thus the same evidence has different value for different justices. The second factor is the common prior  $\rho$ , which as we mentioned in this case “stacks the deck” in favor of the plaintiff. The third is the bias of the justice in question (and in the strategic voting model, also of the remaining justices in the court, through their equilibrium strategy  $s_{-i}^*$ ). The justices’ bias are shown in columns 5 and 6 in the table. Here Justice Breyer is more moderate ( $\pi_{BRE}^{exp} = 0.47$ ) than Justice Souter ( $\pi_{SOU}^{exp} = 0.58$ ). Justice Stevens requires more evidence (a belief of at least 0.63 that the law favors the Plaintiff) to rule in favor of the Plaintiff in this class of cases.

Given these estimates, we can compute our measure of the value of information in the court, FLEX. This is the probability that justice  $i$  votes differently than what she would have voted for in the absence of her private case information. To compute this, we first calculate how each justice would have voted with no private information. From (1), this is simply  $v_{it} = 1$  if  $\rho \geq \pi_i$  and  $v_{it} = 0$  otherwise (vote for the Plaintiff if the public information, as summarized by  $\rho$ , outweighs the private bias  $\pi_i$ ). Then we compare this initial leaning to the probability of voting differently after observing her private information; i.e., FLEX measures the probability that a justice would “change her mind” after observing her private information (vote for the Defendant even when  $\rho \geq \pi_i$ , or for the Plaintiff even when  $\rho < \pi_i$ ):

$$FLEX_i = \begin{cases} \rho\Phi(\theta_i[s_i^* - 1]) + (1 - \rho)\Phi(\theta_i s_i^*) & \text{if } \rho \geq \pi_i \\ \rho[1 - \Phi(\theta_i[s_i^* - 1])] + (1 - \rho)[1 - \Phi(\theta_i s_i^*)] & \text{if } \rho < \pi_i. \end{cases} \quad (12)$$

Note that FLEX is bounded between zero and one, and takes a value of zero for individuals with extremely large biases either for the Plaintiff ( $\pi \rightarrow 0$ ) or for the Defendant ( $\pi \rightarrow 1$ ). Note also that the computation of FLEX for the expressive and strategic voting models differ only in whether we use  $\pi_i^{exp}$  or  $\pi_i^*$  to evaluate whether  $\rho \geq \pi_i$  or  $\rho \leq \pi_i$ . The reason for this is that the equilibrium cutpoint  $s_i^*$  that is recovered from the data is not determined by whether we use the expressive or strategic voting models. Together with the data, the two models imply the same  $s_i^*$  and  $\theta_i$ , and differ only in the biases  $\pi_i$  that rationalize these quantities. As a result, in practical terms this means that the expressive and strategic FLEX scores for any given justice and any given realization of the covariates  $X_t$  are very often identical.<sup>24</sup>

Table 4 contains the average FLEX scores for each justice across different courts and issues. From a casual perusal of the FLEX scores, we see that justices tend to follow their initial leanings: FLEX scores are typically below 1/2, and thus, more often than

---

<sup>24</sup>If instead we were initially given values of  $\{\pi_i, \theta_i\}$  and  $\rho$ , then the two models would imply a different equilibrium cutpoint  $s_i^*$ , and FLEX scores in the two models would differ significantly.

not, justices' votes reflect their initial consideration of the case, based on bias and prior alone. Having said this, FLEX scores are relatively large: on average, the probability of voting differently than what they would have voted for in the absence of case information is about 44%.

[Table 4 about here]

Average FLEX scores display a significant variation across justices. The top five average FLEX scores in our sample are between 48% and 53% and the bottom five average FLEX scores in our sample are between 36% and 39%. Most justices among those with the lowest average FLEX scores are (or were) strong representatives of a liberal judicial philosophy (clearly Ginsburg, Goldberg and Breyer; also arguably Minton, who during his political career defended New Deal legislation with the statement "You cannot eat the Constitution"). Justices with the highest average FLEX scores are not necessarily judicially conservative, but have other noteworthy characteristics. Two of them – Warren and Rehnquist – were chief justices (Chief Justice Burger is also among the highest FLEX scores; Chief Justice Roberts, on the other hand, is in the bottom half of the distribution). A third, Justice Hugo Black, was a very influential advocate of textualism (a legal philosophy restricting heavily the flexibility in the interpretation of the law). A fourth justice in this top class is Justice Whittaker. Charles Whitaker served for a relatively short period of time (from 1957 to 1962) and was considered a "swing vote" on a closely divided Supreme Court. This is also the case of Justice O'Connor, who also has a relatively high average FLEX score (the highest among justices appointed after 1987).

Table 4 also illustrates a substantial variation of FLEX scores between issue areas. This is consistent with the literature and a cursory glance at the raw data, both of which suggest that we should expect the bias and quality of information to vary greatly across different issues. Most notably, the typical FLEX score in Basic Rights cases (38%) is substantially lower than in all other areas, which is consistent with the basic notion that the value of information in the court is lower in areas in which ideological considerations tend to weigh more heavily – such as Basic Rights.<sup>25</sup>

**Politicization of Nominations and the Value of Information in the Court.** We mentioned above that three of the four chief justices in our sample (Warren, Burger and Rehnquist) have relatively high FLEX scores (all three of them are in the top 7 scores).

---

<sup>25</sup>The table shows a relatively large average bias in favor of the plaintiff in Criminal cases. It should be noted, though, that most criminal cases do not have a private Plaintiff pitted against a private Defendant as we are maintaining here for consistency of the comparison. Instead most cases involve either the Federal or a Local Government facing a private party. As we will show below, the distinction turns out to be important, since having the Federal government as Plaintiff increases the bias in criminal cases by around 0.7 (see Figure 3). Note that even with this clarification, Criminal cases show a larger overall bias for one of the sides in the dispute.

The exception is Chief Justice Roberts, who was appointed in 2005. Roberts’ relatively low score obtains even if he has a relatively high quality of information in most issues. Is this a reflection of a wider change in the type of justice appointed to the Supreme Court?

According to many observers, the defeat of Robert Bork’s nomination to the Supreme Court in 1987 marked a significant change in how nominations are considered in the Senate. Epstein, Lindstadt, Segal, and Westerland (2006) summarize this noting that “[a] near-universal consensus exists that the nomination of Robert Bork in 1987 triggered a new regime in the Senate’s voting over presidential nominees – a regime that deemphasizes ethics, competence, and integrity, and stresses instead politics, philosophy and ideology.” Their own analysis of the confirmation decisions of US senators confirms this conventional wisdom. They argue, however, that while the importance of ideology has reached new heights since Bork’s nomination, the Senate’s emphasis on this factor began earlier, in the 1950’s.

Figure 1 allows us to complement this analysis by illustrating the variation in the bias, precision of information, and FLEX scores of Supreme Court justices throughout time.

[Figure 1 about here]

The results in Figure 1 partially confirm these arguments. For economics and federalism cases, FLEX scores have, indeed, fallen about 40% from the Warren to the Roberts courts. For the other two types of cases (basic rights and criminal), however, we see a marked decrease in FLEX only during the Rehnquist and Roberts courts: this decrease was about 16% in criminal cases, and around 20% in basic rights cases. Thus, our results suggest that the increased politicization of the Supreme Court appointment process has become uniformly marked only in the last quarter-century, following the failed nomination of justice Bork (i.e., within the Rehnquist and Roberts courts). Indeed, Justice Stevens, who has served since the Ford presidency, is seen by some as the last link to an era “before the Reagan years, when confirmations became contested territory in the culture wars.”<sup>26</sup> Moreover, for basic rights and criminal cases, FLEX during the current Roberts court is at the same level as in the Warren Court, so that for these two types of cases, the recent decrease in FLEX reversed a trend towards greater value of information in the court which began in the Burger courts.

**Separation of Powers.** In the presentation of results so far, we fixed both Plaintiffs and Defendants to be private parties. Here we focus on whether having the US Government or a Local Government as a Plaintiff affects the value of information in the court, and if so, how. We then also consider here whether cases involving the constitutionality of government acts have important differences from cases of interpretation of laws enacted by Congress.

---

<sup>26</sup>See Jeffrey Toobin, “After Stevens”, in *New Yorker*, cover article (03/22/2010).

Having the Government as Plaintiff changes the results in two ways. The first is by changing justices' common prior beliefs: across all issues, justices attach a higher probability to the government being right than they do a private party. The difference is substantial in Criminal Procedure (0.54) and Basic Rights (0.21), and more modest in Economics and Federalism (0.03).<sup>27</sup> Recall however that not only the prior, but also the bias and quality of information are functions of justice-specific and case-specific characteristics. As a result, a change as the one we are considering here, in the identity of the Plaintiff, also leads to a change in the bias and precision of each justice. For example, within Basic Rights cases (all) justices require more evidence to rule in favor of the Government than what they would require in the case of a private party.

In Figure 2 we generalize this comparison. In order to summarize this information efficiently, we proceed as follows. We first compute, for each justice and issue class, the differential bias and FLEX scores when the Plaintiff is the US Government (or a Local Government) and when the Plaintiff is a private party. We then report in the top two panels of Figure 2 the median and 25-75 percentiles of the distribution of these changes at the individual justice level.

[Figure 2 about here]

The figure shows that in some issue areas, the change in type of Plaintiff has a large effect on justices' preferences. In particular, changing the Plaintiff from a private party to the US (local) Government leads to a median increase of 0.29 (0.41) in  $\pi^{exp}$  in Basic Rights, and to a median increase of 0.72 (0.73) in  $\pi^{exp}$  in Criminal cases. Now recall that our measure of ideological bias ( $\pi_i$ ) is the parameter in justices' preferences that quantifies the relative cost of ruling incorrectly against the defendant ( $\pi_i$ ) vis a vis ruling incorrectly against the plaintiff ( $1 - \pi_i$ ). Moreover  $\pi_i$  also has a second direct interpretation: it is the cutpoint such that justice  $i$  will rule in favor of the plaintiff whenever the probability that the law favors the plaintiff is above  $\pi_i$ . Thus justices appear to have more stringent requirements for the Government, in particular within Criminal Procedure.

In the lower panels of Figure 2, we focus on whether judicial review cases have important differences from cases of statutory interpretation. It is important to keep in mind here that most cases of judicial review are in the Basic Rights and Criminal issues. The change appears to have no effect on either bias or FLEX scores within Criminal Procedure. Within Basic Rights cases, however, justices do appear to be more open-minded in cases that merit judicial review.

---

<sup>27</sup>This is also true for Local Government In Basic Rights (0.16), Economics (0.17), and Criminal Procedure (0.53). In Federalism, however, having a Local Government as plaintiff reduces  $\rho$  by 0.15.

## 6.2 Further results and specification checks

### 6.2.1 Agenda Setting

As we mentioned before, the parameter  $\rho$  describing justices' common prior beliefs will capture both justices' prior beliefs about randomly assigned cases and changes due to endogenous case selection. With this in mind, we included the identify of the chief justice as an additional covariate, thus capturing one important component of agenda-setting: the chief justice's influence on the cases that are taken up by the Supreme Court.

[Figure 3 about here]

Figure 3 shows the difference between the prior  $\rho$  in each issue area in the Burger, Rehnquist and Roberts' courts and the Warren court. If there were no case selection, we would expect these differences to be zero. This is not the case. Table 2 shows that the coefficients for the chief justice dummies are significant. Figure 3 shows that this agenda-setting effect is not negligible, in particular within Economics, where it ranges from a 7% difference (in the Burger court) to a 14% difference (in the Roberts court).

### 6.2.2 Communication

In our formal model, as well as in our empirical exercise, we assumed that justices' private signals are independent conditional on the state. This rules out the possibility of transmission of information through communication prior to the vote. In the Court, however, justices have many opportunities to exchange information about each case. The question here is whether justices, as strategic agents who care about the collective decision, will use pre-vote deliberation to communicate information to their peers, or whether they will use these arguments to try to influence their opinion, possibly not revealing some information that can be harmful to their case, or exaggerating evidence one way or the other. While the incentive to do so is small when interests are well aligned (Coughlan (2000)), this is not the case when there is (interim) disagreement, as in the setting consider here. This makes truthful revelation of information more difficult, as is illustrated in the analysis of Austen-Smith and Feddersen (2005, 2006) (see also Li, Rosen, and Suen (2001) and Doraszelski, Gerardi, and Squintani (2003)). Here we take an empirical perspective on the issue. To consider the possibility of informative pre-vote deliberation in the U.S. Supreme Court, we contrast voting records with the empirical implications of the model assuming that justices do in fact share their private information with one another.

A first, informal test, is to look at unanimous votes: if pooling of private information is strong enough to overwhelm private biases, all justices should vote in the same way, and

thus all (or most) votes should be unanimous.<sup>28</sup> This is not, however, what we observe in the data: more than 72% of the rulings in Criminal issues are decided by non-unanimous votes. Similarly, 60% of the rulings in Basic Rights, 58% of the rulings in Economics, and 56% of the rulings in Federal issues are decided by non-unanimous votes. Moreover, the proportion of cases decided by a small majority (say either a 5-4 vote or a 6-3 vote) is 47% in Criminal issues, 37% in Basic Rights, 34% in Economics, and 28% in Federalism. Thus, while a non trivial number of cases are decided unanimously, this occurrence is far from prevalent.

Using unanimous votes only to gauge the effect of pre-vote deliberation on outcomes is not entirely satisfactory, as it doesn't allow the possibility that after all the information is transmitted, some justices still disagree. We consider this in turn. Suppose that, contrary to our benchmark model, justices' pool their private information through deliberation. Formally, this is equivalent to a model where all justices observe the same public signal prior to their vote. This model has testable implications: in particular, if in a case  $t$  with characteristics  $X_t$  justice  $i$  votes for the Plaintiff and justice  $j \neq i$  votes for the Defendant, this reveals that, given  $X_t$ , justice  $i$  is more biased than justice  $j$  towards the plaintiff. Now consider a second case  $t'$  with the same characteristics,  $X_{t'} = X_t$ , so that the biases of justices  $i$  and  $j$  are the same in case  $t'$  as in case  $t$ . These two justices (i) can vote as they did in  $t$ , or (ii) can both vote for the Plaintiff (if they observed a strong public signal in favor of the Plaintiff) or (iii) can both vote for the Defendant (if they observed a strong public signal in favor of the Defendant). However, it cannot be that justice  $i$  votes for the Defendant and justice  $j \neq i$  votes for the Plaintiff, as this would imply that justice  $i$  is more biased than justice  $j$  towards the defendant, which contradicts her vote in case  $t$ .

Hence, observations of such "pairwise flip-flops" would violate the pre-vote deliberation model, and we search for pairwise violations of this kind. In order to rule out variation that is not related to case-specific information, we consider sets of similar cases (six in all) and fix the composition of the court. To have as many observations as possible, we consider REHN7 (Rehnquist 1994-2004), the natural court with most observations in our sample. The results are presented in Table 5.

[Table 5 about here]

With nine justices in the court, there are  $72 = N \times (N - 1)$  pairs. Most classes of cases present a significant number of violations. Consider for example the first class of cases (private Defendant vs private Plaintiff, lower courts agree, statutory interpretation).

---

<sup>28</sup>An alternative rationale to focus on unanimous votes follows from the analysis of [Gerardi and Yariv \(2007\)](#). If agents send not only relevant information, but also other (random) messages, which the group uses to define correlated voting strategies (here unanimous votes), the set of equilibria expands dramatically. [Gerardi and Yariv \(2007\)](#) show (using in fact unanimous votes at the voting stage) that every outcome that can be implemented with a non-unanimous voting rule  $r$  can also be implemented (as a sequential equilibrium of a cheap talk extension of the voting game) with a non-unanimous rule  $r'$ .



Using the information of the 41 Basic Rights cases in this class, we find 68 pairwise violations among the 72 pairings. Similarly, there are 44 violations in Economics, 24 in Criminal cases, and 56 in Federalism cases. The remaining classes of cases highlight a similar result. All in all, the large number of violations present strong evidence against the hypothesis of substantial sharing of information through deliberation.<sup>29</sup>

### 6.2.3 Does bias capture political ideology?

Throughout this paper, we used the terminology “bias” to denote the parameters  $\pi_{it}$ , describing justice  $i$ ’s propensity to rule for the plaintiff in case  $t$ : the size of the hurdle that information in favor of the Plaintiff has to overcome to lead to a decision in favor of the Plaintiff. Is this bias related at all to political ideology, along the conventional “liberal/conservative” spectrum? To check this, we also estimated a model where we included variables describing whether the Plaintiff’s position in each case is classified as a “liberal” or “conservative” cause. Because under this hypothesis Democratic and Republican justices would be expected to respond differently when voting for the Plaintiff is a liberal vote, we included an interaction term for these variables.

Figure 3 plots the estimated bias of each justice when the Plaintiff’s position is a liberal cause (in the horizontal axis) and when the Plaintiff’s position is a conservative cause (in the vertical axis). The figure reports results for the Basic Rights and Economics cases, where ideological concerns can be expected to play a more prominent role.

[Figure 3 about here]

The filled squares illustrate the bias of republican justices in Economics. The fact that these points are below the 45 degree line indicates that within Economic cases, Republican justices are more inclined to rule in favor of the Defendant when voting for the Defendant is a conservative vote. In particular, the average bias of Republican justices in Economics changes from 0.42 (moderately in favor of the Plaintiff) when voting for the Plaintiff is a conservative vote to 0.62 (moderately in favor of the Defendant) when supporting the Defendant is a conservative vote. In contrast, Democratic justices’ bias in Economic cases is basically unaffected by this distinction. In Basic Rights’ cases, instead, all justices appear to be more inclined to rule in favor of the Plaintiff when voting for the Plaintiff is a liberal vote. This effect is marginally stronger for Democratic justices.<sup>30</sup>

---

<sup>29</sup>A second institutional consideration related to communication is that in contrast with our model – in which we assumed that voting takes place simultaneously – justices vote in sequence, with the chief justice voting first, followed by the associate justices in order of seniority. This could potentially have an effect in terms of transmission of information. The question is more subtle than what it seems at first sight because, as argued in [Dekel and Piccione \(2000\)](#), observing the sequence of votes might not offer any additional *useful* information if voters condition on the event of being pivotal, as in the strategic voting model. In our case, however, justices are heterogeneous, so observing the voting sequence might still reveal useful information. We leave this interesting aspect for future research.

<sup>30</sup>The coefficient estimates for this model are not reported here for the sake of brevity, but are available from the authors upon request.

These results suggest that while in our model bias can reflect a variety of sources, it also partly incorporates the issues associated with political ideology along the conventional “liberal/conservative” spectrum, especially within economics cases. We leave a further exploration of this connection for future research.

### 6.3 Mistakes and Implications for Institutional Design

In any given case, the Court must arrive at a single ruling. The Court being a collective body, this single ruling requires aggregating the individual opinions in one way or the other. The Supreme Court aggregates the individual votes of its members by simple majority rule. In this section we address two questions. First, we provide a measure of performance: what is the probability that the court reaches a decision that is contrary to the true meaning of the law? Second, we ask whether this performance would improve or decline if the court were to use a different mechanism for aggregating the votes of individual justices. In particular, we compare the performance of the court with a counterfactual scenario in which ruling against the Defendant requires the unanimous consent of the justices.

We begin by computing the probability of mistakes in the Court. Note that for any given case characteristics  $X$ , our first stage estimates provide the individual probabilities of ruling for the Defendant when the law favors the Plaintiff ( $1 - \gamma_{i,1}$ ), and for the Plaintiff when the law favors the Defendant, ( $\gamma_{i,0}$ ). (We drop the obvious dependence on  $X$  to simplify notation.) For a simple majority rule, we then use these individual conditional probabilities to compute the probability that the Court will rule for the Defendant when the law favors the Plaintiff,  $\Pr(v = d|\omega = 1)$ , and for the Plaintiff when the law favors the Defendant,  $\Pr(v = p|\omega = 0)$ .<sup>31</sup> Given a prior  $\rho$ , we can then compute the total probability of an incorrect ruling for the Supreme Court as

$$\beta^{SC} = \rho \Pr(v = d|\omega = 1) + (1 - \rho) \Pr(v = p|\omega = 0)$$

The upper panel of Figure 4 shows  $\beta^{SC}$  and  $\Pr(v = p|\omega = 0)$  for Basic Rights issues, for both the strategic and expressive voting models. (As before, for exposition purposes here we focus on private parties, lower courts agree, statutory interpretation.) The figure shows two distinct patterns of mistakes throughout the sample. In the Warren courts and the first Burger court in the sample (BURG2) the total probability of error  $\beta^{SC}$  fluctuates between 2 % and 6 %, and exceeds 10% in some natural courts (WAR1, WAR3, WAR8, WAR9). Instead, in the remaining Burger courts and in the Rehnquist and Roberts courts,  $\beta^{SC}$  is bounded below 2%. The error rate is driven by the probability of ruling incorrectly in favor of the Plaintiff (bars in the figure). This exceeds 10% for several Warren courts, and is bounded below 4% from BURG4 on.

---

<sup>31</sup>Letting  $\mathcal{C}(k)$  denote the set of coalitions with exactly  $k$  members,  $\Pr(v = d|\omega = 1) = \sum_{k=5}^9 \sum_{C \in \mathcal{C}(k)} \prod_{i \in C} (1 - \gamma_{i,1}) \prod_{i \notin C} \gamma_{i,1}$ , and  $\Pr(v = p|\omega = 0) = \sum_{k=5}^9 \sum_{C \in \mathcal{C}(k)} \prod_{i \in C} \gamma_{i,0} \prod_{i \notin C} (1 - \gamma_{i,0})$ .

[Figure 4 about here]

How do these compare with performance under a unanimity rule? To evaluate this, we need to compute the probability of mistakes under unanimity. Now, in the expressive voting model, this is straightforward. Here behavior is unaffected by the aggregation mechanism, and therefore so are the individual strategy cutpoints and conditional probabilities. The only change is in the aggregation rule. Here the probability of the court ruling for the Defendant when the law favors the Plaintiff is  $1 - \prod_{i=1}^9 (1 - \gamma_{i,1})$  and the probability of the court ruling for the Plaintiff when the law favors the Defendant is  $\prod_{i=1}^9 \gamma_{i,0}$ . Thus the total probability of an incorrect ruling for the Supreme Court under unanimity rule in the expressive voting model is  $\beta_{exp}^U$

$$\beta_{exp}^U = \rho \left[ 1 - \prod_{i=1}^9 (1 - \gamma_{i,1}) \right] + (1 - \rho) \left[ \prod_{i=1}^9 \gamma_{i,0} \right]$$

In the strategic voting model, the computation of the total probability of mistakes under unanimity rule requires an additional step because the aggregation mechanism now clearly affects equilibrium behavior. Thus we cannot use the conditional probabilities of ruling for the Defendant recovered from justices' votes, but rather we must recompute the behavioral probabilities that are consistent with equilibrium behavior under unanimity. Fortunately, this is not difficult to do given our previous results. Given our estimates  $\{(\pi_i^*, \theta_i)\}$  we can use Eq. (4) with  $R$  to compute the equilibrium strategy cutpoints  $s_i^{**}$  consistent with unanimity rule. Given  $s^{**}$ , we can then compute  $\gamma_{i,1}^{**} = 1 - \Phi(\theta_i[s_i^{**} - 1])$  and  $\gamma_{i,0}^{**} = 1 - \Phi(\theta_i s_i^{**})$ . Then the total probability of an incorrect ruling for the Supreme Court under unanimity rule in the strategic voting model  $\beta_*^U$  is

$$\beta_*^U = \rho \left[ 1 - \prod_{i=1}^9 (1 - \gamma_{i,1}^{**}) \right] + (1 - \rho) \left[ \prod_{i=1}^9 \gamma_{i,0}^{**} \right]$$

The lower panel of Figure 4 puts everything together (again, for Basic Rights issues). Unanimity rule considerably increases total error rates in comparison to simple majority rule. Consider first the expressive voting model. While here justices do not change their voting behavior in response to the voting rule, unanimity requires the vote of all nine justices to reach a decision in favor of the Plaintiff. As a result, unanimity reduces the probability of ruling incorrectly in favor of the Plaintiff to insignificant levels, but also increases considerably the probability of ruling incorrectly in favor of the Defendant, and leads to a large total probability of error (including, and especially in, the Burger, Rehnquist and Roberts courts). In the strategic voting model, instead, the change in justices' voting strategy in response to the alternative decision rule leads to an increase in the probability of ruling incorrectly in favor of the Plaintiff vis a vis majority rule, and with it to an increase in the total probability of error, especially in the Warren courts.

This basic conclusion holds across all issue areas (results are available from the authors upon request). The results confirm in this application with small committees and heterogeneous agents the asymptotic results of Feddersen and Pesendorfer (1998) and Duggan and Martinelli (2001) (see also Meirowitz (2002)).

## 7 Conclusion

In this paper we have presented results from a voting model for the US Supreme Court in which votes reflect both justices’ personal ideologies, as well as their endeavor to “get it right”: to rule according to an accurate and faithful interpretation of the law as it applies to the specifics of each case. In this context, we study whether case information has enough power to overturn the prior biases and ideological considerations of the justices.

To tackle this question, we introduce a new estimation approach that allows us to handle our model of voting with common values and strategic agents. The model is estimated in two steps. In the first step, using the observed votes, we estimate a “reduced-form” model of justices’ probabilities of voting in favor of the Plaintiff when the law favors the Plaintiff and when the law favors the Defendant. In the second step, we recover the structural parameters characterizing justices’ preferences and information services, using the equilibrium conditions in the voting model.

Our methodology allows us to disentangle the effects of ideology and information for each justice, and then to quantify the tradeoffs between ideology and information in the court. Our results, as encapsulated in our FLEX measure, indicate a substantial value of information: in roughly 44% of cases, justices’ initial leanings – which reflect their priors or their ideological biases – are changed by the case-specific private information of the justices. Moreover, the temporal evolution of FLEX scores suggest that the increased politicization of the Supreme Court appointment process pointed out by the literature has become uniformly marked only in the last quarter-century, following the failed nomination of justice Bork (i.e., within the Rehnquist and Roberts courts).

We close with a remark about the limitations of our results. Possibly the most important of these is that our analysis paints a necessarily incomplete picture of the court. In this paper, we have concentrated exclusively on the collective decision about the *disposition* of the case, but abstracted away from a second important dimension of the justices’ rulings: the opinions written by the court. Because of the principle of *stare decisis*, lower court judges must follow the precedents set by the Supreme Court, which are mainly embodied in the opinions written by the justices to justify their votes; these opinions establish the *legal rules* for lower courts and are therefore particularly important in terms of the long run implications of the court rulings, as well as their written justification.

Clearly our estimates of justices’ bias ( $\pi_{it}$ ), ability to infer the meaning of the law ( $\theta_{it}$ ), and of the value of information in the court (*FLEX*) reflect only the voting deci-

sions of the justices, and do not speak about the determinants of the opinions, or their implications. On the other hand, it is important to clarify that the majority disposition of a case does not require a majority opinion (i.e., the majority for disposition may disagree about the rule that should cover this class of cases); indeed, it is not rare to observe justices voting with the majority, who nevertheless issue a dissenting opinion expounding their differential interpretation of the law underlying the case. Thus, votes and opinions are really two different instruments for achieving the two different goals of disposition and precedence-setting. Because these two goals can be quite independent, we believe that modeling the justices' voting behavior apart from their opinion-writing, while incomplete, suffices to allow us to measure the value of information and perform the various counterfactuals in the paper. Moreover, fully modeling the opinion-writing process, and the assigning of justices to write them, is very complex.<sup>32</sup> We plan to return to this in future research.

---

<sup>32</sup>For more details on the distinction between case disposition and rule creation see [Cameron and Kornhauser \(2008\)](#), where case disposition and rule creation are linked because of the assumption that justices care about the extent of support her opinion attracts.

## 8 Appendix

Suppose that justices care about the outcome of the court ruling and are fully strategic. Any justice  $i$  would then consider the implications of her vote assuming she is pivotal for the decision. In this case the relevant information for justice  $i$  in case  $t$  is not only the private information  $s_{it}$  but also the realization that  $i$  is pivotal for the court's decision, given the equilibrium strategy profile  $\sigma_{-i}$  followed by the remaining justices. Thus in this case it is optimal for  $i$  to rule against the Defendant if and only if

$$J(\sigma_{-i}, s_{it}) \equiv \frac{P_{\sigma_{-i}}(piv_i|P)}{P_{\sigma_{-i}}(piv_i|D)} \frac{f_i(s_{it}|P)}{f_i(s_{it}|D)} \frac{\rho}{1-\rho} - \frac{\pi_i}{1-\pi_i} \geq 0$$

But given the strict MLRP satisfied by the conditional normal distributions, this implies that the best response to any strategy of the remaining justices is a cutoff strategy,

**Lemma 1** *The best response to any strategy of the remaining justices is a cutoff strategy, i.e., in any equilibrium  $\sigma_i$  satisfies*

$$\sigma_i(s_i) = \begin{cases} 1 & \text{if } J(\sigma_{-i}, s_{it}) \geq 0 \\ 0 & \text{if } J(\sigma_{-i}, s_{it}) < 0. \end{cases}$$

**Proof.** The proof follows [Duggan and Martinelli \(2001\)](#), and is included here for convenience only. Suppose that  $\sigma_i$  satisfies the hypothesis. Define the sets  $V \equiv \{s \in S : J(\sigma_{-i}, s) > 0\}$  and  $W \equiv \{s \in S : J(\sigma_{-i}, s) < 0\}$ , and consider  $\sigma'_i$  such that  $\sigma'_i(s) < 1$  for some  $s \in V$  and  $\sigma'_i(s) > 0$  for some  $s \in W$ . Define  $A(s) \equiv U(p|P) \Pr(P|s, piv) + U(p|D) \Pr(D|s, piv)$  and  $B(s) \equiv U(d|P) \Pr(P|s, piv) + U(d|D) \Pr(D|s, piv)$ . The net expected payoff of voting to rule against the Defendant vis a vis voting to rule in favor of the Defendant is

$$\begin{aligned} \Delta(s) &\equiv \int_V A(s) + \int_W B(s) - \int_V [\sigma'_i(s)A(s) + (1 - \sigma'_i(s))B(s)] - \int_W [\sigma'_i(s)A(s) + (1 - \sigma'_i(s))B(s)] \\ &= \int_V (1 - \sigma'_i(s))[A(s) - B(s)]ds + \int_W \sigma'_i(s)[B(s) - A(s)]ds. \end{aligned} \tag{13}$$

Now, since

$$\begin{aligned} A(s) - B(s) &= \Pr(P|s, piv)[U(p|P) - U(d|P)] + \Pr(D|s, piv)[U(p|D) - U(d|D)] \\ &= \Pr(P|s, piv)(1 - \pi_i) - \Pr(D|s, piv)\pi_i, \end{aligned} \tag{14}$$

it follows that  $A(s) - B(s) > 0$  if and only if

$$\frac{\Pr(P|s, piv)}{\Pr(D|s, piv)} - \frac{\pi_i}{1 - \pi_i} > 0 \Leftrightarrow \frac{\Pr(s|P) \Pr(piv|P)p}{\Pr(s|D) \Pr(piv|D)(1 - p)} - \frac{\pi_i}{1 - \pi_i} > 0 \Leftrightarrow J(\sigma_{-i}, s) > 0$$

Now,  $s \in V$  implies that  $J(\sigma_{-i}, s) > 0$ , which in turn implies that  $A(s) - B(s) > 0$ . Similarly,  $s \in W$  implies that  $J(\sigma_{-i}, s) < 0$ , which in turn implies that  $A(s) - B(s) < 0$ . Then  $\Delta > 0$ , which concludes the proof. ■

## References

- AGUIRREGABIRIA, V., AND P. MIRA (2007): “Sequential Estimation of Dynamic Discrete Games,” *Econometrica*, 75, 1–53.
- AUSTEN-SMITH, D., AND J. S. BANKS (1996): “Information Aggregation, Rationality, and the Condorcet Jury Theorem,” *American Political Science Review*, 90, 34–45.
- AUSTEN-SMITH, D., AND T. FEDDERSEN (2005): “Deliberation and Voting Rules,” In *Social Choice and Strategic Decisions: Essays in Honor of Jeffrey S. Banks*, pp. 269–316.
- (2006): “Deliberation, Preference Uncertainty, and Voting Rules,” *American Political Science Review*, 100, 209–218.
- BAJARI, P., L. BENKARD, AND J. LEVIN (2007): “Estimating Dynamic Models of Imperfect Competition,” *Econometrica*, 75, 1331–1370.
- CAMERON, C. M., AND L. A. KORNHAUSER (2008): “Modeling Collegial Courts (3): Judicial Objectives, Opinion Content, Voting and Adjudication Equilibria,” *Law & Economics Research Paper Series*, Working Paper No. 08-54.
- CAMERON, C. M., J. A. SEGAL, AND D. SONGER (2000): “Strategic Auditing in a Political Hierarchy: An Informational Model of the Supreme Court’s Certiorari Decisions,” *American Political Science Review*, 94(1), 101–116.
- CHIANG, C.-F., AND B. G. KNIGHT (2008): “Media Bias and Influence: Evidence from Newspaper Endorsements,” NBER Working Paper No. 14445.
- CILIBERTO, F., AND E. TAMER (2009): “Market Structure and Multiple Equilibria in Airline Markets,” *Econometrica*, 77, 1791–1828.
- CLINTON, J. D., S. JACKMAN, AND D. RIVERS (2004): “The Statistical Analysis of Roll Call Data,” *American Political Science Review*, 55, 355–370.
- CLINTON, J. D., AND A. MEIROWITZ (2003): “Integrating Voting Theory and Roll Call Analysis: A Framework,” *Political Analysis*, 11, 381–396.
- (2004): “Testing Explanations of Strategic Voting in Legislatures: A Reexamination of the Compromise of 1790,” *American Journal of Political Science*, 48(4), 675–689.
- COATE, S., AND M. CONLIN (2004): “A Group Rule-Utilitarian Approach to Voter Turnout: Theory and Evidence,” *American Economic Review*, 94, 1476–1504.
- COATE, S., M. CONLIN, AND A. MORO (2008): “The performance of pivotal-voter models in small-scale elections: Evidence from Texas liquor referenda,” *Journal of Public Economics*, 92, 582–596.
- COUGHLAN, P. (2000): “In Defense of Unanimous Jury Verdicts: Mistrials, Communication, and Strategic Voting,” *American Political Science Review*, 94, 375–393.
- DAUGHETY, A., AND J. REINGANUM (2006): “Speaking Up: A Model of Judicial Dissent and Discretionary Review,” *Supreme Court Economic Review*, 14, 1–51.



- DEGAN, A., AND A. MERLO (2008): “Do Voters Vote Ideologically?,” Working Paper, U. Penn.
- DE PAULA, A., AND A. MERLO (2009): “Identification and Estimation of Preference Distributions when Voters are Ideological,” Working Paper, U. Penn.
- DEKEL, E., AND M. PICCIONE (2000): “Sequential Voting Procedures in Symmetric Binary Elections,” *Journal of Political Economy*, 108, 34–54.
- DORASZELSKI, U., D. GERARDI, AND F. SQUINTANI (2003): “Communication and Voting with Double-Sided Information,” *Contributions to Theoretical Economics*, 3, 1–39.
- DUGGAN, J., AND C. MARTINELLI (2001): “A Bayesian Model of Voting in Juries,” *Games and Economic Behavior*, 37, 259–294.
- EPSTEIN, L., V. HOEKSTRA, J. SEGAL, AND H. SPAETH (1998): “Do Political Preferences Change? A Longitudinal Study of U.S. Supreme Court Justices,” *The Journal of Politics*, 60(3), 801–818.
- EPSTEIN, L., AND J. KNIGHT (1997): *The Choices Justices Make*. CQ Press, New York, NY.
- EPSTEIN, L., R. LINDSTADT, J. A. SEGAL, AND C. WESTERLAND (2006): “The Changing Dynamics of Senate Voting on Supreme Court Nominees,” *The Journal of Politics*, 68, 296–307.
- EPSTEIN, L., AND C. MERSHON (1996): “Measuring Political Preferences,” *American Journal of Political Science*, 40(1), 261–294.
- EPSTEIN, L., T. G. WALKER, N. STAUDT, S. A. HENDRICKSON, AND J. M. ROBERTS (2008): “The U.S. Supreme Court Justices Database (computer file),” <http://epstein.law.northwestern.edu/research/justicesdata.html>.
- FEDDERSEN, T., AND W. PESENDORFER (1997): “Voting Behavior and Information Aggregation in Elections With Private Information,” *Econometrica*, 65, 1029–1058.
- (1998): “Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting,” *American Political Science Review*, 92, 23–35.
- FROEB, L. M., AND B. H. KOBAYASHI (1996): “Naive, Biased, yet Bayesian: Can Juries Interpret Selectively Produced Evidence?,” *Journal of Law, Economics, and Organization*, 12, 257–275.
- GELY, R., AND P. SPILLER (1990): “A Rational Choice Theory of Supreme Court Statutory Decisions with Applications to the ”State Farm” and ”Grove City Cases”,” *Journal of Law Economics & Organization*, 6(2), 263–300.
- GERARDI, D., AND L. YARIV (2007): “Deliberative Voting,” *Journal of Economic Theory*, 134, 317–338.
- HALL, P., AND X.-H. ZHOU (2003): “Nonparametric Estimation of Component Distributions in a Multivariate Mixture,” *Annals of Statistics*, 31, 201–224.

- HECKMAN, J., AND J. J. SNYDER (1997): “Linear Probability Models of the Demand for Attributes with an Empirical Application to Estimating the Preferences of Legislators,” *The RAND Journal of Economics*, 28(0), S142–S189, Special Issue in Honor of Richard E. Quandt.
- HU, Y. (2008): “Identification and Estimation of Nonlinear Models with Misclassification Error Using Instrumental Variables: a General Solution,” *Journal of Econometrics*, 144, 27–61.
- IARYCZOWER, M., G. KATZ, AND S. SAIEGH (2009): “The not-so-popular branch: Bicameralism as a Counter-Majoritarian Device,” HSS, California Institute of Technology.
- KASAHARA, H., AND K. SHIMOTSU (2007): “Nonparametric Identification and Estimation of Multivariate Mixtures,” Queen’s University, Department of Economics.
- KAWAI, K., AND Y. WATANABE (2009): “Inferring Strategic Voting,” Northwestern University, Department of Economics.
- LANDES, W. M., AND R. A. POSNER (2008): “Rational Judicial Behavior: A Statistical Study,” University of Chicago.
- LAX, J. R., AND C. M. CAMERON (2007): “Bargaining and Opinion Assignment on the US Supreme Court,” *Journal of Law, Economics and Organization*, 23, 276–302.
- LI, H., S. ROSEN, AND W. SUEN (2001): “Conflicts and Common Interests in Committees,” *The American Economic Review*, 91, 1478–1497.
- LIM, C. (2008): “Turnover and Accountability of Appointed and Elected Judges,” Typeset, Stanford University.
- LONDREGAN, J. (1999): “Estimating Legislators’ Preferred Points,” *Political Analysis*, 8(1), 35–56.
- MARTIN, A., AND K. QUINN (2002): “Dynamic ideal point estimation via Markov chain Monte Carlo for the US Supreme Court, 1953–1999,” *Political Analysis*, 10(2), 134–153.
- (2007): “Assessing Preference Change on the US Supreme Court,” *Journal of Law, Economics and Organization*, 23(2), 365–385.
- MEIROWITZ, A. (2002): “Informative Voting and Condorcet Jury Theorems with a Continuum of Types,” *Social Choice and Welfare*, 19, 219–236.
- POOLE, K., AND H. ROSENTHAL (1985): “A Spatial Model for Legislative Roll Call Analysis,” *American Journal of Political Science*, 29, 357–384.
- (1991): “Patterns of Congressional Voting,” *American Journal of Political Science*, 35, 228–278.
- SEGAL, J., C. M. CAMERON, AND A. COVER (1990): “Senate Voting on Supreme Court Nominees: A Neoinstitutional Model,” *American Political Science Review*, 84(2), 525–534.
- (1992): “A Spatial Model of Roll Call Voting: Senators, Constituents, Presidents, and Interest Groups in Supreme Court Confirmations,” *American Journal of Political Science*, 36(1), 96–121.

- SEGAL, J., AND A. COVER (1989): “Ideological Values and the Votes of U.S. Supreme Court Justices,” *American Political Science Review*, 83(2), 557–565.
- SEGAL, J., L. EPSTEIN, C. CAMERON, AND H. SPAETH (1995): “Ideological Values and the Votes of U.S. Supreme Court Justices Revisited,” *The Journal of Politics*, 57(3), 812–823.
- SEGAL, J., AND H. SPAETH (1993): *The Supreme Court and the Attitudinal Model*.
- SPAETH, H. J. (2008): “The Original United States Supreme Court Judicial Database, 1953–2007 Terms (computer file), ,” <http://www.cas.sc.edu/poli/juri/sctdata.htm/>.

Court	Years	Basic Rights		Economic		Criminal Procedure		Federal		Total	
		No. of Cases	% for Plaintiff	No. of Cases	% for Plaintiff	No. of Cases	% for Plaintiff	No. of Cases	% for Plaintiff	No. of Cases	% for Plaintiff
WAR1	1953	16	0.56	28	0.46	10	0.40	5	0.40	59	0.47
WAR3	1954 - 56	37	0.65	34	0.65	14	0.64	10	0.50	95	0.63
WAR4	1956	11	0.36	6	1.00	8	0.38	5	0.60	30	0.53
WAR5	1956 - 58	61	0.69	41	0.61	42	0.64	13	0.38	157	0.63
WAR6	1958 - 61	142	0.58	114	0.64	78	0.59	42	0.50	376	0.59
WAR8	1962 - 64	167	0.75	100	0.82	75	0.77	41	0.68	383	0.77
WAR9	1965 - 66	107	0.68	46	0.72	54	0.76	17	0.76	224	0.71
WAR10	1967 - 68	80	0.74	22	0.86	63	0.79	5	0.60	170	0.77
BURG2	1969 - 70	65	0.66	20	0.60	28	0.71	4	0.50	117	0.66
BURG4	1971 - 75	254	0.72	86	0.71	121	0.66	24	0.58	485	0.70
BURG6	1975 - 80	346	0.67	144	0.60	149	0.74	35	0.63	674	0.67
BURG7	1981 - 85	306	0.61	148	0.68	168	0.68	43	0.63	665	0.65
REHN1	1986	58	0.47	28	0.64	39	0.77	15	0.53	140	0.59
REHN3	1987 - 89	141	0.52	69	0.57	85	0.58	30	0.47	325	0.54
REHN4	1990	38	0.61	31	0.74	28	0.57	5	0.80	102	0.65
REHN5	1991 - 92	88	0.64	48	0.46	52	0.69	20	0.80	208	0.63
REHN6	1992 - 93	37	0.38	24	0.63	25	0.48	5	0.60	91	0.48
REHN7	1994 - 04	416	0.62	168	0.63	232	0.68	75	0.57	891	0.63
ROBT1	2005	10	0.70	4	0.50	7	1.00	2	0.50	23	0.74
ROBT2	2005 - 08	67	0.67	42	0.71	56	0.64	10	0.30	175	0.65
Total		2447	0.64	1203	0.66	1334	0.68	406	0.58	5390	0.65

Justice	Year Oath Taken	Year Out of Court	Judicial Experience	Political Party	Proportion of Votes in Favor of the Plaintiff					SC Ideology	SC Qualif.
					Basic Rights	Economic	Criminal Procedure	Federal	All Cases		
Harlan	1955	1971	1	R	0.519	0.488	0.481	0.577	0.501	0.875	0.750
Black	1937	1971	0	D	0.679	0.655	0.707	0.542	0.679	0.875	0.156
Douglas	1939	1975	0	D	0.694	0.630	0.748	0.500	0.691	0.730	0.813
Stewart	1959	1981	4	R	0.637	0.585	0.630	0.613	0.623	0.750	1.000
Marshall	1967	1991	4	D	0.533	0.564	0.548	0.559	0.544	1.000	0.835
Brennan	1957	1990	4	D	0.584	0.630	0.627	0.558	0.606	1.000	1.000
White	1962	1993	0	D	0.644	0.651	0.621	0.649	0.639	0.500	0.500
Warren	1954	1969	0	R	0.710	0.680	0.802	0.601	0.725	0.750	0.855
Clark	1949	1967	0	D	0.553	0.648	0.448	0.549	0.558	0.500	0.125
Frankfurter	1939	1962	0	I	0.577	0.529	0.526	0.547	0.548	0.665	0.965
Whittaker	1957	1962	3	R	0.547	0.503	0.525	0.564	0.527	0.500	1.000
Burton	1945	1958	0	R	0.512	0.569	0.432	0.485	0.513	0.280	0.930
Reed	1938	1957	0	D	0.453	0.588	0.313	0.450	0.482	0.725	0.875
Fortas	1965	1969	0	D	0.749	0.691	0.838	0.591	0.766	1.000	1.000
Goldberg	1962	1965	0	D	0.772	0.780	0.813	0.683	0.784	0.750	0.915
Minton	1949	1956	8	D	0.472	0.516	0.375	0.467	0.475	0.720	0.355
Jackson	1941	1954	0	D	0.500	0.571	0.400	0.600	0.519	1.000	0.915
Burger	1969	1986	13	R	0.668	0.626	0.646	0.623	0.653	0.115	0.960
Blackmun	1970	1994	11	R	0.593	0.574	0.659	0.608	0.606	0.115	0.970
Powell	1972	1987	0	D	0.654	0.635	0.688	0.607	0.658	0.165	1.000
Rehnquist	1972	1986	0	R	0.633	0.601	0.592	0.575	0.615	0.045	0.885
Stevens	1975	-	5	R	0.551	0.554	0.593	0.504	0.563	0.250	0.960
O'Connor	1981	2006	0	R	0.612	0.619	0.640	0.549	0.622	0.415	1.000
Scalia	1986	-	4	R	0.580	0.582	0.588	0.593	0.583	0.000	1.000
Kennedy	1988	-	13	R	0.593	0.619	0.616	0.612	0.606	0.365	0.890
Souter	1990	-	7	R	0.559	0.625	0.675	0.590	0.608	0.325	0.765
Thomas	1991	-	1	R	0.578	0.612	0.565	0.598	0.582	0.156	0.415
Ginsburg	1993	-	13	D	0.547	0.584	0.634	0.554	0.581	0.680	1.000
Breyer	1994	-	14	D	0.617	0.664	0.651	0.540	0.637	0.475	0.545
Roberts	2005	-	2	R	0.662	0.717	0.603	0.500	0.656	0.117	0.970
Alito	2006	-	16	R	0.627	0.690	0.518	0.400	0.606	0.100	0.810

Table 1: Data. Top Panel: Case Information; Lower Panel: Justice Information

Table 2: “First Stage” MLE Estimates: Basic Rights, Economic, and Criminal Issues.

	Basic Rights			Economics			Criminal Procedure			Federalism		
	$\rho$	$\gamma_{\text{no}}$	$\gamma_{\text{at}}$	$\rho$	$\gamma_{\text{no}}$	$\gamma_{\text{at}}$	$\rho$	$\gamma_{\text{no}}$	$\gamma_{\text{at}}$	$\rho$	$\gamma_{\text{no}}$	$\gamma_{\text{at}}$
Case Specific	Constant	0.597 (0.140)	-0.184 (0.512)	7.106 (0.543)	0.341 (0.206)	-0.554 (0.646)	3.051 (0.930)	0.341 (0.386)	-1.627 (0.604)	6.171 (0.821)	0.045 (0.399)	-0.404 (1.264)
	Plaintiff Local Gov.	0.741 (0.133)	-1.420 (0.105)	-0.654 (0.046)	0.930 (0.272)	0.040 (0.151)	-0.599 (0.102)	2.901 (0.424)	-4.398 (0.494)	-0.790 (0.185)	-0.626 (0.386)	0.495 (0.134)
	Plaintiff US Gov.	1.002 (0.177)	-0.329 (0.090)	-0.135 (0.064)	0.159 (0.180)	0.510 (0.073)	0.151 (0.092)	3.077 (0.460)	-2.530 (0.272)	-0.495 (0.189)	0.133 (0.312)	0.667 (0.140)
	Defendant Local Gov.	-0.686 (0.124)	0.863 (0.049)	0.379 (0.072)	-0.439 (0.197)	-0.360 (0.079)	-0.011 (0.114)	-0.814 (0.120)	0.698 (0.186)	1.004 (0.333)	0.459 (0.175)	-0.302 (0.171)
	Defendant US Gov.	-0.823 (0.170)	0.582 (0.064)	-0.469 (0.074)	-0.691 (0.200)	-0.457 (0.088)	-0.456 (0.094)	-0.397 (0.371)	0.388 (0.117)	0.246 (0.229)	-0.630 (0.341)	-0.759 (0.147)
	Lower Courts Agree	-0.084 (0.097)	0.422 (0.039)	-0.189 (0.039)	0.305 (0.142)	-0.673 (0.061)	-0.081 (0.068)	-0.206 (0.174)	0.002 (0.053)	-0.081 (0.056)	-0.251 (0.109)	0.210 (0.129)
	Constitutional Review	-0.137 (0.191)	-0.048 (0.080)	-0.323 (0.062)	-0.062 (0.403)	0.388 (0.142)	1.996 (0.830)	-0.725 (0.292)	-0.312 (0.081)	0.095 (0.097)	0.594 (0.617)	-0.113 (0.220)
	Statutory Interpretation	-0.280 (0.107)	0.037 (0.045)	0.343 (0.042)	0.302 (0.173)	-0.889 (0.066)	-0.375 (0.099)	-0.441 (0.209)	-0.367 (0.063)	-0.018 (0.070)	-0.339 (0.349)	0.280 (0.131)
	CJ Burger	0.246 (0.122)	0.126 (0.112)	0.580 (0.176)	-0.329 (0.177)	-0.227 (0.247)	-0.033 (0.273)	-0.208 (0.196)	-1.327 (0.228)	0.538 (0.202)	-0.013 (0.303)	0.308 (0.613)
	CJ Rehnquist	0.032 (0.128)	-0.139 (0.147)	0.361 (0.193)	-0.502 (0.174)	-0.235 (0.273)	0.758 (0.316)	-0.388 (0.198)	-1.411 (0.265)	0.542 (0.230)	-0.257 (0.309)	0.228 (0.669)
	CJ Roberts	0.239 (0.288)	0.002 (0.191)	0.501 (0.225)	-0.602 (0.348)	0.357 (0.296)	1.383 (0.412)	-0.044 (0.406)	-1.515 (0.286)	0.742 (0.257)	-0.211 (0.796)	0.770 (1.196)
Justice Specific	Prior Judicial Experience		0.012 (0.008)	-0.025 (0.008)		0.014 (0.010)	-0.022 (0.014)		0.056 (0.009)	-0.056 (0.011)		0.058 (0.026)
	Democratic Nominee		0.382 (0.089)	-0.013 (0.092)		0.124 (0.138)	0.564 (0.148)		0.498 (0.104)	-0.179 (0.137)		-0.142 (0.219)
	SC Ideology		0.906 (0.128)	-1.668 (0.129)		0.054 (0.194)	-1.121 (0.202)		1.838 (0.170)	-2.638 (0.192)		-0.100 (0.358)
	SC Qualifications		0.375 (0.157)	-1.089 (0.181)		0.082 (0.224)	-0.536 (0.285)		1.529 (0.177)	-1.267 (0.267)		-0.343 (0.465)
Justice/Case Specific	Average SC Ideology of Remaining Justices		-0.801 (0.304)	1.060 (0.348)		0.058 (0.477)	0.044 (0.593)		-3.265 (0.457)	1.284 (0.439)		-0.922 (1.046)
	Average SC Qualif. of Remaining Justices		-2.953 (0.497)	-5.020 (0.612)		0.013 (0.796)	-0.046 (0.944)		0.448 (0.707)	-2.646 (0.802)		0.134 (1.749)
	Years of Experience in the Court		0.028 (0.004)	-0.036 (0.004)		0.019 (0.005)	-0.044 (0.007)		0.041 (0.005)	-0.052 (0.006)		0.000 (0.010)
												-0.024 (0.012)

Excluded Categories: Plaintiff Private, Defendant Private, Issue Misc., Lower Courts Disagree, Authority Other, CJ Warren, Republican Nominee

Table 3: Rehnquist 1994-2004 Court (REHN7). By Issue. Private vs Private, Statutory Interpretation, Lower Courts Agree.

REHN 7 Court - Rehnquist 1994-2004

Statutory Interpretation, Private Defendant vs Private Plaintiff, Lower Courts Agree

Criminal Procedure													
(prior $\rho$ mean = 0.333, std.error = 0.086)													
Justice	$\gamma_{ro}$	$\gamma_{ri}$	$\theta$	$s^*$	$\pi^{exp}$	$\pi^*$	Justice	$\gamma_{ro}$	$\gamma_{ri}$	$\theta$	$s^*$	$\pi^{exp}$	$\pi^*$
Kennedy	0.203 (0.012)	0.852 (0.008)	1.876 (0.054)	0.444 (0.016)	0.517 (0.020)	0.176 (0.026)	Kennedy	0.323 (0.028)	0.864 (0.022)	1.555 (0.120)	0.295 (0.043)	0.233 (0.030)	0.014 (0.011)
Souter	0.164 (0.009)	0.884 (0.006)	2.175 (0.044)	0.450 (0.011)	0.507 (0.020)	0.176 (0.027)	Souter	0.193 (0.019)	0.908 (0.017)	2.192 (0.120)	0.395 (0.027)	0.232 (0.032)	0.014 (0.011)
Thomas	0.103 (0.009)	0.931 (0.006)	2.743 (0.070)	0.460 (0.012)	0.492 (0.023)	0.175 (0.027)	Thomas	0.063 (0.009)	0.964 (0.008)	3.335 (0.133)	0.459 (0.019)	0.240 (0.039)	0.017 (0.012)
O'Connor	0.231 (0.016)	0.849 (0.011)	1.767 (0.064)	0.415 (0.021)	0.500 (0.021)	0.173 (0.026)	O'Connor	0.289 (0.031)	0.870 (0.023)	1.682 (0.132)	0.331 (0.044)	0.236 (0.032)	0.014 (0.011)
Stevens	0.237 (0.015)	0.847 (0.010)	1.741 (0.059)	0.411 (0.020)	0.499 (0.020)	0.173 (0.023)	Stevens	0.310 (0.030)	0.864 (0.023)	1.594 (0.125)	0.312 (0.044)	0.236 (0.030)	0.014 (0.012)
Breyer	0.216 (0.019)	0.870 (0.011)	1.911 (0.080)	0.411 (0.023)	0.485 (0.022)	0.170 (0.024)	Breyer	0.333 (0.035)	0.875 (0.024)	1.583 (0.137)	0.273 (0.052)	0.220 (0.033)	0.014 (0.011)
Rehnquist	0.197 (0.016)	0.887 (0.010)	2.064 (0.075)	0.414 (0.019)	0.475 (0.023)	0.168 (0.024)	Rehnquist	0.178 (0.023)	0.916 (0.017)	2.299 (0.140)	0.401 (0.030)	0.228 (0.035)	0.014 (0.011)
Scalia	0.153 (0.010)	0.918 (0.006)	2.415 (0.055)	0.423 (0.012)	0.454 (0.022)	0.161 (0.024)	Scalia	0.140 (0.017)	0.946 (0.011)	2.683 (0.122)	0.403 (0.022)	0.198 (0.035)	0.013 (0.011)
Ginsburg	0.326 (0.020)	0.822 (0.011)	1.374 (0.062)	0.329 (0.032)	0.485 (0.020)	0.176 (0.025)	Ginsburg	0.604 (0.034)	0.848 (0.023)	0.764 (0.099)	-0.345 (0.140)	0.234 (0.030)	0.019 (0.013)

Federalism													
(prior $\rho$ mean = 0.570, std.error = 0.069)													
Justice	$\gamma_{ro}$	$\gamma_{ri}$	$\theta$	$s^*$	$\pi^{exp}$	$\pi^*$	Justice	$\gamma_{ro}$	$\gamma_{ri}$	$\theta$	$s^*$	$\pi^{exp}$	$\pi^*$
Kennedy	0.139 (0.013)	0.859 (0.015)	2.158 (0.088)	0.502 (0.020)	0.612 (0.026)	0.455 (0.062)	Kennedy	0.165 (0.027)	0.938 (0.016)	2.513 (0.173)	0.387 (0.033)	0.394 (0.072)	0.070 (0.044)
Souter	0.124 (0.009)	0.892 (0.010)	2.390 (0.070)	0.483 (0.015)	0.587 (0.027)	0.448 (0.061)	Souter	0.196 (0.023)	0.913 (0.014)	2.214 (0.118)	0.387 (0.028)	0.432 (0.053)	0.076 (0.041)
Thomas	0.110 (0.012)	0.933 (0.011)	2.725 (0.110)	0.451 (0.018)	0.520 (0.039)	0.411 (0.066)	Thomas	0.245 (0.041)	0.854 (0.027)	1.741 (0.177)	0.396 (0.053)	0.492 (0.050)	0.080 (0.049)
O'Connor	0.135 (0.012)	0.840 (0.017)	2.098 (0.089)	0.526 (0.021)	0.637 (0.026)	0.461 (0.063)	O'Connor	0.215 (0.029)	0.880 (0.022)	1.966 (0.148)	0.402 (0.038)	0.476 (0.054)	0.080 (0.044)
Stevens	0.157 (0.013)	0.822 (0.016)	1.932 (0.082)	0.522 (0.021)	0.630 (0.025)	0.455 (0.064)	Stevens	0.195 (0.023)	0.880 (0.024)	2.037 (0.145)	0.423 (0.035)	0.491 (0.054)	0.081 (0.047)
Breyer	0.140 (0.017)	0.934 (0.012)	2.592 (0.124)	0.418 (0.022)	0.473 (0.045)	0.391 (0.067)	Breyer	0.159 (0.035)	0.952 (0.017)	2.666 (0.236)	0.375 (0.039)	0.352 (0.102)	0.062 (0.042)
Rehnquist	0.153 (0.015)	0.853 (0.018)	2.075 (0.106)	0.494 (0.023)	0.603 (0.027)	0.452 (0.060)	Rehnquist	0.221 (0.028)	0.822 (0.039)	1.691 (0.180)	0.454 (0.050)	0.538 (0.050)	0.081 (0.049)
Scalia	0.128 (0.010)	0.904 (0.009)	2.438 (0.073)	0.465 (0.015)	0.560 (0.028)	0.436 (0.061)	Scalia	0.195 (0.025)	0.890 (0.020)	2.086 (0.135)	0.412 (0.033)	0.475 (0.053)	0.080 (0.047)
Ginsburg	0.146 (0.017)	0.898 (0.015)	2.326 (0.115)	0.454 (0.022)	0.549 (0.032)	0.433 (0.063)	Ginsburg	0.142 (0.028)	0.966 (0.012)	2.896 (0.213)	0.370 (0.032)	0.309 (0.103)	0.054 (0.038)

Table 4: The Value of Information in the Court: By Justice (Private Parties, Statutory Interpretation, Lower Courts Agree).

Justice	Mid Year of Tenure	Basic Rights ( $\rho = 0.58$ )			Economics ( $\rho = 0.66$ )			Criminal ( $\rho = 0.39$ )			Federalism ( $\rho = 0.61$ )			Average	
		$\pi^{\text{exp}}$	$\theta$	FLEX <sup>exp</sup>	$\pi^{\text{exp}}$	$\theta$	FLEX <sup>exp</sup>	$\pi^{\text{exp}}$	$\theta$	FLEX <sup>exp</sup>	$\pi^{\text{exp}}$	$\theta$	FLEX <sup>exp</sup>	FLEX <sup>exp</sup>	FLEX <sup>exp</sup>
Burton	1952	0.29	2.55	<b>0.38</b>	0.75	1.96	<b>0.63</b>	0.24	2.69	<b>0.51</b>	0.68	2.45	<b>0.59</b>	<b>0.53</b>	
Black	1954	0.40	1.65	<b>0.35</b>	0.73	1.66	<b>0.58</b>	0.34	1.37	<b>0.45</b>	0.70	2.24	<b>0.57</b>	<b>0.49</b>	
Warren	1962	0.37	2.14	<b>0.38</b>	0.78	1.84	<b>0.59</b>	0.30	1.94	<b>0.48</b>	0.64	2.60	<b>0.49</b>	<b>0.49</b>	
Rehnquist	1979	0.52	2.44	<b>0.43</b>	0.61	2.33	<b>0.47</b>	0.25	2.90	<b>0.60</b>	0.56	2.17	<b>0.43</b>	<b>0.48</b>	
Whittaker	1960	0.37	2.41	<b>0.40</b>	0.76	2.04	<b>0.63</b>	0.27	2.31	<b>0.49</b>	0.61	2.81	<b>0.39</b>	<b>0.48</b>	
Frankfurter	1951	0.33	1.65	<b>0.31</b>	0.73	1.73	<b>0.63</b>	0.28	1.19	<b>0.35</b>	0.65	2.71	<b>0.60</b>	<b>0.47</b>	
Burger	1978	0.50	2.44	<b>0.38</b>	0.70	2.00	<b>0.57</b>	0.29	2.63	<b>0.57</b>	0.51	2.88	<b>0.37</b>	<b>0.47</b>	
Blackmun	1982	0.52	2.17	<b>0.43</b>	0.65	2.06	<b>0.51</b>	0.26	2.21	<b>0.55</b>	0.50	2.71	<b>0.40</b>	<b>0.47</b>	
O'Connor	1994	0.54	1.97	<b>0.45</b>	0.61	2.28	<b>0.49</b>	0.26	2.03	<b>0.54</b>	0.48	2.34	<b>0.40</b>	<b>0.47</b>	
Reed	1948	0.32	1.66	<b>0.30</b>	0.73	1.74	<b>0.63</b>	0.28	1.19	<b>0.35</b>	0.65	2.67	<b>0.60</b>	<b>0.47</b>	
Harlan	1963	0.41	1.99	<b>0.38</b>	0.78	1.74	<b>0.57</b>	0.33	1.73	<b>0.47</b>	0.63	2.59	<b>0.47</b>	<b>0.47</b>	
Douglas	1957	0.38	1.49	<b>0.32</b>	0.73	1.57	<b>0.59</b>	0.30	1.07	<b>0.38</b>	0.67	2.55	<b>0.57</b>	<b>0.47</b>	
Stevens	1992	0.54	1.95	<b>0.43</b>	0.60	2.09	<b>0.48</b>	0.26	1.96	<b>0.53</b>	0.49	2.41	<b>0.40</b>	<b>0.46</b>	
Stewart	1970	0.45	1.82	<b>0.38</b>	0.77	1.63	<b>0.54</b>	0.32	1.52	<b>0.48</b>	0.59	2.75	<b>0.44</b>	<b>0.46</b>	
White	1978	0.45	1.98	<b>0.40</b>	0.61	2.23	<b>0.40</b>	0.29	2.10	<b>0.53</b>	0.59	2.34	<b>0.46</b>	<b>0.45</b>	
Clark	1958	0.30	2.57	<b>0.38</b>	0.65	2.34	<b>0.30</b>	0.28	2.85	<b>0.53</b>	0.69	2.32	<b>0.58</b>	<b>0.45</b>	
Thomas	2000	0.51	2.72	<b>0.41</b>	0.43	2.74	<b>0.37</b>	0.26	3.34	<b>0.61</b>	0.48	1.99	<b>0.40</b>	<b>0.45</b>	
Scalia	1997	0.49	2.53	<b>0.41</b>	0.49	2.58	<b>0.38</b>	0.22	2.89	<b>0.59</b>	0.46	2.40	<b>0.41</b>	<b>0.45</b>	
Souter	1999	0.53	2.20	<b>0.43</b>	0.51	2.46	<b>0.39</b>	0.25	2.27	<b>0.55</b>	0.41	2.48	<b>0.39</b>	<b>0.44</b>	
Kennedy	1998	0.55	1.95	<b>0.45</b>	0.54	2.27	<b>0.42</b>	0.25	1.71	<b>0.50</b>	0.38	2.79	<b>0.39</b>	<b>0.44</b>	
Fortas	1967	0.37	1.71	<b>0.35</b>	0.74	2.01	<b>0.64</b>	0.30	1.21	<b>0.39</b>	0.59	3.01	<b>0.38</b>	<b>0.44</b>	
Marshall	1979	0.52	1.15	<b>0.40</b>	0.68	1.87	<b>0.57</b>	0.31	0.72	<b>0.36</b>	0.47	2.83	<b>0.38</b>	<b>0.43</b>	
Powell	1980	0.47	2.27	<b>0.38</b>	0.57	2.50	<b>0.37</b>	0.24	2.58	<b>0.56</b>	0.52	2.59	<b>0.38</b>	<b>0.42</b>	
Roberts	2007	0.45	2.69	<b>0.37</b>	0.29	2.80	<b>0.36</b>	0.23	3.21	<b>0.55</b>	0.35	2.29	<b>0.36</b>	<b>0.41</b>	
Alito	2007	0.47	2.56	<b>0.37</b>	0.30	2.64	<b>0.36</b>	0.24	2.70	<b>0.52</b>	0.28	2.75	<b>0.36</b>	<b>0.40</b>	
Brennan	1974	0.46	1.16	<b>0.35</b>	0.72	1.66	<b>0.57</b>	0.31	0.62	<b>0.29</b>	0.53	2.95	<b>0.38</b>	<b>0.40</b>	
Jackson	1948	0.33	1.45	<b>0.28</b>	0.75	1.70	<b>0.61</b>	0.30	0.82	<b>0.29</b>	0.61	2.80	<b>0.38</b>	<b>0.39</b>	
Breyer	2001	0.50	1.79	<b>0.38</b>	0.35	2.48	<b>0.35</b>	0.26	1.49	<b>0.44</b>	0.32	2.80	<b>0.38</b>	<b>0.39</b>	
Minton	1953	0.29	2.16	<b>0.34</b>	0.69	2.14	<b>0.33</b>	0.27	1.84	<b>0.45</b>	0.61	2.86	<b>0.38</b>	<b>0.37</b>	
Goldberg	1964	0.28	2.10	<b>0.33</b>	0.70	2.21	<b>0.33</b>	0.25	1.84	<b>0.43</b>	0.61	2.88	<b>0.38</b>	<b>0.37</b>	
Ginsburg	2001	0.51	1.31	<b>0.37</b>	0.45	2.30	<b>0.37</b>	0.27	0.74	<b>0.29</b>	0.28	3.07	<b>0.38</b>	<b>0.36</b>	
Average		0.43	2.02	<b>0.38</b>	0.63	2.12	<b>0.48</b>	0.27	1.92	<b>0.47</b>	0.53	2.61	<b>0.44</b>	<b>0.44</b>	

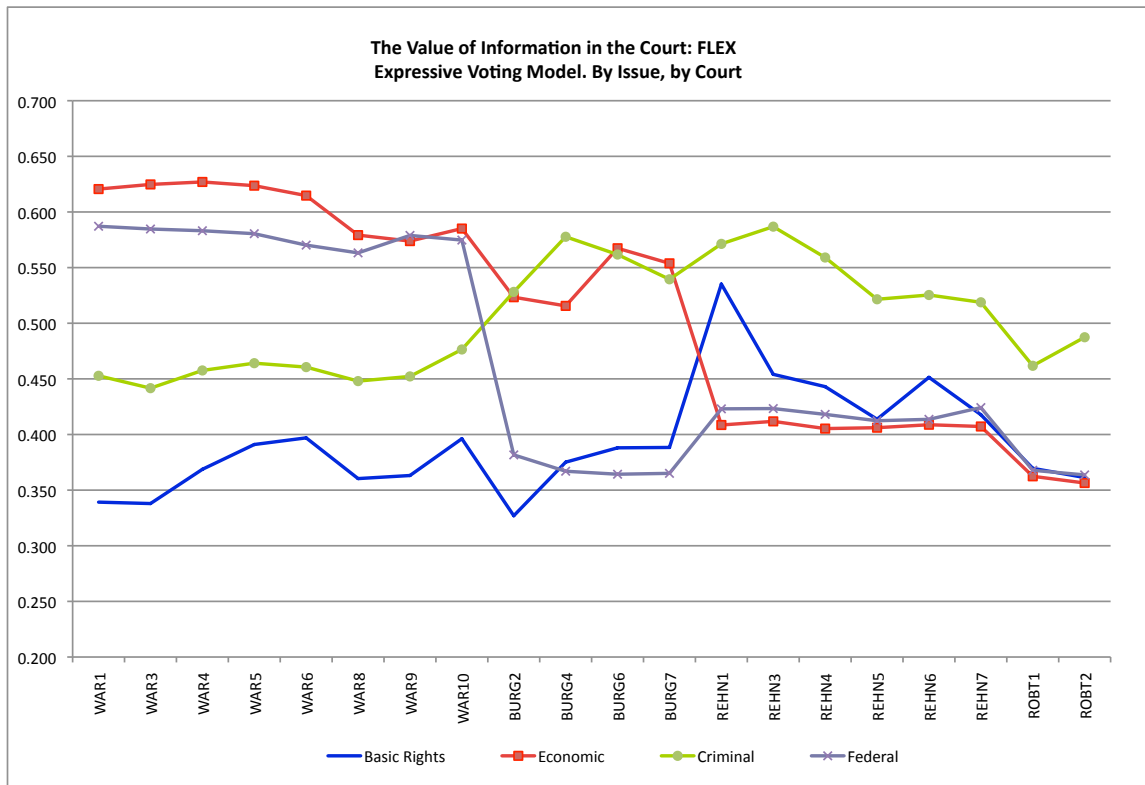


Figure 1: The Value of Information in the Court (FLEX), by Court and Issue Area (Expressive Voting Model, Private Parties, Statutory Interpretation, Lower Courts Agree).



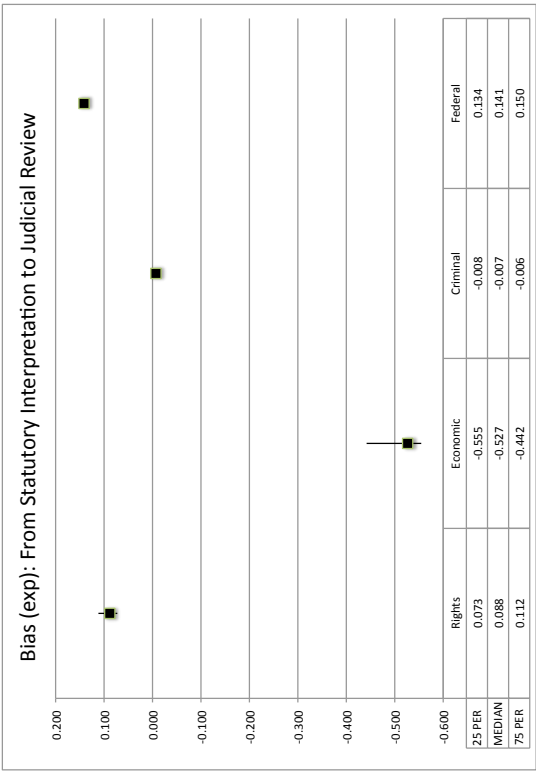
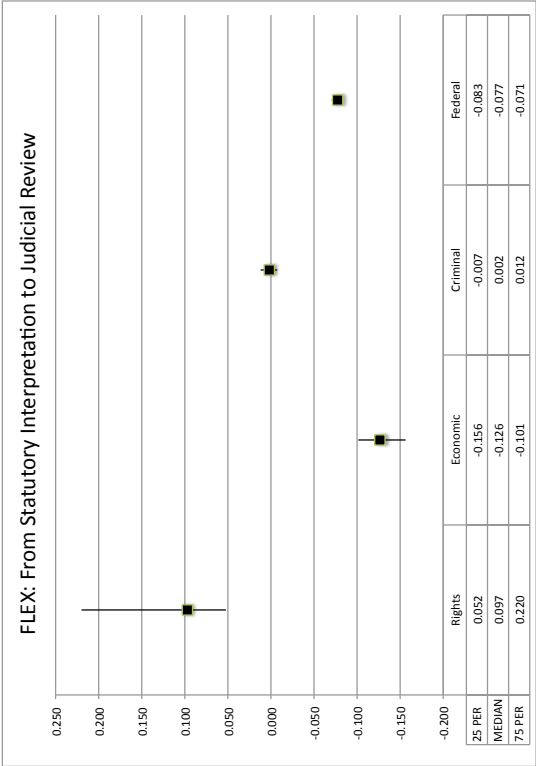
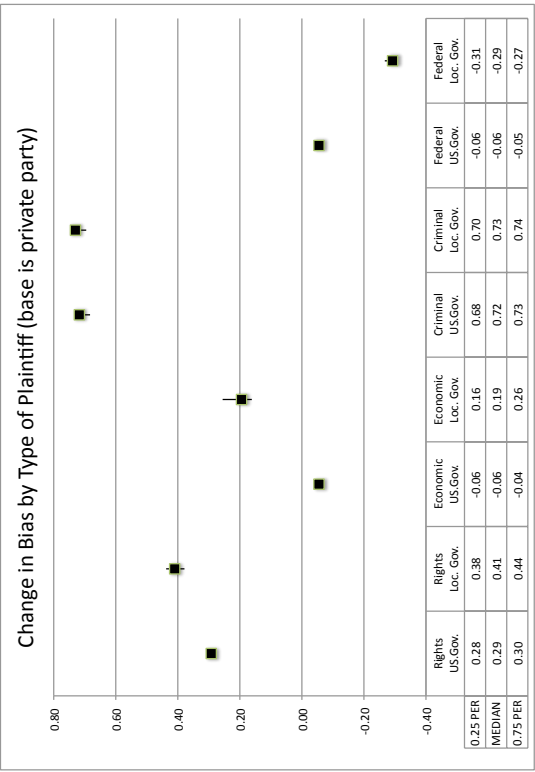
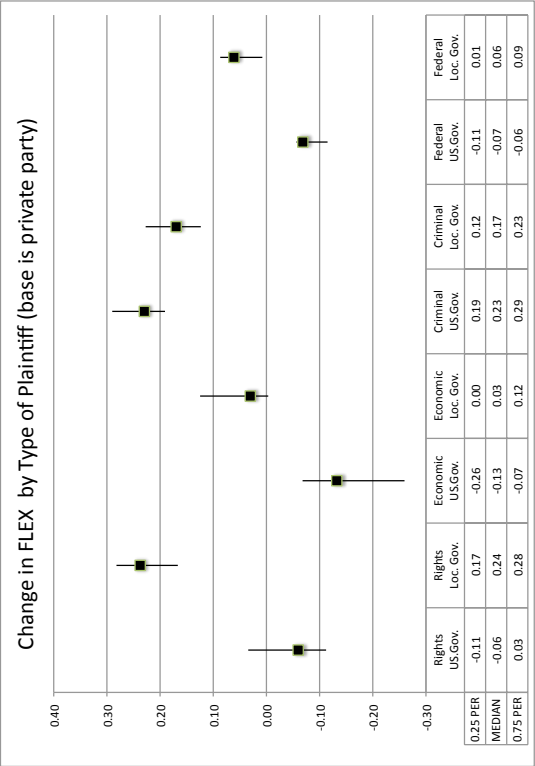


Figure 2: Comparative Statics

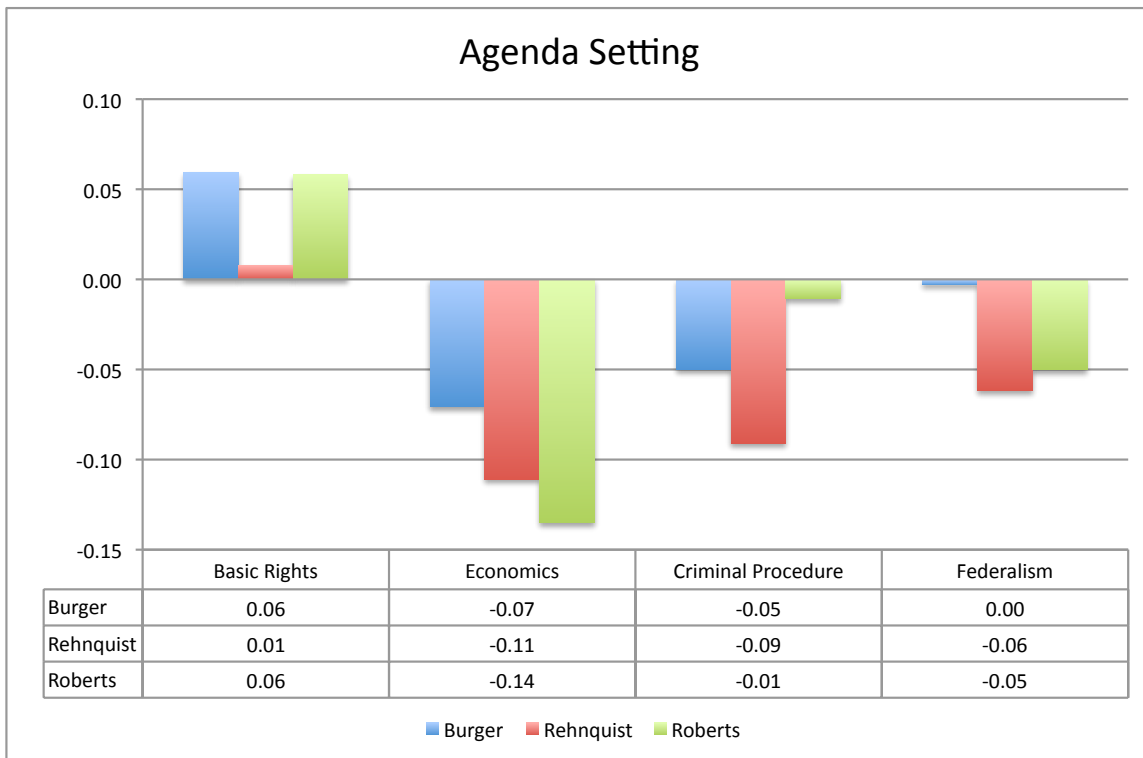


Figure 3: Agenda Setting Power: change in ex ante probability that the law favors the Plaintiff. By Chief Justice, vis a vis Chief Justice Warren. (Private Parties, Statutory Interpretation, Lower Courts Agree).

Rehnquist 1994 - 2004 Court (REHN7)

Private Plaintiff vs Private Defendant, lower Courts Agree, Statutory Interpretation

	Basic Rights	Economics	Criminal	Federalism
No. of Cases	41	31	5	10
No. of Violations	68	44	24	56

US Gov. Plaintiff, vs Private Defendant, lower Courts Agree, Statutory Interpretation

	Basic Rights	Economics	Criminal	Federalism
No. of Cases	12	2	1	9
No. of Violations	30	12	-	16

Private Plaintiff, vs US. Gov. Defendant, lower Courts Agree, Statutory Interpretation

	Basic Rights	Economics	Criminal	Federalism
No. of Cases	9	13	34	3
No. of Violations	48	38	70	0

Local Gov. Plaintiff, vs Private Defendant, lower Courts Agree, Statutory Interpretation

	Basic Rights	Economics	Criminal	Federalism
No. of Cases	12	2	4	2
No. of Violations	4	0	0	0

Private Plaintiff, vs Local Gov. Defendant, lower Courts Agree, Statutory Interpretation

	Basic Rights	Economics	Criminal	Federalism
No. of Cases	12	4	10	6
No. of Violations	12	0	2	0

Private Plaintiff vs Private Defendant, lower Courts Don't Agree, Statutory Interpretation

	Basic Rights	Economics	Criminal	Federalism
No. of Cases	44	46	1	5
No. of Violations	70	60	-	0

Note: First entry in each cell is the number of cases in each subsample, and bottom entry is the number if pairs of justices for which "flip-flopping" (as discussed in Section 6.2.2) is observed in each subsample of cases. Note that, among the nine Supreme Court justices, the maximum number of pairwise violations is  $9 \times 8 = 72$ .

Table 5: Testing the Informative Pre-Vote Deliberation Model: Incidence of "flip-flops" among pairs of justices, in small subsamples of similar cases.

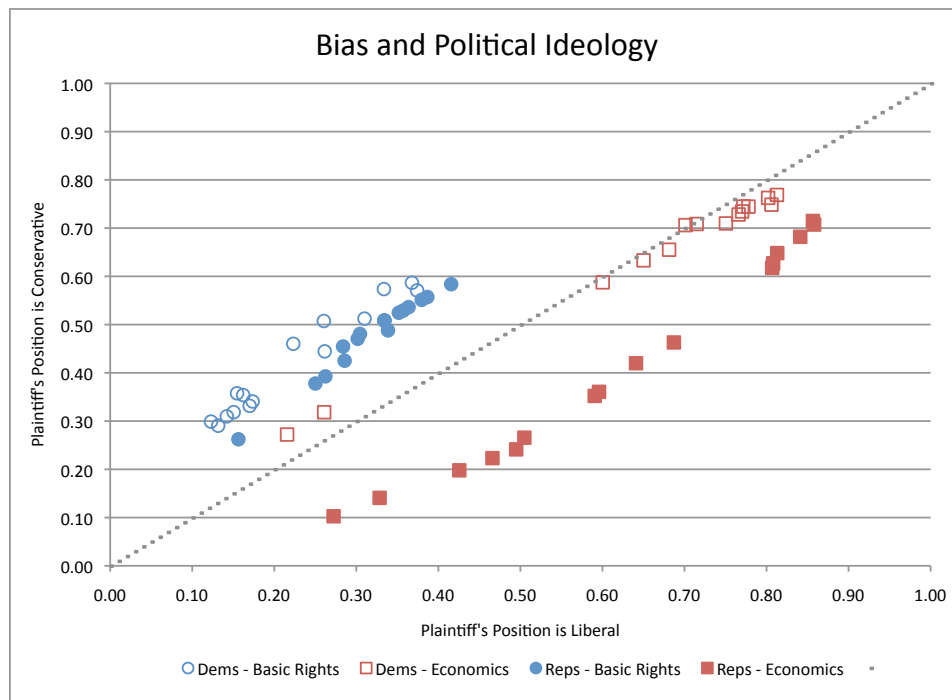


Figure 4: Bias (expressive voting model) per justice, in Economics and Basic Rights' issues. Voting for the Plaintiff is a liberal (conservative) vote in the horizontal (vertical) axis.

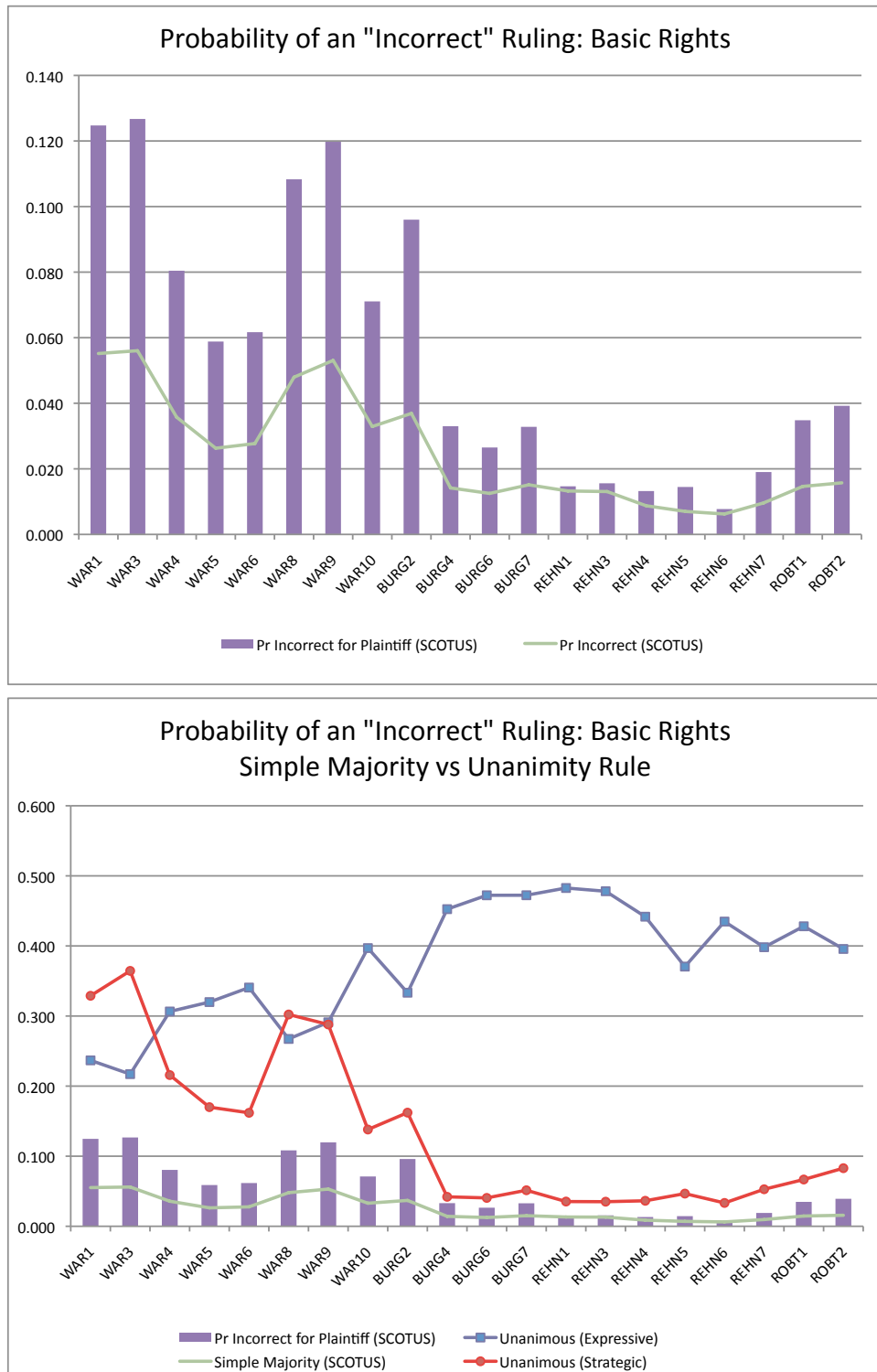


Figure 5: Counterfactual Prediction: Unanimity vs Simple Majority (Basic Rights)