

# Directing Retribution: Mandatory Minimums, Retention, and the Discretion of Trial Judges<sup>1</sup>

Gregory Huber  
Assistant Professor  
Department of Political Science and ISPS  
Yale University  
77 Prospect Street, PO Box 208209  
New Haven, CT 06520  
gregory.huber@yale.edu

Sanford Gordon  
Assistant Professor  
Department of Political Science  
The Ohio State University  
2140 Derby Hall, 154 North Oval  
Columbus, OH 43210  
gordon.256@osu.edu

August 24, 2001

<sup>1</sup>Author order is randomized. Thanks to Jakub Zielinski for indispensable game theory help. Comments welcome.

# 1 Introduction

Trial judges in courts of general jurisdiction are fundamental players in the day-to-day operation of the criminal justice system. Voters and their elected representatives delegate responsibility to these judges to make decisions that, at the extreme, determine whether convicted felons live or die. The exercise of judicial discretion, however, is prone to frequent criticism. Political conservatives allege that liberal judges are “soft on crime” and ignore public preferences about the appropriate punishment of the guilty. The fear that “renegade judges” are overly lenient parallels charges of judicial activism on the part of justices in higher courts. Liberals, on the other hand, claim that judges assign sentences arbitrarily and may knowingly subvert justice in order to maintain public order.

These concerns about the behavior of judges are similar to concerns about numerous government officials to whom the public delegates authority. Federal bureaucrats, for instance, are hypothesized to inflate the costs of production to pad their budgets (Niskanen 1971). Similarly, corrupt lawmakers are accused of guiding government contracts to their friends and political supporters. In the area of criminal justice, district attorneys may knowingly prosecute the innocent or treat defendants differently on the basis of individual characteristics such as race. Police officers may unwarrantedly single out minorities for scrutiny and abuse. As a result of these problems, it is not surprising that effort is invested in controlling the behavior of these officials. Concern about the shirking behavior of bureaucrats leads to the expenditure of considerable resources to identify trustworthy bureaucrats, place constraints on the ability of bureaucrats to arrange sweetheart deals, or subject prosecutors to electoral review to weed out officials who undertake inappropriate activity.

In this paper, we focus on two methods voters and their elected representatives may employ for controlling judicial behavior. The first is ex ante constraints on the sentences judges can assign. The second is ex post review of judges through elections or reappointment proceedings. We derive the optimal range of discretion citizens (or their elected representatives) should give judges to sen-

tence defendants and show how the promise of reelection effectively constrains judicial discretion in a similar, though less immediately efficient manner. We show that ex ante controls on discretion are superior to ex post review when a principal's time horizon is short, because the principal can commit to a set of restrictions that reduce agency loss more than a credible promise to not reelect a judge who assigns a particular sentence.

The remainder of this paper is divided into three sections. Section 2 places our inquiry in the larger literature on the means of controlling officials and on the particular concerns about judicial behavior. Specifically, we discuss the use of agency theory to examine problems of delegation and the basic choice between ex ante and ex post means of control. We also examine citizen preferences for appropriate punishment of the guilty. In Section 3 we present two formal models of judicial control. This section begins with our formulation of actors and preferences. We then present a model of sentencing guidelines, in which a principal must choose the optimal mandatory minimum and maximum sentence given uncertainty about a judge's ideology and a defendant's culpability. Our second modeling effort is an election game between judges and their principals. Section 4 discusses the findings and offers some tentative conclusions. In light of our results, we discuss the advantages of different institutional arrangements for controlling judges and consider several additions to our basic modeling framework. We also relate this paper to our larger effort to understand the operation of the institutions of the criminal justice system.

## **2 Control, Punishment, and Judicial Discretion**

### *The Means of Control*

Voters, legislators, and governors delegate authority to trial judges to conduct specific tasks in the implementation of crime policy. Judges oversee trials, interpret questions of law, approve negotiated guilty pleas, and assign sentences to the convicted. The connections between voters (or their elected representatives) and these officials are special varieties of principal-agent relationships. A principal-agent relationship exists when a principal "contracts" with an agent to fulfill a task. This contract, usually supported by the principal's monitoring efforts, provides incentives

for the agent to act according to the principal's preferences. The most important characteristic of principal-agent relationships is that the principal is usually unable to induce an agent to act exactly according to the former's will (for detailed introductions, see Moe 1984; Milgrom and Roberts 1992). In particular, asymmetric information between the principal and agent will exacerbate "agency loss." One problem of asymmetric information is moral hazard – when the agent takes actions that are unobservable to the principal.

A second problem of asymmetric information, which concerns us in the current context, is adverse selection (Spence 1974). This occurs when the agent possesses private information that she is either unwilling or incapable of transmitting to her superior. Most importantly, the agent knows her own skills, diligence, and preferences. If she is incompetent, lazy, or ideologically remote from her principal, she may act to obscure these facts. At the same time, if she is skilled, diligent, or in agreement with the principal, external factors may obscure those characteristics. In many contracting situations and often in elections, the adverse selection problem is compounded by the fact that replacement agents are *ex ante* observationally equivalent to one another. Each will claim to have talents and skills that are difficult to verify.

Agency theory has provided important insights in a number of areas in political science. Two most directly inform our current research. The first is the relationship between elected officials and the bureaucracy. How should a legislature structure its delegation of policy-making authority to the executive to reduce bureaucratic malfeasance while simultaneously capitalizing on the bureaucracy's expertise or private knowledge? One way is to explicitly shrink bureaucrats' discretion through statutory limits on their behavior (Epstein and O'Halloran 1994, 1996, 1999; Hamilton and Schroeder 1994; Huber and Shipan 2001; Huber, Shipan, and Pfahler 2001; Whitford and Helland 1998; see Lowi 1979 for a normative argument in favor of limiting discretion). For example, by statute the Occupational Safety and Health Administration must respond to worker complaints by initiating a workplace inspection (Huber 2001). Similarly, the Environmental Protection Agency and the states must by law inspect certain facilities responsible for the disposal of hazardous waste every two years (Gordon 2000). Statutory limits on discretion are varieties of *ex ante* controls on

bureaucratic behavior.

Elected officials may also rely on ex post incentives to motivate bureaucratic behavior: Congress retains the “power of the purse” through its residual control of agency resources to counteract detected bureaucratic or executive transgressions (Fenno 1966; Arnold 1979; Kiewiet and McCubbins 1991). The political principal-agent relationship in which ex post incentives are most well known, however, is that between voters and politicians. Some scholars have considered this relationship as a variety of signaling game, the approach we adopt below. Most intriguingly, a voter behaving optimally may inadvertently induce a politician to take actions contrary to the interests of the voter. For example, Austen-Smith (1993) suggests that under some conditions, a legislator’s need to justify her vote to the public prevents her from voting sophisticatedly, even though doing so would further her constituency’s interests. Similarly, in a recent paper, Canes-Wrone, Herron, and Schotts (2001) examine the incentives for presidents to “pander” by going along with popular policies that their private information shows are detrimental to voters’ interests.

We consider the implications of the adverse selection problem in choosing judges. Judges have information about individual cases and their own preferences that their principals lack. Moreover, expertise is essential in discerning the circumstances of particular cases. A competent judge will be better at identifying the offender who needs only a minimal sentence to deter subsequent criminal activity. She will also excel in evaluating the extenuating circumstances that may justify a lesser sentence. For voters, legislators, or governors to gather this information on a case-by-case basis would be prohibitively costly. A judge’s preferences will also affect the sentences she imposes. As we explain below, some judges might be more retributive than others, or have different beliefs about deterrence, rehabilitation, or incapacitation. About 80% of General Social Survey respondents from 1976 to 1994 offered a belief that criminal courts are too lenient, while virtually none believed the courts are too harsh (Warr 1995, 300). It is difficult, however, for voters or elected officials to discern *which* judges (or judicial nominees) depart from public preferences, and by how much.

What types of ex ante and ex post mechanisms might voters or elected officials employ in

an attempt to induce compliant behavior on the part of judges given the adverse selection problem? Consider sentencing behavior, one of the most prominent discretionary roles of the trial judge. Since the 1970s, state legislatures have experimented with a number of institutional remedies to the perceived abuse of sentencing discretion by judges and administrative agencies. The first of these is determinative sentencing, intended to reduce the discretion of parole boards. (In some states, these boards have been abolished outright). The second is an increased reliance on mandatory minimum sentences. Legislatures have increasingly established statutory baseline sentences for particular crimes, most notably those involving drugs. Third, a number of states have created sentencing commissions that are charged with establishing guidelines to structure judicial discretion and monitoring judicial compliance with the guidelines. Fourth, starting in the early 1990s, a number of states have implemented "three strikes" provisions for repeat offenders. Each of these reforms has the effect of reducing the discretion of one party (usually the trial judge) and enhancing that of another (usually the prosecutor).

Federal judges and trial judges in Massachusetts, Rhode Island, and New Hampshire are appointed for life terms. 47 states, however, subject judges to periodic review of their performance in office. In 39, judges must stand for reelection, either via competitive contest or retention vote. In other states, legislatures, governors, or special commissions decide whether to retain incumbent judges. Most trial judges, therefore, run the risk of ex post punishment at the polls if their behavior in office does not conform to the evaluative criteria of a relevant principal.

Overall, this institutional variation motivates our inquiry in this paper. What are the advantages of different means of control? What is lost by adopting each means of controlling judicial malfeasance? Finally, if one had to choose between either ex ante or ex post control (and could not impose both simultaneously), what would the optimal choice of incentives to guide judicial behavior look like?

## *Preference for Punishment*

We wish to compare ex post oversight vs. ex ante control of trial judges. An important first step in modeling these phenomena is to identify how judges' decisions matter to citizens or their elected representatives. In other words, what are the preferences of judges and their principals? We focus on one particularly important judicial decision: the assignment of sentences to the convicted.

Scholars have long identified different motivations for punishment, including retribution, general and specific deterrence, incapacitation, and rehabilitation. We assume that these concerns motivate citizens or their elected representatives to prefer that a defendant of a particular type  $c$  receive a sentence of  $r_i + \alpha c$  (we return to the motivations for this functional form below). In this formulation,  $r_i$  is individual  $i$ 's preference for the *minimal acceptable punishment for a given crime*. The parameter  $c$  is a measure of a defendant's "culpability." Larger  $c$ 's are associated with a higher degree of responsibility and more punishment. The parameter  $\alpha$  is a scaling parameter, shared by all actors in the game, which measures the degree to which culpability affects the preferred punishment (one may also think of  $\alpha$  as the sentence's elasticity). If  $\alpha = 0$ , then  $i$ 's preferred sentence for all defendants is fully independent of culpability. Conversely, if  $\alpha$  is very large, the "baseline" level of punishment,  $r_i$ , contributes very little to the sentence of a "very guilty" defendant ( $c = 1$ ).

Our formulation of preferred punishment can account for many different sources of citizen preferences for punishment. For instance, if we understand retribution to be a belief that all persons who commit a certain crime deserve a certain level of "just deserts," regardless of their degree of culpability, then  $r_i$  captures this value. We will refer to the baseline sentence  $r_i$  as ideology, and call individuals with high  $r_i$ s "conservative" and those with low  $r_i$ s "liberal." The  $\alpha c$  component of an actor's preference for punishment can incorporate concerns about deterrence, incapacitation, and rehabilitation. Efficient deterrence requires that more culpable defendants receive a greater punishment in order to deter them, and similar potential criminals, from (re-) committing the crime. Similarly, incapacitation motivates the formulation that "more dangerous" people are incarcerated

for longer periods of time. These individuals, who commit crimes despite being responsible for their actions, are more dangerous than those who commit crimes only under duress. Finally, rehabilitation drives a preference for more punishment for those who are more responsible. These people, by virtue of their greater likelihood of recidivism, require more assistance and training in becoming responsible citizens.

### 3 Models of Ex Ante and Ex Post Control of Judicial Sentencing Discretion

We examine two ways that citizens and their elected representatives can minimize the loss in delegating sentencing authority to judges. The first is ex ante restrictions on the sentences judges can assign. The second is ex post review of judges, whereby the promise of reappointment is used to encourage judges to behave according to the principal's preferences. In this section, we first describe the basic framework of judicial choice and the determinants of each player's payoffs. Second, we present a model in which ex ante sentencing guidelines are used to constrain judges. The question in this case is how much and what type of discretion should the principal give to the judge. Third, we develop a separate model in which the principal relies solely on ex post incentives to limit agency loss. This is a signaling game, in which the principal must try ascertain the preference of a judge from the assigned sentence.

In both games, a judge assigns a sentence to a criminal following conviction. The principal most prefers that the judge assigns the sentence  $s_p = r_p + \alpha c$ . Here,  $c$ , case-specific information, is a random variable uniformly distributed from zero to one. The principal's first problem is that he cannot observe  $c$ .<sup>1</sup> The judge is privy to this knowledge by virtue of having observed the trial, but her preferences may differ from the principal's. Herein lies the principal's second problem: he doesn't know the judge's preferences, only their distribution. We assume that judges, like defendants, are uniformly distributed on the range  $r_j = [0, 1]$ . Both the judge and the principal

---

<sup>1</sup>The principal could invest in learning the particular facts in every case. This is likely to be expensive, however, and so in the institutional structure described here principals or their elected representatives have delegated the task of ascertaining the culpability of individual defendants to the judge.



have quadratic utility functions, such that  $u_i(S) = -(s_i - S)^2$ , where  $S$  is the assigned sentence and  $s_i$  is individual  $i$ 's ideal sentence given  $c$ . All parties know the principal's ideology with certainty, and all possess the same  $\alpha$ .

### *Statutory Controls on Sentencing Discretion*

The first way principals may constrain judges is by placing limits on the sentences they can assign. For instance, mandatory minimum sentences, statutory maximums, and sentencing guidelines all limit the discretion of judges to sentence defendants. If these rules are binding, a judge cannot sentence less than the minimum or more than the maximum. Likewise, sentencing guidelines require that defendants convicted of the same crime receive different sentences based on other factors deemed relevant to the deserved punishment. Repeat offenders, for example, are given longer sentences in most sentencing systems because they are seen as more culpable (and harder to deter.) Ex ante controls are particularly important when judges are not subject to ex post review, as in the federal courts and the three states mentioned above.<sup>2</sup> With only limited ex post review, ex ante restrictions on sentencing are one of the only ways to limit the loss in delegating sentencing authority to individual judges.

First, consider the judge who has complete discretion and tenure. She would prefer to assign the sentence  $s_j = r_j + \alpha c$ . The principal, however, may prefer a sentence that is larger or smaller (depending on the relative size of his  $r_p$ ). The principal's expected utility across all possible  $r_j$ 's and  $c$ 's is

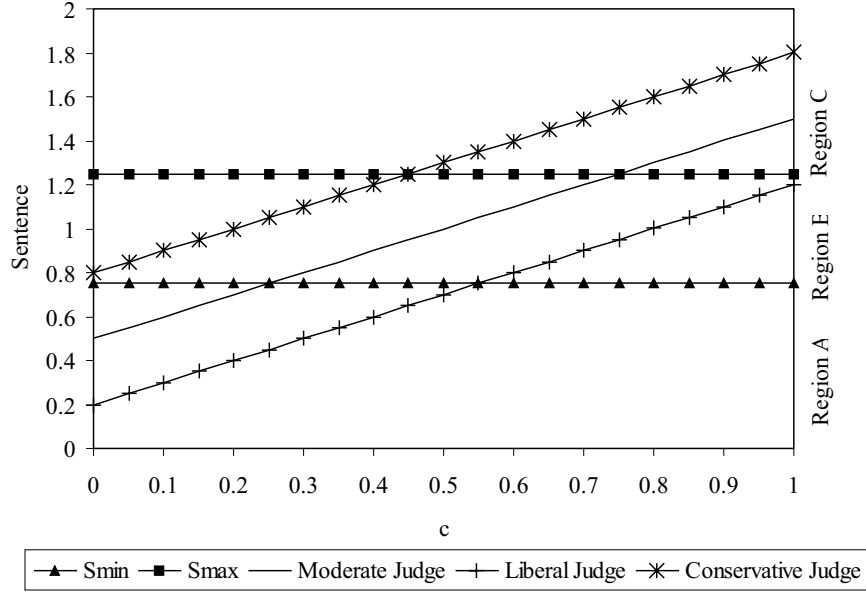
$$\int_0^1 \int_0^1 -(r_p - r_j)^2 dc dr_j = -1/3 + r_p - r_p^2. \quad (1)$$

Now, assume that the principal can choose a mandatory minimum sentence  $S_{min}$  and a mandatory maximum sentence  $S_{max}$  such that the assigned sentence,  $S$  is bounded from below by  $S_{min}$  and from above by  $S_{max}$ . The principal's problem is to choose these values so as to minimize the loss from delegating sentencing authority to the judge. For illustrative purposes, consider the

---

<sup>2</sup>Since the founding, only 13 federal judges have been impeached, and seven convicted. See Volcansek et. al. 1996.

**Figure 1:** The Basic Problem: Setting Constraints for Judges when Preferences and Case-Specific Expertise are Private Information



extreme case where  $\alpha = 0$ . In this situation, the principal would prefer the judge assigned the same sentence  $r_p$  regardless of how culpable a defendant is. Because  $\alpha$  is shared by all players in the game, the judge also prefers to assign a constant sentence to all defendants, but her  $r_j$  may be smaller or larger than the principal's  $r_p$ . In this case, the principal's choice of  $S_{min}$  and  $S_{max}$  is trivial. By setting  $S_{min} = S_{max} = r_p$  the principal assures that all defendants receive the sentence he would assign. Thus, in the extreme case where the judge's private knowledge of  $c$  is worthless, there is no advantage to granting the judge any discretion and there is no agency loss.

By contrast, if  $\alpha > 0$ , then the principal values the information the judge observes about  $c$ . For instance, if  $\alpha = 1$ , and  $r_p = 0.5$ , the principal's ideal sentence ranges from 0.5 to 1.5, depending on  $c$  (See Figure 1). If the principal allows the judge to assign whatever sentence she wants, the principal's expected utility is equation (1) evaluated at  $r_p = 0.5$ , or  $-1/12$ . If the principal grants the judge no discretion, and instead chooses  $S_{mandatory} = S_{min} = S_{max} = 1$  (her average preferred sentence) then her expected utility is

$$\int_0^1 \int_0^1 -(r_p + \alpha c - S_{mandatory})^2 dc dr_j = -\frac{\alpha^2}{3} - (r_p - S)\alpha - (r_p - S_{mandatory})^2. \quad (2)$$

Equation (2) evaluated at  $r_p = 0.5$ ,  $\alpha = 1$ , and  $S_{mandatory} = 1$  is  $-1/12$ . This is equal to the principal's expected loss from the completely unconstrained judge. If  $\alpha > 1$ , then the right side of equation (1) exceeds that of equation 2 for any  $S_{mandatory}$ . If  $\alpha < 1$ , then the right side of equation (1) falls below that of equation (2) for  $S_{mandatory} = r_p + 1/2\alpha$ . Thus, as the value of discerning the characteristics of an individual increases, allowing the judge complete discretion becomes preferable to no discretion at all. Between these extremes of  $S_{min} = 0, S_{max} = 2$  (complete discretion given  $\alpha = 1$ ) and  $S_{min} = S_{max} = r_p + 1/2\alpha$  are moderate restrictions on sentencing. Can these reduce the principal's loss? With non-extreme sentencing constraints, the principal's expected utility is

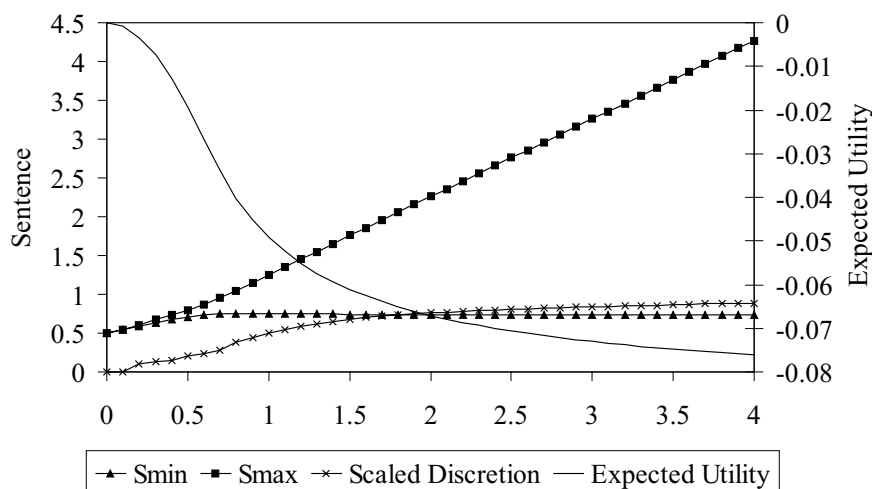
$$\int_0^1 \int_0^1 -(r_p + \alpha c - \min(\max(S_{min}, r_j + \alpha c), S_{max}))^2 dc dr_j. \quad (3)$$

The expression  $\min(\max(S_{min}, r_j + \alpha c), S_{max})$  captures three circumstances:

When  $S_{min} > r_j + \alpha c$ ,  $S = S_{min}$ . This captures  $r_j$ s in Region A in Figure 1, where the judge always assigns the mandatory minimum. When  $S_{max} \geq r_j + \alpha c \geq S_{min}$ ,  $S = r_j + \alpha c$ . This captures judges in region B, who assign a most preferred sentence. When  $r_j + \alpha c > S_{max}$ ,  $S = S_{max}$ . This captures judges in region C, who must always assign the statutory maximum.

Deriving an analytical solution for the maximum of equation 3 is difficult because of the numerous boundary conditions and cases. Below, we derive an analytical solution for the case where  $\alpha = 1$ . For the general case, we use numerical simulations across a range of  $r_p$ 's and  $\alpha$ 's to find the optimal values of  $S_{max}$  and  $S_{min}$ . First, we hold the principal's ideology constant at  $r_p = 1/2$  and allow  $\alpha$  to vary from 0 to 4. Figure 2 displays the  $S_{max}$  and  $S_{min}$  that minimizes the principal's loss across this range of  $\alpha$ s.  $S_{min}$  rises gradually from  $1/2$  when  $\alpha = 0$  to  $0.75$  around  $\alpha = 0.7$ , where it stays for the remainder of the interval. In other words, the optimal  $S_{min}$  is increasing for  $\alpha < 0.7$  and then constant above that point.  $S_{max}$  also increases from  $1/2$  when  $\alpha = 0$  to  $0.95$  when  $\alpha = 0.7$ . Above that value it is approximately equal to  $\alpha + 0.25$ . Why does the optimal  $S_{min}$  remain constant above  $\alpha = 0.7$  while the optimal  $S_{max}$  increases? Foremost, because as  $\alpha$  increases, the range of feasible sentences increases while the minimum

**Figure 2:** Optimal Sentencing Constraints as a Function of  $\alpha$ , Holding the Principal's Ideology constant at 0.5



feasible sentence remains the same. Holding  $S_{max}$  constant would deny judges, even ideologically distant ones, the ability to scale punishment to circumstances. The principal cares about this loss, and so gives more discretion.

Figure 2 also displays a scaled measure of discretion and the principal's expected utility (on the right axis). The standardized measure of discretion is calculated as  $(S_{max} - S_{min})/\alpha$ . Dividing the range of acceptable sentences by  $\alpha$  captures the degree to which discretion is greater or smaller than the range of sentences the principal himself would need to perfectly match sentences to culpability. Scaled discretion increases quickly from 0 to 1, and then quickly flattens out. The principal's expected utility decreases from 0 as  $\alpha$  increases. Thus, the principal cannot recover as much by restricting the judge's range of sentences as the importance of the information observed by the judge increases. Given that the loss from an unconstrained judge when  $r_p = 0.5$  is  $-1/12$  (about  $-0.08333$ ), the principal is gaining only about  $7/1000$  by choosing the optimal  $S_{max}$  and  $S_{min}$  when  $\alpha = 4$ . This is less than 10% of the loss from allowing an unconstrained judge to assign sentences.

Figure 3 display the optimal  $S_{min}$  and  $S_{max}$  for  $\alpha = \{0.5, 1, 2\}$  across the entire range of  $r_p$ 's. In all three cases, both  $S_{min}$  and  $S_{max}$  are roughly parallel and increasing with  $r_p$ . When

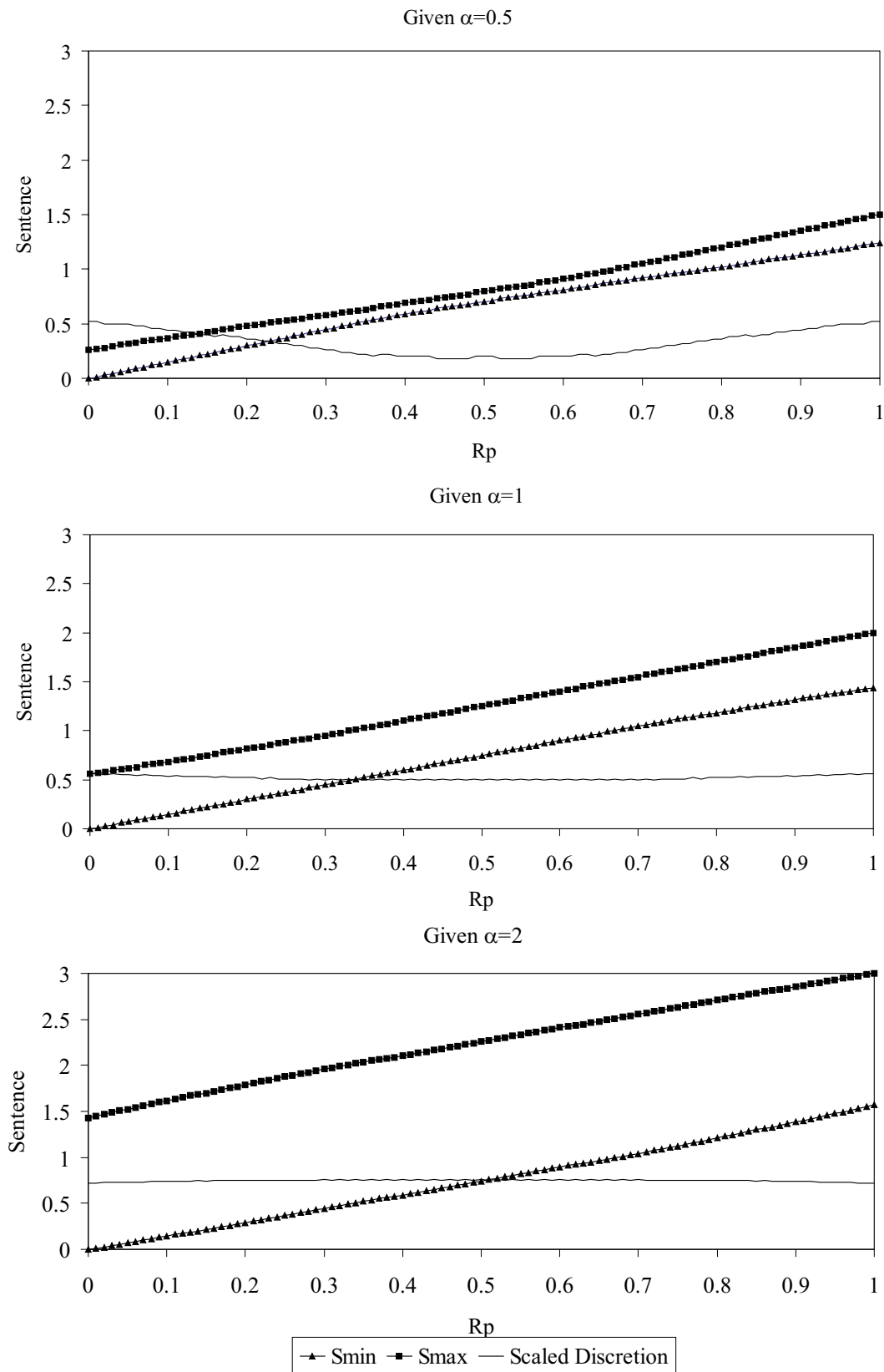
$\alpha = \{0.5, 1\}$  there is some bowing out for extreme values of  $r_p$ . This is reflected in the increase in scaled discretion when  $r_p$  is very small or very large. In contrast, when  $\alpha = 2$ , the gap between  $S_{min}$  and  $S_{max}$  is smallest when  $r_p$  is small or large. In this case, discretion is maximized at  $r_p = 0.5$ .

In addition to these numerical simulations, we also derive an analytical solution to the case where  $\alpha = 1$ . The difficulty with this optimization problem is that a judge may be in one of five states. First, she may be bound entirely by the minimum sentence. For instance, if  $r_j = 0$ ,  $S_{min} = 1.25$ , and  $\alpha = 1$ , the maximum sentence this judge would choose to assign if she were not constrained by a sentencing guideline is 1. When  $S_{min} > r_j + \alpha$ , the judge must in all cases assign  $S_{min}$ . Second, the judge may be partially bound by only the minimum sentence. For instance, if  $S_{min} = 1.25$  and  $\alpha = 1$  but  $r_j = .75$ , then for  $c \leq 0.5$ , the judge will assign sentence  $S_{min}$  and for  $c > .5$  the judge will assign  $r_j + c$ . The point at which the  $S_{min}$  constrain no longer binds is  $c = (S_{min} - r_p)/\alpha$ .

Third, the judge may be bound by both the upper and lower constraint. If  $r_j < S_{min}$  then for  $c < (S_{min} - r_p)/\alpha$  the judge assigns  $S_{min}$ . If  $S_{max} < r_j + \alpha c$  then the judge assigns her most preferred sentence from  $(S_{min} - r_p)/\alpha \leq c \leq (S_{max} - r_p)/\alpha$ . Finally, for  $c > (S_{max} - r_p)/\alpha$  the judge assigns  $S_{max}$ . Fourth, the judge may be bound only by the required maximum sentence. If  $r_j > S_{min}$  but  $r_j < S_{max}$  then for  $c \leq (S_{max} - r_p)/\alpha$  the judge can assign her most preferred sentence, while for  $c > (S_{max} - r_p)/\alpha$  she assigns  $S_{max}$ . The fifth and final case is when neither  $S_{min}$  nor  $S_{max}$  constrains the judge. If  $S_{min} \leq r_j$  and  $S_{max} \geq r_j + \alpha$  then the judge assigns her preferred sentence  $r_j + \alpha c$  for the entire range of  $cs$ .

Examining the middle panel of Figure 3, it is easy to see that when  $\alpha = 1$ , All judges are bound by either  $S_{min}$  or  $S_{max}$ . There are only three circumstances to consider. First, when  $S_{max} \leq 1$ , all judges are bound by the maximum sentence for some values of  $c$  [the left part of the Figure]. Second, when  $S_{min} > 1$ , all judges are bound by the minimum sentence for some values of  $c$  [the right part of the Figure]. Finally, in the middle region of the figure, for certain values of  $c$ , some judges are bound by the upper sentence and some are bound by the minimum sentence. For

**Figure 3:** Optimal Sentencing Constraints as a Function of  $r_p$ , Holding  $\alpha$  constant at 0.5, 1, and 2.



each of these three cases we can define the principal's expected utility.

When  $S_{min} < S_{max} \leq 1$ , the principal's expected utility is

$$EU_{condition1} = \int_0^{S_{min}} \left[ \int_0^{S_{min}-r_j} -(r_p + c - S_{min})^2 dc + \int_{S_{min}-r_j}^1 -(r_p - r_j)^2 dc \right] dr_j + \int_{S_{min}}^{S_{max}} \left[ \int_0^{S_{max}-r_j} -(r_p - r_j)^2 dc + \int_{S_{max}-r_j}^1 -(r_p + c - S_{max})^2 dc \right] dr_j + \int_{S_{max}}^1 \int_0^1 -(r_p + c - S_{max})^2 dc dr_j.$$

When  $1 \leq S_{max} < S_{min} + 1$ , the principal's expected utility is

$$EU_{condition2} = \int_0^{S_{max}-1} \left[ \int_0^{S_{min}-r_j} -(r_p + c - S_{min})^2 dc + \int_{S_{min}-r_j}^1 -(r_p - r_j)^2 dc \right] dr_j + \int_{S_{max}-1}^{S_{min}} \left[ \int_0^{S_{min}-r_j} -(r_p + c - S_{min})^2 dc + \int_{S_{min}-r_j}^{S_{max}-r_j} -(r_p - r_j)^2 dc + \int_{S_{max}-r_j}^1 -(r_p + c - S_{max})^2 dc \right] dr_j + \int_{S_{min}}^1 \left[ \int_0^{S_{max}-r_j} -(r_p - r_j)^2 dc + \int_{S_{max}-r_j}^1 -(r_p + c - S_{max})^2 dc \right] dr_j.$$

When  $1 \leq S_{min} < S_{max} \leq 2$ , the principal's expected utility is

$$EU_{condition3} = \int_0^{S_{min}-1} \int_0^1 -(r_p + c - S_{min})^2 dc dr_j + \int_{S_{min}-1}^{S_{max}-1} \left[ \int_0^{S_{min}-r_j} -(r_p + c - S_{min})^2 dc + \int_{S_{min}-r_j}^1 -(r_p - r_j)^2 dc \right] dr_j + \int_{S_{max}-1}^1 \left[ \int_0^{S_{min}-r_j} -(r_p + c - S_{min})^2 dc + \int_{S_{min}-r_j}^{S_{max}-r_j} -(r_p - r_j)^2 dc + \int_{S_{max}-r_j}^1 -(r_p + c - S_{max})^2 dc \right] dr_j.$$

Differentiating  $EU_{condition1}$  with respect to  $S_{min}$  and solving for  $S_{min}$  allows us to find the  $S_{min}$  that maximizes the principal's utility. This  $S_{min}^*$  is  $3/2r_p$ . Differentiating  $EU_{condition1}$  with

respect to  $S_{max}$  yields

$$dEU_{condition1}/dS_{max} = 1 - 2S_{max} + 2r_p + 2/3S_{max}^3 - r_pS_{max}^2.$$

Solving this equation for  $S_{max}$ , we find that  $S_{max}^*$  is equal to  $3/2r_p + 1/2 + e$ , where  $e$  is an arbitrary decreasing function of  $r_p$ .

Differentiating  $EU_{condition2}$  with respect to  $S_{min}$  and solving for  $S_{min}$  allows us to obtain the mandatory minimum that maximizes the principal's expected utility. Again, this  $S_{min}^*$  is  $3/2r_p$ . Differentiating  $EU_{condition2}$  with respect to  $S_{max}$  and solving for  $S_{max}$  allows us to obtain the statutory maximum that optimizes the principal's utility. This  $S_{max}^* = 3/2r_p + 1/2$ .

Finally, differentiating  $EU_{condition3}$  with respect to  $S_{min}$  yields

$$-1/3 + 2S_{min} - 3S_{min}^2 + 2/3S_{min}^3 - r_pS_{min}^2 + 4r_pS_{min} - 2r_p.$$

It turns out that  $S_{min}^*$  is equal to  $3/2r_p - e$  where  $e$  is an arbitrary increasing function of  $r_p$ .

Differentiating  $EU_{condition3}$  with respect to  $S_{max}$  and solving for  $S_{max}$  allows us to obtain the  $S_{max}$  that maximizes the principal's utility. Again, this  $S_{max}^* = 3/2r_p + 1/2$ .

Overall, for all values of  $r_p$  and given  $\alpha = 1$ , the optimal  $S_{max}$  and  $S_{min}$  are:

For  $0 \leq r_p < 1/3$ :  $S_{min}^* = 3/2r_p$  and  $S_{max}^*$  is slightly larger than  $3/2r_p + 1/2$ .

For  $1/3 \leq r_p < 2/3$ :  $S_{min}^* = 3/2r_p$  and  $S_{max}^* = 3/2r_p + 1/2$ .

For  $2/3 \leq r_p \leq 1$ :  $S_{min}^*$  is slightly smaller than  $3/2r_p$  and  $S_{max}^* = 3/2r_p + 1/2$ .

Note that for every value of  $r_p > 0$ ,  $S_{min} > r_p$  and  $S_{max} < r_p + 1$ . In other words, even a judge with the exact same ideology as the principal would be unable to sentence as the principal desired in certain cases. Because the principal is uncertain about the type of the judge who will impose sentences in a particular case, ex ante he is better off denying all judges this discretion. Herein is the principal's basic conflict: Restricting discretion reduces the loss from ideologically distant judges, but it also weakens the ability of judges to use their case-specific information to tailor sentences to particular defendants. The principal reduces discretion to minimize the loss from ideologically distant judges until the loss in the "appropriateness" of imposed sentences is equal to this gain.



## *Sentencing and Judicial Retention*

Next, we consider a world in which there are no ex ante institutional constraints on judicial behavior. Instead, the judge is given discretionary sentencing power by a principal who after observing the judge's behavior must decide whether to retain the judge or replace her with a draw from the underlying distribution of replacements. We maintain most of the model specification and assumptions from the previous section. In particular, both the judge and principal have quadratic utility over sentences,  $u_i(s) = -(r_i + \alpha c - s)^2$ . Case-specific circumstances,  $c$ , and judge's baseline sentences,  $r_j$ , are drawn from independent uniform  $[0, 1]$  distributions. The drawn values of  $c$  and  $r_j$  are unknown to the principal but known with certainty by the judge. As before, all players share a common  $\alpha$ . The major departure in the current setting is that the judge values retaining office in addition to the ability to match sentences to her own ideal.

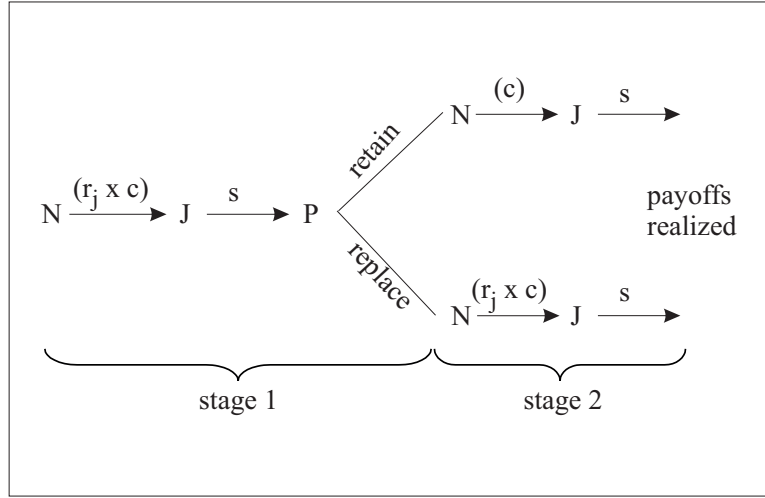
We model the retention decision as a variant of a two-stage signaling model. In the first period, Nature draws a judge with baseline sentence  $r_j$  and a circumstance  $c$ . The judge then sets a sentence, and with this act, reveals information about her type. Because the type-space is two dimensional, signals will almost never fully reveal  $r_j$ . Instead, the principal must make inferences about the likely ideology of the judge given the observed sentence, and decide accordingly whether to retain or replace her. In the second period, if the judge is retained, she receives a new case and, now safely elected, can impose her ideal sentence. Retaining office is worth  $V$  to the judge. If the judge is replaced, both a new judge and new case are drawn from their underlying distributions, and the new judge implements her ideal sentence.<sup>3</sup> The sequence of events in the signaling game is displayed in Figure 4.

A strategy for the judge,  $S_t$ , is a mapping in each period  $t$  of defendant culpability and judicial ideology to a sentence,  $S : \mathbb{R}^1 \times \mathbb{R}^1 \rightarrow \mathbb{R}^1$ . In the absence of institutional constraints,

---

<sup>3</sup>At this point, we are ignoring the fact that the judge would eventually stand for reappointment once again at the end of the second stage. This is clearly an unwarranted assumption, and one we hope to relax in the future. For now, this specification is sufficient for modeling the principal's response when he strictly desires a judge for the next period with  $r_j$  closer to  $r_p$ .

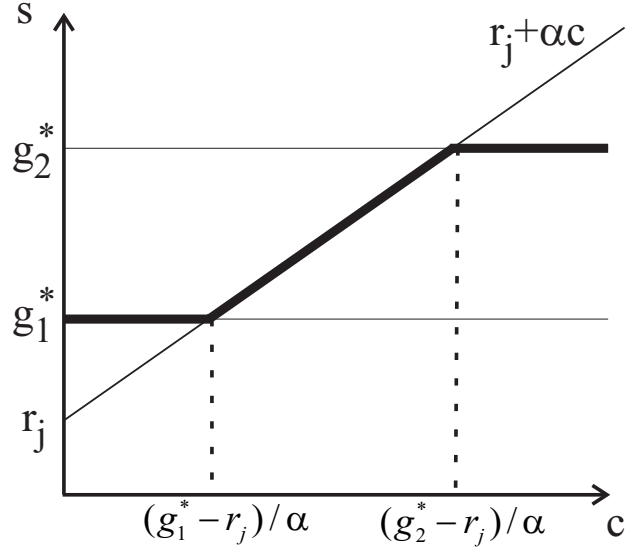
**Figure 4:** Judicial Retention Game: Extensive Form.



the practical range of sentences along  $\mathfrak{R}^1$  is  $S \in (0, 1 + \alpha)$ . Even the most conservative judge ( $r_j = 1$ ) will never find it in her interest to assign a sentence higher than  $1 + \alpha$ , and even the most conservative principal will not find it in his interest to reward it. A strategy for the principal,  $G$ , is a mapping of the judge’s sentence and the principal’s baseline sentence  $r_p$  to a retention decision,  $G : \mathfrak{R}^1 \times \mathfrak{R}^1 \rightarrow \{0, 1\}$ .

For purposes of comparison with the mechanism design model of the previous section, and for purposes of analytic tractability, we make an important simplifying assumption for the time being: that the value of  $V$  is “large.” By this we mean that the benefit of retention is sufficiently high that judges of all types will do what is necessary to be reelected. As an empirical matter, a large majority of judges standing for retention maintain their posts (Aspin 1999). We also assume that  $\alpha \geq 1$ . Under these assumptions, the principal’s equilibrium strategy will be characterized by a set of cutpoints along the sentence space,  $g_1^*$  and  $g_2^*$ . Upon observing sentences between and including those points, the principal will opt to retain the incumbent judge. Because the value of holding office is large, the judge will sometimes impose her ideal sentence,  $r_j + \alpha c$ , and for relatively extreme values of  $c$ , impose either  $g_1^*$  or  $g_2^*$ . Figure 5 displays the sentencing strategy of one judge as a function of case-specific information.

**Figure 5:** Judicial Sentencing Strategy as a Function of Case-Specific Information.



In the figure,  $c$  is broken into three distinct regions. If  $V$  were smaller, however,  $c$  could be broken into as many as five regions. The middle three would resemble those depicted in the figure. In the two outmost regions, however, the disutility to the judge of departing from her ideal sentence would outweigh the benefit of holding office. In those cases, the judge would simply impose her ideal sentence and fail to be retained for the next period.

**Equilibrium.** The following strategies constitute a Bayesian perfect Nash equilibrium:

1. Principal's strategy:  $G^* = 1$  iff  $g_1^* \leq S \leq g_2^*$ , where a)  $g_1^* = \min(1, 2r_p - \sqrt{4r_p^2 + 2 - 6r_p})$  and b)  $g_2^* = \max(1, 2r_p + \alpha - 1 + \sqrt{4r_p^2 - 2r_p})$ .

2. Judge's strategy:

$$S_{t=1}^* = \begin{cases} g_1^* & : r_j + \alpha c < g_1^* \\ r_j + \alpha c & : g_1^* \leq r_j + \alpha c \leq g_2^* \\ g_2^* & : r_j + \alpha c > g_2^* \end{cases}$$

$$S_{t=2}^* = r_j + \alpha c$$

To solve the game, we start with the action of the judge in the last period,  $t = 2$ . The judge implements her ideal sentence,  $s_{t=2} = r_j + \alpha c$ , which minimizes her expected loss. If the principal

opted to replace at the end of the first period, both the baseline sentence of the new judge  $r_j$  and the case-specific information  $c$  are drawn anew from their standard uniform prior distributions. The expected payoff to the principal of replacing the judge is therefore equation 1 from the previous section.

Given a sentence between  $g_1^*$  and  $g_2^*$ , the principal will strictly prefer retaining the incumbent judge, while outside that range, he will favor replacing the judge. This implies that at the two cutpoints, the principal will be indifferent between retaining and replacing. The indifference conditions may be expressed as:

$$-\int_0^1 \int (r_j - r_p)^2 f_{r_j}^p(r_j|s = g_1^*) f_c(c) dF_{r_j}^p(r_j) dc = -\frac{1}{3} + r_p - r_p^2 \quad (4a)$$

$$-\int_0^1 \int (r_j - r_p)^2 f_{r_j}^p(r_j|s = g_2^*) f_c(c) dF_{r_j}^p(r_j) dc = -\frac{1}{3} + r_p - r_p^2, \quad (4b)$$

where  $f_{r_j}^p(r_j|s = x)$  is the principal's posterior distribution on  $r_j$  given an observed sentence  $s = x$ .

**Lemma 1.** The principal's posterior distribution on  $r_j$  given an observed sentence  $g_1^*$  is decreasing right triangular on the interval  $(0, g_1^*)$ . This gives the pdf

$$f_{r_j}^p(r_j|s = g_1^*) = -\frac{2r_j}{g_1^{*2}} + \frac{2}{g_1^*}.$$

**Proof.** First, note that irrespective of the value of  $c$ , no judge with  $r_j > g_1^*$  ever imposes the cutpoint sentence. This is apparent in Figure 5. Only those judges whose baseline sentence falls at or below  $g_1^*$  will ever find it in their interest to impose that punishment. Next, note that the judge's ideal sentence  $r_j + \alpha c$  crosses  $g_1^*$  at  $c = (g_1^* - r_j)/\alpha$ . Judges with draws of  $c$  higher than this quantity do not impose the cutpoint sentence. Given that  $c$  is uniformly distributed on  $(0, 1)$ , that quantity represents the probability that a judge of type  $r_j \leq g_1^*$  imposes  $g_1^*$ . Applying Bayes theorem, we have

$$\begin{aligned}
f_{r_j}^p(r_j|s = g_1^*) &= \frac{\Pr(s = g_1^*|r_j)f(r_j)}{\Pr(s = g_1^*)} \\
&= \frac{\int_0^{\frac{g_1^*-r_j}{\alpha}} f(c)dc f(r_j)}{\int_0^{g_1^*} \int_0^{\frac{g_1^*-r_j}{\alpha}} f(c)f(r_j)dc dr_j} \\
&= -\frac{2r_j}{g_1^{*2}} + \frac{2}{g_1^*}.
\end{aligned}$$

□

A more intuitive way of thinking about the posterior is to mentally shift the  $r_j + \alpha c$  line in Figure 5 upward toward  $g_1^*$ . As  $r_j$  increases, the probability that a judge of that type was responsible for the observed cutpoint sentences decreases linearly.

We may now plug the posterior distribution into condition (4a). Expanding integrals gives

$$\begin{aligned}
-\frac{g_1^{*2}}{6} + \frac{2r_p g_1^*}{3} - r_p^2 &= -\frac{1}{3} + r_p - r_p^2 \\
g_1^* &= 2r_p \pm \sqrt{4r_p^2 + 2 - 6r_p}.
\end{aligned}$$

(We can eliminate the larger root of the quadratic, which always exceeds one. If  $g_1^*$  exceeded one, the principal would always select judges more conservative than he.) Several features of the lower cutpoint emerge. First, it is a function solely of the principal's ideology, and not the elasticity of the sentence to culpability. Second, the quantity within the radical decreases from  $r_p = 0$  to  $r_p = 0.75$ , becoming negative at 0.5. From  $r_p = 0.5$  to one, the principal's equilibrium cutpoint therefore holds constant at one (plus an imaginary function of  $r_p$ ).

**Lemma 2.** The principal's posterior distribution on  $r_j$  given an observed sentence  $g_2^*$  is increasing right triangular on the interval  $(g_2^* - \alpha, 1)$ . This gives the pdf

$$f_{r_j}^p(r_j|s = g_2^*) = \frac{2r_j}{(1 - g_2^* + \alpha)^2} - \frac{2(g_2^* - \alpha)}{(1 - g_2^* + \alpha)^2}.$$

**Proof.** Completeness requires considering several separate conditions, but for purposes of exposition, we restrict our attention to the case where  $g_2^* > \alpha \geq 1$ . In this situation, only judges with

$r_j \in [g_2^* - \alpha, 1]$  ever impose the sentence  $g_2^*$ ; they do so when  $(g_2^* - r_j)/\alpha \leq c \leq 1$ . Using Bayes theorem as above produces the posterior on  $r_j$  given a sentence of  $g_2^*$ .  $\square$

Substituting the posterior distribution into condition (4b) and expanding integrals gives

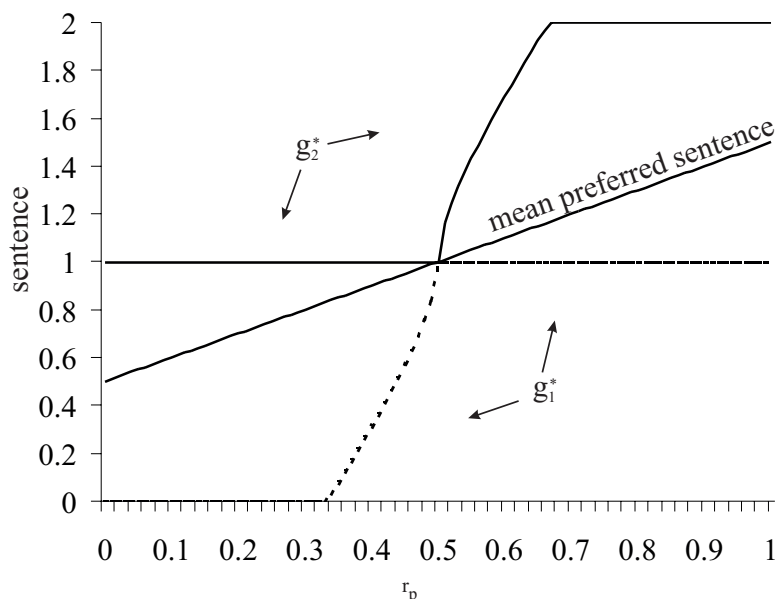
$$-\frac{g_2^{*2}}{6} + \frac{2r_p g_2^*}{3} + \frac{\alpha g_2^*}{3} - \frac{g}{3} - \frac{\alpha^2}{6} - \frac{2r_p \alpha}{3} - r_p^2 + \frac{4r_p}{3} + \frac{\alpha}{3} - \frac{1}{2} = -\frac{1}{3} + r_p - r_p^2$$

$$g_2^* = 2r_p + \alpha - 1 \pm \sqrt{4r_p^2 - 2r_p}.$$

We may eliminate the smaller root, which decreases with  $r_p$ . Unlike the lower cutpoint, the upper increases as a function of  $\alpha$ . The quantity within the radical is nonnegative for  $r_p \geq 0.5$ . For more liberal principals ( $r_p < 0.5$ ), the cutpoint “tops out” at  $1 + \alpha$ . Figure 6 displays equilibrium cutpoints given the principal’s ideology and  $\alpha = 1$ . We have constrained those cutpoints falling below zero at zero and above  $1 + \alpha$  at  $1 + \alpha$ , as the constrained cutpoints are substantively identical. Two features of the figures are noteworthy. First, the threat of ex post electoral sanction has at first glance a similar effect to imposing legislatively enacted mandatory minimum and maximum sentences: If the value of holding office is sufficiently high, it serves to bind sentences between *effective*, if not actual sentencing constraints.

Second, the degree of effective sentencing discretion decreases precipitously as the principal becomes more *moderate*. This seems counterintuitive at first: Wouldn’t a moderate principal have the most to gain by freeing the hands of his agent to take advantage of her private information? This logic, which holds in the context of the mechanism design model of the previous section, fails to hold in the context of ex post evaluation. The reason for this is surprisingly simple: When the preferences of a replacement judge exactly mirror ones own in expectation, even the slightest departure from the mean expected sentence of a judge with those preferences implies, even if ever so slightly, that the replacement would likely be an improvement. In such a case, all judges pool by imposing the same sentence, irrespective of the value of  $c$ .

**Figure 6:** Strategy for Retention: Observed Sentence Cutpoints as a Function of Principal's Preferences.



#### 4 Discussion and Conclusion

Our examination of ex ante and ex post means of controlling judges provides interesting insights into the problem of the design and operation of the criminal justice system. In this section, we discuss the relative merits of ex ante and ex post controls, both within the constraints of our modeling efforts and in more general terms. We then consider the implications of simultaneously using both methods to reduce agency loss, and speculate on several worthwhile extensions to our basic framework.

Our analysis of both types of strategies suggests that these techniques are similar in how they reduce the loss from judges with different preferences for punishment. With sentencing guidelines, voters and their representatives forgo monitoring of judges by instead constraining their sentencing decisions ahead of time. Judges can no longer sentence below the minimum or above the maximum prescribed penalty. In the signaling game, the principal chooses credible sentencing thresholds such that even though a judge would prefer to sentence below the lower threshold or above the upper threshold, she instead sentences at one or the other to retain office. Each reduces

the agency loss for a range of judges, and each setup offers particular advantages.

The largest advantage to adopting *ex ante* means of control is that the political principal can choose an upper and lower sentence boundary that he might not be able to commit to after the fact. In other words, whereas in the signaling game the voter must be able to demonstrate he will not reelect a judge who sentences below the lower threshold, this is not the case with the mandatory minimum. Mandatory sentences tie the hands of political principals ahead of time, allowing them to commit even if they would have a difficult time doing so later.

Moreover, in our signaling setup, we have only considered the case in which the value of holding office is large enough to induce every judge to sentence at the upper or lower barrier offered by the voter. If the value of office is small, then ideologically extreme judges may give up on holding office next time around in order to sentence how they see fit in the present. These are the very judges that, if they sentence solely according to their ideology, will cause the voter the greatest loss. A sentencing guideline can mitigate this loss, because all judges – especially those extreme ones – are bound by it.

Given these advantages, does it make sense that so many jurisdictions continue to rely primarily on the electoral mechanism to motivate judicial compliance with public preferences? The short answer is no. The advantages that accrue to *ex ante* controls in the first period are largely dissipated when we think about how the game would unfold over multiple periods. Consider the case in which the value of office is insufficient to motivate judges with extreme preferences to moderate their positions. When this is the case, *ex post* evaluation serves a function that *ex ante* controls do not: it is a sorting mechanism, by which the principal can categorize judges as “too liberal,” or “too conservative.” In the short run, the principal suffers the loss of extreme sentencing. In the long run, an efficient sorting mechanism allows the principal to capitalize on the discretion of judges who demonstrate their fidelity. In the steady state, only compliant judges are left holding office.

In the models outlined above, case-specific information  $c$  is unavailable to the principal. One can imagine a model, however, in which  $c$  is revealed probabilistically after a sentence is



handed down. In cases where culpability is revealed, the judge's ideology would be revealed with certainty. Under such a situation, ex post evaluation would have clear advantages over ex ante controls. Most obviously, judges could be sorted perfectly upon revelation of *c*. Consider, by contrast, the disadvantage of ex ante controls in this context. To be sure, a moderate principal, upon discovering that a relatively inculpable defendant received a harsh sentence from a strongly retributive judge, could readjust the law to remedy the perceived injustice. However, Constitutional limitations on ex post facto laws would forbid the principal from adjusting a sentence upward in the case where a judge handed down a sentence perceived as too lenient.

The foregoing analysis suggests a number of future directions. First, a more realistic model might have a principal who simultaneously sets ex ante controls on behavior *and* decides whether to retain the judge after observing sentencing behavior. In Connecticut, South Carolina, Vermont, and Virginia, the state legislature plays such a role, as its members, and not the voters, decide whether to reappoint incumbent trial judges. Ex ante controls wielded by a principal with an electoral sanctioning mechanism would necessarily be less stringent than those arrived at by one with ex ante power alone. Mandatory minimum and maximum sentences would need to be relaxed to preserve the sorting nature of the electoral mechanism. In such a situation, the principal would set the sentencing guidelines such that they would preserve sorting while still constraining judges unwilling to alter their behavior to guarantee reelection.

A second extension would be to consider the problem of multiple principals – one of whom has ex ante control, and one of whom exercises ex post evaluation authority. Consider the problem of the median legislator in a state with district-wide retention elections. The legislator's constituents want him or her to set mandatory minimum and maximum sentences to constrain judges that will eventually be evaluated not by those citizens, but by their more liberal or conservative counterparts in neighboring districts. Sharing the mechanisms of control makes the problem of oversight of trial judges a vastly more complicated matter.

A third direction is to consider the relative value of investing in ex ante review of potential judges to reduce the problem of adverse selection. In the institution-free setup, where judges can

sentence as they see fit, this is a voter's best chance for reducing agency loss. It is unclear, however, how valuable ex ante review of judges is when voters and elected officials can also turn to sentencing guidelines and ex post reviews to evaluate judicial performance. How precisely must a selection mechanism measure a judge's preferences in order to make investing in it worthwhile if the judge's decisions will be restricted by mandatory maximum and minimum sentences? Similarly, if the principal can just "throw the bastards out" at the next election if they perform poorly, what is the advantage to keeping an ideologically distant judge off the bench if she can be removed at the end of her first term? Of course, this depends on the principal's time horizon as well as the accuracy of the vetting and ex post review mechanisms.

Overall, this paper represents a first step in examining the design of institutions for the control of judges. We focus on one aspect of judicial decision making, assigning sentences, and explore the relative strengths of ex ante and ex post constraints on discretion. Given the political rhetoric in policy debates about the advantages of restrictions on sentencing and the need to evaluate incumbent judges, there is relatively little work on the institutional determinants of trial judge behavior. Our modeling effort suggests this approach has promise in helping to guide the development and assist in understanding the operation of the modern criminal justice system.

## References

- Arnold, R. Douglas. 1979. *Congress and the Bureaucracy: A Theory of Influence*. New Haven: Yale University Press.
- Aspin, Larry. 1999. "Trends in Judicial Retention Elections, 1964-1998." *Judicature* 83, 79-81.
- Austen-Smith, David. 1993. "Explaining the Vote: Constituency Constraints on Sophisticated Voting." *American Journal of Political Science* 36, 68-95.
- Canes-Wrone, Brandice, Michael C. Herron, and Kenneth W. Schotts. 2001. "Leadership and Pandering: A Theory of Executive Policymaking." *American Journal of Political Science* 45, 532-550.
- Epstein, David, And Sharyn O'Halloran. 1994. "Administrative Procedures, Information, and Agency Discretion." *American Journal of Political Science* 38, 697-722.
- Epstein, David, And Sharyn O'Halloran. 1996. "Divided Government and the Design of Administrative Procedures: A Formal model and Empirical Test." *Journal of Politics* 58, 373-397.
- Epstein, David, And Sharyn O'Halloran. 1999. *Delegating Powers: A Transaction Cost Politics Approach to Policy Making under Separate Powers*. New York: Cambridge University Press..
- Fenno, Richard F. 1966. *The Power of the Purse: Appropriations Politics in Congress*. Boston: Little Brown.
- Gordon, Sanford C. 2000. "The Allocation of Bureaucratic Resources: A Stochastic Process Model of Regulatory Targeting." Typescript.
- Hamilton, James T., and Chris Schroeder. 1994. "Strategic Regulators and the Choice of Rule-making Procedures: The Selection of Formal vs. Informal Rules in Regulating Hazardous Waste." *Law and Contemporary Problems* 57, 111-160.
- Huber, Gregory. 2001. "Interests and Influences: Explaining Patterns of Enforcement in Government Regulation of Occupational Safety." Ph.D. dissertation, Princeton University.
- Huber, John D., and Charles R. Shipan. 2001. Political Control of the State in Modern Democracies. Unpublished book manuscript.
- Huber, John D., Charles R. Shipan, and Madelaine Pfahler. 2001. "Legislatures and Statutory Control of Bureaucracy." *American Journal of Political Science* 45, 330-345.
- Kiewiet, D. Roderick, and Mathew D. McCubbins 1991. *The Logic of Delegation: Congressional Parties and the Appropriations Process*. Chicago: University of Chicago Press.
- Lowi, Theodore J. 1979. *The End of Liberalism, 2nd edition*. New York: Norton.
- Milgrom, Paul, and John Roberts. 1992. *Economics, Organization, and Management*. Englewood Cliffs, NJ: Prentice Hall.
- Moe, Terry M. 1984. "The New Economics of Organization." *American Journal of Political Sci-*

ence 28, 739-777.

Niskanen, William A., Jr. 1971. *Bureaucracy and Representative Government*. Chicago: Aldine-Atherton.

Spence, A. Michael. 1974. *Market Signaling: Informational Transfer in Hiring and Related Screening Processes*. Cambridge: Harvard University Press.

Volcansek, Mary L., Maria Elisabetta de Franciscis, and Jacqueline Lucienne Lafon. *Judicial Misconduct: A Cross-National Comparison*. Gainesville: University of Florida Press.

Warr, Mark. 1995. "Poll Trends: Public Opinion on Crime and Punishment." *Public Opinion Quarterly* 59, 296-310.

Whitford, Andrew B., and Eric Helland. 1998. "If I Had a Hammer: The Legislative Design of Superfund." Paper presented at the annual meeting of the Midwest Political Science Association.