Journal of Public Economics 97 (2013) 230–244

Contents lists available at SciVerse ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

To elect or to appoint? Bias, information, and responsiveness of bureaucrats and politicians[☆]

Matias Iaryczower [a,*], Garrett Lewis [b], Matthew Shum [c]

[a] Department of Politics, Princeton University, 037 Corwin Hall, Princeton NJ 08544, United States
[b] Department of Politics, Princeton University, Princeton NJ 08544, United States
[c] Division of Humanities and Social Sciences, California Institute of Technology, 301D Baxter Hall, Pasadena, CA 91125, United States

ARTICLE INFO

Article history:
Received 19 May 2011
Received in revised form 1 August 2012
Accepted 6 August 2012
Available online 25 August 2012

Keywords:
Structural estimation
Strategic voting
Common values
Bureaucrats
Politicians

ABSTRACT

In this paper, we address empirically the trade-offs involved in choosing between bureaucrats and politicians. In order to do this, we map institutions of selection and retention of public officials to the type of public officials they induce. We do this by specifying a collective decision-making model, and exploiting its equilibrium information to obtain estimates of the unobservable types. We focus on criminal decisions across US states' Supreme Courts. We find that justices that are shielded from voters' influence ("bureaucrats") on average (i) have better information, (ii) are more likely to change their preconceived opinions about a case, and (iii) are more effective (make less mistakes) than their elected counterparts ("politicians"). We evaluate how performance would change if the courts replaced majority rule with unanimity rule.

1. Introduction

The basic principle of representative democracy dictates that all legislative and top executive positions in public office are to be occupied by elected representatives (politicians). But besides this broad guiding principle, the idea of representation in the operation of government is much more muddled. In all modern democracies, a number of public positions of great influence are held by non-elected officials (bureaucrats). Examples for the US include the Supreme Court, the Federal Reserve Board, and federal agencies.

The different methods of selection and retention of public officials induce differences in the performance of government. Working well, elections may induce public officials to act in the public interest, even when their preferences are not aligned with those of the public; this is the *disciplining* role of elections. Working badly, elections can also induce an official who has more information than the public to *pander* to the public, choosing not the appropriate action, but instead the most popular action; elections can also induce officials to divert resources away from developing expertise.

Given these various competing effects, it is ultimately an empirical question how politicians and bureaucrats differ in type and performance. Do voters select different *types* of public officials – more or less biased, better or worst at gathering and processing information – than government officials? Do reelection concerns induce public officials to improve their proficiency to deal with the flow of information of each decision? Are bureaucrats more *effective* than politicians?

In this paper, we tackle these questions. We build on the foundations laid by a large literature, which provides overwhelming evidence that bureaucrats and politicians produce different public policy outcomes. Our starting premise is that in order to understand the trade-offs involved in choosing between bureaucrats and politicians, we need to map institutions to the type of public officials they induce. The difficulty, of course, is that this type is unobservable. We bridge this gap by specifying a model of voting in committees, and using equilibrium information to recover the unobservable types. The main idea is to exploit the information contained in the joint observation of the individual decisions of members of committees that deal with issues involving both ideological considerations and common values. The underlying common value induces correlation in votes in equilibrium, which allows us to disentangle bias and quality of information.

We focus here on criminal decisions in US states' Supreme Courts. The application suits the approach perfectly for two reasons. First, selection and retention methods vary across states: while in some states Supreme Court justices are elected, in others they are appointed by elected officials. Moreover, non-elected justices are appointed for life in some states, but must face a political reappointment or an

☆ We thank Juliana Bambaci, Charles Cameron, Yosh Halberstam, Jean-Laurent Rosenthal, and participants in seminars at NYU, Northwestern-Kellogg, UC Davis and UPenn for useful comments to a previous version of this paper. Financial support from NSF grants SES-1061326 (Iaryczower) and SES-1061266 (Shum) is gratefully acknowledged.
* Corresponding author.
E-mail addresses: miaryc@princeton.edu (M. Iaryczower), glewis@Princeton.EDU (G. Lewis), mshum@caltech.edu (M. Shum).

0047-2727/$ – see front matter
http://dx.doi.org/10.1016/j.jpubeco.2012.08.007

up-or-down retention election by voters in other states. Second, as other high courts, state Supreme Courts are committees making decisions on issues in which there is an underlying common value component; a correct decision under the law, even if this can be arbitrarily hard to grasp.[1]

Incorporating elements of common values does not mean ruling out disagreement. Without full certainty in how the law applies to the particulars of each case, the decision of the court will typically balance the members' goal of reaching a correct decision, with conflict among them in terms of what is the correct decision in each case. This conflict arises naturally in the relatively complex cases considered by the high courts because of differences in the information processed by each justice, because of differences in their ability to produce and evaluate case-specific information, and because of idiosyncratic biases in how justices approach different cases.[2]

In the model, we assume that the goal of any justice i in any given case t is to rule according to i's own best understanding of how the law applies to the particulars of the case. Specifically, we assume that in each case t, a justice's understanding of the particulars of the case is summarized by a private signal, with precision θ_{it}. The imprecision of the signal leaves room for interpretation, which in turn allows justices' idiosyncratic biases to come into play. In the model, individual i's bias reflects the different weights that i gives to different types of decision-making errors in case t:π_{it} is the cost for i of wrongly overturning the decision of the lower court, and $(1-\pi_{it})$ the cost of wrongly upholding the decision of the lower court. In this case justice i prefers to overturn in case t if and only if the probability that the law favors the Petitioner is at least π_{it}. Information precision and bias then interact to produce outcomes. Higher precision means that it is typically more clear for the justice whether the court should overturn or uphold the decision of the lower court. A larger bias means that despite her case-specific information, a justice persists in going with her preconception of how to rule in a case like this.

In this framework, electoral institutions can sway a judge's vote by changing the θ or π with which she makes her decision. Whether electoral concerns affect θ or π more prominently is an important distinction; for instance, from the point of view of committee design, it is important to know whether electoral concerns cause judges to vote less informedly (i.e., lower θ) or become more inclined to uphold or overturn the decisions of the lower courts (i.e., increase or decrease π, respectively).

Using a structural estimation approach, we disentangle the effects of electoral concerns on bias π and quality of information θ. In particular, we recover the values of $(\theta_{it},\pi_{it})|X_t$ for each justice i conditioning on observable characteristics of the cases and the justices, including experience variables (prior judicial and political experience, experience in the state Supreme Court), context variables (measures of the political preferences of voters and politicians at the time of appointment and at the time of decision), and, most importantly, the institutional variables (whether the justice was elected, appointed for an original term subject to a political reappointment or a retention election, or appointed for life). We do this for two variants of the model: the expressive voting model (where justices care about getting their decision own 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.[3] Given our estimates of θ and π, we can also simulate effects of counterfactual voting rules and electoral institutions on vote outcomes.

The main results clarify the trade-offs inherent in choosing between bureaucrats and politicians. First, justices that are shielded from voters' evaluations ("bureaucrats") on average have higher quality of information than justices that face either reelection or retention elections ("politicians"). In fact, the quality of information of justices that are shielded from voters' influence is on average 33% larger than that of justices facing retention elections, and 39% larger than that of justices that are elected. Institutions of selection and retention of justices also affect justices' bias. In particular, we find that elected justices are also typically more inclined to overturn the decision of the lower court than those who do not face a voter evaluation after being appointed.

These two components of justices' type – quality and bias – affect how justices' information is reflected in their voting behavior. We find that justices who are shielded from voters' evaluation not only have better information, but are also more likely than elected justices to change their preconceived opinions about a case. We quantify the flexibility of a judge to incorporate case-specific information with the FLEX measure introduced in Iaryczower and Shum (2012). This is the probability that a judge votes differently than what she would have voted for in the absence of case-specific information. We show that the average FLEX scores for elected justices (0.37) and justices facing retention elections (0.36) are significantly lower than the corresponding FLEX scores for appointed justices facing political reappointment (0.48), and for justices appointed for life (0.60).

Our estimation approach also allows us to assess the effect of institutions on the performance of the court, as measured by the probability that the court reaches an incorrect decision. While these error rates are small overall, we find that justices appointed for life and appointed justices with a political reappointment on average have a lower probability of reaching an incorrect decision (0.1%) than both justices that face retention elections (0.5%), and justices that are elected (0.3%). The pattern of mistakes, moreover, is highly asymmetric. At both the individual level and at the court level, on average justices tend to wrongly overturn more often than wrongly uphold lower courts' decisions.

The asymmetry in mistakes begs the question of how the effectiveness of the courts would change if simple majority rule were replaced by unanimity rule, requiring the consent of all Supreme Court justices to overturn the decision of the lower court. Our results imply that this change would have major consequences for the effectiveness of the courts, particularly in the expressive voting model. When justices care about their vote only, the shift to unanimity achieves the purpose of reducing the probability of overturning incorrectly, but only by dramatically increasing the probability of incorrectly upholding the lower court (reaching 43.6% for elected justices, and 39.4% for non-elected justices facing retention elections). The strategic voting model predicts relatively large changes (although less dramatic) in the opposite direction. Because strategic justices would modify their voting strategy in response to the change in the voting rule, becoming more inclined to overturn, changing from majority to unanimity rule would actually increase the probability that the court incorrectly overturns the decision of the lower court (reaching 2.8% for elected justices, and 2% for justices facing retention elections).

The rest of the paper is organized as follows. Section 2 contains a literature review. Section 3 introduces the theoretical model of collective decision-making in the court and characterizes equilibrium outcomes. Section 4 describes the estimation procedure. Section 5

[1] Decision-making in the court is different than in a legislature. As Supreme Court 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." (From the statement submitted to the Senate Committee on the Judiciary by Justice Ruth Ginsburg.) This distinction is also emphasized by Cameron and Kornhauser (2008), among others.

[2] Justices' biases can, but do not necessarily reflect ideological considerations. These preconceptions about how the law maps to the particulars of each case can also reflect ingrained theoretical arguments about the law, personal experiences, and other determinants for a non-neutral approach to this case.

[3] In the law and economics literature, this distinction is referred to as whether judges are consequentialist or non-consequentialist (see Cameron and Kornhauser (2008)).

presents the main results. In Section 6, we show that the main results of the paper remain qualitatively unchanged if we reestimate the model based on an alternative coding of votes, as either in favor or against the State (in this case the bias parameter is interpreted as the relative preference against the State). Section 6 concludes.

2. Related literature

The theoretical literature on bureaucrats and politicians builds on the seminal contributions of Barro (1973) and Ferejohn (1986), which provide the foundations of the theory of elections as disciplining device. Barro (1973) introduces the main idea that voters can limit (but not eliminate) rent extraction by elected politicians by making their reelection conditional on observed behavior. Ferejohn (1986) formalizes a similar idea within a moral-hazard framework, and derives the optimal retrospective voting rule. Banks and Sundaram (1998) study the optimal retention rule for voters in a model that incorporates both moral hazard and adverse selection. Canes-Wrone et al. (2001) consider a model in which elected officials have the same preferences as the electorate, and the incumbent attempts to signal talent (e.g., more precise information). They conclude that elected officials will pander (choose the popular, ex ante preferred action) only under some limited conditions. Canes-Wrone and Shotts (2007) show that elected officials will be more inclined to pander when there is uncertainty regarding their congruence with the electorate.

Maskin and Tirole (2004) and Alesina and Tabellini (2007, 2008) explicitly compare bureaucrats and politicians. Maskin and Tirole (2004) introduce a lack of congruence between voters and public officials. They conclude that non-elected officials (bureaucrats, or “judges”) are preferred when the public is poorly informed about what the optimal action is, and when feedback about the quality of the decision is limited. Alesina and Tabellini (2007) models career concerns of bureaucrats (appointed officials) and politicians (elected representatives). They conclude that bureaucrats are preferred in technical tasks for which ability is more important than effort, or when there is large uncertainty about whether the policymaker has the required abilities to fulfill her task.

A key common element in these theoretical approaches is that the type of government official is unknown. The key factor driving the results is the amount of information that is revealed in different institutional settings about unobservable characteristics of public officials (their preferences, their competence, or the readiness to exert effort). To the best of our knowledge, this key feature has not been incorporated into applied research on the topic. There is, however, a wealth of empirical research motivated by the same underlying questions that inspired the theoretical literature.

First, a number of papers show that elected and appointed government officials do in fact behave differently. There is substantial evidence documenting this finding for the case of elected and appointed regulators (see Besley and Coate, (2003) and Besley and Case, (2003) for a survey). There is also a relatively large literature documenting this finding for the case of elected and appointed judges in the US states. Hanssen (2000) shows that states with elected judges have significantly smaller bureaucracies, and interprets this as evidence that elected judges are more independent. Hanssen (2004) shows that institutions that diminish the ability of politicians to determine whether a judge remains in office are associated with closer competition between political parties, and with larger differences in party platforms, while the least independence-enhancing institutions are associated with a stronger single party control. Besley and Payne (2005) show that states that appoint their judges have lower levels of discrimination charges compared to those that use some form of election. Gordon and Huber (2007) analyze the sentencing behavior of district court judges that are elected and appointed (facing a subsequent retention election) in the state of Kansas. They show that close to the elections, elected judges are harsher in sentencing relative to appointed judges.[4]

Choi et al. (2010) also focus on state Supreme Court judges, and share our emphasis on measuring the effects of the judicial selection process on non-ideological characteristics of the judges. Their methodology is quite different, as they focus on *opinions* instead of voting, and use observable proxies of judges' quality, including the number of opinions written, and the number of times a judge's opinions were cited by other courts at the same level of the judicial hierarchy.[5]

Also taking a structural approach, Lim (2011) estimates a model that fully incorporates career concerns into judges' behavior, using sentencing data from Kansas[6] (see Diermeier et al., (2005) for a similar approach in Congress). Differently than in our paper, Lim's model does not allow the possibility of common values and dispersed information, which seem central to the nature of decision-making in the court.[7] Here we allow both ideology and common values in the context of equilibrium behavior. Our model of collective decision-making builds on Austen-Smith and Banks (1996), and Feddersen and Pesendorfer (1997, 1998), and is closest to that of Duggan and Martinelli (2001). The empirical approach builds on Iaryczower and Shum (2012).[8]

An interesting issue in connection to strategic voting in this setting is the possible impact of pre-vote deliberation on outcomes. The main question is whether strategic agents will use pre-vote deliberation to communicate information to their peers, or whether they will use these arguments to try to influence their opinion, possibly not revealing some information that can be harmful to their case, or exaggerating evidence one way or the other. While the incentive to do so is small when interests are well aligned (Coughlan (2000)), this is not the case when there is (interim) disagreement, as in the setting considered here. This makes truthful revelation of information more difficult, as is illustrated in the analysis of Austen-Smith and Feddersen (2005, 2006) (see also Li et al., (2001) and Doraszelski et al., 2003).[9] Visser and Swank (2007) consider pre-vote deliberation when committee members want to signal their ability to a principal. Reputation concerns here induce committee members to misrepresent their information in deliberations. In spite of this, committee members vote unanimously. This is because in this setting, disagreement signals lack of competence. Visser and Swank's basic logic – that information is reflected in the variation of the justices' votes – also underlies the

[4] More broadly, there is overwhelming evidence showing that judges are sensitive to the political environment. See Brace and Hall (1990, 1993, 1997) for US states, Gely and Spiller (1990), Spiller and Gely (1992) for the US Supreme Court, Helmke (2002) and Iaryczower et al. (2002, 2006) for the Supreme Court in Argentina, along many others.

[5] Choi et al. (2010) conclude that appointed judges write higher quality opinions than elected judges do, but elected judges write more opinions.

[6] Lim shows that the sentencing behavior of elected judges is in fact an important determinant of their reelection, and that while the sentencing behavior of appointed judges does not vary much with the political orientation of the district, elected justices tend to be more lenient in liberal leaning districts.

[7] For structural estimation of ideological models of voting in committees (that do not directly incorporate career concerns) see Poole and Rosenthal (1985, 1991), Heckman and Snyder (1997), Londregan (1999), Clinton et al. (2004) – for the US Congress – and Martin and Quinn (2002, 2007) – for the US Supreme Court. Degan and Merlo (2009) and DePaula 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 (forthcoming) also perform structural estimation of strategic voting (i.e. “pivotal voting”) models with ideological voters.

[8] With common values and dispersed information, strategic considerations come into play. Our methodology deals with these strategic considerations. For a connected approach, emphasizing the bicameral structure of Congress, see Iaryczower et al. (2012).

[9] If agents send not only relevant information, but also other (random) messages, which the group uses to define correlated voting strategies, more can be done. Gerardi and Yariv (2007) show that every outcome that can be implemented with a non-unanimous voting rule r can also be implemented (as a sequential equilibrium of a cheap talk extension of the voting game) with a non-unanimous rule r'. This obviously enlarges the set of possible equilibrium outcomes for each given voting rule.

identification of the key model parameters from the observed vote data (see Section 4). However, in our setting votes provide information about not only justices' ability, but also their bias.

3. A model of decision-making in the court

In this section, we describe the model of collective decision-making in the court. In doing so, we take the parameters of the problem as given, and their dependence on publicly observable characteristics of the choice situation as understood. We make this relation explicit in Section 4.

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 vote to uphold or overturn the decision of the lower court. We denote this vote by $v_i^t\in\{0,1\}$, with $v_i^t=0$ indicating a vote to uphold and $v_i^t=1$ a vote to overturn the decision of the lower court. The court aggregates the decisions of the individual justices by simple majority rule; i.e. overturns ($v_t=1$) if $\sum_i v_i^t\geq\frac{n+1}{2}$ and upholds ($v_t=0$) otherwise.

We consider two related models of individual behavior. In the *expressive* 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 decision of the court. We assume that the goal of any justice i in any given case t is that she (in the expressive voting model) or the court (in the strategic voting model) rules according to i's own best understanding of how the law applies to the particulars of the case.

Specifically, before ruling in each case t, each justice i observes a private signal $s_{it}=\omega_t+\sigma_{it}\varepsilon_{it}$, where $\varepsilon_{it}\sim\mathcal{N}(0,1)$. Here $\omega_t\in\{0,1\}$ is an unobservable variable – for both the econometrician and the justices – indicating whether the decision of the lower court should be overturned ($\omega_t=1$) or upheld ($\omega_t=0$) according to the law, and $\theta_{it}=1/\sigma_{it}$ is a scale parameter that quantifies the informativeness of i's signals.[10] This parameterization of the information structure satisfies the Monotone Likelihood Ratio Property (MLRP), which is important in what follows.

Justices care about this information because their payoffs are state dependent. In particular, we assume that given $\pi_{it}\in(0,1)$, justice i has a payoff of $-\pi_{it}$ when she/the court incorrectly overturns the lower court ($v_t=1$ when $\omega_t=0$) and of $-(1-\pi_{it})$ when she/the court incorrectly upholds the lower court ($v_t=0$ when $\omega_t=1$).[11] 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 overturn in case t if and only if $\Pr^i(\omega_t=1|E)\geq\pi_{it}$. Equivalently, Justice i votes to overturn in case t given E if and only if the likelihood ratio $\Pr^i(E|\omega_t=1)/\Pr^i(E|\omega_t=0)$ is larger than $\frac{\pi_{it}}{1-\pi_{it}}\frac{1-\rho_t}{\rho_t}$, where $\rho_t\equiv\Pr(\omega_t=1)$ denotes justices' common prior probability of the unobserved state 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 i and j can disagree almost always (with $\pi_{it}\approx 0$ and $\pi_{it}\approx 1$). On the other hand, when $\pi_{it}=1/2$ for all i, the setting boils down to an unbiased, pure common values model.[12]

[10] We write θ_{it} and not simply θ_i, invariant in t, because in the estimation we will allow the precision of information to depend on characteristics of the case. With identical observable characteristics across cases we would have $\theta_{it}=\theta_i$ for all t. The same remark applies to the bias π_{it} below.

[11] Thus, $\pi_{it}\neq 1/2$ reflects a bias toward upholding or overturning the lower court in case t. This bias can reflect a variety of factors inducing a non-neutral approach to this case, such as ingrained theoretical arguments about the law, personal experiences, or ideological considerations.

[12] In our setting, justices share common priors, but their biases are captured by the π_{it} parameters. See Froeb and Kobayashi (1996) for a model where justices' biases are manifested in their priors. Moreover, our attempt to estimate a model where priors ρ, as well as bias π, differed across justices i and cases t resulted in poorly behaved estimates. See footnote 18 below for an explanation.

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 decision, and therefore vote based on their own information s_{it}, i.e., vote to overturn whenever $\Pr^i(\omega_t=1|s_{it})\geq\pi_{it}$. Then E consists only of s_{it}, and i votes to overturn if

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

Let s_{it}^{exp} denote the value of s_{it} that solves Eq. (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 votes to overturn whenever $s_{it}\geq s_{it}^{exp}$, and to uphold otherwise. These cutoff points s_{it}^{exp} for $i=1,\ldots,n$ completely characterize 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

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

In the *strategic voting model*, justices care about the decision of the court. As a result, any Justice i then considers the implications of her vote assuming that she is pivotal for the decision. (This supposition is not correct when the justice is not in fact pivotal, but for the same reason these mistakes have no cost for the outcome-oriented justice.) Here, the relevant information for Justice i in case t is not only her private information s_{it}, but also the equilibrium information contained in the event that i is pivotal for the court's decision, given the equilibrium strategy profile followed by the remaining justices. Let $\mu_{jt}:\mathbf{R}\to[0,1]$ denote the strategy of Justice j in case t, where $\mu_{jt}(s_{jt})\equiv\Pr(v_{jt}=1|s_{jt})$. Then Eq. (1) becomes

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

As before, the MLRP implies that i's best response to any strategy $\mu_{-i,t}$ of the remaining justices is a cutoff strategy, such that i votes to overturn ($\mu_{i,t}(s_{it})=1$) if s_{it} satisfies Eq. (3), and votes to uphold ($\mu_{it}(s_{it})=0$) otherwise. This in turn implies that all responsive equilibria are cutoff equilibria; i.e., that any equilibrium is characterized by cutpoints s_{it}^{st} for each Justice $i=1,\ldots,n$ such that Justice i votes to overturn if and only if $s_{it}\geq s_{it}^{st}$. Now, given cutoff strategies, $\Pr(v_{it}=1|\omega_t)=\int\mu_{it}(s)\phi(\theta_{it}[s-\omega_t])ds=[1-\Phi(\theta_{it}[s_{it}^{st}-\omega_t])]$. Therefore from Eq. (3), and letting $\mathcal{C}_{R-1}^{i}$ denote the set of coalitions $C\subset N\setminus i$ with $R-1$ members, $\{s_{it}^{st}\}_{i=1}^{n}$ is given by the n equations

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

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

$$\Pr(v_t)\equiv\sum_{\omega_t}\Pr(\omega_t)\prod_{i=1}^{n}\left[1-\Phi\left(\theta_{it}\left[s_{it}^{st}-\omega_t\right]\right)\right]^{v_{it}}\Phi\left(\theta_{it}\left[s_{it}^{st}-\omega_t\right]\right)^{1-v_{it}}. \tag{5}$$

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^{st} for the strategic model.[13]

[13] Equilibrium in the strategic voting model might not be unique. We assume that if there are multiple equilibria, justices consistently play the same equilibrium whenever the characteristics of the problem are unchanged. It should be noted, however, that in the estimation, for any vector of conditional voting probabilities in the first stage (see Section 4) we recover the types (θ_i,π_i) uniquely.

4. Estimation

The estimation procedure has two parts, which we describe below.

4.1. Estimation: first step

We introduce the following notation:

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

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:

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

Because 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 Respondent $\gamma_{it,1}$ and $\gamma_{it,0}$. Accordingly, we also parameterize these probabilities to depend upon X_t (covariates for case t) and Z_i (covariates for Justice i) in the following way, which also restricts $\gamma_{i,t,1} \geq \gamma_{i,t,0}$, for all X_t:

$$\begin{aligned} \gamma_{i,0}(\zeta,\eta) &= \frac{\exp(Z_i'\zeta + X_t'\eta)}{1+\exp(Z_i'\zeta+X_t'\eta)}, \quad \in[0,1]; \\ \gamma_{i,1}(\zeta,\eta,\alpha,\delta) &= \frac{\gamma_{i,0}+\exp(Z_i'\alpha+X_t'\delta)}{1+\exp(Z_i'\alpha+X_t'\delta)}, \quad \in\left[\gamma_{i,0}(\zeta,\eta),1\right]. \end{aligned} \tag{6}$$

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, which corresponds to the reduced-form likelihood function for bids in both the expressive and strategic voting models:

$$\begin{aligned} \max_{\alpha,\beta,\zeta,\eta,\delta} \sum_t \log\Bigg[\rho(X_t;\beta)\cdot\prod_{i=1}^n \left\{\gamma_{i,1}(\zeta,\eta,\alpha,\delta)^{v_{it}}\left(1-\gamma_{i,1}(\zeta,\eta,\alpha,\delta)\right)^{1-v_{it}}\right\} \\ +(1-\rho(X_t;\beta))\cdot\prod_{i=1}^n\left\{\gamma_{i,0}(\zeta,\eta)^{v_{it}}\left(1-\gamma_{i,0}(\zeta,\eta)\right)^{1-v_{it}}\right\}\Bigg]. \end{aligned} \tag{7}$$

Given the MLE estimates of $\hat{\zeta},\hat{\eta},\hat{\alpha},\hat{\delta}$, we can compute the corresponding priors $\hat{\rho}\equiv\rho\left(X_t,\hat{\beta}\right)$ as well as vote probabilities $\hat{\gamma}_{i,0}\equiv\gamma_{i,0}\left(\hat{\zeta},\hat{\eta}\right)$ and $\hat{m}a_{i,1}\equiv\gamma_{i,1}\left(\hat{\zeta},\hat{\eta},\hat{\alpha},\hat{\beta}\right)$ for any vector of covariates (X_t,Z_t).

4.2. 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 equilibrium strategies and the two structural parameters, π_i and βeta_i, for each Justice i. Recall our earlier assumptions that Justice i's private information is $s_{it}=\omega_t+\frac{1}{\theta_i}\varepsilon_{it}$, with $\varepsilon_{it}\sim\mathcal{N}(0,1)$. Then $\gamma_{i,1}\equiv 1-\Phi(\theta_i[s_i^*-1])$) and $\gamma_{i,0}\equiv(1-\Phi(\theta_i s_i^*))$. Solving these equations for 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[14]

$$\hat{\theta}_i=\Phi^{-1}\left(1-\hat{\gamma}_{i,0}\right)-\Phi^{-1}\left(1-\hat{\gamma}_{i,1}\right);\hat{s}_i=\frac{\Phi^{-1}\left(1-\hat{\gamma}_{i,0}\right)}{\Phi^{-1}\left(1-\hat{\gamma}_{i,0}\right)+\Phi^{-1}\left(\hat{\gamma}_{i,1}\right)}. \tag{8}$$

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

$$\frac{\phi(\theta_i[\hat{s}_i-1])}{\phi(\theta_i\hat{s}_i)}=\frac{\hat{\pi}_i^{exp}}{1-\hat{\pi}_i^{exp}}\frac{1-\hat{\rho}}{\hat{\rho}}, \tag{9}$$

while in the strategic voting model this is given by the system of Eq. (4). 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. 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.

Strikingly, 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 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.

Moreover, the estimation logic here implies that, when we consider only full-court cases, the expressive and strategic voting models are observationally equivalent, in that they generate the same likelihood function. When there is exogenous variation in the court size, however, then there is some possibility for distinguishing between the two models. Specifically, consider an ideal experiment where a set of identical cases was heard by 7 vs. 5 judges (and the 5 judges are a subset of the 7-judge court). Under the expressive voting model, the individual voting probabilities for each judge in the 5-judge court should not vary depending on the court size, while in the strategic voting model, these voting probabilities would vary (because the pivotal event changes with the court size). Based on this intuition, we consider, in online Appendix B, a procedure to assess quantitatively whether the expressive or strategic voting model is more appropriate for the data analyzed in this paper.

4.3. Two-step approach: a remark

Note that because of our two-step estimation approach – in which we recover the values of θ and π separately for each set of case covariates X – *by construction*, for all X observed in the data, the values of $\pi(X),\theta(X)$ are consistent with the priors $\rho(X)$ and vote probabilities $(\gamma_1(X),\gamma_0(X))$ as defined in Eq. (9). In other words, if we started with priors $\rho(X)$ and parameters $(\theta(X),\pi(X))$, and computed the equilibrium vote probabilities, they must coincide with $\gamma(X)$ as defined in Eq. (9). On the other hand, the particular logit functional form which we used for the vote probabilities in Eq. (9) imposes some structure

[14] Note that the estimate of i's information quality is increasing in the probability of correctly ruling in favor of the Petitioner $(\gamma_{i,1})$, and decreasing in $\gamma_{i,0}$, which is the probability of incorrectly ruling against the Respondent. The estimate of the equilibrium cutpoint, instead, is a decreasing function of the ratio between $\Phi^{-1}\left(\hat{\gamma}_{i,1}\right)$ and $\Phi^{-1}\left(1-\hat{\gamma}_{i,0}\right)$. Thus $\hat{s}_i$ is (roughly) decreasing in the ratio of the probability of voting correctly in favor of the Petitioner $(\gamma_{i,1})$ *relative* to the probability of correctly voting in favor of the Respondent $(1-\gamma_{i,0})$.

on the problem, and is not completely without loss of generality. Note that in the true model, any X is associated with parameters $(\theta(X),\pi(X),\rho(X))$, which are then associated with equilibrium voting probabilities $(\gamma_0(X),\gamma_1(X))$. Let $g(\cdot)=(g_0(\cdot),g_1(\cdot))$ denote the function mapping values of X to these "structural" equilibrium voting probabilities. In our estimation procedure, we impose a flexible parametric form on $g(\cdot)$, assuming that as a function of X, $g_0(\cdot)$ and $g_1(\cdot)$ have a logit functional form, as in Eq. (9). Our MLE estimates then best approximate $g(\cdot)$ within the class of logistic functions. This entails some loss, of course – as $g(\cdot)$ could be non-logistic – but because the logit functional form is very flexible, the cost of this assumption is likely to be small.

4.4. Identification

Clearly, identification of model parameters hinges on the identification of the reduced-form parameters from the first-stage MLE. This in turns relies crucially on the mixture structure of the votes, which are unconditionally dependent due to the unobserved state ω_t. Specifically, consider a state Supreme Court with $n=9$ justices (such as Texas). In this case, the vote vector v_t can take 2^9 values, and 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, thus satisfying a necessary condition for identification.[15,16]

At a more intuitive level, the key for identification is that the common value induces a correlation of votes in equilibrium: all justices tend to receive larger signals when the decision of the lower court should be overturned than when it should be upheld.

Suppose first that cases are homogeneous, so that justices' types and prior are invariant across cases. If justices' quality of information were large relative to their bias, and the prior relatively uninformative (say $\pi_i\approx 1/2$ for all i and $\rho\approx 1/2$), the court would "flip–flop" evenly between unanimous decisions to uphold and overturn. Now suppose that instead $\rho\approx 2/3$, so that the court should overturn the previous ruling more frequently. Then justices will tend to receive large signals more frequently, and moreover justices will use strategies that lean more toward overturning. As a result, the majority of the court would vote to overturn more often than before. This illustrates a first intuition: the frequency in which the majority decision is to overturn the decision of the lower court tracks the prior ρ; a larger frequency corresponds to a larger estimated prior ρ.

Now suppose that we change the bias of one Justice i in our previous example so that her bias is large relative to the quality of her information. Then while all other justices will alternate between sometimes overturning and sometimes upholding, i will stay put in one decision. This illustrates the second principle at work: absence of variability in individual decisions signals large bias. Finally, return to the previous example in which all justices are moderate. As we pointed out before, if the quality of information is sufficiently high for all justices, then we would expect these to be unanimous votes. But as the quality of information of some justices is lower, these justices would disagree with the majority more often. This suggests the third principle: justices with variable voting records who tend to be in the minority are associated with a low quality of information.

Now, as it is, this identification scheme appears to penalize "maverick" justices who go against the grain by assigning them a low precision parameter. However, in the empirical work, we control for many case-specific covariates, and take into account inherent differences among justices due to political ideology, judicial experience, etc. Therefore, justices with low θ's are those who have attributes that characterize justices who vote inconsistently, *even after* taking characteristics of the case into account: these are not maverick justices, but erratic ones.

5. Bureaucrats and politicians

Having characterized equilibrium behavior (Section 3) and having described our estimation procedure (Section 4), we can now begin to uncover the differences in type and performance of bureaucrats and politicians. In order to do so, we apply our method to decisions on criminal cases by US states' Supreme Courts. The variability in selection and retention methods across states and the common task across courts (after controlling for case-specific heterogeneity) allows us to pin down the selection and incentive effects of institutions on justices' unobservable types.

5.1. Data and specification

The data for this project has primarily been collected from the State Court Data Project (Brace et al. (2000)), with additional information obtained from the Court Statistics Project at the National Center for State Courts, *Marquis' Who's Who*, and the updated version of Berry et al. (1998). The State Court Data Project (SCDP) provides a detailed compilation of data for state Supreme Court cases in all 50 states of the United States during the years 1995 through 1998. The database contains a case-level dataset that describes the particulars of each case during this time frame, including the decision of each justice of the relevant court. The SCDP also includes a justice-level dataset, that provides data for each of the 520 justices that served on some court during the period observed, including whether the justice was elected or appointed, and whether the justice served for life or faced either reelection or reappointment to the bench. *Marquis' Who's Who* provided additional biographical information on each justice.

The courts themselves are described in depth in the Court Statistics Project (CSP), which collects data related to the administrative and legal structures of the state Courts in the United States. The basic layout shared across every state includes at least one trial court, one or more appeals courts, and a court of last resort (generally the Supreme Court).[17] For the purposes of this paper, the term "Supreme Court" refers to the court of last resort as it pertains to a given case.

Within our data, we retained those cases that were complete in their information and in which the justices sat en banc.[18] This left a total of 5958 criminal cases. We then pool the data across all natural courts according to the following specification.

The main variable in the analysis is voting data per se. In our main analysis, we code justices' votes as either in favor of overturning or

[15] 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. For more details, see Hall and Zhou (2003) or the discussion in Iaryczower and Shum (2012).

[16] Eq. (7) also shows that the mixture structure of the likelihood, and hence the identification argument, would be lost if the priors ρ were allowed to be heterogeneous across justices. This explains the poor results we obtained from an alternative specification in which justices have heterogeneous priors (see footnote 14 above). Intuitively, since the priors ρ_i and the voting probabilities $\gamma_{i,1}$ and $\gamma_{i,0}$ vary across justices, these two types of components cannot be disentangled nonparametrically, so that they are identified only due to the specific functional form assumptions that we make (i.e., the logit probabilities (Eq. (6))). This suggests that successful estimation of this alternative model may require a larger dataset of cases, with substantial variation in covariates which could be plausibly excluded from judge's priors, but affect voting probabilities.

[17] Exceptions to the basic design first include New York, in which the Supreme Court acts as an appeals court and the Court of Appeals acts as the court of last resort, and second, Oklahoma, where there are two courts of last resort dedicated to criminal and civil cases, respectively.

[18] 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 Petitioner 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 A.7).

upholding the decision of the lower court. This coding follows from the fact that the key consideration at the Supreme Court level is not determining the guilt or innocence of the accused, but instead to assess whether or not errors have been committed at trial. In Section 5.6 we report the results of reestimating the model based on an alternative coding of votes, in favor or against the state (FAST).

As case-specific covariates, we included basic information about the case, the parties involved, and the legal issue under consideration. These include the manner in which the state Supreme Court takes jurisdiction (appeal or original or habeas corpus), whether the Petitioner is a Person (the original defendant) or the State, the class of legal issues under consideration (issues of evidence, sentencing and jury instruction, and others), and whether a formal opinion was issued with the case as opposed to a per curiam opinion. It is possible that courts in some states might have a more difficult task ahead of them than others as a result of differences in the mix of cases varying in complexity. These differences might be particularly relevant between murder cases and lesser offenses, and for cases involving constitutional challenges. To account for this possibility, we include as an additional covariate whether the original crime considered was murder or not, as well as whether the death penalty was imposed by lower courts or not. We also include whether the case involved a challenge of a law based on the US or State Constitutions, and the number of legal issues considered by the Supreme Court in each case. On the whole, these variables summarize (in admittedly reduced-form manner) the complex appeal process leading to the heterogeneous set of cases handled by state Supreme Courts.

Table A.7 in the Appendix summarizes the case-specific data, including the proportion of unanimous and minimal winning votes in each state. (Table 3 in online Appendix D provides summaries per state.) While a majority of cases are decided by unanimous decisions, there is also a sizable fraction of non-unanimous verdicts. Moreover, on average, there is a smaller fraction of unanimous verdicts in courts composed of elected judges (69%), than in those composed of appointed judges (over 80% of cases), a pattern which is somewhat at odds with the (Visser and Swank, 2007) model.

5.1.1. Justice-specific covariates

We include three classes of justice-specific covariates: *experience* variables, *institutional* variables, and *context* variables. Covariates which describe the justice before she became a state SC judge control for the selection effect, while covariates which vary across time as the justice is in office control for incentive effects.

Experience variables include the number of years of prior judicial experience, whether each justice had prior political experience or not, and the number of years serving in the state Supreme Court. Institutional variables describe the selection and retention methods in the state in which the justice serves. In particular, we capture here whether the justice was elected or appointed, and in this case, whether she was appointed for life by elected officials, appointed for one term by elected officials with a possible reappointment by the same elected officials, or appointed for one term by elected officials with a possible reappointment depending on an up-or-down decision by voters in a retention election.[19] As an additional selection covariate, we include Brace et al. (2000)'s party-adjusted judicial ideology (PAJID) score for each Justice at the time of appointment. The context variables include the interaction of the institutional variables with the (updated version of) Berry et al.'s (1998) citizen (CIT) and government (GOV) ideology for the relevant state in the year in which the decision was made. For both PAJID and CIT, larger values denote a more liberal stance.[20] The justice-specific data is summarized in Table A.8 in the Appendix. (Table 4 in online Appendix D provides summaries by state.)

5.2. First stage coefficients

This section has two purposes. We begin by discussing the first-stage estimates. We then present a full example of our second stage estimates to aid the interpretation of the general results. We leave the discussion of the general substantive results and the "economic" significance of covariates for the next section.

Table 1 presents the "first stage" MLE estimates of the coefficients of the common prior function $\rho(X_t)$, and of the state-contingent probabilities of voting to overturn correctly ($\gamma_1(X_t,Z_{it})$) and incorrectly ($\gamma_0(X_t,Z_{it})$).

First note that all but one of the case-specific covariates have a statistically significant effect on either justices' prior belief about the case, or their conditional voting probabilities. This suggests that our case-specific covariates capture significant variation due to heterogeneity in case-selection across states. In particular, we find that on average justices make better decisions in cases considered on appeal, voting more often in favor of the Petitioner when it should win, and against it when it should lose. This is intuitive, as the decision of the lower court provides valuable information. Justices are also more likely to overturn an incorrect decision of the lower court, and to uphold a correct decision, when the original ruling is in favor of the State (Petitioner Person). Thus, ex post monitoring is worse when the original ruling is against the state. In addition, the combined coefficients of murder and death penalty imply that justices are significantly less likely to overturn a death penalty conviction, in particular when they should in fact do so.

Consider next the justice-specific covariates. The variable PAJID captures the political "preferences" of the relevant principal in the selection of each justice (be it voters or elected officials) at the time of appointment. This captures a *selection* effect. We find that a higher value of PAJID (a more liberal principal) on average makes judges less likely to overturn the decision of the lower court, both when this should be upheld and (especially) when it should not. There is also an additional selection effect associated with competitive elections. On average, elected judges tend to make more mistakes, overturning more often when they should uphold and less often when they should overturn. Moreover, all experience variables (judicial experience, political experience, and experience in the court) have a statistically significant effect on the conditional voting probabilities $\gamma_0(\cdot)$ and $\gamma_1(\cdot)$. More years of experience in the court and more political experience induce a lower probability of overturning in both states, and additional years of prior judicial experience increase the probability of correctly overturning and upholding the decision of the lower court.

The interaction of the institutional variables with the context variables CIT and GOV at the time of the decision captures *incentive* effects. Relative to its negligible effect in retention elections systems, higher CIT values (more liberal voters) are associated with both a higher probability of correctly upholding ($1-p_0$) and correctly overturning (p_1) lower courts in competitive elections. This suggests that more liberal voters tend to have better judges when they use competitive elections, while more conservative voters have more competent judges when they use retention elections. Moreover, judges facing more liberal voters

[19] There is further variability within these classes. In all states in which justices are originally appointed and later face a retention election, the appointment is made by the Governor from nominees selected by a nominating commission. However, the term of the appointments can vary. In other states, the Governor's appointment requires the confirmation Senate, and in others the appointment is a legislative action. Terms also vary. For more details, see http://www.judicialselection.us/.

[20] CIT is the measure of citizen ideology proposed by Berry et al. The measure infers the ideological position of the electorate from the ideological orientations of members of Congress, as operationalized by interest-group ratings. Berry et al.'s measure of state elite's ideology is defined similarly, using the ideology of members of congress to estimate the ideological positions of state legislators and the Governor. Brace et al.'s PAJID variable builds on Berry et al.'s measures of citizen and elite ideology, but also incorporates information about the political party of each judge.

Table 1
"First stage" MLE estimates (standard errors in parentheses).

		ρ	γ_{it0}	γ_{it1}
Case specific	Constant	0.416	−0.475	0.771
		(0.170)	(0.121)	(0.168)
	Appeal	−1.449	−1.262	0.536
		(0.146)	(0.086)	(0.099)
	Petitioner Person	1.390	−0.342	1.502
		(0.112)	(0.061)	(0.086)
	Murder case	0.009	−0.142	0.128
		(0.010)	(0.060)	(0.050)
	Issues	0.080	−0.064	0.037
		(0.023)	(0.017)	(0.012)
	Evidence	0.179	−0.061	0.016
		(0.075)	(0.037)	(0.012)
	Jury instruction	0.186	−0.127	−0.030
		(0.080)	(0.057)	(0.021)
	Death penalty	−0.084	0.130	−0.288
		(0.086)	(0.077)	(0.075)
	Formal opinion	0.259	0.120	0.021
		(0.089)	(0.046)	(0.021)
	Jud. Review (US)	−0.011	−0.157	−0.362
		(0.511)	(0.159)	(0.131)
	Jud. Review (State)	0.060	0.818	0.064
		(0.149)	(0.129)	(0.091)
Justice/case specific	Years of experience in the court		−0.001	−0.002
			γ_{it0}	γ_{it1}
Justice specific	PAJID		−0.001	−0.004
			(0.001)	(0.001)
	Elected		0.546	−0.269
			(0.102)	(0.093)
	Life		0.453	−0.025
			(0.672)	(0.096)
	Reappt		−0.237	0.871
			(0.338)	(0.145)
	CIT		0.000	0.000
			(0.000)	(0.000)
	GOV		0.000	0.000
			(0.000)	(0.000)
	CIT∗Elected		−0.010	0.017
			(0.002)	(0.003)
	CIT∗Reappt.		−0.020	0.000
			(0.011)	(0.001)
	CIT∗Life		−0.052	−0.002
			(0.020)	(0.004)
	GOV∗Elected		0.006	−0.013
			(0.001)	(0.001)
	GOV∗Reappt.		0.010	−0.001
			(0.003)	(0.001)
	GOV∗Life		0.009	0.003
			(0.011)	(0.002)
	Judicial experience		−0.032	0.009
			(0.006)	(0.004)
	Political experience		−0.349	−0.472
			(0.113)	(0.107)

tend to incorrectly overturn more often when they are subject to reappointment than if they are appointed for life, and even more so when they face competitive elections and retention elections. On the other hand, judges facing more liberal voters correctly overturn more often if they are elected, but not if they are appointed for life, reappointed, or face retention elections. The coefficients of the GOV measure of elite ideology imply that judges in retention election systems are as insensitive to politicians' ideology measures as judges appointed for life. Instead, judges facing more liberal politicians incorrectly overturn more often when they are to be reappointed than when they are appointed for life. Finally, we also find that GOV has a differential effect on elected judges. Interestingly, the effect of politicians' ideology goes in the direction of compensating the corresponding effect of voters' ideology in competitive election systems.

All in all, the results of the first-stage are very compelling, and provide strong evidence of a statistically and substantively significant effect of political institutions on justices' prior beliefs and equilibrium conditional voting probabilities.

5.3. Second stage estimates: three sample courts

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 $\gamma_{i,t,1}=\gamma_1(X_t,Z_{it})$ and $\gamma_{i,t,0}=\gamma_0(X_t,Z_{it})$ that a justice with characteristics Z_{it} in case t votes to overturn correctly and incorrectly. For a given court composition C, we can then use the predicted values of $\gamma_{i,t,1}$ and $\gamma_{i,t,0}$ for each member i of C to recover the case and justice specific values of s^*_{it}, and the "deep parameters" θ_{it} and π_{it}. To describe the main results we fix all case-specific covariates at the state-specific sample means, and use the justice-specific covariates of the justices sitting in each state's Supreme Court.[21]

We begin by presenting the complete set of estimates for three sample courts – the LNCs of California, Connecticut and Texas – to aid the interpretation of the general results.

In the table, we indicate the MLE estimate of the common prior probability that the decision of the lower court should (according to the law) be overturned in the average case in each state. In these examples, this prior probability that overturning the decision of the lower court is required is $\rho=0.72$ for California, $\rho=0.65$ for Connecticut, and $\rho=0.69$ for Texas. This indicates that given their specific case selection, in all three states the common prior belief favors overturning the decision of the lower court.

The first two columns present the MLE estimates of the probability that Justice i votes to overturn when she should uphold (γ_{it0}) and when she should in fact overturn (γ_{it1}). Thus, for example, Justice Marvin Baxter of California had a probability of $\gamma_{it1}=0.93$ of correctly voting to overturn in instances in which he should overturn, and a probability of $1-\gamma_{it0}=1-0.10=0.90$ of correctly voting to uphold when he should uphold. Similarly, Justice Robert Berdon of Connecticut had a probability of $\gamma_{it1}=0.94$ of correctly voting to overturn, and a probability of $1-\gamma_{it0}=1-0.02=0.98$ of correctly voting to uphold.

Column 3 presents the estimate of the quality of the information of each justice. The higher quality-of-information estimate for Justice Berdon (3.59) vis a vis that of Justice Baxter (2.76) reflects mainly Berdon's lower probability of (incorrectly) voting to overturn when the law favors the Respondent (0.02 vs. 0.10). Column 4 presents the equilibrium cutpoint. This is the signal threshold s^*_i such that Justice i votes to uphold whenever she observes a signal below s^*_i and to overturn otherwise. Thus, while Justice Berdon would vote to uphold after observing a signal below $s^*_{BER}=0.57$, it would take a signal below $s^*_{BAX}=0.47$ for Justice Baxter to do the same. Equilibrium strategies are then a result of the prior belief, the quality of information and the preferences of each justice. Justices' bias are shown in columns 5 and 6 in the table. Note that in both the strategic and the expressive voting models, Justice Baxter (CA) is more inclined to overturn than Justice Berdon (CT). In the expressive voting model, for example, Justice Baxter requires less evidence to overturn (a belief of at least $\pi^{exp}_{BAX}=0.67$ that the standing decision is incorrect) than Justice Berdon ($\pi^{exp}_{BER}=0.82$).

Using these estimates we can compute a measure of the value of information in the court, as introduced in Iaryczower and Shum (2012). The measure, FLEX, is the probability that Justice i votes differently than what she would have voted for in the absence of her private case information:

$$\mathrm{FLEX}_{it}=\begin{cases}\rho_t\Phi(\theta_{it}[s^*_{it}-1])+(1-\rho_t)\Phi(\theta_{it}s^*_{it}) & \text{if } \rho_t\geq\pi_{it}\\ \rho_t[1-\Phi(\theta_{it}[s^*_{it}-1])]+(1-\rho_t)[1-\Phi(\theta_{it}s^*_{it})] & \text{if } \rho_t\geq\pi_{it}.\end{cases} \quad (10)$$

FLEX is bounded between zero and one, and takes a value of zero for individuals with extremely large biases toward overturning ($\pi\to0$) or

[21] When there is more than one court composition (natural court) per state in the data (as it usually is the case), we report results for the largest natural court (LNC); i.e. the court that decided more cases than any other natural court of the same state.

Table 2
Full set of estimates for California, Connecticut and Texas. Case-specific covariates fixed at state sample average; justices evaluated at individual-specific covariates. Standard errors in parentheses.

	Justice	γ_{it0}	γ_{it1}	θ	S^*	π^{exp}	π^{ST}
California (ρ = 0.72)	Baxter, M.R.	0.098 (0.010)	0.929 (0.009)	2.762 (0.085)	0.469 (0.017)	0.668 (0.019)	0.635 (0.073)
	Chin, MW.	0.084 (0.009)	0.936 (0.009)	2.899 (0.086)	0.476 (0.016)	0.675 (0.018)	0.636 (0.077)
	George, R.M.	0.080 (0.009)	0.932 (0.009)	2.893 (0.085)	0.486 (0.016)	0.694 (0.016)	0.645 (0.076)
	Kennard, J.L.	0.095 (0.010)	0.929 (0.009)	2.779 (0.084)	0.472 (0.017)	0.673 (0.018)	0.637 (0.073)
	Mosk, S.	0.049 (0.008)	0.857 (0.025)	2.720 (0.132)	0.608 (0.027)	0.850 (0.024)	0.718 (0.064)
	Werdegar, K.M.	0.095 (0.010)	0.930 (0.009)	2.784 (0.084)	0.470 (0.017)	0.669 (0.018)	0.636 (0.076)
	Brown, J.R.	0.100 (0.010)	0.934 (0.009)	2.787 (0.086)	0.461 (0.017)	0.652 (0.020)	0.629 (0.075)
Connecticut (ρ = 0.65)	Berdon, R.I.	0.021 (0.007)	0.939 (0.012)	3.586 (0.167)	0.568 (0.024)	0.820 (0.055)	0.891 (0.106)
	Borden, D.M.	0.024 (0.008)	0.954 (0.009)	3.669 (0.170)	0.541 (0.024)	0.768 (0.054)	0.877 (0.107)
	Callahan, R.J.	0.030 (0.010)	0.953 (0.009)	3.550 (0.168)	0.528 (0.025)	0.731 (0.049)	0.864 (0.112)
	Katz, J.	0.046 (0.015)	0.958 (0.008)	3.407 (0.170)	0.494 (0.028)	0.638 (0.035)	0.844 (0.118)
	Norcott Jr., F.L.	0.034 (0.011)	0.961 (0.007)	3.581 (0.166)	0.509 (0.025)	0.680 (0.040)	0.851 (0.113)
	Peters, E.A.	0.048 (0.016)	0.947 (0.010)	3.286 (0.175)	0.508 (0.029)	0.673 (0.038)	0.853 (0.101)
	Palmer, R.N.	0.048 (0.016)	0.948 (0.010)	3.291 (0.175)	0.505 (0.029)	0.668 (0.037)	0.852 (0.107)
Texas (ρ = 0.69)	Baird, C.F.	0.176 (0.018)	0.917 (0.011)	2.317 (0.103)	0.402 (0.020)	0.568 (0.027)	0.082 (0.040)
	Clinton, S.H.	0.176 (0.018)	0.919 (0.010)	2.325 (0.104)	0.400 (0.020)	0.565 (0.030)	0.081 (0.038)
	Keller, S.	0.180 (0.018)	0.925 (0.010)	2.351 (0.104)	0.389 (0.019)	0.547 (0.033)	0.079 (0.057)
	Maloney, F.	0.177 (0.018)	0.918 (0.010)	2.322 (0.103)	0.400 (0.020)	0.565 (0.026)	0.081 (0.025)
	Mansfield, S.	0.180 (0.018)	0.925 (0.010)	2.351 (0.104)	0.389 (0.019)	0.546 (0.030)	0.079 (0.063)
	McCormick, M.J.	0.176 (0.018)	0.919 (0.010)	2.325 (0.103)	0.400 (0.020)	0.565 (0.029)	0.081 (0.030)
	Meyers, L.E.	0.162 (0.016)	0.925 (0.010)	2.426 (0.103)	0.406 (0.018)	0.562 (0.029)	0.081 (0.037)
	Overstreet, M.L.	0.159 (0.016)	0.919 (0.010)	2.398 (0.102)	0.417 (0.019)	0.580 (0.028)	0.083 (0.091)
	White, B.M.	0.175 (0.017)	0.916 (0.011)	2.314 (0.103)	0.404 (0.020)	0.571 (0.028)	0.082 (0.026)

upholding the decision of the lower court ($\pi \to 1$).[22] In general, FLEX scores integrate information about the quality of information and bias of each justice.

5.4. Main results

Proceeding as in Section 5.3, we can compute the prior ρ, voting probabilities $(\gamma_{i,0},\gamma_{i,1})$, equilibrium strategy cutpoint s_i^*, and the justice type (θ_i,π_i) for the court in each state. As in Table 2, we fix all case-specific covariates at their state-specific sample mean, and use the justice-specific covariates of the justices sitting in each state's Supreme Court. This allows us to exclude from this comparison the heterogeneity in the type of cases considered by each court that is captured in the first stage coefficient estimates.[23]

The full results are presented in Table 1 in online Appendix D. Table 3 below summarizes the results by type of institution. As we described before, the states are arranged in four groups. The first is the group of states in which justices are elected in competitive plurality elections. The second group includes states in which justices are originally appointed by elected officials, but face an up-or-down decision by voters in a retention election to retain their position in the court. The third group includes states in which justices are appointed by elected officials, and considered for reappointment after a first term also by elective officials. The fourth group includes states in which justices are appointed by elected officials for life.[24]

Table 3 shows that differences in the institutions of selection and retention of justices are associated with systematic differences in the quality, preferences, and value of information in the court. First, justices that do not face any kind of voter evaluation after being appointed on average have higher quality of information than justices that face reelection or retention elections. In fact, the information quality for justices appointed for life and justices that are appointed and reappointed is on average 33% larger than that of justices facing retention elections, and 39% larger than that of justices that are elected. Moreover, elected justices are also typically more inclined to overturn the decision of the lower court than those who do not face a voter evaluation after being appointed. In the expressive voting model, for instance, the average elected justice would vote to overturn the decision of the lower court if her belief that the lower court's decision is incorrect is above $E[\pi^{exp}|elected]=0.60$. Instead, the average justice subject to reappointment would vote to overturn only if her posterior belief is above $E[\pi^{exp}|Reapp]=0.69$, while a justice appointed for life would do so only if her belief is at least $E[\pi^{exp}|life]=0.90$. Table 3 shows, moreover, that differences in information quality across institutional environments trump differences in bias. In fact, the average FLEX scores for elected justices (0.37) and justices facing retention elections (0.36) are significantly lower than the corresponding FLEX scores for appointed justices facing political reappointment (0.48), and for justices appointed for life (0.60).

The differences in the quality and bias of elected and appointed judges persist even if we suppress differences across states in case selection, or in voters' and politicians' ideology. Fig. 1 plots the average bias and quality of information for each individual state.

The top-left panel shows the benchmark results, with case-specific covariates fixed at each state sample average, and individual justices evaluated at their own justice-specific covariates. The top-right panel plots the estimates of bias and quality of information for each state with the CIT and GOV variables set at their overall sample averages for all justices. Eliminating the differences in citizens' and elite's ideologies has the effect of making the group of "elected" states more homogeneous, but otherwise preserves differences in types across systems. The bottom-left panel plots the estimates of bias and quality of information for each state with the case covariates set at the overall sample averages. Compared to the benchmark, eliminating differences in case selection across states makes "elected", and in particular also "retention" states more homogeneous, but again preserves differences in

[22] Note 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 $\rho \leq \pi_i$. This is because the equilibrium cutpoint s_i^* recovered from the data is the same in the expressive or strategic voting models. 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. If instead we were initially given values of $\{\pi_i,\theta_i\}$ and ρ, the two models would imply a different cutpoint s_i^*, and FLEX scores in the two models would differ significantly.

[23] The substantial variation in the estimated priors across states suggests that in fact we are able to control for a significant amount of heterogeneity in case-selection across states.

[24] We include New Jersey in this group because upon being reappointed, justices are appointed for life. Illinois, New Mexico and Pennsylvania have up-or-down retention elections for reappointment.

Table 3
Type, prior, strategy and conditional voting probabilities. Average across states per electoral institution.

Institution	ρ	γ_{it0}	γ_{it1}	θ	S^*	$\pi^{\exp}$	π^{st}	FLEX$^{\exp}$	FLEXst
Elected	0.674	0.118	0.919	2.614	0.458	0.606	0.372	0.371	0.343
Appointed, with voter retention	0.702	0.101	0.921	2.734	0.475	0.661	0.528	0.359	0.352
Appointed, with reappointment	0.680	0.038	0.961	3.571	0.504	0.692	0.678	0.481	0.505
Appointed for life	0.662	0.014	0.912	3.733	0.628	0.900	0.984	0.609	0.609
All states	0.683	0.091	0.926	2.895	0.484	0.661	0.520	0.404	0.393

types across systems. The figure on the bottom-right plots bias and quality of information with the CIT and GOV variables and the case covariates set at the overall sample averages. As in the previous cases, the differences in types across systems persist.

5.5. Effectiveness of bureaucrats and politicians

Table 3 focused on the correlation between selection and retention procedures and justices's bias and quality of information. In our next results, we switch attention from the type of justices to their performance: is there a systematic difference in the effectiveness of elected and appointed justices? A natural measure of performance in our model is the probability of a mistake in the decision of the court. In this section we use the estimated individual conditional voting probabilities to compute this probability.

Note that for any given case characteristics X, our first stage estimates provide the probability that a member i of a given court j votes to uphold when the decision of the lower court should be overturned $1-\gamma_{i,1}$, and to overturn when the decision of the lower court should be upheld, $\gamma_{i,0}$ (we drop the obvious dependence on X to simplify notation). We can then use these voting probabilities to compute the probability that court j will incorrectly uphold and overturn the decision of the lower court under a simple majority rule, $Pr(v_j=0|\omega=1)$, and $Pr(v_j=1|\omega=0)$. Given a prior ρ_j, we can then compute the probability of an incorrect decision for court j,

$$\beta_j^{SM} = \rho_j \Pr\left(v_j = 0|\omega = 1\right) + \left(1-\rho_j\right)\Pr\left(v_j = 1|\omega = 0\right).$$

Fig. 2 plots β_j^{SM} for each state, together with the type I and type II errors. The total probability of an incorrect ruling β_j^{SC} (the bars in the figure) ranges from under 0.1% for the top five states – New York, Maine, Virginia, Connecticut and Oklahoma – to between 0.7% and 3.6% for the bottom five states – Colorado, Nevada, Tennessee, Louisiana and New Mexico. Thus, even when individual members have a much larger probability of making a wrong decision (see Table 3), the "wisdom of the majority" implies that state Supreme Courts have a relatively low total error rate.

The pattern of mistakes, however, is highly asymmetric. State courts tend to incorrectly overturn the decision of the lower courts more frequently than to incorrectly uphold the lower courts (e.g., 1.8% vs.

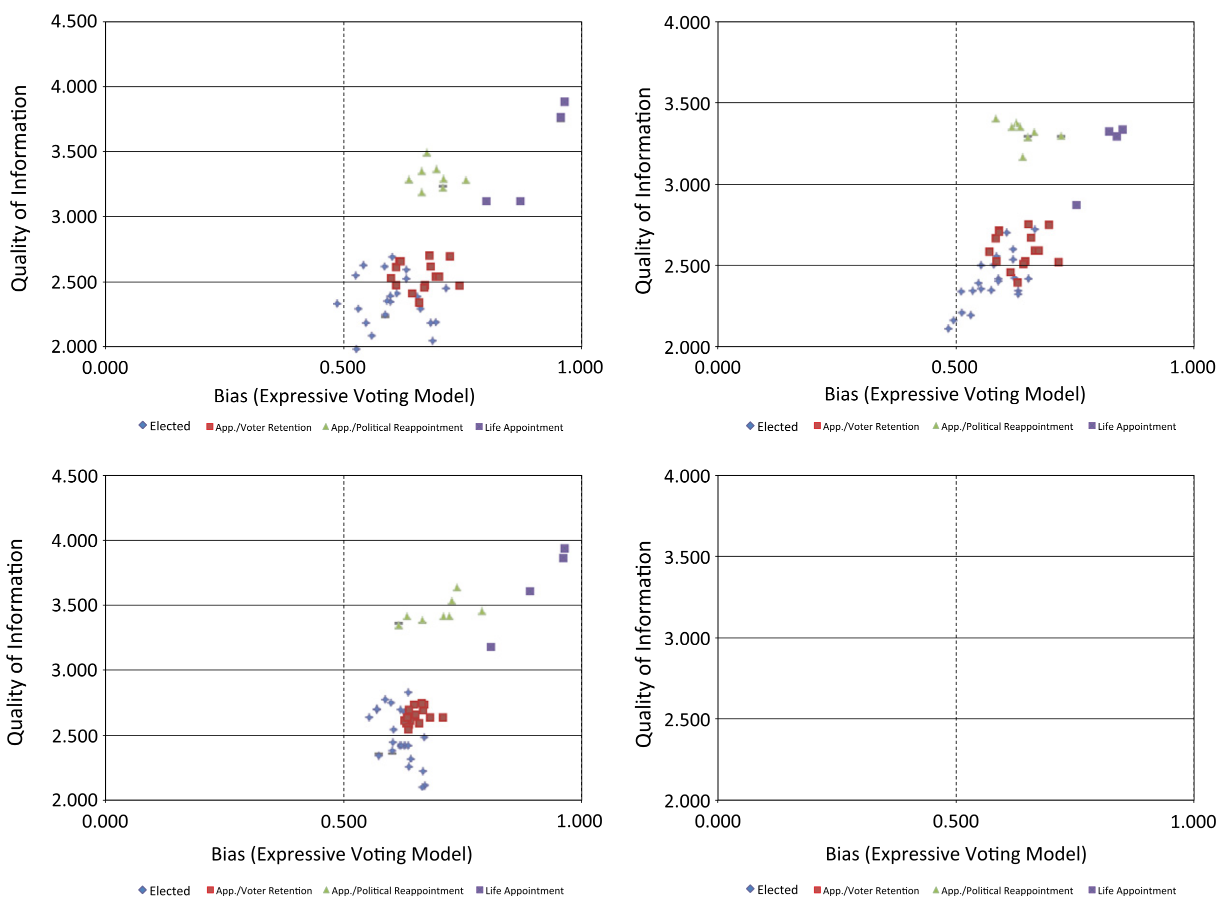


Fig. 1. Bias (expressive voting) and quality of information, per state. Top-left: benchmark results; top-right: CIT/GOV at sample averages; bottom-Left: case covariates at sample averages; and bottom-right: CIT/GOV and case covariates at sample averages.

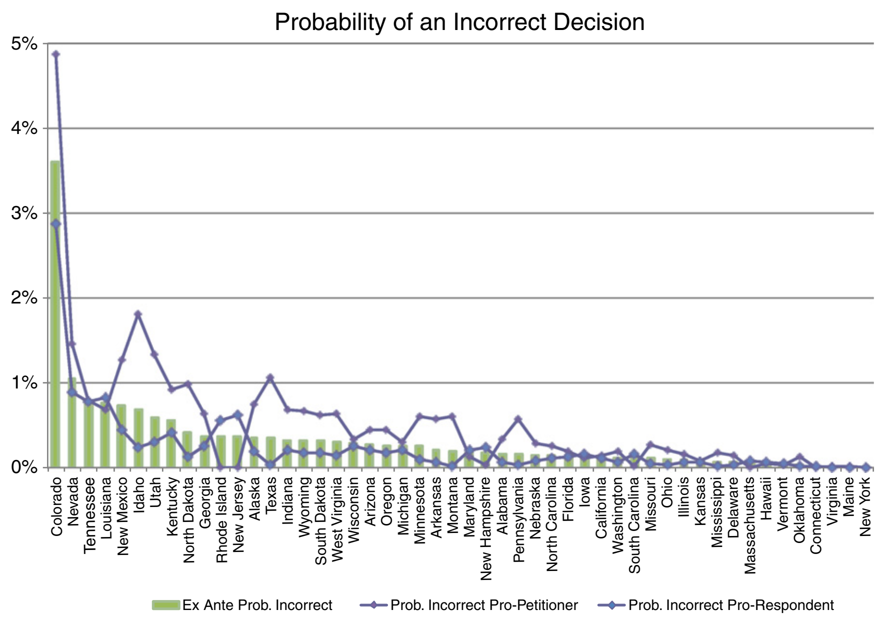


Fig. 2. Probability of an incorrect decision at the court level. Type I and type II errors.

0.2 % in Idaho, 1.5% vs. 0.9% in Nevada, 1.3% vs. for 0.3% in Utah, 1.1% vs. 0.03% in Texas). This is true in particular for judges that are elected or face retention elections. In fact, all but one of the twenty two states in which justices are elected (95%), and all but two of the sixteen states in which judges face retention elections (88%) have this property. Instead, only three of the eight states with reappointment (38%) and none of the four states with life appointment have a similar feature. This asymmetry in mistakes at the level of the court is in line with our earlier results at the individual level, which showed that judges that are elected and judges that face retention elections vote to overturn incorrectly at a higher rate than to uphold incorrectly (Table 3).

Aggregating the effectiveness results by institution type reinforces the conclusions in Section 5.4. We showed before that appointed judges on average have higher quality of information than elected judges. We also showed that appointed judges on average are more inclined to change their mind from the position they would have taken without hearing the particulars of the case than elected judges. In addition, judges that are shielded from voters' influence on average also have a lower probability of reaching an incorrect decision (0.1%) than justices that face retention elections (0.5%), and elected justices (0.3%). The effect is larger when we consider the probability of incorrectly overturning the decision of the lower courts. While judges that are shielded from voters' influence incorrectly overturn the lower court very infrequently (0.03%), the corresponding probabilities are 0.7% for justices facing retention elections and 0.6% for justices facing competitive reelections.

5.5.1. Counterfactuals: can unanimity rule improve performance?

Since state courts tend to wrongly overturn more frequently than to wrongly uphold the decisions of the lower courts, it is interesting to compare the performance of the courts under the current rules with a counterfactual scenario in which overturning lower courts' decisions requires the unanimous consent of all members of the court.

To evaluate this counterfactual scenario, we need to compute the probability of mistakes under unanimity. 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 upholding incorrectly is $1-\prod_{i=1}^{n_j}(1-\gamma_{i,1})$ and the probability of the court overturning incorrectly is $\prod_{i=1}^{9}\gamma_{i,0}$. Thus the total probability of an incorrect ruling for the Supreme Court under unanimity rule in the expressive voting model is $\beta_j^{U,exp}$

$$\beta_j^{U,exp}=\rho_j\left[1-\prod_{i=1}^{n_j}\left(1-\gamma_{i,1}\right)\right]+\left(1-\rho_j\right)\left[\prod_{i=1}^{n_j}\gamma_{i,0}\right].$$

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

$$\beta_j^{U,st}=\rho_j\left[1-\prod_{i=1}^{n_j}\left(1-\gamma_{i,1}^{**}\right)\right]+\left(1-\rho_j\right)\left[\prod_{i=1}^{n_j}\gamma_{i,0}^{**}\right].$$

Table 4 shows the results per state, grouped as before by class of political institution. The results show that introducing the change to unanimity rule would have major consequences on policy outcomes and the effectiveness of the court.

In the expressive voting model, where justices care about their own vote only, replacing majority rule by unanimity rule does achieve the purpose of reducing the probability that the court incorrectly overturns lower courts' decisions (column 4 in Table 4). However, it does so only by dramatically increasing the probability of incorrectly upholding lower courts' decisions (reaching 43.6% for elected justices, and 39.4% for non-elected justices facing retention elections).

The change to unanimity rule also increases the probability of errors in the strategic voting model, although in a less dramatic fashion. Here, however, the changes occur in the *opposite* direction. As a result of the move to unanimity, strategic justices who care about the decision of the court would modify their voting strategy. Because being pivotal (all other n_j-1 members voting to overturn) now carries more favorable information in favor of overturning the decision of the

Table 4
Probability of mistakes: majority rule and unanimity rule.

Institution	Simple majority (actual)			Unanimity (expensive)			Unanimity (strategic)		
	Pr. wrong overturn	Pr. wrong uphold	β_j^{SM}	Pr. wrong overturn	Pr. wrong uphold	$\beta_j^{U,exp}$	Pr. wrong overturn	Pr. wrong uphold	$\beta_j^{U,st}$
Elected	0.6%	0.2%	0.3%	0.0%	43.6%	29.0%	6.1%	1.4%	2.8%
Appointed, with voter retention	0.7%	0.4%	0.5%	0.0%	39.4%	27.3%	2.6%	1.8%	2.0%
Appointed, with reappointment	0.0%	0.0%	0.0%	0.0%	21.1%	14.2%	0.3%	0.5%	0.4%
Appointed for life	0.0%	0.4%	0.2%	0.0%	42.1%	27.3%	0.0%	2.8%	1.8%
All states	0.5%	0.2%	0.3%	0.0%	38.5%	26.0%	3.6%	1.5%	2.1%

lower court, in equilibrium all justices must become less inclined to uphold the decision of the lower court (see Feddersen and Pesendorfer, 1998). We estimate that the strategic effect associated with the change to unanimity significantly increases the probability of a mistaken decision to overturn, reaching 2.8% for elected justices, and 2% for justices facing retention elections. Moreover, because of the relatively extreme bias of justices appointed for life, the change to unanimity rule also increases their probability of mistakes to about 1.8% on average.

5.6. Alternative specification: decisions for and against the state

In the benchmark specification of the model, we coded justices' votes as either in favor of overturning or upholding the decision of the lower court. This follows from the logic that the key consideration at the Supreme Court level is not determining the guilt or innocence of the accused, but instead to assess whether or not errors have been committed at trial. In this benchmark specification, then, the bias parameter $\pi_i(X)$ must be interpreted as capturing a preference toward upholding or overturning the lower court given characteristics X: $\pi_i(X)>1/2$ provides a hurdle to overturn the lower court's decision, while $\pi_i(X)<1/2$ provides a hurdle to uphold the lower court's decision.

One could argue, however, that judges might instead base their decisions on the underlying issue, so that a coding of votes as in favor or against the State (the prosecution) would more accurately capture the true dimension of conflict. With this alternative coding of votes (i.e. $v_i=1$ denotes a vote for the prosecution), the bias parameter $\pi_i(X)$ has to be interpreted as measuring the relative preference against the State in a case with characteristics X; i.e., a judge with bias $\pi_i(X)$ votes in favor of state if and only if her belief that the prosecution should win the case based on the information available is above $\pi_i(X)$.

In this section we report the results of reestimating the model based on coding votes as in favor or against the state ("FAST"). Below we discuss the first-stage coefficient estimates, the estimates of bias and quality of information, and the probability of mistakes.

5.6.1. First-stage coefficients

The first-stage coefficient estimates in the FAST specification are qualitatively very similar to the coefficient estimates in the benchmark uphold/overturn (UO) specification. The results are presented in Table 5.

As before, all but one of the case-specific covariates have a statistically significant effect on either justices' prior belief about the case, or their conditional voting probabilities. As in the UO specification, we find that on average justices make better decisions in cases considered on appeal, and when the original ruling is in favor of the State. And as before, judges are more likely to vote for the prosecution (less likely to overturn) after a death penalty conviction in the lower court.

The same is true for justice-specific covariates. A higher value of PAJID on average makes judges more likely to rule in favor of the State, both when they should, and when they shouldn't. As before, there is also an additional selection effect, with elected judges making more mistakes on average. Moreover, as in the benchmark UO specification, the experience variables have a statistically significant effect. Judges with more political experience vote more often in favor of the State, those with more experience in the court vote less often in favor of the State, and as before, additional years of prior judicial experience reduces the probability of voting in favor of the State

Table 5
FAST specification: "first stage" MLE estimates (standard errors in parentheses).

		ρ	γ_{it0}	γ_{it1}
	Appeal	−1.391	−0.736	0.129
		(0.116)	(0.167)	(0.067)
	Petitioner Person	0.054	−1.396	0.749
		(0.040)	(0.078)	(0.088)
	Murder case	1.075	−1.617	0.882
		(0.113)	(0.063)	(0.084)
	Issues	−0.220	0.032	0.089
		(0.089)	(0.025)	(0.058)
	Evidence	−0.098	0.021	0.162
		(0.022)	(0.010)	(0.021)
	Jury instruction	−0.294	−0.312	−0.232
		(0.077)	(0.055)	(0.057)
	Death penalty	−0.220	−0.234	−0.283
		(0.082)	(0.058)	(0.061)
	Formal opinion	0.358	0.344	0.315
		(0.114)	(0.076)	(0.106)
	Jud. Review (US)	0.023	0.202	−0.150
		(0.029)	(0.069)	(0.053)
	Jud. Review (State)	−0.267	−0.012	0.437
		(0.274)	(0.085)	(0.394)
Justice/case specific	Years of experience in the court		−0.001	−0.002
			(0.001)	(0.001)
			γ_{it0}	γ_{it1}
Justice specific	PAJID		0.008	0.008
			(0.002)	(0.002)
	Elected		0.264	−0.629
			(0.108)	(0.114)
	Life		0.219	1.388
			(0.477)	(1.865)
	Reappt		−0.547	0.504
			(0.607)	(0.459)
	CIT		0.000	0.000
			(0.000)	(0.000)
	GOV		0.000	0.000
			(0.000)	(0.000)
	CIT∗Elected		−0.017	0.010
			(0.003)	(0.003)
	CIT∗Reappt.		−0.015	0.006
			(0.018)	(0.008)
	CIT∗Life		−0.020	−0.035
			(0.017)	(0.098)
	GOV∗Elected		0.012	−0.009
			(0.001)	(0.001)
	GOV∗Reappt.		0.007	−0.005
			(0.004)	(0.003)
	GOV∗Life		0.017	0.155
			(0.005)	(0.186)
	Judicial experience		−0.012	0.032
			(0.005)	(0.006)
	Political experience		0.595	0.708
			(0.098)	(0.144)

Table 6
FAST specification: types, strategies, and voting probabilities.

Institution	ρ	γ_{it0}	γ_{it1}	θ	S^*	π^{exp}	π^{st}	FLEX$^{\text{exp}}$	FLEX$^{\text{st}}$
Elected	0.280	0.051	0.852	2.727	0.610	0.466	0.903	0.294	0.276
Appointed, with voter retention	0.276	0.060	0.899	2.930	0.552	0.391	0.677	0.334	0.313
Appointed, with reappointment	0.278	0.022	0.955	3.754	0.543	0.414	0.748	0.320	0.333
Appointed for life	0.277	0.043	0.995	5.440	0.333	0.035	0.002	0.653	0.694
All states	0.278	0.049	0.895	3.173	0.559	0.399	0.734	0.339	0.330

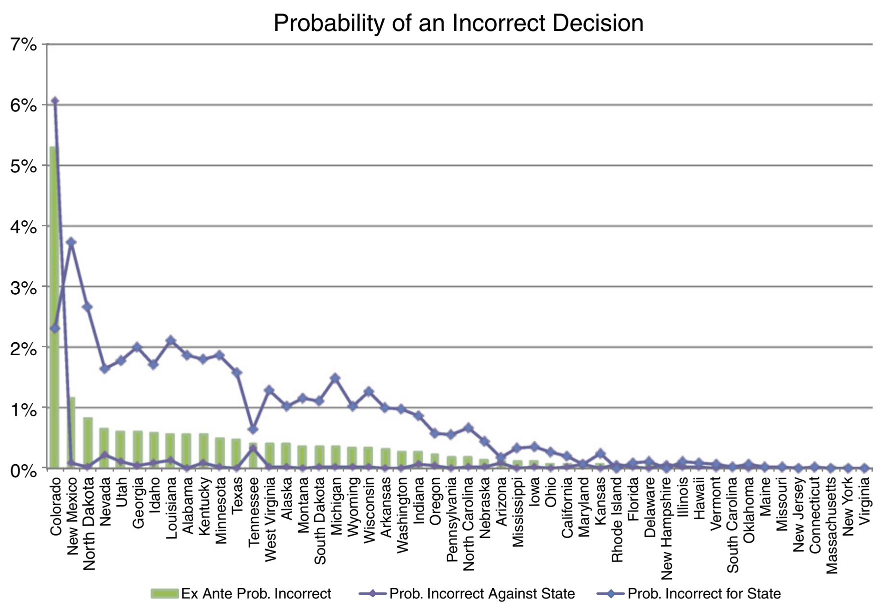


Fig. 3. FAST specification: prob. of incorrect decisions in the court.

when it should lose and against the State when it should win. Incentive effects are also similar as before. In particular, we find that judges' voting behavior is insensitive to the voters' ideology scores in retention election systems, and for judges who are isolated from voters, but higher CIT values (more liberal voters) are associated with both a higher probability of correctly voting against the State $(1-p_0)$ and correctly voting for the State (p_1) in competitive elections.

5.6.2. Bias and quality of information

The main effect of changing from the benchmark UO specification to the FAST specification is to change the results and the interpretation of the bias estimates. In FAST, the bias parameter π_i captures i's relative preference against the State: i votes in favor of the State if and only if her belief that the prosecution should win the case based on the information available is above π_i.

Table 2 in online Appendix D presents the state averages of the estimates of type, strategy and voting probabilities in the FAST specification. Table 6 below summarizes the result by type of institution.

First, the differences in quality of information across systems of selection and retention in the benchmark UO specification are qualitatively unchanged in the FAST specification. In fact, FAST amplifies the differences in quality between systems. On average, the quality of information for justices appointed for life is 45% larger than that of justices subject to reappointment, and $E[\theta|Reapp]$ in turn is 28% higher than $E[\theta|Retention]$ and 37% higher than $E[\theta|Elected]$. Moreover, the estimates of justices' prior beliefs given state-average case characteristics indicate that on information grounds justices are typically predisposed against the state. And in the strategic voting model, justices are also typically predisposed against the state in terms of preferences. On average, an elected judge votes for the state if her belief that the prosecution should win is at least 90%. This threshold goes down to 75% for judges subject to reappointment and to 68% for judges in retention election systems. In the expressive voting model, instead, our estimates imply that justices are typically moderately biased in favor of the State. On average, an elected judge votes for the state if her belief that the prosecution should win is at least 47%. This threshold goes down to 41% for judges subject to reappointment and to 39% for judges in retention election systems.[25] The difference reflects the informational content of being decisive. Given equilibrium strategies (typically biased against the state), being pivotal carries favorable information for the state's case.

5.6.3. Mistakes

Fig. 3 presents the probability of mistakes in the Court, by state, in the FAST specification. The results are qualitatively similar to the corresponding results in the benchmark UO specification. Moreover, here too the errors are asymmetric, although in this case in favor of the state. While in all but one state (Colorado) the probability of an incorrect decision against the state is negligible, the probability of an incorrect decision in favor of the state is at least 1% for twenty states, at least 1.5% for twelve states, and at least 2% for four states. As in the benchmark specification, the asymmetry is most pronounced in elected states, where the probability of an incorrect decision in favor of the state is at least 1% for fifteen of the twenty two states.

[25] Justices appointed for life on average have a large bias in favor of the state in both the strategic and expressive voting models.

6. Conclusion

What separates bureaucrats from politicians? This fundamental question for representative democracy has three parts. First, do voters select a different type of public official – more or less biased, better or worst at gathering and processing information –than government officials? Second, do reelections induce public officials to improve their proficiency to deal with the flow of information of each decision? Do they induce them to be more responsive to the public? Third, are bureaucrats more effective than politicians?

In order to answer these questions, we need to map institutions to the type of public officials they induce. The difficulty, of course, is that this type is unobservable. The contribution of this paper is to bridge this gap by specifying a decision-making model, and using equilibrium information to recover the unobservable types.

We focus on criminal decisions in US states' Supreme Courts. The main results we obtain clarify the trade-offs inherent in choosing between bureaucrats and politicians. First, justices that are shielded from voters' evaluations on average have higher quality of information than justices that face either reelection or retention elections. In fact, the information quality for justices that are shielded from voters' influence is on average 33% larger than that of justices facing retention elections, and 39% larger than that of justices that are elected. Institutions of selection and retention of justices also affect justices' bias: elected justices are also typically more inclined to overturn the decision of the lower court than those who do not face a voter evaluation after being appointed. However, this effect is more modest in magnitude. As a result, differences in information quality across jurisdictions trump differences in bias, and justices who are shielded from voters not only have better information, but are also more likely than elected justices to change their preconceived opinions about a case, and have a lower probability of making incorrect decisions than elected justices.

Finally, we show that while the pattern of mistakes of state Supreme Courts is highly asymmetric – with the courts more frequently wrongly overturning than wrongly upholding lower courts' decisions – changing the voting rule to a rule making it harder to overturn lower courts' decisions would produce major consequences to public outcomes and the effectiveness of the courts. Thus any such change should be considered with great care.

Appendix A. Additional Tables

Table A.7
Case-specific data.

	Obs	Mean	Std. dev.	Min	Max
Appeal	5958	0.910	0.286	0	1
Petitioner Person	5947	0.832	0.374	0	1
Formal opinion	5938	0.872	0.334	0	1
Jud. review (US Const.)	5925	0.032	0.175	0	1
Jud. review (St. Const.)	5929	0.042	0.200	0	1
Murder case	5958	0.373	0.484	0	1
No. legal issues	5958	2.310	2.106	0	28
Evidence	5958	0.579	0.494	0	1
Jury instruction	5958	0.376	0.484	0	1
Death penalty	5958	0.166	0.372	0	1
No. of justices per court	5958	6.559	1.277	5	9
Prop. votes to overturn	5958	0.358	0.430	0	1
Unanimous to overturn	5958	0.246	0.431	0	1
Unanimous to uphold	5958	0.520	0.500	0	1
Minimal winning	5958	0.080	0.272	0	1

Table A.8
Justice-specific data. Justices in largest natural courts (LNC) in each state. Average values of justice-specific covariates, per state (324 justices).

State	Elected	Apptd for life	Appt. w/ pol. reappointment	Prior judicial experience	Prior political experience	Years of experience	PAJID at appointment	CIT at decision	GOV at decision
N	520	520	520	507	507	510	453	520	520
Mean	0.49	0.06	0.16	6.87	0.15	5.50	39.49	45.85	41.02
Std. dev.	0.50	0.24	0.36	7.06	0.36	7.34	22.59	14.97	22.88
Min.	0.00	0.00	0.00	0.00	0.00	0.00	1.25	9.25	1.64
Max.	1.00	1.00	1.00	35.00	1.00	62.70	96.62	86.47	93.88

Appendix B. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jpubeco.2012.08.007.

References

Alesina, A., Tabellini, G., 2007. Bureaucrats or politicians? Part I: a single policy task. American Economic Review 97 (1), 169–179 (April).
Alesina, A., Tabellini, G., 2008. Bureaucrats or politicians? Part II: multiple policy tasks. Journal of Public Economics 92 (3–4), 426–447 (April).
Austen-Smith, D., Banks, J.S., 1996. Information aggregation, rationality, and the Condorcet Jury Theorem. American Political Science Review 90, 34–45 (Mar.).
Austen-Smith, D., Feddersen, T., 2005. Deliberation and Voting Rules. Social Choice and Strategic Decisions: Essays in Honor of Jeffrey S. Banks, pp. 269–316.
Austen-Smith, D., Feddersen, T., 2006. Deliberation, preference uncertainty, and voting rules. American Political Science Review 100, 209–218 (May).
Banks, J., Sundaram, R., 1998. Optimal retention in agency problems. Journal of Economic Theory 82 (2), 293–323 (October).
Barro, R., 1973. The control of politicians: an economic model. Public Choice 14 (1), 19–42 (March).
Berry, W., Ringquist, E., Fording, R., Hanson, R., 1998. Measuring citizen and government ideology in the American states, 1960–93. American Journal of Political Science 42 (1), 327–348 (January).
Besley, T., Case, A., 2003. Political institutions and policy choices: evidence from the United States. Journal of Economic Literature 41 (1), 7–73 (March).
Besley, T., Coate, S., 2003. Elected versus appointed regulators: theory and evidence. Journal of the European Economic Association 1, 1176–1206.

Besley, T.J., Payne, A.A., 2005. Implementation of anti-discrimination policy: does judicial selection matter? ISE STICERD Research Paper No. PEPP04.
Brace, P., Hall, M., 1990. Neo-institutionalism and dissent in State Supreme Courts. The Journal of Politics 52 (1), 54–70 (February).
Brace, P., Hall, M.G., 1993. Integrated models of judicial dissent. The Journal of Politics 55 (4), 914–935 (November).
Brace, P., Hall, M., 1997. The interplay of preferences, case facts, context, and rules in the politics of judicial choice. The Journal of Politics 59 (4), 1206–1231 (November).
Brace, P., Langer, L., Hall, M., 2000. Measuring the preferences of State Supreme Court judges. The Journal of Politics 66 (2), 387–413 (May).
Cameron, C.M., Kornhauser, L.A., 2008. Modeling collegial courts (3): judicial objectives, opinion content, voting and adjudication equilibria, law & Economics Research Paper Series. Working Paper No. 08–54.
Canes-Wrone, B., Shotts, K., 2007. When do elections encourage ideological rigidity? American Political Science Review 101 (2), 273–288 (May).
Canes-Wrone, B., Herron, M., Shotts, K., 2001. Leadership and pandering: a theory of executive policymaking. American Journal of Political Science 45 (3), 532–550 (July).
Choi, S.J., Gulati, G.M., Posner, E.A., 2010. Professionals or politicians: the uncertain empirical case for an elected rather than appointed judiciary. Journal of Law, Economics, and Organization 26 (2), 290–336.
Clinton, J.D., Jackman, S., Rivers, D., 2004. The statistical analysis of roll call data. American Political Science Review 55, 355–370 (May).
Coate, S., Conlin, M., 2004. A group rule-utilitarian approach to voter turnout: theory and evidence. American Economic Review 94, 1476–1504.
Coate, S., Conlin, M., Moro, A., 2008. The performance of pivotal-voter models in small-scale elections: evidence from Texas liquor referenda. Journal of Public Economics 92, 582–596.
Coughlan, P., 2000. In defense of unanimous jury verdicts: mistrials, communication, and strategic voting. American Political Science Review 94, 375–393 (Jun.).
Degan, A., Merlo, A., 2009. Do voters vote ideologically? Journal of Economic Theory 144, 1869–1894.
DePaula, A., Merlo, A., 2009. Identification and estimation of preference distributions when voters are ideological. Working Paper. U. Penn.
Diermeier, D., Keane, M., Merlo, A., 2005. A political economy model of congressional careers. American Economic Review 347–373 (March).
Doraszelski, U., Gerardi, D., Squintani, F., 2003. Communication and voting with double-sided information. Contributions to Theoretical Economics 3, 1–39.
Duggan, J., Martinelli, C., 2001. A Bayesian model of voting in juries. Games and Economic Behavior 37, 259–294.
Feddersen, T., Pesendorfer, W., 1997. Voting behavior and information aggregation in elections with private information. Econometrica 65, 1029–1058.
Feddersen, T., Pesendorfer, W., 1998. Convicting the innocent: the inferiority of unanimous jury verdicts under strategic voting. American Political Science Review 92, 23–35 (Mar.).
Ferejohn, J., 1986. Incumbent performance and electoral control. Public Choice 50, 5–25 (December).
Froeb, L.M., Kobayashi, B.H., 1996. Naive, biased, yet Bayesian: can juries interpret selectively produced evidence? Journal of Law, Economics, and Organization 12, 257–275.
Gely, R., Spiller, P., 1990. A rational choice theory of Supreme Court Statutory decisions with applications to the "State Farm" and "Grove City" cases. Journal of Law, Economics and Organization 6 (2), 263–300 (October).
Gerardi, D., Yariv, L., 2007. Deliberative voting. Journal of Economic Theory 134, 317–338 (May).
Gordon, S., Huber, G., 2007. The effect of electoral competitiveness on incumbent behavior. Quarterly Journal of Political Science 107–138 (April).
Hall, P., Zhou, X.-H., 2003. Nonparametric estimation of component distributions in a multivariate mixture. The Annals of Statistics 31, 201–224.
Hanssen, A., 2000. Independent courts and administrative agencies: an empirical analysis of the states. Journal of Law, Economics and Organization 16 (2), 534–571 (January).
Hanssen, A., 2004. Is there a politically optimal level of judicial independence? American Economic Review 94 (3), 712–729 (June).
Heckman, J., Snyder, J.J., 1997. Linear probability models of the demand for attributes with an empirical application to estimating the preferences of legislators. The Rand Journal of Economics 28 (0), S142–S189 (January, special Issue in Honor of Richard E. Quandt).
Helmke, G., 2002. The logic of strategic defection: court-executive relations in Argentina under dictatorship and democracy. American Political Science Review 96 (2), 291–303 (June).
Iaryczower, M., Shum, M., 2012. The value of information in the Court. Get it right, keep it tight. American Economic Review 102, 202–237 (February).
Iaryczower, M., Spiller, P., Tommasi, M., 2002. Judicial independence in unstable environments, Argentina 1935–1998. American Journal of Political Science 46 (4), 699–716 (October).
Iaryczower, M., Spiller, P., Tommasi, M., 2006. Judicial lobbying: the politics of Labor Law, constitutional interpretation. American Political Science Review 100, 85–97 (February).
Iaryczower, M., Katz, G., Saiegh, S., forthcoming. Voting in the bicameral Congress: largemajorities as a signal of quality. Journal of Law, Economics and Organization. http://dx.doi.org/10.1093/jleo/ews022.
Kawai, K., Watanabe, Y., forthcoming. Inferring strategic voting. American Economic Review.
Li, H., Rosen, S., Suen, W., 2001. Conflicts and common interests in committees. American Economic Review 91, 1478–1497.
Lim, C., 2011. Turnover and Accountability of Appointed and Elected Judges, Typeset. Stanford University.
Londregan, J., 1999. Estimating legislators' preferred points. Political Analysis 8 (1), 35–56 (January).
Martin, A., Quinn, K., 2002. Dynamic ideal point estimation via Markov Chain Monte Carlo for the US Supreme Court, 1953–1999. Political Analysis 10 (2), 134–153 (January).
Martin, A., Quinn, K., 2007. Assessing preference change on the US Supreme Court. Journal of Law, Economics and Organization 23 (2), 365–385 (January).
Maskin, E., Tirole, J., 2004. The politician and the judge: accountability in government. American Economic Review 94, 1034–1054 (January).
Poole, K., Rosenthal, H., 1985. A spatial model for legislative roll call analysis. American Journal of Political Science 29, 357–384.
Poole, K., Rosenthal, H., 1991. Patterns of congressional voting. American Journal of Political Science 35, 228–278.
Spiller, P., Gely, R., 1992. Congressional control or judicial independence: the determinants of U.S. Supreme Court labor-relations decisions, 1949–1988. The Rand Journal of Economics 23 (4), 463–492 (December).
Visser, B., Swank, O., 2007. On committees of experts. Quarterly Journal of Economics 122, 337–372.

To Elect or to Appoint? Appendix B: Expressive and Strategic Voting.

Matias Iaryczower, Garrett Lewis and Matthew Shum

June 3, 2012

In this appendix we consider a procedure to assess quantitatively whether the expressive or strategic voting model is more appropriate for the state supreme court voting data used in this paper. As we remarked in the main text, for a given set of cases, the first-stage likelihood function for the votes is identical for both the expressive and strategic voting models, and in this sense the two models are observationally equivalent. Here, we bring additional data to bear which was not used in our main analysis, in order to discriminate between these models.

Specifically, in the paper we only considered cases in which all members of the court vote. Because of this, we did not use the subsample of cases for which only $k < n$ of the court members (an "incomplete court") voted. This is the additional subsample which we use to compare the strategic and expressive voting models. In particular, using the estimates from the main analysis, we construct predicted voting probabilities for these "incomplete court" cases, under both the strategic and expressive voting hypotheses. We then compare the two models based on which one generates the higher likelihood according to these predicted voting probabilities. We proceed in the following sequence of steps.

1. We select four states (Massachusetts, Connecticut, Montana and Pennsylvania) in which the size of the full court is $n = 7$, and for which there are a relatively large number of cases in which only $k = 5$ justices vote. This gives us a total of 317 cases, with 50 from Massachusetts, 116 from Connecticut, 125 from Montana, and 26 from Pennsylvania. (See Table 1.)

2. For each case $t \in S$, let $d(t)$ be the set of 5 justices voting in case t, and let $e(t)$ be a complete (7 member) court observed in the data that contains d_t. Using our estimates,

we compute for each case $t \in S$ (i) the conditional voting probabilities $(\gamma_0(t), \gamma_1(t))$, (ii) the prior $\rho(t)$ and the (iii) types consistent with the expressive voting model $(\theta(t), \pi^{exp}(t))$ and (iv) strategic voting model $(\theta(t), \pi^*(t))$, *for the counterfactual in which all of $e(t)$ justices – the "complete court" – voted in case t.*

3. We can now compute our estimates of the conditional voting probabilities for each member $j \in d(t)$ in each case $t \in S$ for the actual "incomplete" court $d(t)$ observed in case t. For the expressive voting model, the voting probabilities for each judge is invariant to whether the court is complete or incomplete, so that the likelihood for the observed votes in case t under the expressive voting model is

$$L^S_{exp} = \rho(t) \prod_{i \in d(t)} \left[\gamma_{i,1}(t)^{v_{it}} (1 - \gamma_{i,1}(t))^{1-v_{it}}\right] + (1-\rho) \prod_{i \in d(t)} \left[\gamma_{i,0}(t)^{v_{it}} (1 - \gamma_{i,0}(t))^{1-v_{it}}\right]$$

In the strategic voting model, however, a judge's equilibrium voting probabilities depend on the number and characteristics of the other judges, so that the voting probabilities in the complete and incomplete courts will differ. To compute the equilibrium voting probabilities for the actual "incomplete" court in case t consistent with the strategic voting model, then, involves recomputing the equilibrium strategies given the estimated preference parameters. Specifically, we: (i) take the estimates of the prior ρ_t and the types consistent with the strategic voting model (θ_t, π^*_t) for the counterfactual in which all of $e(t)$ justices voted in case t, as computed in (ii) and (iv) above (these are contingent on X_t, and therefore typically different in each case t); (ii) using the estimates of $(\theta_t, \pi^*_t, \rho_t)$ for each case t, compute the equilibrium strategies consistent with the five member court $d(t)$, say $\tilde{s}(t)$, according to the equilibrium conditions of the strategic voting model, as in Eq. (4) of the main text; and finally, (iii) use $\tilde{s}(t)$ to compute the conditional voting probabilities consistent with the strategic voting model for the five member court $d(t)$ in each case $t \in S$, say $(\tilde{\gamma}_0(t), \tilde{\gamma}_1(t))$. Then the likelihood for the strategic voting model in sample S is

$$L^S_{st} = \rho(t) \prod_{i \in d(t)} \left[\tilde{\gamma}_{i,1}(t)^{v_{it}} (1 - \tilde{\gamma}_{i,1}(t))^{1-v_{it}}\right] + (1-\rho) \prod_{i \in d(t)} \left[\tilde{\gamma}_{i,0}(t)^{v_{it}} (1 - \tilde{\gamma}_{i,0}(t))^{1-v_{it}}\right]$$

We can then compare the likelihood for the strategic and expressive voting models on S. A finding that $L^S_{st} > L^S_{exp}$ provides evidence in favor of the strategic voting model, while $L^S_{exp} < L^S_{st}$ provides evidence in favor of the expressive voting model.

The results from this exercise, shown in the bottom of Table 1, indicate that these two likelihoods are virtually identical: $L^S_{st} = 57.78 \simeq 59.17 = L^S_{exp}$. Thus, this test slightly favors the expressive model, but this small difference in the likelihood functions would not be statistically significant given the modest sample size.[1] Hence, it appears difficult to distinguish between the strategic and expressive voting models using the data employed in this paper.

State	**# cases**	**full court size**	**reduced court size**
Massachusetts	43	7	5
Connecticut	109	7	5
Montana	124	7	5
Pennsylvania	25	7	5
Total:	301		
Log-likelihoods:			
Expressive model	59.17		
Strategic model	57.78		

Table 1: Specification test results

[1]Given the test outcome, we did not attempt to approximate the sampling distribution of these statistics; adequately accommodating the sampling error in both the parameter estimates used in the exercise, as well as the intrinsic sampling error in the subsample of "incomplete court" cases, would require a multi-step bootstrapping procedure which is computationally burdensome.

To Elect or to Appoint?
Appendix C: Unobserved Heterogeneity.

Matias Iaryczower, Garrett Lewis and Matthew Shum

June 1, 2012

In this appendix we consider potential sources of unobserved heterogeneity in our empirical analysis. While we cannot provide an omnibus solution to this problem, we consider various possible unobserved regional factors shaping institutional choice.

1. Differences in the law across states due to "when states entered the union" or to "former Spanish heritage". Figure 1 below plots the estimates of bias and quality of information (the state averages) as a function of the date in which the state entered the Union. Figure 2 presents the estimates of bias and quality of information for states of hispanic heritage and of non-hispanic heritage. We included five states as having hispanic heritage: Arizona, California, Florida, New Mexico and Texas. In neither figure do we see any systematic differences in estimated bias and quality between states based on their date of entry to the union, or their hispanic heritage.

2. Differences on "which cases are subject to mandatory [vs discretionary] review" across states. The first point to note here is that all states have mandatory review in death penalty cases, which we distinguish in our main specification. More generally, states differ in whether they have mandatory or discretionary review for different kind of cases. Table 1 below shows this information for each state.[1] The table distinguishes between mandatory review of criminal appeals, discretionary review of criminal appeals, and an intermediate (mixed) category. Using this information, we evaluate two considerations.

[1] This information can be obtained in the "State Court Organization 2004" report by the Bureau of Justice Statistics, which is available online.

The first is whether states with different electoral institutions tend to have different type of review (this is a necessary condition for the argument to go forward). This information is presented in in the upper panel of Figure 3. This shows that states in which justices are elected, face voter retention, or are appointed for life are evenly split between mandatory, discretionary and mixed review systems. In other words, conditioning on the state electing its Supreme Court justices, it is equally likely to have mandatory or discretionary review. The same is true for systems with Voter Retention, or Life appointments. The only exception is in the case of states in which justices face political reappointment, which typically have mandatory review (six out of eight cases).

The second and more important consideration is whether there is some relationship between type of review and our estimates of bias and quality of information. The lower panel of Figure 3 shows these estimates (the state-level averages) distinguishing between type of review. The figure suggests that there is no systematic pattern between type of review and our estimates of bias and quality of information.

3. More homogeneous courts will tend to bring up cases in which they agree, and avoid cases in which they don't.

Evaluating this argument fully is challenging because we only observe data on cases which were heard by the court, and thus it is simply not feasible to estimate a model of case selection. However, we argued that the prior ρ will incorporate both justices' prior beliefs (about randomly assigned cases) and endogenous case selection. Thus, to consider this possibility, we plot the estimated prior ρ per state together with the within-court heterogeneity in the bias estimates (the standard deviation of the bias estimates within each court). If the hypothesis is true (and if ρ captures case selection), we should observe a negative relationship: less heterogeneous courts should have a larger prior ρ, indicating that they are ex ante more favorable to overturning. The figure suggests that there is no systematic pattern between heterogeneity in the court (as measured by the standard deviation of the bias estimates within each court) and case selection (as measured by ρ).

Table 1: Mandatory and Discretionary Review (source: "State court organization 2004", Bureau of Justice Statistics.)

Mandatory and Discretionary Jurisdiction of State Supreme Courts
(1 if mandatory, 0 if mixed, -1 if discretionary)

Electoral System	State	Criminal Appeals	Death Penalty Cases	Quality of Information	Bias (expressive)
Appointed, for Life	Massachusetts	0	-	4.144	0.967
	New Hampshire	-1	1	3.376	0.802
	New Jersey	0	1	3.380	0.873
	Rhode Island	1	-	4.030	0.958
Appointed, with Reappointment	Connecticut	0	1	3.482	0.711
	Delaware	1	1	3.449	0.666
	Hawaii	1	-	3.556	0.713
	Maine	1	-	3.609	0.668
	New York	1	1	3.757	0.678
	South Carolina	1	1	3.540	0.761
	Vermont	1	-	3.626	0.698
	Virginia	-1	1	3.545	0.640
Appointed, with Voter Retention	Alaska	1	-	2.874	0.614
	Arizona	0	1	2.957	0.727
	California	-1	1	2.803	0.697
	Colorado	0	1	1.638	0.609
	Florida	0	1	2.735	0.675
	Indiana	0	1	2.879	0.687
	Iowa	0	-	2.792	0.603
	Kansas	0	1	2.965	0.683
	Maryland	-1	1	2.719	0.671
	Missouri	-1	1	2.800	0.704
	Nebraska	-1	1	2.736	0.614
	Oklahoma	1	1	2.730	0.747
	South Dakota	1	1	2.921	0.621
	Tennessee	-1	1	2.603	0.661
	Utah	1	1	2.669	0.647
	Wyoming	1	1	2.915	0.622
Elected	Alabama	1	1	2.448	0.551
	Arkansas	-1	1	2.673	0.615
	Georgia	0	1	2.443	0.686
	Idaho	1	1	2.654	0.602
	Illinois	0	1	2.854	0.636
	Kentucky	0	1	2.309	0.690
	Louisiana	0	1	2.242	0.531
	Michigan	-1	-	2.594	0.489
	Minnesota	0	-	2.604	0.601
	Mississippi	1	1	2.645	0.658
	Montana	1	1	2.806	0.530
	Nevada	1	1	2.450	0.696
	New Mexico	0	1	2.613	0.594
	North Carolina	-1	1	2.709	0.718
	North Dakota	1	-	2.883	0.545
	Ohio	0	1	2.880	0.589
	Oregon	0	1	2.554	0.666
	Pennsylvania	0	1	2.779	0.635
	Texas	0	1	2.348	0.563
	Washington	-1	1	2.507	0.591
	West Virginia	-1	-	2.951	0.605
	Wisconsin	-1	-	2.557	0.533

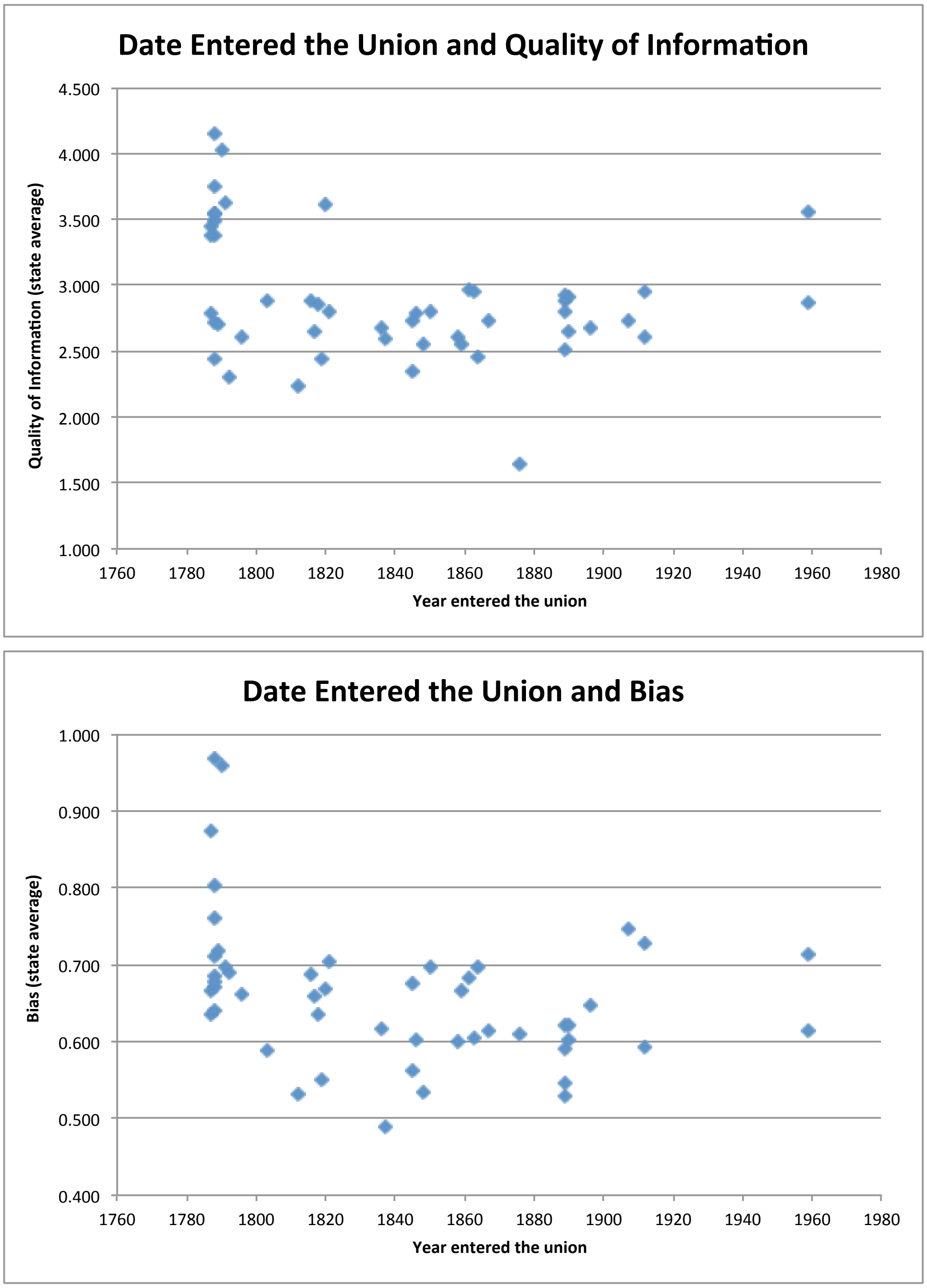


Figure 1: Date entered the Union, Bias and Quality of Information

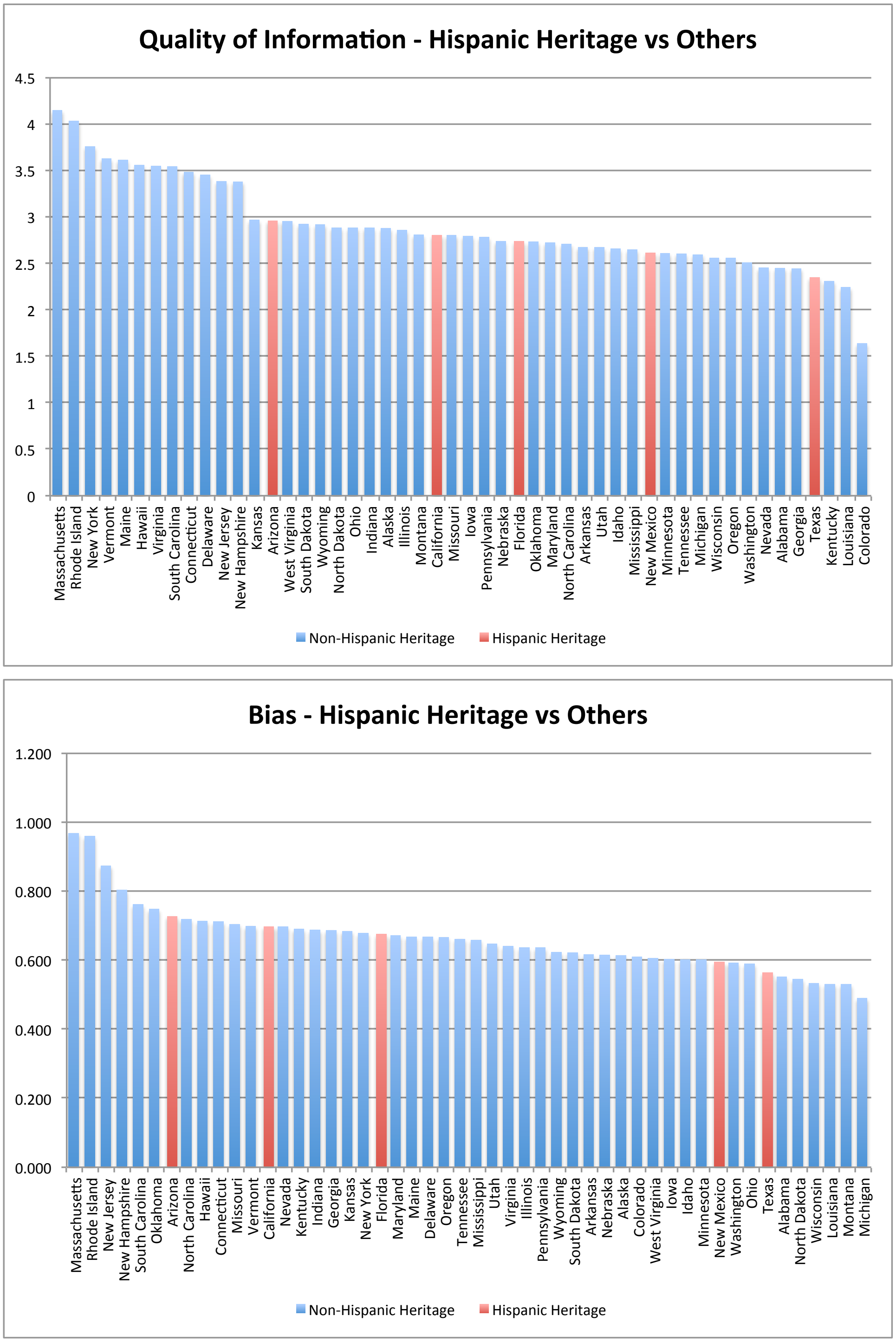
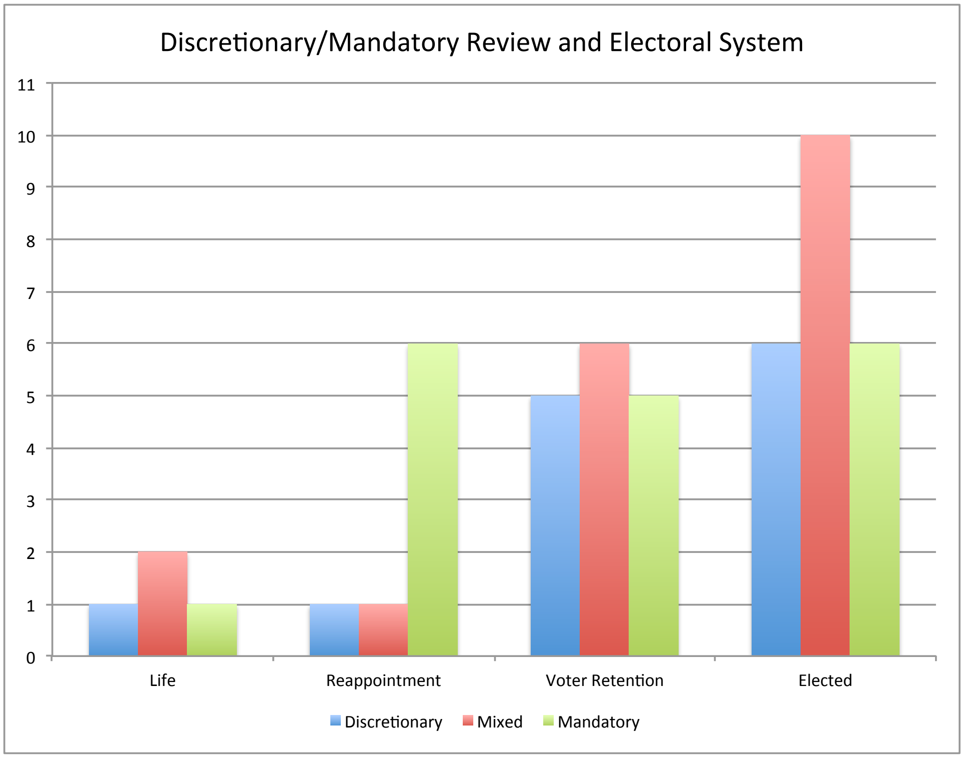


Figure 2: Spanish Heritage, Bias and Quality of Information



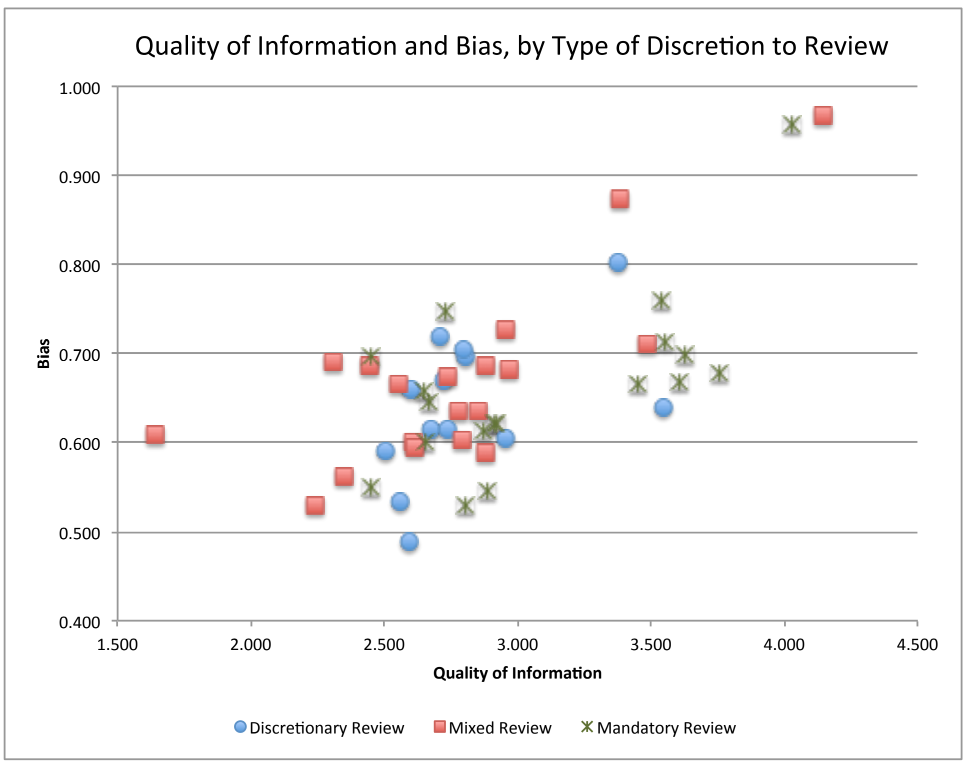


Figure 3: Mandatory and Discretionary Review. Upper Panel plots the number of states with Discretionary, Mixed, and Mandatory Review for Criminal Appeals, by electoral system (source: "State court organization 2004", Bureau of Justice Statistics.). Lower Panel plots the estimates of bias and quality of information by type of discretion to review.

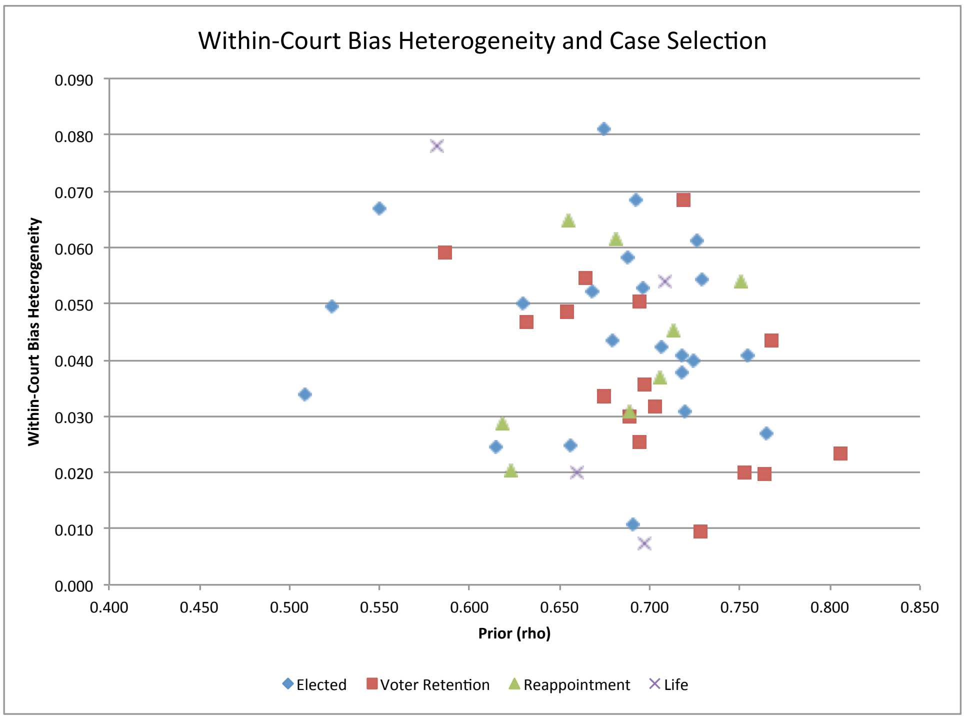


Figure 4: Within-court bias heterogeneity and Case Selection. Larger values of ρ indicates justices are ex ante more favorable to overturning.

To Elect or to Appoint?
Appendix D: State Tables.

Matias Iaryczower, Garrett Lewis and Matthew Shum

June 1, 2012

Table 1: Benchmark (UO) Specification: Type, Prior, Strategy and Conditional Voting Probabilities. Average Across states per Electoral Institution (case-specific covariates fixed at state sample average; justices at own individual-specific covariates).

	State	ρ	γ_{it0}	γ_{it1}	θ	$s*$	π^{exp}	π^{st}	FLEX^{exp}	FLEXst
Elected	Alabama	0.615	0.134	0.909	2.448	0.455	0.551	0.247	0.390	0.390
	Arkansas	0.718	0.123	0.934	2.673	0.436	0.615	0.233	0.295	0.295
	Georgia	0.720	0.126	0.901	2.443	0.470	0.686	0.537	0.367	0.316
	Idaho	0.718	0.132	0.935	2.654	0.425	0.602	0.297	0.292	0.292
	Illinois	0.679	0.087	0.931	2.854	0.478	0.636	0.477	0.377	0.340
	Kentucky	0.724	0.140	0.888	2.309	0.470	0.690	0.557	0.356	0.319
	Louisiana	0.524	0.129	0.864	2.242	0.508	0.531	0.563	0.486	0.514
	Michigan	0.509	0.103	0.907	2.594	0.489	0.489	0.422	0.486	0.488
	Minnesota	0.688	0.124	0.925	2.604	0.445	0.601	0.282	0.365	0.325
	Mississippi	0.729	0.116	0.924	2.645	0.454	0.658	0.294	0.374	0.295
	Montana	0.693	0.124	0.949	2.806	0.413	0.530	0.088	0.305	0.305
	Nevada	0.726	0.121	0.897	2.450	0.479	0.696	0.632	0.448	0.315
	New Mexico	0.656	0.115	0.920	2.613	0.460	0.594	0.430	0.357	0.357
	North Carolina	0.754	0.099	0.920	2.709	0.477	0.718	0.584	0.334	0.282
	North Dakota	0.675	0.106	0.946	2.883	0.436	0.545	0.252	0.384	0.327
	Ohio	0.668	0.094	0.940	2.880	0.459	0.589	0.295	0.341	0.341
	Oregon	0.706	0.115	0.910	2.554	0.472	0.666	0.509	0.412	0.324
	Pennsylvania	0.765	0.122	0.947	2.779	0.419	0.635	0.154	0.247	0.247
	Texas	0.690	0.174	0.920	2.348	0.400	0.563	0.081	0.311	0.311
	Washington	0.629	0.119	0.905	2.507	0.474	0.591	0.396	0.451	0.386
	West Virginia	0.696	0.091	0.945	2.951	0.454	0.605	0.374	0.314	0.314
	Wisconsin	0.549	0.106	0.901	2.557	0.492	0.533	0.484	0.465	0.457
	Average	**0.674**	**0.118**	**0.919**	**2.614**	**0.458**	**0.606**	**0.372**	**0.371**	**0.343**
Appointed, with Voter Retention	Alaska	0.697	0.096	0.941	2.874	0.455	0.614	0.397	0.315	0.315
	Arizona	0.764	0.080	0.939	2.957	0.477	0.727	0.621	0.264	0.264
	California	0.718	0.086	0.921	2.803	0.491	0.697	0.647	0.351	0.314
	Colorado	0.632	0.224	0.807	1.638	0.466	0.609	0.526	0.411	0.407
	Florida	0.689	0.090	0.918	2.735	0.491	0.675	0.618	0.427	0.340
	Indiana	0.753	0.092	0.940	2.879	0.461	0.687	0.500	0.270	0.270
	Iowa	0.586	0.078	0.914	2.792	0.509	0.603	0.664	0.460	0.568
	Kansas	0.694	0.072	0.932	2.965	0.496	0.683	0.654	0.466	0.331
	Maryland	0.654	0.084	0.907	2.719	0.511	0.671	0.734	0.509	0.622
	Missouri	0.768	0.100	0.934	2.800	0.459	0.704	0.430	0.320	0.259
	Nebraska	0.674	0.102	0.928	2.736	0.465	0.614	0.380	0.341	0.341
	Oklahoma	0.806	0.108	0.932	2.730	0.454	0.747	0.362	0.228	0.228
	South Dakota	0.694	0.090	0.942	2.921	0.461	0.621	0.425	0.318	0.318
	Tennessee	0.665	0.099	0.902	2.603	0.499	0.661	0.657	0.461	0.461
	Utah	0.728	0.117	0.930	2.669	0.446	0.647	0.421	0.291	0.291
	Wyoming	0.703	0.092	0.943	2.915	0.457	0.622	0.406	0.310	0.310
	Average	**0.702**	**0.101**	**0.921**	**2.734**	**0.475**	**0.661**	**0.528**	**0.359**	**0.352**
Appointed, with Reappointment	Connecticut	0.655	0.036	0.951	3.482	0.522	0.711	0.862	0.595	0.636
	Delaware	0.750	0.055	0.967	3.449	0.465	0.666	0.430	0.260	0.260
	Hawaii	0.689	0.035	0.959	3.556	0.509	0.713	0.766	0.601	0.671
	Maine	0.713	0.041	0.969	3.609	0.484	0.668	0.487	0.355	0.298
	New York	0.619	0.026	0.965	3.757	0.518	0.678	0.836	0.607	0.607
	South Carolina	0.623	0.026	0.943	3.540	0.552	0.761	0.936	0.598	0.598
	Vermont	0.682	0.034	0.963	3.626	0.506	0.698	0.735	0.530	0.667
	Virginia	0.706	0.047	0.968	3.545	0.475	0.640	0.371	0.303	0.303
	Average	**0.680**	**0.038**	**0.961**	**3.571**	**0.504**	**0.692**	**0.678**	**0.481**	**0.505**
Appointed, for Life	Massachusetts	0.697	0.004	0.926	4.144	0.650	0.967	1.000	0.647	0.647
	New Hampshire	0.708	0.033	0.936	3.376	0.547	0.802	0.936	0.672	0.672
	New Jersey	0.582	0.015	0.872	3.380	0.656	0.873	1.000	0.513	0.513
	Rhode Island	0.659	0.004	0.913	4.030	0.659	0.958	1.000	0.603	0.603
	Average	**0.662**	**0.014**	**0.912**	**3.733**	**0.628**	**0.900**	**0.984**	**0.609**	**0.609**

Table 2: FAST Specification: Type, Prior, Strategy and Conditional Voting Probabilities. Average Across states per Electoral Institution (case-specific covariates fixed at state sample average; justices at own individual-specific covariates)

	State	ρ	γ_{it0}	γ_{it1}	θ	$s*$	π^{exp}	π^{st}	FLEX^{exp}	FLEXst
Elected	Alabama	0.299	0.052	0.802	2.483	0.658	0.527	0.992	0.276	0.276
	Arkansas	0.301	0.042	0.857	2.810	0.619	0.523	0.969	0.287	0.287
	Georgia	0.279	0.061	0.826	2.491	0.622	0.447	0.925	0.275	0.275
	Idaho	0.305	0.047	0.871	2.826	0.598	0.489	0.858	0.298	0.298
	Illinois	0.257	0.048	0.921	3.110	0.540	0.345	0.670	0.327	0.273
	Kentucky	0.266	0.077	0.831	2.413	0.599	0.396	0.838	0.326	0.278
	Louisiana	0.213	0.086	0.823	2.321	0.597	0.316	0.750	0.298	0.243
	Michigan	0.227	0.056	0.841	2.602	0.615	0.393	0.915	0.234	0.234
	Minnesota	0.259	0.050	0.828	2.622	0.634	0.464	0.958	0.252	0.252
	Mississippi	0.309	0.047	0.864	2.811	0.604	0.506	0.981	0.299	0.299
	Montana	0.315	0.032	0.850	2.925	0.640	0.598	0.991	0.290	0.290
	Nevada	0.289	0.064	0.871	2.692	0.573	0.408	0.703	0.365	0.298
	New Mexico	0.291	0.045	0.830	2.651	0.640	0.521	0.921	0.274	0.274
	North Carolina	0.237	0.047	0.871	2.834	0.596	0.406	0.915	0.305	0.243
	North Dakota	0.303	0.030	0.847	2.957	0.646	0.602	0.968	0.278	0.278
	Ohio	0.271	0.039	0.900	3.072	0.579	0.438	0.918	0.272	0.272
	Oregon	0.324	0.062	0.876	2.714	0.572	0.449	0.844	0.366	0.326
	Pennsylvania	0.295	0.041	0.878	2.915	0.599	0.490	0.952	0.287	0.287
	Texas	0.297	0.059	0.810	2.440	0.641	0.494	0.983	0.282	0.282
	Washington	0.276	0.056	0.827	2.567	0.628	0.466	0.980	0.269	0.269
	West Virginia	0.302	0.029	0.883	3.112	0.614	0.564	0.946	0.287	0.287
	Wisconsin	0.250	0.059	0.846	2.625	0.606	0.411	0.901	0.311	0.256
	Average	**0.280**	**0.051**	**0.852**	**2.727**	**0.610**	**0.466**	**0.903**	**0.294**	**0.276**
Appointed, with Voter Retention	Alaska	0.363	0.033	0.892	3.096	0.597	0.588	0.930	0.344	0.344
	Arizona	0.237	0.045	0.941	3.264	0.520	0.279	0.391	0.257	0.257
	California	0.258	0.054	0.906	2.994	0.549	0.363	0.720	0.320	0.274
	Colorado	0.206	0.241	0.817	1.642	0.439	0.181	0.104	0.529	0.641
	Florida	0.293	0.054	0.923	3.043	0.529	0.352	0.586	0.413	0.309
	Indiana	0.264	0.040	0.899	3.031	0.578	0.423	0.801	0.267	0.267
	Iowa	0.274	0.046	0.892	2.953	0.577	0.423	0.887	0.332	0.278
	Kansas	0.234	0.039	0.900	3.075	0.577	0.390	0.893	0.306	0.241
	Maryland	0.290	0.068	0.927	2.973	0.505	0.301	0.335	0.412	0.360
	Missouri	0.251	0.054	0.936	3.152	0.514	0.284	0.389	0.330	0.276
	Nebraska	0.277	0.046	0.886	2.906	0.583	0.435	0.906	0.278	0.278
	Oklahoma	0.302	0.052	0.908	2.962	0.549	0.398	0.837	0.311	0.311
	South Dakota	0.308	0.032	0.890	3.088	0.602	0.538	0.922	0.296	0.296
	Tennessee	0.265	0.073	0.908	2.818	0.523	0.306	0.398	0.360	0.294
	Utah	0.285	0.050	0.870	2.775	0.593	0.451	0.823	0.284	0.284
	Wyoming	0.312	0.032	0.892	3.103	0.599	0.536	0.917	0.300	0.300
	Average	**0.276**	**0.060**	**0.899**	**2.930**	**0.552**	**0.391**	**0.677**	**0.334**	**0.313**
Appointed, with Reappointment	Connecticut	0.243	0.027	0.948	3.596	0.541	0.350	0.771	0.318	0.250
	Delaware	0.276	0.023	0.948	3.653	0.549	0.420	0.768	0.278	0.278
	Hawaii	0.255	0.023	0.955	3.698	0.539	0.371	0.672	0.261	0.261
	Maine	0.305	0.016	0.941	3.737	0.577	0.557	0.982	0.298	0.298
	New York	0.296	0.015	0.957	3.892	0.558	0.501	0.954	0.294	0.294
	South Carolina	0.278	0.032	0.972	3.788	0.492	0.260	0.212	0.536	0.707
	Vermont	0.306	0.018	0.956	3.829	0.551	0.476	0.839	0.305	0.305
	Virginia	0.264	0.020	0.962	3.841	0.536	0.380	0.791	0.269	0.269
	Average	**0.278**	**0.022**	**0.955**	**3.754**	**0.543**	**0.414**	**0.748**	**0.320**	**0.333**
Appointed, for Life	Massachusetts	0.284	0.039	1.000	6.340	0.279	0.000	0.000	0.688	0.688
	New Hampshire	0.318	0.040	1.000	5.374	0.328	0.004	0.000	0.655	0.655
	New Jersey	0.208	0.045	0.979	3.870	0.447	0.136	0.009	0.600	0.761
	Rhode Island	0.297	0.046	1.000	6.177	0.278	0.000	0.000	0.671	0.671
	Average	**0.277**	**0.043**	**0.995**	**5.440**	**0.333**	**0.035**	**0.002**	**0.653**	**0.694**

Table 3: Case-Specific Data

State	Appeal	Petitioner Person	Formal Opinion	Jud. Review (US Const.)	Jud. Review (St. Const.)	Murder Case	No. Legal Issues	Evidence	Jury Instruction	Death Penalty	No. of Justices	Prop. Votes to Overturn	Unanimous to Overturn	Unanimous to Uphold	Minimal Winning
Alabama	1.00	0.73	0.87	0.00	0.00	0.33	1.53	0.40	0.47	0.33	9.00	0.44	0.27	0.33	0.13
Alaska	1.00	1.00	1.00	0.10	0.00	0.00	1.40	0.40	0.20	0.00	5.00	0.46	0.40	0.50	0.10
Arizona	0.89	0.88	1.00	0.05	0.11	0.54	4.86	0.70	0.49	0.50	5.00	0.51	0.39	0.40	0.08
Arkansas	0.94	0.94	0.91	0.03	0.07	0.33	2.03	0.66	0.24	0.06	7.00	0.20	0.13	0.72	0.06
California	0.90	0.81	0.99	0.03	0.00	0.48	3.47	0.54	0.56	0.37	7.00	0.37	0.19	0.42	0.13
Colorado	0.34	0.23	0.51	0.04	0.03	0.05	1.22	0.23	0.14	0.02	7.00	0.77	0.70	0.18	0.06
Connecticut	0.86	0.65	0.95	0.05	0.05	0.33	2.00	0.52	0.33	0.05	7.00	0.50	0.14	0.19	0.10
Delaware	0.83	0.90	0.97	0.13	0.10	0.50	2.50	0.60	0.57	0.37	5.00	0.27	0.27	0.70	0.00
Florida	0.82	0.84	0.33	0.01	0.05	0.63	2.18	0.51	0.55	0.59	7.00	0.42	0.28	0.47	0.09
Georgia	0.96	0.93	0.99	0.02	0.03	0.63	2.28	0.66	0.37	0.07	7.00	0.26	0.21	0.69	0.03
Hawaii	0.93	0.78	0.99	0.05	0.03	0.23	2.67	0.63	0.30	0.00	5.00	0.44	0.36	0.52	0.03
Idaho	0.87	0.89	0.96	0.03	0.09	0.14	2.05	0.39	0.35	0.05	5.00	0.22	0.16	0.71	0.06
Illinois	0.93	0.76	1.00	0.11	0.15	0.57	2.56	0.58	0.44	0.50	7.00	0.35	0.21	0.47	0.10
Indiana	0.92	0.95	1.00	0.00	0.02	0.69	2.75	0.73	0.54	0.12	5.00	0.25	0.20	0.70	0.02
Iowa	1.00	0.67	0.67	0.00	0.00	0.00	2.00	0.33	0.33	0.00	7.00	0.33	0.33	0.67	0.00
Kansas	0.96	0.78	1.00	0.02	0.02	0.47	2.95	0.62	0.54	0.00	7.00	0.33	0.30	0.64	0.02
Kentucky	0.90	0.87	0.95	0.03	0.05	0.43	2.74	0.69	0.33	0.12	7.00	0.42	0.19	0.32	0.19
Louisiana	0.99	0.46	0.68	0.05	0.08	0.43	1.90	0.54	0.35	0.26	7.00	0.65	0.48	0.25	0.04
Maine	0.95	0.91	1.00	0.00	0.01	0.06	1.86	0.66	0.41	0.00	7.00	0.30	0.28	0.70	0.00
Maryland	0.95	0.78	0.98	0.03	0.04	0.24	1.63	0.49	0.26	0.08	7.00	0.56	0.41	0.28	0.09
Massachussets	0.93	0.86	0.91	0.00	0.00	0.38	1.86	0.64	0.36	0.00	7.00	0.30	0.24	0.60	0.07
Michigan	1.00	0.44	0.85	0.01	0.00	0.21	1.33	0.51	0.22	0.00	7.00	0.63	0.24	0.08	0.36
Minnesota	0.91	0.79	0.97	0.03	0.01	0.54	1.98	0.66	0.31	0.00	7.00	0.32	0.25	0.58	0.08
Mississippi	0.96	0.98	0.95	0.01	0.03	0.32	2.37	0.62	0.37	0.11	9.00	0.36	0.14	0.38	0.17
Missouri	0.87	0.92	0.90	0.06	0.19	0.53	4.16	0.58	0.59	0.48	7.00	0.31	0.23	0.60	0.03
Montana	0.95	0.88	0.99	0.01	0.04	0.04	1.85	0.60	0.15	0.03	7.00	0.32	0.22	0.55	0.15
Nebraska	0.94	0.81	0.84	0.06	0.05	0.31	1.99	0.51	0.35	0.04	7.00	0.36	0.28	0.57	0.07
Nevada	0.85	0.93	0.39	0.02	0.01	0.41	2.77	0.65	0.38	0.24	5.00	0.44	0.26	0.39	0.16
New Hampshire	0.93	0.92	0.97	0.01	0.03	0.07	1.57	0.64	0.28	0.00	5.00	0.27	0.21	0.66	0.07
New Jersey	0.89	0.46	0.91	0.09	0.06	0.33	1.89	0.46	0.52	0.13	7.00	0.59	0.46	0.26	0.07
New Mexico	1.00	0.81	1.00	0.06	0.00	0.25	1.75	0.38	0.44	0.00	5.00	0.39	0.19	0.50	0.06
New York	1.00	0.76	0.91	0.02	0.03	0.23	1.43	0.43	0.14	0.00	7.00	0.38	0.28	0.50	0.04
North Carolina	1.00	0.94	1.00	0.02	0.02	0.87	4.14	0.83	0.69	0.39	7.00	0.21	0.18	0.75	0.01
North Dakota	0.98	0.85	1.00	0.01	0.01	0.06	1.72	0.61	0.28	0.00	5.00	0.25	0.21	0.72	0.01
Ohio	0.92	0.77	0.71	0.03	0.03	0.47	2.81	0.48	0.42	0.40	7.00	0.29	0.17	0.58	0.07
Oklahoma	0.68	0.97	0.91	0.02	0.01	0.58	2.97	0.66	0.52	0.54	5.03	0.23	0.18	0.66	0.02
Oregon	0.87	0.87	1.00	0.00	0.00	0.27	1.47	0.67	0.20	0.27	7.00	0.70	0.47	0.07	0.13
Pennsylvania	0.76	0.90	1.00	0.03	0.01	0.54	2.78	0.63	0.32	0.43	7.00	0.38	0.18	0.46	0.15
Rhode Island	0.93	0.84	0.47	0.01	0.01	0.11	1.89	0.54	0.26	0.00	5.00	0.31	0.26	0.64	0.02
South Carolina	0.99	0.72	0.91	0.02	0.01	0.33	1.83	0.45	0.35	0.17	5.00	0.49	0.42	0.42	0.05
South Dakota	0.95	0.90	0.86	0.02	0.04	0.14	1.88	0.53	0.32	0.01	5.00	0.30	0.22	0.55	0.12
Tennessee	0.88	0.71	1.00	0.05	0.06	0.33	1.87	0.57	0.30	0.19	5.00	0.41	0.33	0.47	0.10
Texas	0.85	0.81	0.88	0.06	0.06	0.30	1.80	0.44	0.28	0.21	9.00	0.45	0.16	0.22	0.17
Utah	0.80	0.83	1.00	0.00	0.02	0.24	2.30	0.50	0.28	0.04	5.00	0.31	0.17	0.50	0.17
Vermont	0.95	0.90	0.66	0.02	0.05	0.13	1.76	0.52	0.34	0.00	5.00	0.29	0.19	0.62	0.06
Virginia	0.95	0.86	0.98	0.19	0.19	0.36	2.86	0.57	0.34	0.31	7.00	0.35	0.31	0.57	0.03
Washington	0.99	0.73	0.93	0.04	0.01	0.13	1.57	0.57	0.33	0.01	9.00	0.51	0.31	0.28	0.12
West Virginia	0.91	0.94	0.45	0.05	0.04	0.31	1.76	0.55	0.36	0.00	5.00	0.35	0.28	0.55	0.03
Wisconsin	1.00	0.54	0.98	0.07	0.05	0.02	1.39	0.44	0.22	0.00	7.00	0.57	0.37	0.24	0.24
Wyoming	0.99	0.93	1.00	0.03	0.04	0.07	2.14	0.55	0.30	0.00	5.00	0.15	0.11	0.81	0.04
Obs	5958	5947	5938	5925	5929	5958	5958	5958	5958	5958	5958	5958	5958	5958	5958
Mean	0.91	0.83	0.87	0.03	0.04	0.37	2.31	0.58	0.38	0.17	6.56	0.36	0.25	0.52	0.08
Std.Dev.	0.29	0.37	0.33	0.17	0.20	0.48	2.11	0.49	0.48	0.37	1.28	0.43	0.43	0.50	0.27
Min	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5.00	0.00	0.00	0.00	0.00
Max	1.00	1.00	1.00	1.00	1.00	1.00	28.00	1.00	1.00	1.00	9.00	1.00	1.00	1.00	1.00

Table 4: Justice-Specific Data. Justices in Largest Natural Courts (LNC) in Each State. Average Values of Justice-Specific Covariates, per State (324 Justices)

State	Elected	Apptd for life	Appt. w/ Pol. Reappointment	Prior Judicial Experience	Prior Political Experience	Years of Experience	PAJID at appointment	CIT at decision	GOV at decision
Alabama	1.00	0.00	0.00	5.13	0.00	12.15	33.59	41.37	45.13
Alaska	0.00	0.00	0.00	2.40	0.00	15.34	41.77	22.34	33.87
Arizona	0.00	0.00	0.00	9.20	0.00	5.71	25.32	40.26	1.64
Arkansas	1.00	0.00	0.00	5.00	0.14	5.36	39.23	47.00	46.83
California	0.00	0.00	0.00	5.43	0.14	8.80	29.22	55.65	44.08
Colorado	0.00	0.00	0.00	3.83	0.14	10.63	42.56	42.48	55.92
Connecticut	0.00	0.00	1.00	10.29	0.14	7.29	57.87	58.98	42.62
Delaware	0.00	0.00	1.00	8.40	0.00	6.41	42.79	43.79	53.75
Florida	0.00	0.00	0.00	8.71	0.00	9.68	49.50	44.60	55.27
Georgia	1.00	0.00	0.00	11.83	0.00	2.92	45.49	41.77	85.34
Hawaii	0.00	0.00	1.00	7.20	0.00	4.44	82.22	79.81	93.88
Idaho	1.00	0.00	0.00	7.60	0.20	4.73	30.74	20.35	2.38
Illinois	1.00	0.00	0.00	13.83	0.14	13.84	44.93	59.93	36.68
Indiana	0.00	0.00	0.00	1.00	0.00	6.22	50.25	40.30	51.84
Iowa	0.00	0.00	0.00	9.86	0.14	13.93	25.27	40.77	20.92
Kansas	0.00	0.00	0.00	9.83	0.14	9.18	21.20	36.08	6.08
Kentucky	1.00	0.00	0.00	9.29	0.14	7.71	37.83	36.46	73.96
Louisiana	1.00	0.00	0.00	9.00	0.14	10.82	29.80	35.13	39.58
Maine	0.00	0.00	1.00	6.29	0.00	7.34	61.16	52.20	48.38
Maryland	0.00	0.00	0.00	9.00	0.14	11.42	78.04	58.37	90.17
Massachusetts	0.00	1.00	0.00	5.86	0.00	11.24	55.51	86.47	79.78
Michigan	1.00	0.00	0.00	7.43	0.14	11.54	47.79	48.45	15.77
Minnesota	1.00	0.00	0.00	2.50	0.14	5.06	46.35	49.55	40.83
Mississippi	1.00	0.00	0.00	6.33	0.33	7.74	30.75	23.78	27.88
Missouri	0.00	0.00	0.00	3.29	0.14	6.54	24.97	46.33	70.91
Montana	1.00	0.00	0.00	0.86	0.14	6.83	32.68	41.62	7.50
Nebraska	0.00	0.00	0.00	2.71	0.00	6.76	38.82	32.68	53.00
Nevada	1.00	0.00	0.00	3.00	0.60	9.56	27.14	39.51	50.91
New Hampshire	0.00	1.00	0.00	5.00	0.25	10.09	4.23	38.55	43.80
New Jersey	0.00	0.71	0.29	5.00	0.14	12.67	39.06	61.72	23.43
New Mexico	1.00	0.00	0.00	8.00	0.00	5.18	45.09	44.83	49.75
New York	0.00	0.00	1.00	13.14	0.00	7.45	56.16	64.28	43.68
North Carolina	1.00	0.00	0.00	6.29	0.14	8.38	35.00	42.29	60.33
North Dakota	1.00	0.00	0.00	2.60	0.20	8.29	32.83	51.87	18.53
Ohio	1.00	0.00	0.00	7.71	0.14	6.58	36.99	48.07	15.00
Oklahoma	0.00	0.00	0.00	6.44	0.00	17.08	39.01	9.25	10.54
Oregon	1.00	0.00	0.00	8.71	0.29	6.91	60.50	56.54	55.79
Pennsylvania	1.00	0.00	0.00	4.86	0.00	7.86	43.74	56.29	28.23
Rhode Island	0.00	1.00	0.00	12.25	0.40	6.01	41.10	77.11	71.08
South Carolina	0.00	0.00	1.00	11.40	1.00	5.92	35.00	41.37	24.45
South Dakota	0.00	0.00	0.00	7.20	0.00	6.39	23.66	42.13	9.00
Tennessee	0.00	0.00	0.00	6.80	0.20	6.13	48.20	32.02	24.18
Texas	1.00	0.00	0.00	0.89	0.00	6.96	34.14	34.80	31.33
Utah	0.00	0.00	0.00	3.20	0.00	12.50	28.66	36.77	5.30
Vermont	0.00	0.00	1.00	6.80	0.20	9.46	66.85	75.46	83.99
Virginia	0.00	0.00	1.00	8.43	0.00	12.81	34.56	36.51	25.51
Washington	1.00	0.00	0.00	6.88	0.11	7.19	52.57	50.96	53.79
West Virginia	1.00	0.00	0.00	8.20	0.20	2.60	38.23	70.14	54.62
Wisconsin	1.00	0.00	0.00	9.71	0.29	9.57	30.55	52.42	35.50
Wyoming	0.00	0.00	0.00	4.40	0.00	9.58	29.45	31.20	4.17
Average in LNCs	0.47	0.07	0.15	6.77	0.13	8.66	40.72	46.02	41.08

Full Sample									
N	520	520	520	507	507	510	453	520	520
Mean	0.49	0.06	0.16	6.87	0.15	5.50	39.49	45.85	41.02
Std.Dev.	0.50	0.24	0.36	7.06	0.36	7.34	22.59	14.97	22.88
Min.	0.00	0.00	0.00	0.00	0.00	0.00	1.25	9.25	1.64
Max.	1.00	1.00	1.00	35.00	1.00	62.70	96.62	86.47	93.88