The Value of Information in the Court.

Get it Right, Keep it Tight.*

Matias Iaryczower and Matthew Shum†

September 30, 2009

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 that in roughly 44% of cases, justices' initial leanings are changed by their personal assessments of the case. Our results also indicate that the value of information has decreased over time. We perform counterfactual simulations to draw implications for institutional design. *JEL Classification*: D72, D78, C13.

*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). We thank Ernesto Dal Bo, John Matsusaka, Jean-Laurent Rosenthal, and participants at the 21^{st} Stony Brook Game Theory Festival, and the Empirical-Micro Workshops at Claremont-McKenna College, USC, and Vanderbilt University for useful comments to previous versions of the paper.

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

> "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*

> "Even if he were mediocre, there are a lot of mediocre judges and people and lawyers. They are entitled to a little representation, aren't they, and a little chance? We can't have all Brandeises, Frankfurters and Cardozos." *U.S. Senator Roman L. Hruska.*[1]

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 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."[2] 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 the vote of Senators of whether to confirm or not a nominee (Segal et al. (1990, 1992), Epstein et al. (2006)). Qualifications matter because the decision-making process in the Court is not *only* ideological.

[1] The Supreme Court: A Seat for Mediocrity? Monday, Mar. 30, 1970. In reference to President Nixon's Supreme Court nominee Judge G. Harrold Carswell.

[2] From the statement submitted to the Senate Committee on the Judiciary by Justice Ruth Ginsburg.

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.[3] 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, each justice observes a private signal, which reflects her understanding of the particulars of the case. The precision (θ_i) of each justice's signal measures her ability to map the specifics of the case 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 cases. In the model, this bias or ideology boils down to a threshold π_i such that the justice prefers to rule for the Plaintiff if and only if the probability that the law favors the Plaintiff is at least

[3]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).

π_i. 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 $\pi_i \approx 0$ (or $\pi_i \approx 1$), justice i will vote almost completely in line with her ideology. On the other hand, when $\pi_i = 1/2$ for all i, the setting boils down to an unbiased, pure common values model.

In the estimation, we recover the values of $(\theta_i, \pi_i)|X$ for each justice i conditioning on observable covariates X of the cases (type of issue, characteristics of the Plaintiff and the Defendant, disposition of the lower courts) and the justices (judicial experience, court experience, political party, etc). 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 $\{(\theta_i, \pi_i)\}_{i=1}^n|X$, 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 that what he or she would vote 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. On the flip side, we also show that since the Warren court, Supreme Court justices in general have become less open-minded: average FLEX scores decreased 21 % in Economic cases, 36% in Criminal Procedure, and 30% in Federalism. Only in basic rights cases did FLEX scores increase over time, by a modest 8%. These trends are consistent with the increasing politicization of confirmations of nominees to the Supreme Court in recent years, as emphasized by the literature (see in particular Epstein et al. (2006)).

In other results, we compare the value of information across justices, issues, and periods of the court. In particular, we show that the value of information in the court is lower in areas in which ideological considerations tend to weight more heavily, such as Basic Rights and Criminal Procedure, and higher in Federalism and Economics. 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.[4] Note that in a small committee such as the Court, composed of an heterogeneous group of individuals in terms of both preferences and abilities, majority rule does not generically dominate unanimity rule (see for example Feddersen and Pesendorfer (1998), Duggan and Martinelli (2001), and Meirowitz (2002)). In fact we show that for Basic Rights cases, majority rule outperforms unanimity rule in the Warren courts, but would have reduced total error rates in several Rehnquist courts. Similar conclusions can be drawn in other issues.

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 et al. (1995) expands the coverage of the original SC scores,

[4]This is not only an interesting theoretical exercise, but also an initial exploration of the effects of a potentially desirable institutional change. Increasing the threshold to overturn the standing decision would tend to increase the number of cases that the court hears, up from the currently 1% of the requests that it receives each year (approximately 80 out of 8000). Changing the majority rule would of course also change the decisions of the court. This second element is what we address here.

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 et al. (1998) argue that the preferences of justices - as measured by the proportion of liberal votes on civil liberties cases - changes through time.[5] Finally, Landes and Posner (2008) 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 et al. (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 et al. (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.[6,7]

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.[8] In this paper, we therefore structurally estimate a model that

[5] See also Segal and Spaeth (1993) and Epstein and Knight (1997).

[6] 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 et al. (2008), and Kawai and Watanabe (2009) also perform structural estimation of strategic voting (ie. "pivotal voting") models, with ideological voters.

[7] See Lim (2008) for structural estimation of a model that incorporates career concerns into judges' behavior.

[8] The SVM does not preclude, however, a publicly known valence differential between alternatives. See Londregan

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.[9,10] 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 et al. (2009), who model strategic voting and common values in Congress.[11] 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, $i = 1, \ldots, n$, who consider T independent cases, $t = 1, \ldots, T$. 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

(1999).

[9] 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.

[10] 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); Spiller and Gely (1992)). 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 Daughety and Reinganum (2006) for an information-based model explaining how appellate judges may influence the cases that the SC chooses to hear.

[11] In Chiang and Knight (2008), common values enter a non-strategic model in which voters gain information about candidates from newspaper endoresements.

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 $s_{it} = \omega_t + \sigma_i \varepsilon_t$, where $\varepsilon_t \sim \mathcal{N}(0,1)$. Here $\omega_t \in \{0,1\}$ in 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 $\theta_i = 1/\sigma_i$ 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 $\pi_i \in (0,1)$, justice i has a payoff of $-\pi_i$ 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 $-(1-\pi_i)$ 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 $\Pr^i(\omega_t = 1|E) \geq \pi_i$. 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 $\frac{\pi_i}{1-\pi_i}\frac{1-\rho}{\rho}$, where $\rho \equiv \Pr(\omega_t = 1)$ denotes justices' common prior probability of the unobserved state ω_t. Note that since ω_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. In particular, with $\pi_i \approx 0$ (or $\pi_i \approx 1$), justice i is almost always ideological. On the other hand, when $\pi_i = 1/2$ for all i, the setting boils down to an unbiased, pure common values model.[12]

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 $\Pr^i(\omega_t = P|s_{it}) \geq \pi_i$. Then E consists only of s_{it}, and i votes to rule against the Defendant if

$$\frac{\Pr(s_{it}|\omega_t = 1)}{\Pr(s_{it}|\omega_t = 0)} = \frac{\phi(\theta_i[s_{it} - 1])}{\phi(\theta_i s_{it})} \geq \frac{\pi_i}{1-\pi_i}\frac{1-\rho}{\rho} \tag{1}$$

Let s_i^{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 $s_{it} \geq s_i^{exp}$ and in favor of the Defendant otherwise. This cutoff point s_i^{exp} completely characterizes

[12] In our setting, justices share common priors, but their ideological biases are captured by the π_i parameters. See Froeb and Kobayashi (1996) for a model where justices' biases are manifested in their priors.

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

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

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

$$\frac{P_{\sigma_{-i}}(piv_i|\omega_t = 1)}{P_{\sigma_{-i}}(piv_i|\omega_t = 0)} \frac{\phi(\theta_i[s_{it} - 1])}{\phi(\theta_i s_{it})} \geq \frac{\pi_i}{1 - \pi_i} \frac{1 - \rho}{\rho} \tag{3}$$

As before, the MLRP implies that the best response to any strategy σ_{-i} of the remaining justices is a cutoff strategy, such that i rules against the Defendant ($\sigma_i(s_i) = 1$) if s_{it} implies (3), and in favor of the Defendant ($\sigma_i(s_i) = 0$) otherwise.[13] This in turn implies that all responsive equilibria are cutoff equilibria; i.e., that any equilibrium is characterized by cutpoints s_i^* for each justice $i = 1, \ldots, n$ such that justice i votes against the Defendant if and only if $s_{it} \geq s_i^*$. Now, given cutoff strategies, $\Pr(v_{it} = 1|\omega_t) = \int \sigma_i(s)\phi(\theta_i[s - \omega_t])ds = [1 - \Phi(\theta_i[s_i^* - \omega_t])]$. Therefore from (3), and letting $\mathcal{C}_{R-1}^i$ denote the set of coalitions $C \subset N \setminus i$ with $R - 1$ members, $\{s_i^*\}_{i=1}^n$ is given by the n equations

$$\frac{\sum_{C \in \mathcal{C}_{R-1}} \left(\prod_{j \in C} [1 - \Phi(\theta_j[s_j^* - 1])] \right) \left(\prod_{j \neq i, j \notin C} \Phi(\theta_j[s_j^* - 1]) \right)}{\sum_{C \in \mathcal{C}_{R-1}} \left(\prod_{j \in C} [1 - \Phi(\theta_j s_j^*)] \right) \left(\prod_{j \neq i, j \notin C} \Phi(\theta_j s_j^*) \right)} \frac{\phi(\theta_i[s_i^* - 1])}{\phi(\theta_i s_i^*)} = \frac{\pi_i}{1 - \pi_i} \frac{1 - \rho}{\rho} \tag{4}$$

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

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

[13] The proof of this result follows Duggan and Martinelli (2001), and is included in the online appendix for convenience.

The likelihood functions for the expressive and the strategic models (Eqs. 2,5) are almost identical, except for the cutoff points: s^{exp} for the expressive model, and s^* for the strategic model.

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 (2008)). The second is a database of justice specific information comes from the *United States Supreme Court Justice Database* (Epstein et al. (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.[14] 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 Judical Power), *Economic* (includes Economic Activity, Unions and Attorneys) and *Federal* (includes Federalism, Interstate Relations, and Federal Taxation).[15] 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,

[14] 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 analu = 0 (this conforms to standard practice).

[15] A fifth residual category groups Miscellaneous cases (Spaeth's issues 980-99).

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.[16] The lower panel of Table 1 (in the Appendix) summarizes this information for each justice in our data.

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 modelling 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. In this section, we present an argument as to the identification of these model parameters from the observed vote data.

Our 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 ρ, 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 some covariates X, and in our empirical work below, we will control for case heterogeneity using a rich set of covariates.

First step. We introduce the following notation:

[16] 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 et al. (1990, 1992), and Epstein et al. (2006)).

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

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:

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

Conditional on the state ω_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 ρ. Identification of the state-specific voting probabilities $\{\gamma_{i,1}, \gamma_{i,0}\}_{i=1}^n$ and the mixing probability ρ are available in, eg., Hall and Zhou (2003), Hu (2008), and Kasahara and Shimotsu (2007).

In the case of the US Supreme Court, where there are $n = 9$ justices, 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 ρ) to estimate. Since $2^9 >> 19$, the relevant necessary condition for identification is satisfied. 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.[17]

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, π_i and θ_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 $\varepsilon_{it} \sim \mathcal{N}(0, 1)$.

[17] Moreover, the inequality $\gamma_{i,1} > \gamma_{i,0}$, which is implied by the monotone likelihood ratio property, is crucial for identification: without this assumption, the voting probabilities would only be identified up to an arbitrary classification of ω_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.

Then $\gamma_{i,1} \equiv 1 - \Phi\left(\theta_i[s_i^* - 1])\right)$ and $\gamma_{i,0} \equiv (1 - \Phi(\theta_i s_i^*))$. Solving these equations for θ_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 [18]

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

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 (γ_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 π_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

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

while in the strategic voting model this is given by

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

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

Note that, in recovering θ_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 π_i. This distinction between θ_i and π_i is a remarkable property of this problem. It implies that

[18] Note that for each justice, we use the estimates of $\gamma_{i,0}, \gamma_{i,1}$ to recover the two quantities θ_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.

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 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 ρ_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 which 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 (β,δ,η) as well as the justice-specific variables (α_i,ζ_i) for $i = 1,\ldots,n$. For this, we maximize the following likelihood function

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

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 θ_{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 θ and π.

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.[19]

In our empirical model, this endogenous case selection is not explicitly accommodated, and will thus be captured in the parameter ρ 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 β 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 ρ 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 examing the coefficients on these variables.

For convenience, we will continue to refer to ρ as justices' prior beliefs in discussing the estimation results, but the ambiguity in identification should be kept in mind.

[19] See Daughety and Reinganum (2006) for many details on the SCOTUS case selection process.

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.[20] The covariates 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.[21] 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

[20] 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 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).

[21] 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).

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), merits 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]

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" θ_{it} and π_{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-lived natural court in our data: REHN7, with chief justice Rehnquist, between 1994 and 2004.

[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 ($\rho = 0.333$).[22] 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 (γ_{it0}) and when the law favors the Plaintiff (γ_{it1}). Thus, taking Economics for example, justice Breyer had a probability of $\gamma_{it1} = 0.869$ of correctly ruling for the Plaintiff, and a probability of $1-\gamma_{it0} = 1-0.110 = 0.890$ 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 (0.869 vs 0.685) and a smaller probability of ruling for the Plaintiff when the law favors the Defendant (0.110 vs 0.124).

The fourth column presents the equilibrium cutpoint. Thus, taking again Economics for example, Justice Breyer would vote for the Plaintiff after observing a signal $s_i \geq 0.523$, but it would take more evidence (a signal above 0.613) 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 0.706). 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 ρ, 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 ($\pi^{exp}_{BRE} = 0.640$) than Justice Souter ($\pi^{exp}_{SOU} = 0.722$). Justice Stevens requires more evidence (a belief of at least 0.731 that the law

[22] 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.

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 ρ, outweights the private bias π_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 $\rho < \pi_i$):

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

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 ($\pi \to 0$) or for the Defendant ($\pi \to 1$). Note also that the computation of FLEX for the expressive and strategic voting models differ only in whether we use π_i^{exp} or π_i^{st} 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 θ_i, and differ only in the biases π_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.[23]

The FLEX scores for the expressive and strategic voting models are presented in columns 8 and 9 of the table. To continue with the Economics example, note that the rankings of precision and equilibrium cutpoints among Breyer, Souter and Stevens that we discussed above ($\theta_{BRE} > \theta_{SOU} > \theta_{STE}$, and $1/2 < s^*_{BRE} < s^*_{SOU} < s^*_{STE}$) lead to a similar ranking of FLEX scores: 57% for Breyer, 52% for Souter, and 47% for Stevens. Note also that while Souter has a higher precision than Ginsburg (2.115 vs. 2.057) Ginsburg has a higher FLEX score because Souter's larger bias leads to a more demanding threshold s_i^*.

[23] If instead we were initially given values of $\{\pi_i, \theta_i\}$ and ρ, then the two models would imply a different equilibrium cutpoint s_i^*, and FLEX scores in the two models would differ significantly.

Armed with our measure of the value of information in the court for all justices and case characteristics, we can now address several key questions about the determinants of supreme court decisions.

How has the open-mindedness of justices changed over time? Figure 1 shows the median value of FLEX, precision, and bias in the expressive and strategic model per court. The results show that, from the beginning of the Warren Court to the second Roberts court, median FLEX scores decreased 21% in Economic issues, 36% in Criminal Procedure, and 30 % in Federalism.[24] This is not an artifact of the beginning-to-end comparison. In Economic issues, FLEX decreased by 10% between the Warren and the Burger courts, an additional 5% between the Burger and the Rehnquist courts, and 5% more between the Rehnquist and Roberts court. Similarly, in Criminal Procedure the corresponding numbers were 4%, 8% and 33%. In Federalism, FLEX decreased 35% between the Warren and the Burger courts. The 14% increase between the Burger and the Rehnquist courts was not enough to reverse this trend (FLEX decreased an additional 3% between the Rehnquist and Roberts courts). The sole increase in FLEX is within Basic Rights. Here FLEX increased a total of 8% beginning to end, because of a 9% increase between the Warren and Burger courts and a 21% increase between the Burger and Rehnquist courts. However, it again decreased 22% between the Rehnquist and the Roberts courts. Thus at least compared to the Rehnquist court, the Roberts court is less open minded uniformly in all issue areas.

At the macro level, the FLEX measure moves in an opposite direction as justices' biases (π). From Table 4 below, we see that for basic rights cases, π has become closer to 0.5 over time, implying that justices have become less biased, which is in line with the finding that FLEX has gone up over time for these cases. For the other three types of cases, π has stayed either unchanged or gotten farther from 0.5, implying that justices' votes have become more ideological over time. This is in line with the downward trend in FLEX over time for these cases.

On the whole, these findings confirm the arguments of some Supreme Court scholars that the confirmation process for nominees have become more politicized over time. In particular, Epstein et al. (2006) argue that the trend toward greater attention to ideology started not with (the failed confirmation of) Robert H. Bork in 1987, but instead earlier, with appointments to the Warren

[24]These numbers use the expressive voting model. The corresponding figures for the strategic voting model are 23%, 36% and 32%).

Court. Our results in Figure 1 are in line with these arguments.

[Figure 1 about here]

Variations in open-mindedness across justices and issues Having considered the macro-level trends in FLEX over time, we next explore how our measures of FLEX vary across justices, and across cases. Table 4 contains the average FLEX scores for each justice across different courts and issues.

[Table 4 about here]

Note that with the exception of justice Burton, all FLEX scores are below 1/2. That is, in general, it is more likely than not that their vote will echo their initial leanings, based on their bias and priors. Having said this, the 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%.

In addition, both the literature and a cursory glance at the data suggests that we should expect both the bias and possibly also the precision estimates to vary greatly across different issue areas. The results confirm the expectations. The typical FLEX score in Economic issues (53%) is much larger than in the remaining areas. In fact, this dominance also holds regarding the entire distribution: the 15 % percentile of Economic FLEX scores is also above the average FLEX scores in other areas. Federalism follows with an average FLEX score close to the overall average (45 %). FLEX is lowest for Basic Rights (40 %) and Criminal Procedure (39%). These results are consistent with the basic intuition that the value of information in the court is lower in areas in which ideological considerations tend to weight more heavily, such as Basic Rights and Criminal Procedure, and higher in issues in which ideology is relatively less important, as Federalism and Economics.[25]

[25]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.

At the same time, there is also significant variation among areas at the individual justice level. However, this variation does not necessarily coincide with the ranking of the averages; for example justices Clark, Powell and Minton have a larger FLEX score in Criminal Procedure than in Economic issues.

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.

The first way in which the Government affects the results 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).[26] But a Government Plaintiff does not only affect the priors. In the top two panels of Figure 3 we 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 figure the median and 25-75 percentiles in the distribution of these changes at the individual justice level.

[Figure 3 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 0.38 (0.28) in π^{exp} in Basic Rights, and to a median increase of 0.74 (0.72) in π^{exp} in Criminal cases. Now recall that our measure of ideological bias (π_i) is the parameter in justices' preferences that quantifies the relative cost of ruling incorrectly against the defendant (π_i) vis a vis ruling incorrectly against the plaintiff ($1-\pi_i$). Moreover π_i also has a second direct interpretation: it is the cutpoint such that justice i will rule in favor of

[26] 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 ρ by 0.15.

the plaintiff whenever the probability that the law favors the plaintiff is above π_i. Thus justices appear to have more stringent requirements for the Government, in particular within Criminal Procedure.

In the lower panels of Figure 3, 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.

Agenda Setting. As we mentioned before, the parameter ρ 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 4 shows the difference between the prior ρ 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 4 shows that the 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).

[Figure 4 about here]

A second institutional detail related to agenda setting more broadly defined 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.[27] To address this

[27]The question is more subtle than what it seems at first sight. Because even in the simultaneous voting game individuals condition on the event of being pivotal (in the strategic voting model) observing the sequence of votes might not offer any additional *useful* information. This is the main argument in Dekel and Piccione (2000), who then show under some conditions an equivalence of the set of equilibria between the simultaneous and sequential

possibility, we also estimated an alternative version of the model in which justices vote sequentially. The results are remarkably similar to those obtained from the simultaneous voting model, and are not reported here.[28]

6.2 Mistakes and Implications for Institutional Design

In any given case, the Court comes up with 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, $\Pr(v=p|\omega=0)$.[29] Given a prior ρ, we can then compute the total probability of an incorrect ruling for the Supreme Court, β^{SC}

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

voting games. In our case, however, justices are heterogeneous, so observing the voting sequence might reveal useful information.

[28] A related issue is the possibility of information transmission in deliberation (see Coughlan (2000), Austen-Smith and Feddersen (2005, 2006), and Gerardi and Yariv (2007)). While this is certainly an issue in pure common value settings, it is less clear that in a environment in which individuals can disagree significantly deliberation will convey relevant information in equilibrium. In this paper we do not allow for this possibility. We plan to address this important issue in future research.

[29] 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 $\Pr(v=p|\omega=0) = \sum_{k=5}^{9}\sum_{C\in\mathcal{C}(k)}\prod_{i\in C}\gamma_{i,0}\prod_{i\notin C}(1-\gamma_{i,0})$.

The upper panel of Figure 5 shows β^{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.) Both the total probability of error β^{SC} and the conditional probability of error $\Pr(v = p|\omega = 0)$ are around 6% throughout our sample, with significant variations across different courts.

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 β^{U}_{exp}

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

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 $\{(\pi_i^{st}, \theta_i)\}$ 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 β^{U}_{ST} is

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

The lower panel of Figure 5 puts everything together (again, for Basic Rights issues). In the expressive voting model, unanimity with heterogeneous justices leads to very large error rates, almost entirely due to rulings that incorrectly favor the Defendant. In the strategic voting model, the comparison is more involved. Unanimity rule here sometimes leads to a lower total error, because of the lower probability of incorrectly ruling for the Defendant than simple majority rule.

The reason why majority rule outperforms unanimity rule in the Warren courts is that justices' biases are larger here than in the Burger, Rehnquist or Roberts courts (see Figure 2). This increases the probability of ruling incorrectly for the Plaintiff in both majority rule and unanimity, but it has a far more devastating effect in unanimity. The lower median bias in the later courts reduces this effect. On the other hand, unanimity diminishes the probability of ruling incorrectly for the Defendant. In the Rehnquist courts, this effect outweights the first and generates a lower error rate than simple majority rule.

A similar analysis for the remaining issues confirm these conclusions (see Figure 5). In particular, we find that in general we cannot rank unanimity and simple majority in terms of the overall error rate. In Economics, as in Basic Rights, simple majority dominates in all courts in the expressive voting model, but this conclusion does not extend to the strategic voting model. In Criminal procedure, in turn, simple majority rule dominates in all courts in the strategic voting model, but this conclusion does not extend to the expressive voting model. Only in Federalism does simple majority rule lead to a smaller error rate in all courts, in both the strategic and the expressive voting models. All in all, we conclude that when justices are very biased, unanimity rule leads to large errors, and majority rule dominates. When justices are less biased, one rule or the other can be better depending on the parameters of the problem. With heterogeneous individuals in a small committee, majority rule does not generically dominate unanimity rule.[30]

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

[30]See for example Feddersen and Pesendorfer (1998), Duggan and Martinelli (2001), and Meirowitz (2002).

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. On the flip side, we also show that since the Warren court, Supreme Court justices in general became less open-minded: average FLEX scores decreased 21 % in Economic cases, 36% in Criminal Procedure, and 30% in Federalism (and increased 8% in Basic Rights). These trends are consistent with the politicization of confirmations of nominees to the supreme court emphasized by the literature.

In closing, we want to emphasize some of the limitations of our results. Possibly the most important of these is that our analysis paints a necessarily incomplete picture of the court. While the final up or down decision on which we focus is undoubtedly an important part of court rulings, a second element is also crucial: the opinions of the court. Because of the principle of *stare decisis*, lower court judges must follow the precedents set by the Supreme Court, as well as their written justification. Thus the opinions are particularly important in terms of the long run implications of the court rulings. Clearly our estimates of bias, ability to infer the meaning of the law, and of the value of information in the court are limited to the voting decisions of the justices, and do not speak about the determinants of the opinions, or their implications. We plan to address this in future research.

References

Austen-Smith, David and Jeffrey S. Banks, "Information Aggregation, Rationality, and the Condorcet Jury Theorem," *American Political Science Review*, 1996, *90*, 34–45.

_ **and Tim Feddersen**, "Deliberation and Voting Rules," 2005. In Social Choice and Strategic Decisions: Essays in Honor of Jeffrey S. Banks, pp. 269–316.

_ **and** _ , "Deliberation, Preference Uncertainty, and Voting Rules," *American Political Science Review*, May 2006, *100*, 209–218.

Chiang, Chun-Fang and Brian G. Knight, "Media Bias and Influence: Evidence from Newspaper Endorsements," 2008. NBER Working Paper No. 14445.

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

_ **and** _ , "Testing Explanations of Strategic Voting in Legislatures: A Reexamination of the Compromise of 1790," *American Journal of Political Science*, 2004, *48* (4), 675–689.

_ **, Simon Jackman, and Douglas Rivers**, "The Statistical Analysis of Roll Call Data," *American Political Science Review*, May 2004, *55*, 355–370.

Coate, Stephen and Michael Conlin, "A Group Rule–Utilitarian Approach to Voter Turnout: Theory and Evidence," *American Economic Review*, 2004, *94*, 1476–1504.

_ **,** _ **, and Andrea Moro**, "The performance of pivotal-voter models in small-scale elections: Evidence from Texas liquor referenda," *Journal of Public Economics*, 2008, *92*, 582–596.

Coughlan, Peter, "In Defense of Unanimous Jury Veredicts: Mistrials, Communication, and Strategic Voting," *American Political Science Review*, June 2000, *94*, 375–393.

Daughety, Andrew and Jennifer Reinganum, "Speaking Up: A Model of Judicial Dissent and Discretionary Review," *Supreme Court Economic Review*, 2006, *14*, 1–51.

Degan, Arianna and Antonio Merlo, "Do Voters Vote Ideologically?," 2008. Working Paper, U. Penn.

de Paula, Aureo and Antonio Merlo, "Identification and Estimation of Preference Distributions when Voters are Ideological," 2009. Working Paper, U. Penn.

Dekel, Eddie and Michele Piccione, "Sequential Voting Procedures in Symmetric Binary Elections," *Journal of Political Economy*, 2000, *108*, 34–54.

Duggan, John and Cesar Martinelli, "A Bayesian Model of Voting in Juries," *Games and Economic Behavior*, 2001, *37*, 259–294.

Epstein, Lee and Carol Mershon, "Measuring Political Preferences," *American Journal of Political Science*, Feb 1996, *40* (1), 261–294.

_ **and Jack Knight**, *The Choices Justices Make*, New York, NY: CQ Press, 1997.

_ **, Rene Lindstadt, Jeffrey A. Segal, and Chad Westerland**, "The Changing Dynamics of Senate Voting on Supreme Court Nominees," *The Journal of Politics*, 2006, *68*, 296–307.

_ **, Thomas G. Walker, Nancy Staudt, Scott A. Hendrickson, and Jason M. Roberts**, "The U.S. Supreme Court Justices Database (computer file)," 2008. http://epstein.law.northwestern.edu/research/justicesdata.html.

_ , **Valerie Hoekstra, Jeffrey Segal, and Harold Spaeth**, "Do Political Preferences Change? A Longitudinal Study of U.S. Supreme Court Justices," *The Journal of Politics*, Aug 1998, *60* (3), 801–818.

Feddersen, Timothy and Wolfgang Pesendorfer, "Voting Behavior and Information Aggregation in Elections With Private Information," *Econometrica*, September 1997, *65*, 1029–1058.

_ **and** _ , "Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting," *American Political Science Review*, March 1998, *92*, 23–35.

Froeb, Luke M. and Bruce H. Kobayashi, "Naive, Biased, yet Bayesian: Can Juries Interpret Selectively Produced Evidence?," *Journal of Law, Economics, and Organization*, 1996, *12*, 257–275.

Gely, Rafael and Pablo Spiller, "A Rational Choice Theory of Supreme Court Statutory Decisions with Applications to the "State Farm" and "Grove City Cases"," *Journal of Law, Economics & Organization*, Oct 1990, *6* (2), 263–300.

Gerardi, Dino and Leeat Yariv, "Deliberative Voting," *Journal of Economic Theory*, May 2007, *134*, 317–338.

Hall, Peter and Xiao-Hua Zhou, "Nonparametric Estimation of Component Distributions in a Multivariate Mixture," *Annals of Statistics*, 2003, *31*, 201–224.

Heckman, James and James Snyder, "Linear Probability Models of the Demand for Attributes with an Empirical Application to Estimating the Preferences of Legislators," *The RAND Journal of Economics*, Jan 1997, *28* (0), S142–S189. Special Issue in Honor of Richard E. Quandt.

Hu, Yingyao, "Identification and Estimation of Nonlinear Models with Misclassification Error Using Instrumental Variables: a General Solution," *Journal of Econometrics*, 2008, *144*, 27–61.

Iaryczower, Matias, Gabriel Katz, and Sebastian Saiegh, "The not-so-popular branch: Bicameralism as a Counter-Majoritarian Device.," 2009. HSS, California Institute of Technology.

Kasahara, Hiroyuki and Katsumi Shimotsu, "Nonparametric Identification and Estimation of Multivariate Mixtures," Dec 2007. Queen's University, Department of Economics.

Kawai, Kei and Yasutora Watanabe, "Inferring Strategic Voting," May 2009. Northwestern University, Department of Economics.

Landes, William M. and Richard A. Posner, "Rational Judicial Behavior: A Statistical Study," April 2008. University of Chicago.

Lim, Claire, "Turnover and Accountability of Appointed and Elected Judges," 2008. Typeset, Stanford University.

Londregan, John, "Estimating Legislators' Preferred Points," *Political Analysis*, Jan 1999.

Martin, Andrew and Kevin Quinn, "Dynamic ideal point estimation via Markov chain Monte Carlo for the US Supreme Court, 1953-1999," *Political Analysis*, Jan 2002.

_ **and** _ , "Assessing preference change on the US Supreme Court," *Journal of Law, Economics and Organization*, Jan 2007, *23.*

Meirowitz, Adam, "Informative Voting and Condorcet Jury Theorems with a Continuum of Types," *Social Choice and Welfare*, Aug 2002, *19*, 219–236.

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

_ and _ , "Patterns of Congressional Voting," *American Journal of Political Science*, 1991, *35*, 228–278.

Segal, Jeffrey and Albert Cover, "Ideological Values and the Votes of U.S. Supreme Court Justices," *American Political Science Review*, Jun 1989, *83* (2), 557–565.

_ and Harold Spaeth, *The Supreme Court and the Attitudinal Model* 1993.

_ , Charles M. Cameron, and Albert Cover, "Senate Voting on Supreme Court Nominees: A Neoinstitutional Model," *American Political Science Review*, Jun 1990, *84* (2), 525–534.

_ , _ , and _ , "A Spatial Model of Roll Call Voting: Senators, Constituents, Presidents, and Interest Groups in Supreme Court Confirmations," *American Journal of Political Science*, Feb 1992, *36* (1), 96–121.

_ , Lee Epstein, Charles Cameron, and Harold Spaeth, "Ideological Values and the Votes of U.S. Supreme Court Justices Revisited," *The Journal of Politics*, Aug 1995, *57* (3), 812–823.

Spaeth, Harold J., "The Original United States Supreme Court Judicial Database, 1953-2007 Terms (computer file), ," 2008. http://www.cas.sc.edu/poli/juri/sctdata.htm/.

Spiller, Pablo and Rafael Gely, "Congressional Control or Judicial Independence: The Determinants of U.S. Supreme Court Labor-Relations Decisions, 1949-1988," *The RAND Journal of Economics*, Dec 1992, *23* (4), 463–492.

		Basic Rights		Economic		Criminal Procedure		Federal		Total	
Court	Years	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		
		ρ	γ_{it0}	γ_{it1}	ρ	γ_{it0}	γ_{it1}	ρ	γ_{it0}	γ_{it1}	ρ	γ_{it0}	γ_{it1}
	Constant	0.597	-0.184	7.106	0.341	-0.554	3.051	0.341	-1.627	6.171	0.045	-0.404	0.733
		(0.140)	(0.512)	(0.543)	(0.206)	(0.646)	(0.930)	(0.386)	(0.604)	(0.821)	(0.399)	(1.264)	(1.515)
Case Specific	Plaintiff Local Gov.	0.741	-1.420	-0.654	0.930	0.040	-0.599	2.901	-4.398	-0.790	-0.626	0.495	0.437
		(0.133)	(0.105)	(0.046)	(0.272)	(0.151)	(0.102)	(0.424)	(0.494)	(0.185)	(0.386)	(0.134)	(0.275)
	Plaintiff US Gov.	1.002	-0.329	-0.135	0.159	0.510	0.151	3.077	-2.530	-0.495	0.133	0.667	-0.133
		(0.177)	(0.090)	(0.064)	(0.180)	(0.073)	(0.092)	(0.460)	(0.272)	(0.189)	(0.312)	(0.140)	(0.160)
	Defendant Local Gov.	-0.686	0.863	0.379	-0.439	-0.360	-0.011	-0.814	0.698	1.004	0.459	-0.302	-0.302
		(0.124)	(0.049)	(0.072)	(0.197)	(0.079)	(0.114)	(0.379)	(0.120)	(0.186)	(0.333)	(0.175)	(0.171)
	Defendant US Gov.	-0.823	0.582	-0.469	-0.691	-0.457	-0.456	-0.397	0.388	0.246	-0.630	-0.759	-1.196
		(0.170)	(0.064)	(0.074)	(0.200)	(0.088)	(0.094)	(0.371)	(0.117)	(0.229)	(0.341)	(0.147)	(0.170)
	Lower Courts Agree	-0.084	0.422	-0.189	0.305	-0.673	-0.081	-0.206	0.002	-0.081	0.451	-0.251	0.210
		(0.097)	(0.039)	(0.039)	(0.142)	(0.061)	(0.068)	(0.174)	(0.053)	(0.056)	(0.241)	(0.109)	(0.129)
	Constitutional Review	-0.137	-0.048	-0.323	-0.062	0.388	1.996	-0.725	-0.312	0.095	0.594	-0.113	-0.004
		(0.191)	(0.080)	(0.062)	(0.403)	(0.142)	(0.830)	(0.292)	(0.081)	(0.097)	(0.617)	(0.220)	(0.306)
	Statutory Interpretation	-0.280	0.037	0.343	0.302	-0.889	-0.375	-0.441	-0.367	-0.018	0.043	-0.339	0.280
		(0.107)	(0.045)	(0.042)	(0.173)	(0.066)	(0.099)	(0.209)	(0.063)	(0.070)	(0.349)	(0.131)	(0.184)
	CJ Burger	0.246	0.126	0.580	-0.329	-0.227	-0.033	-0.208	-1.327	0.538	-0.013	0.308	-0.168
		(0.122)	(0.112)	(0.176)	(0.177)	(0.247)	(0.273)	(0.196)	(0.228)	(0.202)	(0.303)	(0.613)	(0.785)
	CJ Rehnquist	0.032	-0.139	0.361	-0.502	-0.235	0.758	-0.388	-1.411	0.542	-0.257	0.228	-0.329
		(0.128)	(0.147)	(0.193)	(0.174)	(0.273)	(0.316)	(0.198)	(0.265)	(0.230)	(0.309)	(0.669)	(0.880)
	CJ Roberts	0.239	0.002	0.501	-0.602	0.357	1.383	-0.044	-1.515	0.742	-0.211	0.770	-0.115
		(0.288)	(0.191)	(0.225)	(0.348)	(0.296)	(0.412)	(0.406)	(0.286)	(0.257)	(0.796)	(0.741)	(1.196)
Justice Specific	Prior Judicial Experience		0.012	-0.025		0.014	-0.022		0.056	-0.056		-0.028	0.058
			(0.008)	(0.008)		(0.010)	(0.014)		(0.009)	(0.011)		(0.018)	(0.026)
	Democratic Nominee		0.382	-0.013		0.124	0.564		0.498	-0.179		-0.142	0.291
			(0.089)	(0.092)		(0.138)	(0.148)		(0.104)	(0.137)		(0.219)	(0.238)
	SC Ideology		0.906	-1.668		0.054	-1.121		1.838	-2.638		-0.100	0.427
			(0.128)	(0.129)		(0.194)	(0.202)		(0.170)	(0.192)		(0.358)	(0.400)
	SC Qualifications		0.375	-1.089		0.082	-0.536		1.529	-1.267		-0.343	0.840
			(0.157)	(0.181)		(0.224)	(0.285)		(0.177)	(0.267)		(0.465)	(0.444)
Justice/Case Specific	**Average SC Ideology of Remaining Justices**		**-0.801**	**1.060**		**0.058**	**0.044**		**-3.265**	**1.284**		**-0.922**	**-0.810**
			(0.304)	**(0.348)**		**(0.477)**	**(0.593)**		**(0.457)**	**(0.438)**		**(1.046)**	**(1.433)**
	Average SC Qualif. of Remaining Justices		**-2.953**	**-5.020**		**0.013**	**-0.048**		**0.448**	**-2.646**		**0.134**	**0.599**
			(0.497)	**(0.612)**		**(0.796)**	**(0.944)**		**(0.707)**	**(0.802)**		**(1.749)**	**(2.029)**
	Years of Experience in the Court		0.028	-0.036		0.019	-0.044		0.041	-0.052		0.000	-0.024
			(0.004)	(0.004)		(0.005)	(0.007)		(0.005)	(0.006)		(0.010)	(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

	Justice	γ_{it0}	γ_{it1}	θ	$s*$	$\pi^{\exp}$	π^{ST}	$L(piv)$	$FLEX^{ST}$	$FLEX^{exp}$
Basis Rights (ρ = 0.566)	Kennedy	0.230	0.811	1.622	0.455	0.537	0.268	0.315	0.441	0.441
	Souter	0.187	0.849	1.921	0.463	0.532	0.269	0.324	0.438	0.438
	Thomas	0.119	0.906	2.497	0.472	0.522	0.271	0.340	0.435	0.435
	O'Connor	0.261	0.810	1.517	0.421	0.521	0.264	0.330	0.428	0.428
	Stevens	0.267	0.808	1.492	0.416	0.520	0.264	0.332	0.427	0.427
	Breyer	0.245	0.834	1.659	0.416	0.509	0.261	0.342	0.422	0.422
	Rehnquist	0.223	0.853	1.812	0.420	0.501	0.258	0.347	0.420	0.420
	Scalia	0.176	0.891	2.167	0.430	0.484	0.251	0.356	0.419	0.419
	Ginsburg	0.362	0.785	1.141	0.309	0.504	0.269	0.362	0.399	0.399
Economics (ρ = 0.610)	Breyer	0.110	0.869	2.347	0.523	0.640	0.954	11.781	0.573	0.573
	Thomas	0.086	0.866	2.476	0.552	0.683	0.958	10.612	0.562	0.562
	Scalia	0.101	0.814	2.170	0.589	0.703	0.961	10.355	0.536	0.536
	Ginsburg	0.115	0.804	2.057	0.584	0.690	0.960	10.793	0.535	0.535
	Souter	0.097	0.793	2.115	0.613	0.722	0.962	9.858	0.522	0.522
	Kennedy	0.110	0.740	1.872	0.656	0.730	0.962	9.485	0.494	0.494
	Rehnquist	0.121	0.732	1.788	0.655	0.719	0.962	9.774	0.493	0.493
	O'Connor	0.106	0.712	1.805	0.691	0.744	0.963	8.914	0.476	0.476
	Stevens	0.124	0.685	1.637	0.706	0.731	0.961	9.162	0.466	0.466
Criminal (ρ = 0.333)	Thomas	0.156	0.947	2.630	0.385	0.184	0.005	0.023	0.581	0.581
	Scalia	0.309	0.929	1.968	0.254	0.161	0.004	0.020	0.485	0.485
	Rehnquist	0.373	0.899	1.597	0.202	0.190	0.004	0.018	0.452	0.452
	Souter	0.396	0.892	1.498	0.176	0.194	0.004	0.017	0.439	0.439
	O'Connor	0.527	0.870	1.058	-0.065	0.210	0.004	0.016	0.359	0.359
	Stevens	0.552	0.868	0.989	-0.131	0.212	0.004	0.016	0.343	0.343
	Kennedy	0.567	0.871	0.959	-0.177	0.211	0.004	0.016	0.332	0.332
	Breyer	0.578	0.882	0.987	-0.199	0.202	0.004	0.017	0.321	0.321
	Ginsburg	0.807	0.901	0.420	-2.063	0.241	0.006	0.020	0.162	0.162
Federalism (ρ = 0.570)	Rehnquist	0.106	0.868	2.365	0.528	0.609	0.293	0.267	0.460	0.540
	Thomas	0.116	0.892	2.432	0.491	0.557	0.288	0.322	0.441	0.441
	Stevens	0.092	0.915	2.706	0.492	0.556	0.292	0.331	0.439	0.439
	O'Connor	0.103	0.914	2.632	0.481	0.538	0.287	0.346	0.435	0.435
	Scalia	0.092	0.923	2.752	0.482	0.537	0.288	0.349	0.434	0.434
	Souter	0.092	0.940	2.884	0.462	0.491	0.275	0.392	0.425	0.425
	Kennedy	0.076	0.959	3.167	0.452	0.451	0.256	0.419	0.421	0.421
	Breyer	0.072	0.969	3.327	0.440	0.406	0.233	0.444	0.417	0.417
	Ginsburg	0.065	0.978	3.526	0.430	0.357	0.204	0.461	0.414	0.414

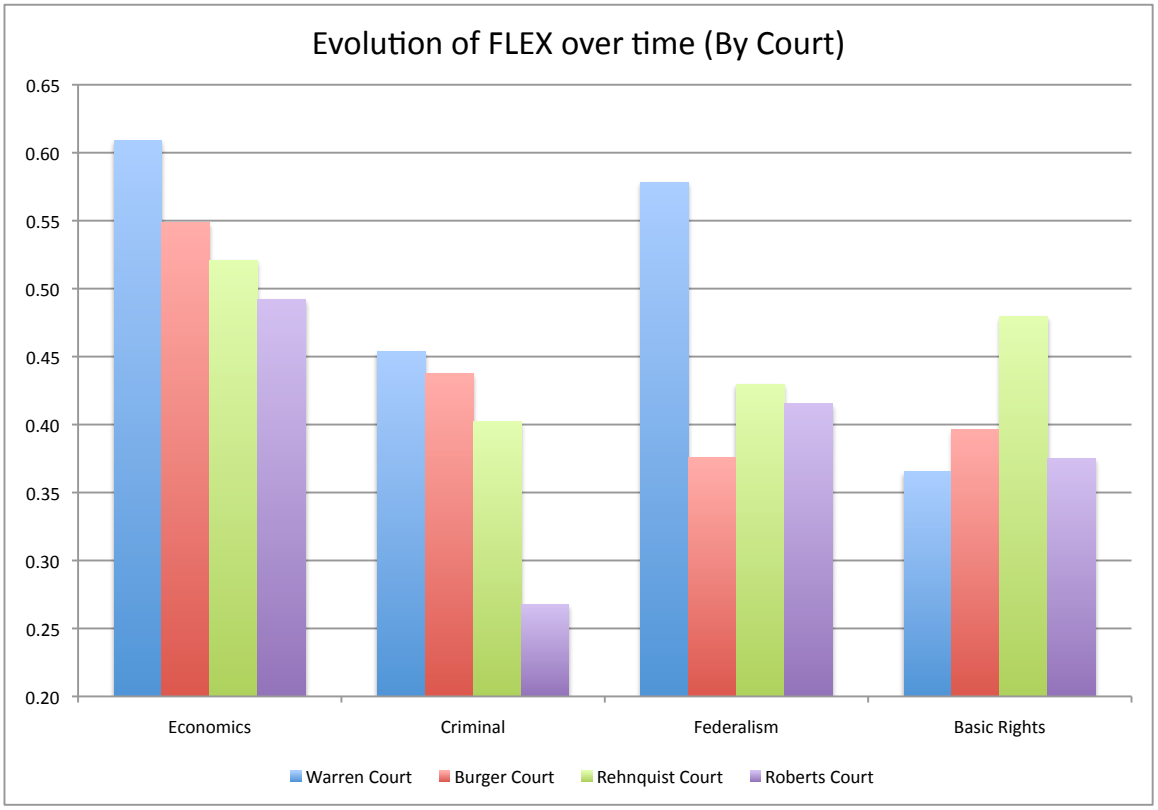


Figure 1: The Value of Information in the Court (FLEX), by Court and Issue Area (Expressive Voting Model, Private Parties, Statutory Interpretation, Lower Courts Agree).

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 (ρ = 0.58)			Economics (ρ = 0.66)			Criminal (ρ = 0.39)			Federalism (ρ = 0.61)			Average	
			π^{exp}	θ	FLEX^{exp}	π^{exp}	θ	FLEX^{exp}	π^{exp}	θ	FLEX^{exp}	π^{exp}	θ	FLEX^{exp}	FLEX^{exp}	SC-Ideology
Early Period	Reed	1948	0.32	1.66	**0.30**	0.73	1.74	**0.63**	0.28	1.19	**0.35**	0.65	2.67	**0.60**	**0.47**	0.73
	Jackson	1948	0.33	1.45	**0.28**	0.75	1.70	**0.61**	0.30	0.82	**0.29**	0.61	2.80	**0.38**	**0.39**	1.00
	Frankfurter	1951	0.33	1.65	**0.31**	0.73	1.73	**0.63**	0.28	1.19	**0.35**	0.65	2.71	**0.60**	**0.47**	0.67
	Burton	1952	0.29	2.55	**0.38**	0.75	1.96	**0.63**	0.24	2.69	**0.51**	0.68	2.45	**0.59**	**0.53**	0.28
	Minton	1953	0.29	2.16	**0.34**	0.69	2.14	**0.33**	0.27	1.84	**0.45**	0.61	2.86	**0.38**	**0.37**	0.72
	Black	1954	0.40	1.62	**0.35**	0.73	1.66	**0.58**	0.34	1.30	**0.45**	0.70	2.27	**0.57**	**0.49**	0.88
	Douglas	1957	0.39	1.44	**0.32**	0.73	1.58	**0.59**	0.30	0.99	**0.35**	0.67	2.60	**0.59**	**0.46**	0.73
	Clark	1958	0.30	2.57	**0.38**	0.65	2.34	**0.30**	0.28	2.85	**0.53**	0.69	2.32	**0.58**	**0.45**	0.50
	Whittaker	1960	0.37	2.41	**0.40**	0.76	2.04	**0.63**	0.27	2.31	**0.49**	0.61	2.81	**0.39**	**0.48**	0.50
	Warren	1962	0.37	2.14	**0.38**	0.78	1.84	**0.59**	0.30	1.94	**0.48**	0.64	2.60	**0.49**	**0.49**	0.75
	Harlan	1963	0.42	1.94	**0.38**	0.78	1.75	**0.57**	0.32	1.64	**0.47**	0.63	2.62	**0.49**	**0.48**	0.88
Middle Period	Goldberg	1964	0.28	2.10	**0.33**	0.70	2.21	**0.33**	0.25	1.84	**0.43**	0.61	2.88	**0.38**	**0.37**	0.75
	Fortas	1967	0.37	1.71	**0.35**	0.74	2.01	**0.64**	0.30	1.21	**0.39**	0.59	3.01	**0.38**	**0.44**	1.00
	Stewart	1970	0.47	1.68	**0.41**	0.78	1.66	**0.53**	0.30	1.27	**0.41**	0.60	2.85	**0.45**	**0.45**	0.75
	Brennan	1974	0.47	1.05	**0.36**	0.74	1.63	**0.55**	0.31	0.52	**0.22**	0.54	3.09	**0.38**	**0.38**	1.00
	Burger	1978	0.56	2.10	**0.48**	0.71	2.07	**0.57**	0.23	1.87	**0.47**	0.54	3.11	**0.37**	**0.47**	0.12
	White	1978	0.48	1.76	**0.41**	0.67	2.15	**0.47**	0.25	1.63	**0.43**	0.62	2.57	**0.51**	**0.46**	0.50
	Rehnquist	1979	0.56	2.15	**0.45**	0.71	2.16	**0.55**	0.20	2.16	**0.51**	0.59	2.51	**0.48**	**0.50**	0.05
	Marshall	1979	0.55	0.92	**0.42**	0.73	1.79	**0.53**	0.30	0.45	**0.22**	0.50	3.08	**0.39**	**0.39**	1.00
	Powell	1980	0.54	1.95	**0.40**	0.62	2.48	**0.42**	0.19	1.83	**0.45**	0.55	2.86	**0.39**	**0.41**	0.17
	Blackmun	1982	0.56	1.87	**0.45**	0.71	1.92	**0.53**	0.22	1.52	**0.42**	0.53	3.02	**0.41**	**0.45**	0.12
Late Period	Stevens	1992	0.57	1.65	**0.48**	0.73	1.86	**0.50**	0.23	1.25	**0.38**	0.54	2.78	**0.43**	**0.45**	0.25
	O'Connor	1994	0.57	1.68	**0.48**	0.74	2.03	**0.50**	0.22	1.32	**0.40**	0.52	2.72	**0.42**	**0.45**	0.41
	Scalia	1997	0.52	2.25	**0.43**	0.69	2.31	**0.54**	0.17	2.10	**0.48**	0.52	2.81	**0.43**	**0.47**	0.00
	Kennedy	1998	0.57	1.66	**0.46**	0.72	1.98	**0.50**	0.23	1.03	**0.33**	0.44	3.20	**0.42**	**0.43**	0.37
	Souter	1999	0.55	1.90	**0.45**	0.72	2.18	**0.52**	0.21	1.48	**0.41**	0.48	2.90	**0.42**	**0.45**	0.32
	Thomas	2000	0.54	2.42	**0.44**	0.68	2.49	**0.55**	0.19	2.50	**0.53**	0.55	2.43	**0.46**	**0.50**	0.16
	Ginsburg	2001	0.53	1.03	**0.38**	0.69	2.02	**0.52**	0.28	0.36	**0.13**	0.36	3.50	**0.41**	**0.36**	0.68
	Breyer	2001	0.53	1.47	**0.39**	0.65	2.24	**0.55**	0.25	0.79	**0.24**	0.43	3.27	**0.41**	**0.40**	0.47
	Roberts	2007	0.50	2.34	**0.37**	0.68	2.59	**0.54**	0.17	2.19	**0.43**	0.49	2.81	**0.41**	**0.44**	0.12
	Alito	2007	0.52	2.20	**0.38**	0.68	2.41	**0.53**	0.19	1.69	**0.36**	0.41	3.27	**0.41**	**0.42**	0.10
	Average		0.45	1.85	**0.39**	0.71	2.02	**0.53**	0.25	1.54	**0.40**	0.57	2.82	**0.45**	**0.44**	0.51

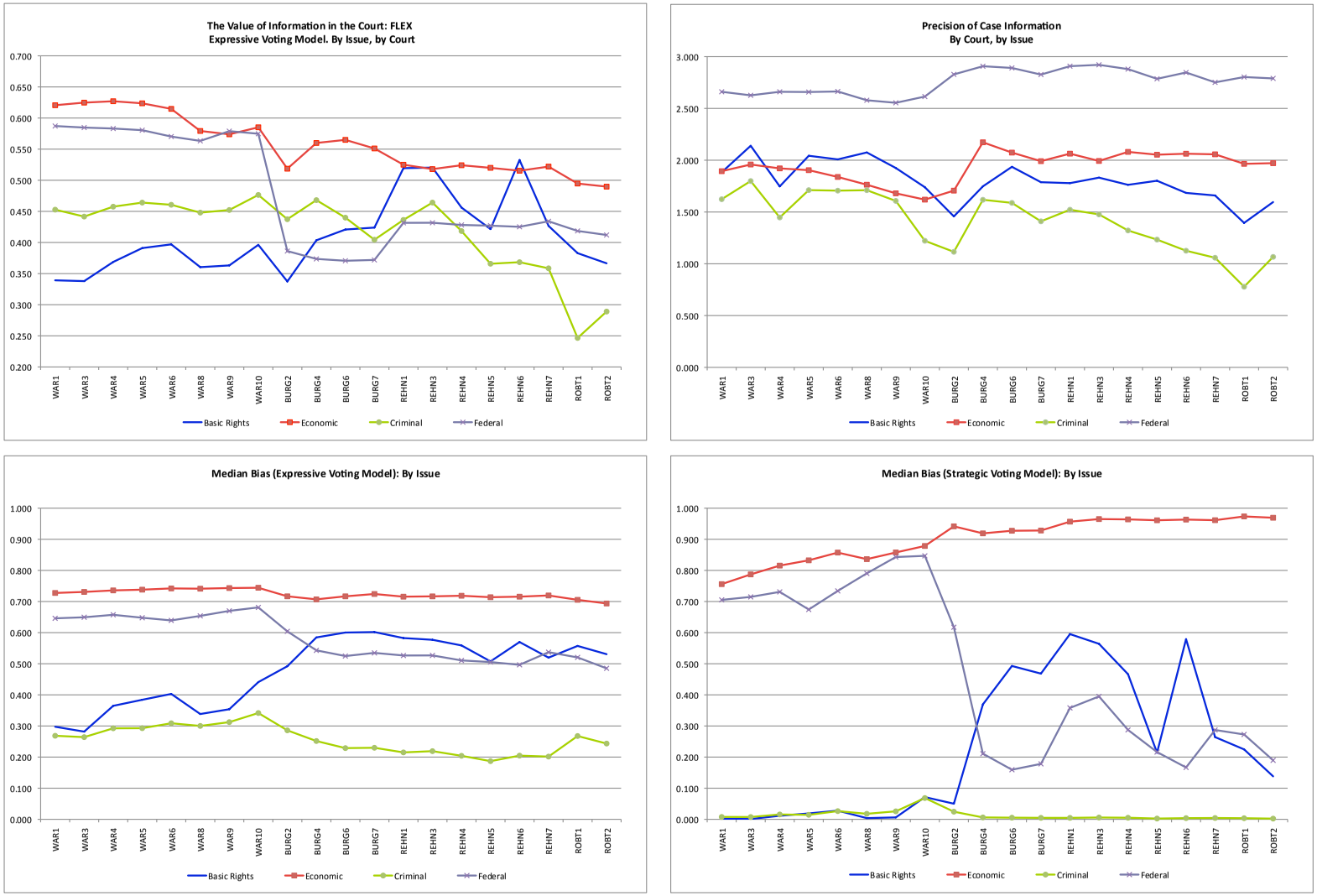
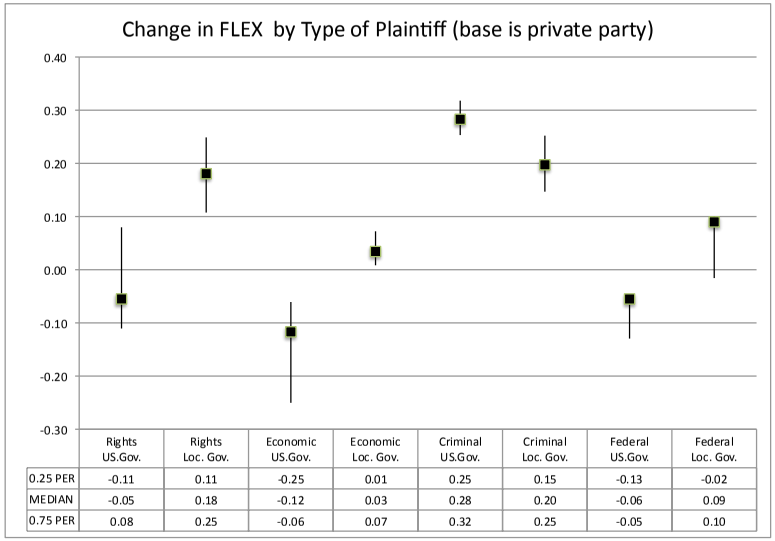
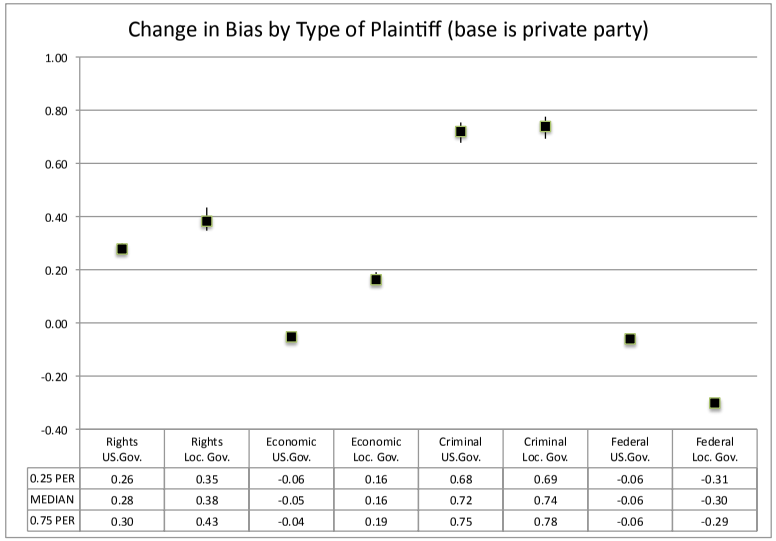


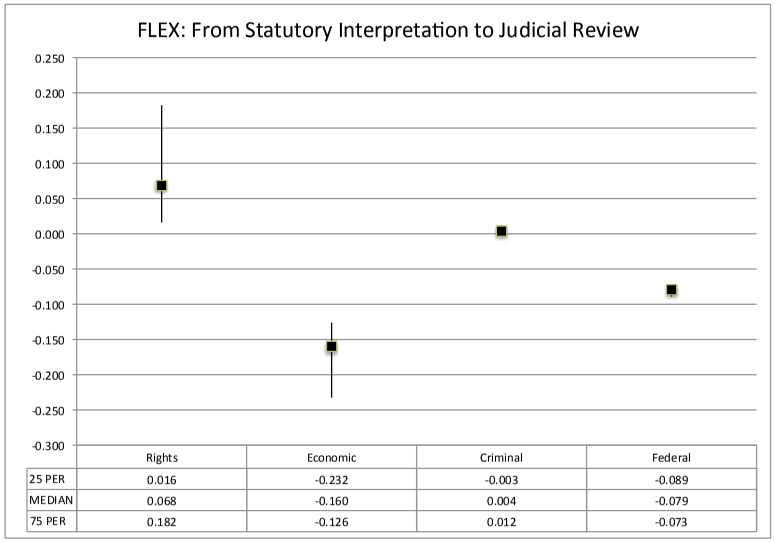
Figure 2: Median values of FLEX scores, Precision, and Bias in the Expressive and Strategic Model, by Court (Private Plaintiff vs Private Defendant, Statutory Interpretation, Lower Courts Agree.)



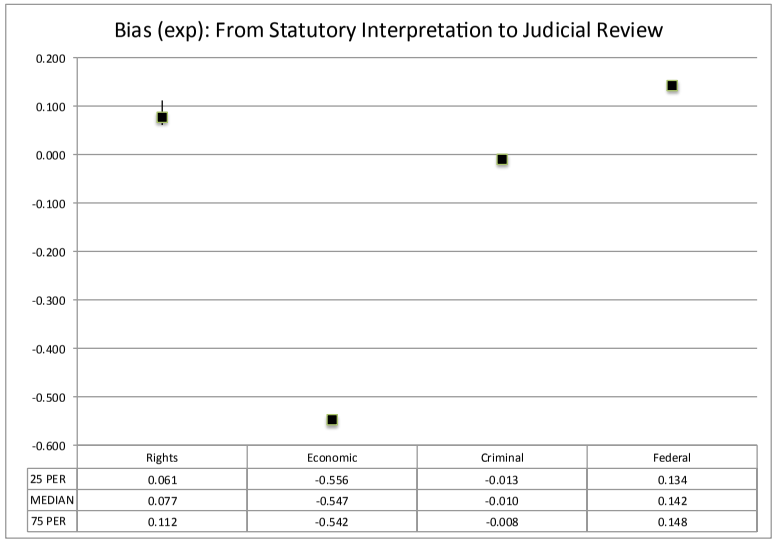
	Rights US.Gov.	Rights Loc. Gov.	Economic US.Gov.	Economic Loc. Gov.	Criminal US.Gov.	Criminal Loc. Gov.	Federal US.Gov.	Federal Loc. Gov.
0.25 PER	-0.11	0.11	-0.25	0.01	0.25	0.15	-0.13	-0.02
MEDIAN	-0.05	0.18	-0.12	0.03	0.28	0.20	-0.06	0.09
0.75 PER	0.08	0.25	-0.06	0.07	0.32	0.25	-0.05	0.10



	Rights US.Gov.	Rights Loc. Gov.	Economic US.Gov.	Economic Loc. Gov.	Criminal US.Gov.	Criminal Loc. Gov.	Federal US.Gov.	Federal Loc. Gov.
0.25 PER	0.26	0.35	-0.06	0.16	0.68	0.69	-0.06	-0.31
MEDIAN	0.28	0.38	-0.05	0.16	0.72	0.74	-0.06	-0.30
0.75 PER	0.30	0.43	-0.04	0.19	0.75	0.78	-0.06	-0.29

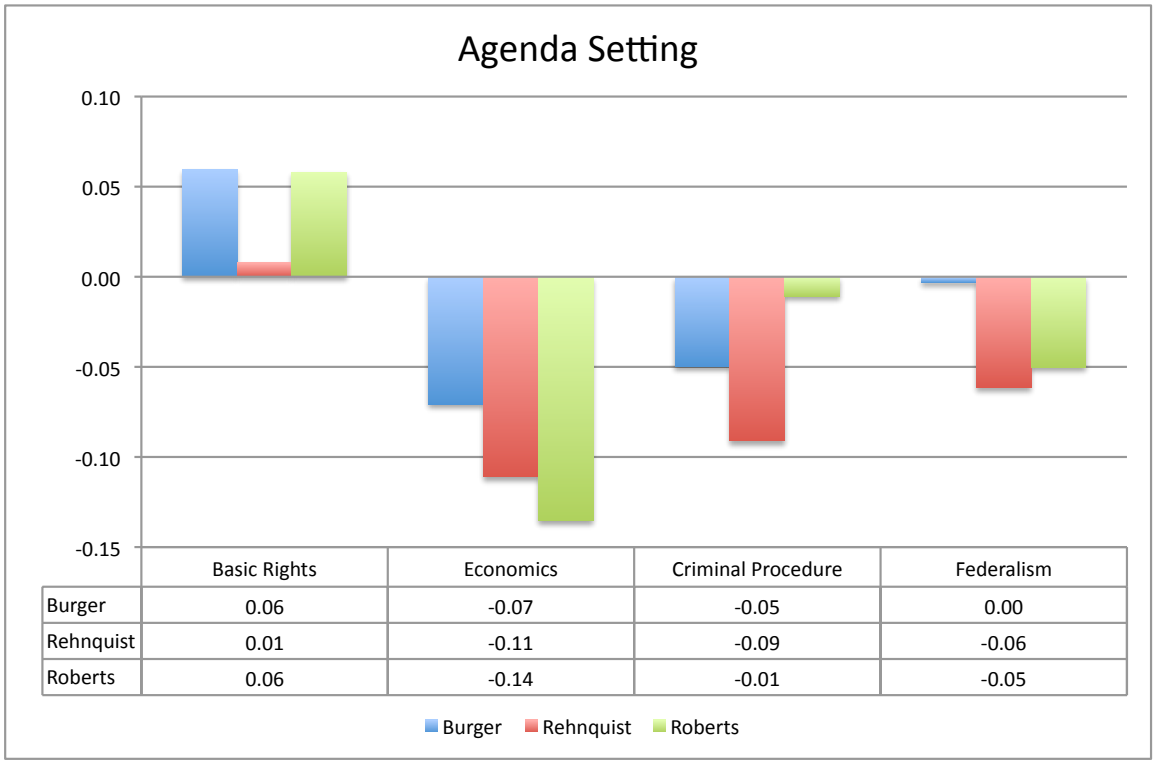


	Rights	Economic	Criminal	Federal
25 PER	0.016	-0.232	-0.003	-0.089
MEDIAN	0.068	-0.160	0.004	-0.079
75 PER	0.182	-0.126	0.012	-0.073



	Rights	Economic	Criminal	Federal
25 PER	0.061	-0.556	-0.013	0.134
MEDIAN	0.077	-0.547	-0.010	0.142
75 PER	0.112	-0.542	-0.008	0.148

Figure 3: Comparative Statics



	Basic Rights	Economics	Criminal Procedure	Federalism
Burger	0.06	-0.07	-0.05	0.00
Rehnquist	0.01	-0.11	-0.09	-0.06
Roberts	0.06	-0.14	-0.01	-0.05

Figure 4: 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).

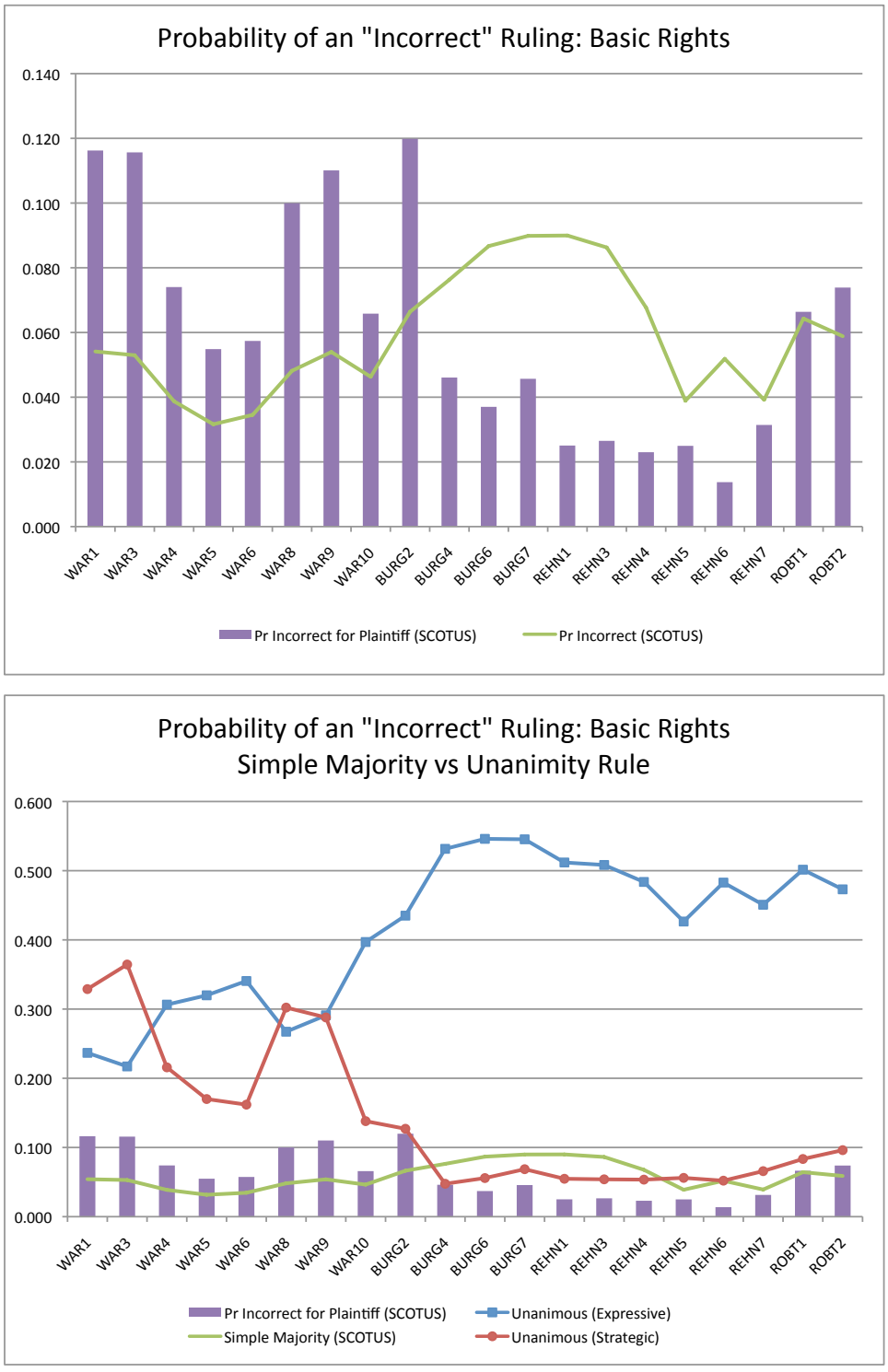


Figure 5: Counterfactual Prediction: Unanimity vs Simple Majority (Basic Rights)

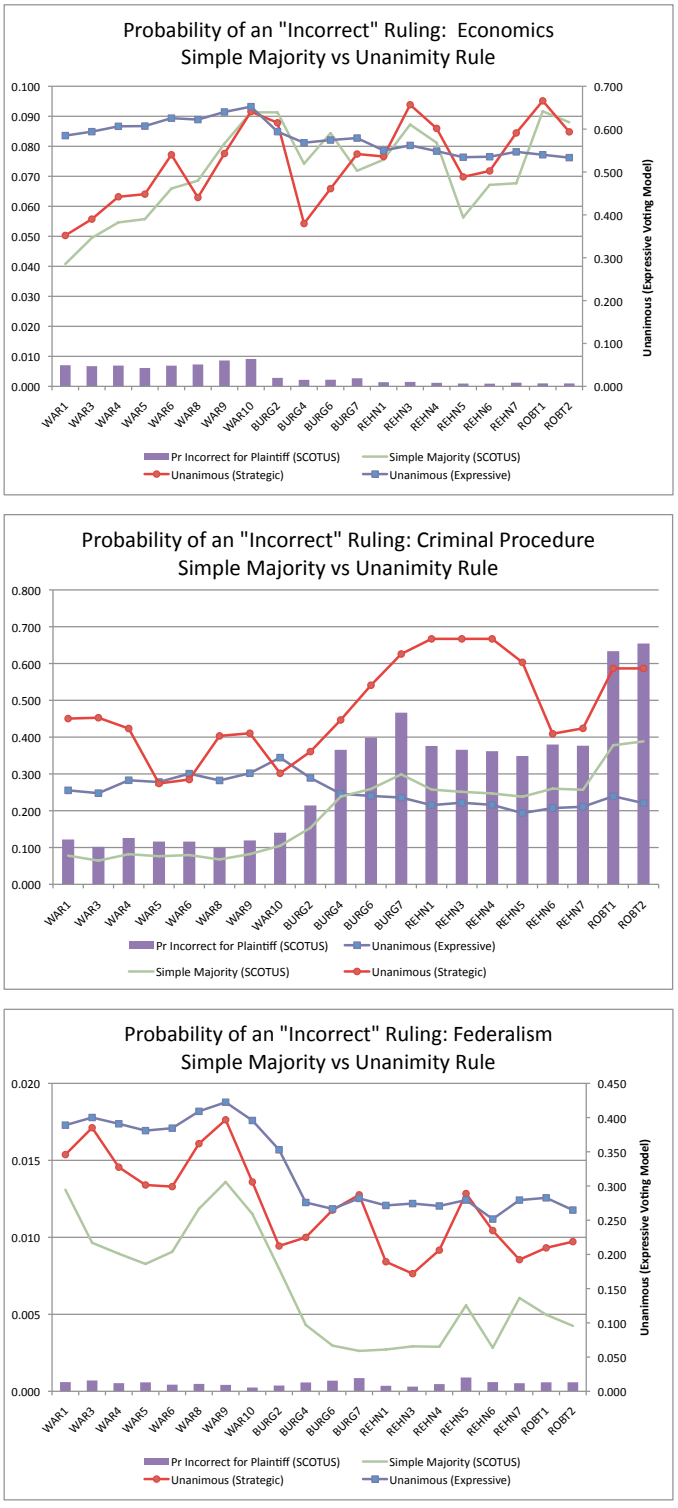


Figure 6: Unanimity vs Simple Majority in Economics, Criminal Procedure, and Federalism