

**Precedent – Use It or Lose It?:  
An Informational Model of Judicial Decision-Making\***

by

Ethan Bueno de Mesquita and Matthew Stephenson

Department of Government, Harvard University

---

\*We are indebted to Scott Ashworth, Bruce Bueno de Mesquita, Eric Dickson, Gilles Serra, Ken Shepsle, and the participants in the Harvard Rational Choice Lunch Group, the Harvard Political Economy Research Workshop, and the Texas A&M Conference on Institutional Games and the U.S. Supreme Court for valuable comments. Financial support from the John M. Olin Center for Law, Economics, and Business is gratefully acknowledged. Comments may be sent to [mesquita@fas.harvard.edu](mailto:mesquita@fas.harvard.edu).

# Precedent – Use It or Lose It?:

An Informational Model of Judicial Decision-Making

## **Abstract**

We develop an informational model of judicial decision-making in which deference to precedent is useful to policy-oriented appellate judges because it improves the accuracy with which they can communicate legal rules to trial judges. Our simple model yields a host of interesting implications and hypotheses regarding conditions under which judges will maintain or break with precedent, the constraining effect that precedent has on judicial decision-making, the voting behavior of Supreme Court Justices, the relationship between a precedent's age and its authority, the effect of legal complexity on the level of deference to precedent, the relative stability of rules and standards, and long-term patterns of legal evolution. Perhaps most importantly, we demonstrate that “legalist” features of judicial decision-making are consistent with an assumption of policy-oriented judges.

## 1 Introduction

One of the most perplexing and controversial problems in the social-scientific study of judicial behavior is the debate over the relative importance of legal and policy concerns in judicial decisions. Scholars subscribing to policy-oriented models of judicial behavior argue that judges are concerned with the external effects of their rulings on allocations of risk, wealth, power, or opportunity.<sup>1</sup> Judges may be motivated by policy concerns because of partisan loyalties, a sincere desire to effect particular changes in the world, or pursuit of promotion or re-election. Whatever the reason, the policy-oriented judge cares about actual judicial “outputs” more than any particular method of arriving at those outputs. On the other hand, some scholars argue that judges are concerned with “legalism”, that is with correctly following the rules and norms of proper judicial reasoning. A legalist judge maximizes her utility by adhering faithfully to these internal rules, regardless of the external result. The purely policy-oriented judge and the purely legalist judge are of course ideal types, and few scholars, if any, believe that judges are motivated solely by concern with external effects or by fidelity to internal norms. Nonetheless, these two sets of factors are often presented as competing explanations for judicial behavior, and their relative importance is the

---

<sup>1</sup>Both the “attitudinal model” (Segal and Spaeth 1993 ) and the “strategic model” (Knight and Epstein 1998) of judicial behavior are “policy-oriented”, in that both models assume judges are primarily concerned with substantive outcomes.

subject of heated debate (Posner 1993, Segal and Spaeth 1996a, Segal and Spaeth 1996b, Brisbin 1996, Knight and Epstein 1996, Songer and Lindquist 1996, Revesz 1997, Cross and Tiller 1998, Edwards 1998, Revesz 1999).

We contribute to the discussion of the nature of judicial decision-making by demonstrating that “legalist” principles are, at least in some cases, compatible with – and in fact explained by – judges’ concern with the external policy effects of their rulings. Focusing on one central principle of judicial decision-making in Anglo-American legal systems – *stare decisis*, or deference to precedent – we show that purely policy-oriented judges will often defer to legal precedent, even when doing so requires them to issue decisions that deviate from the holding they otherwise would prefer. The reason for this is that appellate judges can use prior cases to increase the accuracy of their communication with trial judges. Often, a judge is willing to modify her substantive ruling in order to purchase this increased accuracy in communication. Thus, in the model we present, it is not the case that policy-oriented judges ignore precedent, nor is it the case that judges care about precedent instead of, or in addition to, caring about policy. Rather, judges care about precedent *because* they care about policy.

In addition to demonstrating this basic idea, our model has the advantage of being able to account systematically for both adherence to and departure from precedent. Specifically, our comparative static analysis shows how variation in four parameters – the distance between the existing legal rule and the deciding judge’s ideal, the

age of the existing precedent, the difficulty of intelligibly integrating existing precedent with new rulings, and the precision or imprecision of communication between judges – affects the relative likelihoods that a judge will adhere to or break with existing legal precedent. Our model also has implications for how much judges are able to change the substantive law without ever openly breaking from established precedent. Further, our informational model of *stare decisis* sheds light on other important empirical puzzles in the study of judicial decision-making. We reconcile the seemingly contradictory observations that arguments from precedent play a major role in Supreme Court deliberation and adjudication, and yet Justices consistently vote their preferences rather than following established precedents. Our approach also offers an explanation for why long lines of cases might evoke both deference and skepticism. Finally, our model yields several novel hypotheses, including predictions regarding the types of legal issues in which long lines of precedent will emerge, the relative stability of rules versus standards, and patterns of long-term legal evolution.

The paper proceeds as follows. Section 2 outlines the scholarly debate on the nature of judicial decision making and makes more explicit our contribution to this debate. Section 3 sets out the structure of our formal model, while Section 4 analyzes the appellate judge’s choice between following and breaking with established precedent. In Section 5 we perform comparative static analysis on the parameters of the model to show how they affect the appellate judge’s decision. In this section we also

examine the degree to which judges can change the law without ever publicly breaking from established precedent. In Section 6, we discuss our main results, showing how our model can account for seemingly contradictory empirical observations and developing hypotheses that could be subjected to future empirical testing and which have implications for important debates in the legal literature.

## 2 Competing Theories of Judicial Decision-Making

We focus on the principle of *stare decisis*,<sup>2</sup> which dictates that judges ought to apply rules and principles laid down in prior cases because it is, or at least is claimed to be, one of the most important principles of judicial decision-making in the Anglo-American common law system. Judges are not, under ordinary circumstances, supposed to overturn “settled law” (Nelson 2001).<sup>3</sup> Our focus on *stare decisis* is also due to its salience for the more general question of the nature of judicial preferences. Judicial deference to established precedent is a focus of the debate between the policy-oriented and legalist models of judicial decision-making because these models seem

---

<sup>2</sup>*Stare decisis et non quieta movere* – “stand by the thing decided and do not disturb the calm.”

<sup>3</sup>One aspect of this principle is the idea that lower courts are supposed to follow the precedents set by higher courts (vertical *stare decisis*). This is not much different from the principle in many hierarchical organizations that subordinate units are supposed to follow the directions and guidelines laid down by their superiors. A more interesting aspect of the *stare decisis* principle – and one more unique to judicial decision-making – is the principle that courts are supposed to follow their own prior decisions (horizontal *stare decisis*).

to offer such different predictions for how important precedent will be in practice. A legalist judge is expected to place great weight on the *stare decisis* principle and consequently is expected to defer to prior decisions even when the judge herself would have decided the precedent-setting case differently.<sup>4</sup>

By contrast, it is not clear why a policy-oriented judge would ever defer to constraints imposed by prior decisions. It may be that judges often agree with the principles laid down in old cases and follow them for that reason, but if this is the case then *stare decisis* is merely a description of – rather than a reason for – patterns of judicial decision-making. To the extent that a legal precedent exerts a causal influence on at least some judicial decisions, it must be the case that in these decisions judges would prefer to issue a different decision if the precedent did not exist. Thus, a policy-oriented judge would be expected to attach little importance to established precedent when making her decisions.

Scholars have offered a number of reasons why even a policy-oriented judge might want to respect established legal precedent. One argument – the one most familiar to lawyers – is that stability in the law is in itself a valued policy goal, and judges would therefore be willing to defer to an established legal rule because the act of deference

---

<sup>4</sup>This might not always be the case since there might be some other “legal” principle, e.g. the proper interpretation of a statute, that could trump the principle of deference to decided cases even for a legalist judge. Nonetheless, as a general rule in most cases it is safe to assume that legalist judges in a common law system would attach substantial weight to precedent.

itself advances their policy preference for stability. But this explanation has a difficult time accounting for adherence to precedent in areas of law where stability and the need for long-term planning are less salient. Perhaps more importantly, stability in the law is a collective good; for a judge to sacrifice other policy goals for the sake of stability, she must believe that other judges will also value stability sufficiently highly that they will not overturn precedent. But, if she believes that other judges do place a high value on stability, she herself may be tempted to break with precedent and establish a new legal rule, since it will be respected by future courts with little overall loss in legal stability. Additionally, the value of legal stability is considerably reduced to a judge if the stable legal rule is objectionable (Kornhauser 1989). While the need for stability may be enough when the judge is indifferent between legal rules, it is a less plausible basis for deference to precedent when the judge has strong substantive preferences between rules.

Another suggestion as to why policy-oriented judges might respect the principle of *stare decisis* is that judges want their own precedents followed, and therefore follow precedents set by others (Landes and Posner 1976; Rasmusen 1994). Because judges prefer a world in which all precedents, including their own, are respected to a world in which no precedents are respected, they will attempt to enforce universal respect for precedent by punishing judges who “defect”.<sup>5</sup> While there may be some of this tit-for-

---

<sup>5</sup>If there is not specific, targeted retaliation against particular judges for failing to follow prece-

that in the real world, this explanation has several problems as a general account of the practice of *stare decisis*. First, it relies on the empirically dubious assumption that judges look to other judges' respect for legal precedent when determining whether to follow precedent set by those judges. Second, it has trouble accounting for why the precedents of retired judges are ever followed without reference to even more complicated, and empirically problematic, punishment mechanisms. Third, in the absence of such mechanisms, this explanation also does not explain well why some judges break with some precedents but follow others. In a simple model where judges expect general retaliation for any break with existing precedent, if they break once they have no reason not to break always.

A third possibility is that policy-oriented judges do not care about precedent *per se*, but recognize the need to preserve their institutional power and legitimacy. Because this legitimacy derives in part from a public belief that judges apply a specialized set of legal skills, including the ability to interpret and apply established legal precedents, as well as a belief that judges are politically-neutral interpreters of law, policy-oriented judges will want to perpetuate the belief that they make decisions according to precedent (Cox 1976, Maltz 1980). This may in turn lead them to modify their decisions sometimes to show more respect for precedent, in order to dent, the general judicial interest in precedent-following alone is insufficient, because respect for precedents is a collective good and individual judges' dominant strategy would be to free-ride (Macey 1989).

enhance their institutional power. However, the collective action objection applies to this explanation as well; the overall effect of any particular decision on institutional legitimacy is likely to be small, while the policy ramifications of the case itself often are large. Moreover, the effect of a decision on public perception of the court's legitimacy may have more to do with the content of the decision than whether it involved overturning a precedent (Hyde 1983, Rehnquist 1986, Nelson 2001).

Finally, several scholars have pointed out that deference to precedent may be valuable even to policy-oriented judges because of the valuable informational function that judicial precedents serve (Shapiro 1972, Rehnquist 1986, Kornhauser 1989). The informational perspective comprises two types of explanations for deference to precedent. First, reasoning from precedents may improve communication between appeals courts, allowing for judicial specialization and error-correction (Shapiro 1972, Kornhauser 1989, Macey 1989). A more pessimistic version of the same basic argument is that the practice of *stare decisis* is essentially an "information cascade" in which rational agents ignore their own information and imitate the behavior of preceding decision-makers, often leading to uncorrected inefficient results (Talley 1999).

We suggest a second informational function which the doctrine of *stare decisis* might serve. This function involves communication between high courts and lower courts. The basic idea is that the development of lines of cases can communicate a legal principle better than any individual case could. An initial case may invoke a

general phrase or principle, such as “due process”, “reasonable”, “compelling interest”, and the like; future cases develop and give meaning to these inherently vague phrases. Hence, a lower court can learn more about the appellate court’s view of the proper interpretation of, say, a due process balancing test by examining a line of ten cases in which the same test was applied than by reading the first (or the last) ruling the appeals court issued. Similarly, an initial case might declare a bright-line rule that, though clear, is both over- and under-inclusive.<sup>6</sup> Further cases can carve out exceptions and make qualifications so that the line of cases applying the rule offers lower courts a much more nuanced test than that announced by the original decision.

Our model explores this second type of informational use of legal precedent, a use that we believe has been neglected in the scholarly discussion of judicial decision-making, and which has not been formally developed or rigorously analyzed. In so doing, we provide an account of judicial decision-making that explains why policy-oriented judges are expected to be influenced by precedent, while also capturing cases where judges break from precedent.

---

<sup>6</sup>A bright line rule is a rule that minimizes ambiguity by setting well-defined and simple categories of prohibited and permissible behavior.

### 3 The Model

Consider a simple judicial system consisting of an appellate judge and a trial judge. The appellate judge hears a limited number of cases, and, through her decisions in these cases, she can announce how the law ought to be interpreted. The trial judge decides the vast majority of cases, and in making his decisions he attempts to apply the law established by the appellate judge as best he can.<sup>7</sup> However, the trial judge’s understanding of the appellate judge’s rulings is imperfect, and he often does not do exactly as the appellate judge would have done if she were deciding the case. The appellate judge cannot correct all these “mistakes” by altering or reversing the trial court’s decision on appeal. This may be because the appellate court’s jurisdiction is discretionary and many appeals are not heard, or because many cases are not appealed at all. Thus, the appellate court judge, in order to successfully influence the application of the law to the majority of cases, needs to communicate her preferred interpretation of the law to the trial court as accurately as possible, subject to time and resource constraints.

The appellate judge has preferences over the legal rule defined on a unidimensional continuum.<sup>8</sup> Denote the appellate judge’s ideal point  $j \in \mathbb{R}$ . The policy continuum

---

<sup>7</sup>There are many possible substantive defenses of this assumption that could ultimately be modelled explicitly. These include promotion being dependent on faithful enactment of established law or explicit selection of trial judges inclined towards respect for established law.

<sup>8</sup>We will refer to the judge’s preferences over this continuum as her policy preferences; however,

might represent, for example, the level of care exercised by a defendant in a civil action, with  $j$  representing the minimum amount of care the appellate judge believes a defendant must exercise to avoid liability. That is, the appellate judge in this example believes that all defendants who exercise a level of care less than  $j$  ought to be liable, and all defendants who exercise a level of care above  $j$  ought to escape liability. As such,  $j$  is the appellate judge's ideal legal rule. The trial judge has no preference over the policy dimension per se; rather, he tries to implement whatever legal rule he believes has been announced by the appellate judge.

Every time the appellate judge decides a case, she issues a ruling with two components. First, she announces a substantive holding on the proper application of the legal rule to the case at hand. This substantive holding corresponds to a point in the policy space, denoted by  $r \in \mathbb{R}$ . Second, the appellate judge declares whether her substantive holding is consistent with existing precedent or whether she is breaking with precedent. "Precedent" here means the line of appellate cases on the relevant legal issue that have been decided prior to the present case and that have never been overruled by a subsequent appellate case. If the appellate judge declares that her ruling is consistent with precedent, her case is added to the relevant line of cases, and the trial judge will interpret her substantive holding in the context of the other we use policy broadly, to reflect public-policy preferences, normative judgements regarding fairness or justice, or a weighted combination of various factors.

substantive holdings in the line in order to ascertain the legal rule the appellate judge wishes to enact. If, on the other hand, the appellate judge declares that she is breaking with precedent, then a new line of cases is established, and the trial judge will treat the most recent appellate decision as the exclusive statement of the legal rule. The number of cases in the line of precedent, including the decision being made in the current round, is denoted by  $t$ . Thus a line of precedent is a series of substantive rulings  $(r_1, r_2, \dots, r_{t-1})$ , where  $r_{t-1}$  is the most recent ruling and the current judge issues ruling  $r_t$ .

The appellate judge's announcement of whether she is maintaining or breaking from precedent is communicated perfectly to the trial judge. However, the communication of the substantive holding of each ruling is inherently noisy. Thus, when a trial judge attempts to understand an appellate ruling, he observes  $t$  signals, each drawn from a normal distribution with variance  $\sigma^2$ . The means of these distributions are the  $r$ 's associated with the substantive holdings of the different cases in the line of precedent. The trial court judge averages these signals to form his understanding of the legal rule.<sup>9</sup> This is equivalent to the trial court judge observing a single signal,

---

<sup>9</sup>The trial judge could in theory aggregate his signals in some other way. For example, he might take a weighted average of the signals, giving the greatest weight to the most recent ruling. However, a different aggregation technique would in turn change the optimal decision of the appellate judge. As it is not clear what aggregation mechanism would be most efficient, we defer that question to future research and assume for simplicity that trial judges average their signals.

denoted by  $x$ , drawn from a normal distribution with mean  $\mu_t = \frac{\sum_{i=1}^t r_i}{t}$  and variance  $\frac{\sigma^2}{t}$ . If the appellate judge breaks with precedent, there is only one decision in the line of cases, so  $t = 1$  and the trial judge receives a signal drawn from a normal distribution with mean  $r_{t=1}$ , the most recent appellate holding, and variance  $\sigma^2$ . The mean of the distribution of the signal in the period prior to the current appellate judge's decision is denoted  $\mu_{t-1}$ , and, we refer to this value as “existing precedent”. Because  $x$  (the value of the draw from the distribution that the trial court observes) is the only information that the trial judge has, it is his best guess as to what the appellate judge wants. Consequently he will treat  $x$  as the controlling legal rule when making his decisions.

The appellate judge's utility has two components. First, she would like the decisions of the trial court judges to be as close as possible to her ideal point. That is, she wishes to minimize  $|x - j|$ . The reason for this is that any cases that fall in this interval are cases that the trial judge will get “wrong” from the appellate judge's point of view. Again consider a case where the policy dimension represents the level of care taken by a civil defendant. If the defendant exercised a level of care less than the minimum of  $x$  and  $j$ , the trial court will correctly find the defendant liable. Similarly, if the defendant exercised a level of care above the maximum of  $x$  and  $j$ , the trial court will correctly find the defendant not liable. However, when the defendant exercised a level of care in the interval between  $x$  and  $j$ , the trial court will

rule incorrectly. In this event, if  $x > j$ , the trial court will mistakenly find liability, while if  $x < j$ , the trial court will mistakenly find no liability. The larger the size of the interval,  $|x - j|$ , the larger the number of cases that will be decided incorrectly. We assume that the appellate judge's utility function is quadratic and so consider the square of the expected distance between  $x$  and  $j$ , which is equal to  $\left((\mu_t - j)^2 + \frac{\sigma^2}{t}\right)$ .

Second, if the appellate judge declares that her decision is consistent with precedent, it is costly to offer a ruling,  $r_t$ , that is substantively different from the existing precedent,  $\mu_{t-1}$ . We are agnostic as to the relative importance of these two factors in her utility calculation; we weight the importance of the latter by the parameter  $\alpha \geq 0$  on which we perform comparative static analysis.<sup>10</sup> Thus, the utility cost associated with changing the law while claiming to adhere to precedent is  $\alpha (\mu_{t-1} - r_t)^2$ . If the appellate judge breaks with precedent, this cost does not apply.

This cost arises because writing an opinion that intelligibly integrates existing precedent with a change in the substance of the legal rule becomes increasingly difficult as the distance between precedent and the substantive holding grows. Thus, there are real costs in intellectual effort and research associated with such a decision. If these costs are not invested, that is if the appellate judge were to declare that her decision were consistent with precedent without explaining how the substance of her

---

<sup>10</sup>If  $\alpha > 1$  this implies that the cost of writing the decision is weighted more heavily than how close the trial court's decision is to the appellate court's ideal point. If  $0 < \alpha < 1$  the closeness of the trial court decision to the appellate judge's ideal point is weighted more heavily.

ruling and prior rulings could be intelligibly integrated, then the trial judge would be unable to make decisions with reference to the full line of cases.

It may at first seem implausible to assume that judges bear this cost when maintaining precedent but do not bear a cost when breaking with precedent. There are two reasons for this assumption. First, our analytic results are unchanged by the addition of a fixed cost for breaking with precedent. Such a cost would change the actual point at which appellate judges break, but would not affect the comparative statics. Second, any cost from breaking with precedent, beyond the loss of information already modelled, would arise from legalist values. We omit legalist values from the judicial utility function in order to see whether the norm of *stare decisis* is consistent even with purely policy-oriented judges.

The expected utility to the appellate judge is therefore:

$$EU = - \left( (\mu_t - j)^2 + \frac{\sigma^2}{t} \right) - \tilde{\alpha} (\mu_{t-1} - r_t)^2$$

where  $\tilde{\alpha} = \alpha$  if the judge maintains precedent and  $\tilde{\alpha} = 0$  if the judge departs from precedent.

#### 4 The Analysis

The appellate judge has two choices to make. She must select her substantive holding ( $r_t$ ) and she must decide whether to break with precedent or to maintain precedent. These two decisions are interrelated. In order to determine whether or not to break

with precedent, the judge must first know the expected outcomes, which can only be determined if the judge already knows the content of the ruling she will give either if she breaks with precedent or maintains precedent. Thus, we work backwards, first determining the appellate judge's optimal substantive holding contingent on whether she has broken with or maintained precedent.

If the judge breaks with precedent the second term of her expected utility function is zero, because  $\tilde{\alpha} = 0$ . Further, because she has broken with existing precedent,  $t$  reverts to 1. Therefore, the line of cases considered by the trial judge contains only this appellate judge's most recent decision, that is,  $\mu_{t=1} = r_{t=1}$ . It is obvious that the appellate judge's expected utility in this case is maximized when  $r_t = \mu_t = j$ . Thus, if a judge breaks with precedent, her substantive holding will be her ideal legal rule, and her expected utility  $EU(break)$  will be  $-\sigma^2$ .

If the appellate judge does not break with precedent, her decision is more complicated. She would like to move precedent towards her ideal point, but she is constrained by the cost of deviating too much from the line of precedent of which her decision becomes a part. The judge thus chooses  $r_t$  by solving the following maximization problem:

$$\max_{r_t} \left\{ - \left[ (\mu_t - j)^2 + \frac{\sigma^2}{t} \right] - \alpha (\mu_{t-1} - r_t)^2 \right\} \quad (1)$$

Note that  $\mu_t$ , the mean of the distribution from which the trial judge will draw his understanding of the legal rule, is a function of  $r_t$ . In particular  $\mu_t = \frac{(t-1)\mu_{t-1} + r_t}{t}$ .

Thus, we can rewrite the maximization problem as:

$$\max_{r_t} \left\{ - \left[ \left( \frac{(t-1)\mu_{t-1} + r_t}{t} - j \right)^2 + \frac{\sigma^2}{t} \right] - \alpha (\mu_{t-1} - r_t)^2 \right\} \quad (2)$$

Solving for the first order condition yields:

$$r_t^* = \mu_{t-1} + \frac{t(j - \mu_{t-1})}{1 + \alpha t^2}$$

Knowing the optimal substantive holding,  $r_t^*$ , allows us to calculate the mean of the new distribution from which the trial court judge will draw his understanding of the proper legal rule. This new mean,  $\mu_t^*$ , is given by:

$$\mu_t^* = \frac{(t-1)\mu_{t-1} + r_t}{t} = \frac{j + \alpha t^2 \mu_{t-1}}{1 + \alpha t^2} \quad (3)$$

Now, in order to find the appellate judge's expected utility from issuing a decision consistent with precedent,  $EU(\textit{maintain})$ , we substitute  $r_t^*$  and  $\mu_t^*$  into the expected utility function. This yields:

$$\begin{aligned} EU(\textit{maintain}) &= - \left[ \left( \frac{j + \alpha t^2 \mu_{t-1}}{1 + \alpha t^2} - j \right)^2 + \frac{\sigma^2}{t} \right] \\ &\quad - \alpha \left[ \mu_{t-1} - \left( \mu_{t-1} + \frac{t(j - \mu_{t-1})}{1 + \alpha t^2} \right) \right]^2 \end{aligned} \quad (4)$$

which simplifies to:

$$EU(\textit{maintain}) = - \frac{\alpha t^2}{1 + \alpha t^2} (j - \mu_{t-1})^2 - \frac{\sigma^2}{t} \quad (5)$$

In our model the central benefit to appellate judges of maintaining precedent is informational. The more cases the trial judge has to refer to, the more accurately he

will understand the legal rule for which that line of cases stands. Thus, the appellate judge wishes to maintain precedent because it makes her communication with trial courts less noisy. However, the use of precedent comes at a price. In particular, judges bear a cost for deviating substantively too far from the line of precedent which they claim to uphold. This constrains judges who are maintaining precedent from implementing a legal rule which matches their personal ideal, as can be seen in the model. When precedent is maintained,  $\mu_t^* = \frac{j + \alpha t^2 \mu_{t-1}}{1 + \alpha t^2} \neq j$  (unless  $\alpha = 0$  or  $\mu_{t-1} = j$ ). The rule applied by the trial court when precedent is maintained is biased away from the current appellate judge's ideal point ( $j$ ) in the direction of the old precedent ( $\mu_{t-1}$ ).

On the other hand, while the appellate judge can make sure that the mean of the trial court's signal is equal to her ideal point if she breaks with precedent, the variance of this signal, i.e. the noisiness of communication, will be higher. Again, this can be seen clearly in the model, since  $\sigma^2 > \frac{\sigma^2}{t}$  as long as  $t > 1$ . The appellate judge thus faces a trade-off between the accuracy with which she communicates the legal rule to the trial courts and the proximity of that rule to her ideal. She determines whether or not to break with precedent by comparing the expected utilities associated with each choice; she breaks with precedent if and only if:

$$\begin{aligned}
 EU(\textit{break}) - EU(\textit{maintain}) &= \sigma^2 \frac{1-t}{t} + \frac{\alpha t^2}{1 + \alpha t^2} (j - \mu_{t-1})^2 > 0 \iff \\
 \frac{\alpha t^3}{(t-1)(1 + \alpha t^2)} (j - \mu_{t-1})^2 - \sigma^2 &> 0
 \end{aligned} \tag{6}$$

## 5 Results

### 5.1 Comparative Statics

Comparative static analysis on the parameters of this model yields a number of interesting results regarding how these parameters affect the relative desirability of maintaining or breaking with existing precedent.

It is clear from equation (6) that the desirability of breaking with precedent decreases as  $\sigma^2$ , the noisiness of each individual signal, increases. The intuition is that as communication between appellate and trial courts becomes less precise, the extra information provided by situating a decision in a line of precedent becomes more valuable to the appellate judge. On the other hand, equation (6) implies that increasing  $|j - \mu_{t-1}|$ , the distance between the appellate judge's ideal point and existing precedent, increases the attractiveness of breaking with precedent. Because a large distance between existing precedent and the appellate judge's ideal point constrains how close she can move the expected trial court decision to her ideal point, she is less willing to trade-off control over the substantive rule for increased accuracy of transmission.

Equation (6) also shows that breaking with precedent becomes more attractive as  $\alpha$  increases. The reason for this is that, when  $\alpha$  is close to 0, the appellate judge can move  $\mu_t$  very close to her ideal point,  $j$ , even when she maintains precedent. Thus,

the informational benefit of situating her decision in a long line of cases comes at very little cost in terms of substance. However, as  $\alpha$  grows, the appellate judge's ability to move the legal rule close to her ideal point becomes more constrained, making adherence to existing precedent less attractive. One can see this by examining equation (3). Assuming that precedent is maintained, the distance that the legal rule will be moved is given by  $|\mu_t^* - \mu_{t-1}| = \left| \frac{j - \mu_{t-1}}{1 + \alpha t^2} \right|$  which is clearly decreasing in  $\alpha$ .

Deriving the comparative statics on  $t$ , the number of cases in the line of precedent, is more complicated.  $t$  is a discrete variable, so rather than using differential calculus, we take the first differences of equation (6) with respect to  $t$ :

$$\begin{aligned} & [EU_{t+1}(\text{break}) - EU_{t+1}(\text{maintain})] - [EU_t(\text{break}) - EU_t(\text{maintain})] \equiv FD(t) \\ = & -(j - \mu_{t-1})^2 \alpha \frac{\alpha t^4 + 2\alpha t^3 + \alpha t^2 - 2t^3 + 2t + 1}{(1 + \alpha t^2 + 2\alpha t + \alpha)(1 + \alpha t^2)t(t-1)} \end{aligned} \quad (7)$$

A little algebra demonstrates that this first-difference can be positive or negative, depending on  $\alpha$  and  $t$ . In particular

$$\begin{aligned} FD(t) & < 0 \text{ if } \frac{2t^3 - 2t - 1}{t^4 + 2t^3 + t^2} \equiv \bar{\alpha}(t) < \alpha \\ FD(t) & > 0 \text{ if } \frac{2t^3 - 2t - 1}{t^4 + 2t^3 + t^2} \equiv \bar{\alpha}(t) > \alpha \end{aligned} \quad (8)$$

Increasing  $t$  decreases the desirability of breaking with precedent when  $\alpha > \bar{\alpha}(t)$ ; otherwise increasing  $t$  increases the desirability of breaking with precedent. It is important to note that this threshold,  $\bar{\alpha}(t) = \frac{2t^3 - 2t - 1}{t^4 + 2t^3 + t^2}$ , is itself a decreasing function

of  $t$ . There are two cases to consider in understanding this result. The first is when  $\alpha > \bar{\alpha}(2)$  (the lowest possible values of  $\bar{\alpha}(t)$  when the appellate judge chooses to maintain existing precedent). In this case, for all values of  $t$ , the attractiveness of breaking with precedent is decreasing as  $t$  increases. That is, for a sufficiently large  $\alpha$ , older precedents are always less vulnerable than younger precedents. This case is illustrated in figure (1). If, however,  $\alpha < \bar{\alpha}(2)$ , then increases in  $t$  increase the attractiveness of breaking with precedent for a certain number of periods. Specifically, increases in  $t$  will increase the likelihood of breaking as long as  $\alpha$  is below the threshold value of  $\bar{\alpha}(t)$ . But, as  $t$  increases, this threshold value decreases, meaning that  $\alpha$  will eventually be greater than the threshold. At that point, the effect of increasing  $t$  switches so that increases in  $t$  decrease the desirability of breaking with precedent. This case is illustrated in figure (2).

In order to understand the intuition behind these comparative statics, recall that an increase in  $t$  has two effects on the desirability to the judge of maintaining the line of precedent. On the one hand, an increase in  $t$  increases the constraint on how much the judge can move the substantive legal rule, making older precedents less attractive. On the other hand, an increase in  $t$  improves the accuracy with which the legal rule is communicated to the trial court, making older precedents more attractive. When  $\alpha$

Figure 1:  $EU(\text{Break})-EU(\text{Maintain})$  with  $\alpha > \bar{\alpha}(2)$

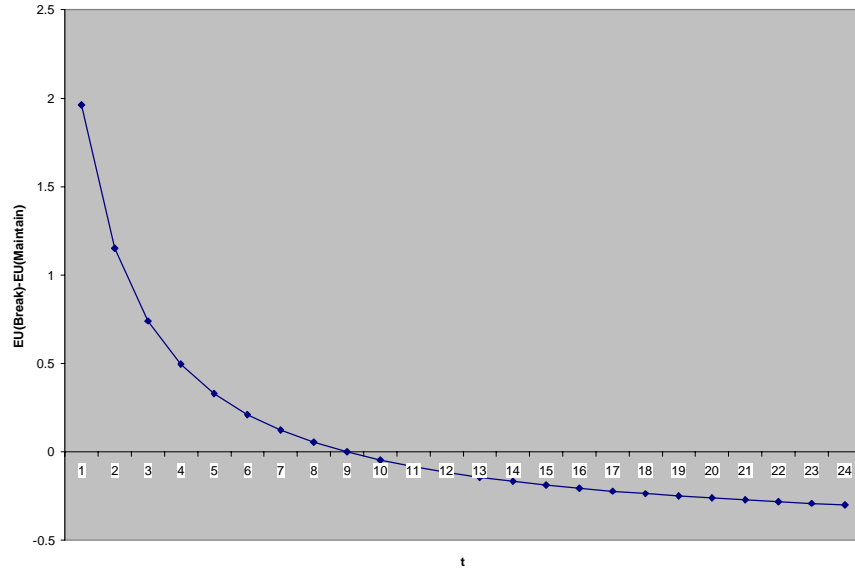
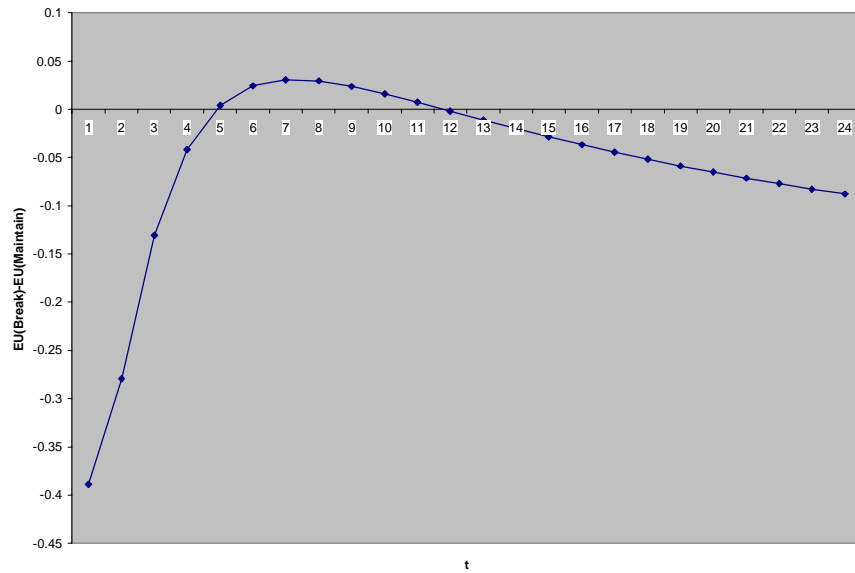


Figure 2:  $EU(\text{Break})-EU(\text{Maintain})$  with  $\alpha < \bar{\alpha}(2)$



is sufficiently high, the information effect *always* overwhelms the constraining effect. The reason for this is that when  $\alpha$  is high, the judge's ability to move the legal rule is already so constrained that the marginal effect of an increase in  $t$  on this constraint is negligible. However, when  $\alpha$  and  $t$  are sufficiently low, the constraining effect of an increase in  $t$  is more important to the judge than the effect on information, thereby making an increase in  $t$  increase the attractiveness of breaking with precedent. As  $t$  increases further, the information effect of marginal increases in  $t$  will eventually overwhelm the constraining effects; no matter how low  $\alpha$  is, when  $t$  is sufficiently high the judge is already so constrained that the marginal constraining effect of an increase in  $t$  is again negligible.

## 5.2 Limits of Legal Change

An interesting question in the context of our model is, how far on the substance dimension can a legal rule move without a judge ever claiming to have broken from precedent? Is this distance bounded or unbounded? That is, how far would precedent move if, over an infinite series of turns, each appellate judge's ideal point was such that she was exactly indifferent between moving existing precedent and breaking, leading these judges to move  $\mu_t$  as far from  $\mu_{t-1}$  as is possible (always in the same direction) without declaring a break from existing precedent? We can find the maximum distance the legal rule can move under these conditions (i.e.,  $\max |\mu_\infty - \mu_1|$ ) by

solving the following constrained infinite sum:

$$\sum_{t=2}^{\infty} \{\mu_t^* - \mu_{t-1}\} = \sum_{t=2}^{\infty} \left\{ \frac{j + \alpha t^2 \mu_{t-1}}{1 + \alpha t^2} - \mu_{t-1} \right\} = \sum_{t=2}^{\infty} \left\{ \frac{j - \mu_{t-1}}{1 + \alpha t^2} \right\}$$

$$s.t. \frac{(j - \mu_{t-1})^2 \alpha t^3}{(1 + \alpha t^2)(t-1)} - \sigma^2 = 0 \quad (9)$$

By making the constraint an equality, we impose the condition that in each round the judge has preferences that lead her to move the legal rule as far as possible without preferring to break with precedent. Solving the constraint for  $(j - \mu_{t-1})$  and substituting, we find that the sum is equal to:

$$\frac{\sigma}{\sqrt{\alpha}} \sum_{t=2}^{\infty} \frac{1}{t} \frac{\sqrt{t-1}}{\sqrt{t}\sqrt{1+\alpha t^2}} \quad (10)$$

Notice that if  $\alpha \neq 0$  this can be bounded above as follows:

$$\frac{\sigma}{\sqrt{\alpha}} \sum_{t=2}^{\infty} \frac{1}{t} \frac{\sqrt{t-1}}{\sqrt{t}\sqrt{1+\alpha t^2}} < \frac{\sigma}{\sqrt{\alpha}} \sum_{t=2}^{\infty} \frac{1}{t} \frac{1}{\sqrt{\alpha t^2}} = \frac{\sigma}{\alpha} \sum_{t=2}^{\infty} \frac{1}{t^2} = \left( \frac{\pi^2}{6} - 1 \right) \frac{\sigma}{\alpha} \quad (11)$$

Thus, the distance that can be moved is finite, even over an infinite number of cases. The legal rule cannot move a distance greater than  $\left( \frac{\pi^2}{6} - 1 \right) \frac{\sigma}{\alpha}$  from the decision of the first case in the line of precedent without at least one judge deciding to break with precedent. We can find comparative statics on the actual point of convergence by referring directly to equation (10). Notice that if  $\alpha$  is very small, so that judges are relatively unconstrained by the precedential line in which they are writing, this distance is quite large, whereas as  $\alpha$  gets large the distance shrinks.

Similarly, as  $\sigma$  becomes large, the distance that can be moved gets larger because judges are willing to write opinions in line with precedent even when the substantive holdings of those decisions are very far away from the substantive position of the existing precedent.

## **6 Discussion**

### **6.1 Resolution of Some Empirical Puzzles**

Our model provides a new perspective on when, why, and how judicial decisions are constrained by existing precedents. This perspective helps account for empirical observations of judicial behavior that are otherwise difficult to reconcile. To take a striking example, consider the role of deference to precedent in Supreme Court decision-making. On the one hand, a rigorous analysis of voting patterns on the Court finds that justices who vote against a “landmark” case – that is, a case that establishes an important precedent – tend not to switch their voting pattern in subsequent cases, even though the “landmark” case decision ought to constitute a legal precedent (Segal and Spaeth 1996a). This seems to be strong evidence that judges, at least at the Supreme Court level, do not feel constrained by legal precedents as such. On the other hand, if it is really the case that the justices do not attach much importance to precedent, then it is hard to explain why they devote so much time and

intellectual energy to it in their deliberations, and why they place so much emphasis on it in most of their decisions. Analysis of the U.S. Supreme Court decision-making process, after all, reveals that arguments from precedent vastly outnumber all other kinds of arguments in attorneys' written briefs, the Court's written opinions, and the justices' arguments in conference discussions (Phelps and Gates 1991, Knight and Epstein 1996). If it were really the case that judges cared about case outcomes rather than precedent, then all the emphasis on arguments from precedent would seem an irrational waste of resources. But if precedent were really influencing justices' decisions, then the persistent patterns of continuous dissent from landmark decisions is difficult to explain.

Our model is able to account quite simply for this seeming contradiction. Justices care about precedent precisely because they care about policy; if they can sufficiently improve their communication of the proper legal rule by integrating their decision with an existing line of cases, they will do so, even if it means somewhat modifying the legal rule they announce and expending some energy on writing a compelling and coherent opinion that integrates seemingly divergent rulings. Thus, though our model does not formalize the process of formulating a judicial opinion, it is entirely consistent with the observation that judges put a lot of time and attention into trying to integrate their preferred outcome into an established line of cases.

However, in our model, if a judge decides that the additional accuracy is not worth

the sacrifice in terms of substantive policy, then she will vote to break with precedent. If she dissents in a landmark case, she will usually continue to dissent in subsequent cases. After all, in our model judges do not care about precedent per se, so there is no reason for a judge to vote to uphold a legal rule that is far from her ideal simply because that rule had been established in a prior case. She will only vote to adhere to precedent if she can, in the process, shift the legal rule sufficiently closer to her ideal point that the gain in terms of accuracy is worth the cost in terms of substance.

A second empirical anomaly that our model illuminates is the seemingly schizophrenic attitude of judges toward long-established precedents. On the one hand, many would quote approvingly Justice Holmes' quip that "it is revolting to have no better reason for a rule of law than that it was laid down in the time of Henry IV. It is still more revolting if the grounds upon which it was laid down have vanished long since, and the rule simply persists from blind imitation of the past" (1897). On the other hand, it is often thought that a legal rule with a long history is worthy of particular deference. Thus, lawyers and judges sometimes argue, implicitly or explicitly, against tampering with long-established legal rules, even while disagreeing with their substantive content. Again, our model suggests a simple reconciliation of these apparently contradictory notions. Recall from the comparative statics that increasing  $t$  has two effects. It constrains the judge's ability to affect the substance of the legal rule, leading to the frustration embodied in Holmes' remark. However, an old precedent –

i.e. a long line of cases – though difficult to move, has a great deal of informational value. The consequence is that old precedents become entrenched so that even when judges disagree with the substantive rules they are reluctant to overrule them.

## 6.2 New Hypotheses

In addition to offering new insights into these important empirical puzzles, our model also suggests a number of new hypotheses regarding patterns of judicial decision-making. First, recall the comparative statics on the parameter  $\sigma$  which measures the inherent difficulty in communication between appellate and trial judges. The model demonstrated that as  $\sigma$  increases the attractiveness of breaking with precedent decreases. Substantively, this implies that areas of law that are highly complex and not amenable to simple legal regulation are more likely to develop long lines of cases, with both high levels of deference to precedent and evolution and change of legal rules within the precedent. However, when the legal issue is quite simple, making communication of the legal rule less difficult, judges have little use for precedent. They will either follow the old rule exactly or change it completely.

Next our model offers a new perspective on the long-standing debate about the relative merits of rules and standards (Schlag 1985, Kaplow 1992, Sullivan 1992, Korobkin 2000). Typically, standards contain general principles — for example, “due process” or negligence” – whereas rules use more specific and precise language. Strict

liability tort regimes and statutes of limitations are examples of such legal rules. In terms of our model, rule-like holdings might correspond to relatively higher values of  $\alpha$  whereas standards might correspond to lower values of  $\alpha$  (recall that  $\alpha$  measures the costs to judges of effecting substantive legal changes while claiming consistency with an existing line of precedent). This is because it is easier for judges to adapt the broader, more general language of standards, while claiming fidelity to the original principle than it is for judges to alter an unambiguous legal rule. As shown above, as  $\alpha$  increases the distance that a judge will move the substance of existing precedent while maintaining that precedent decreases but, as a result, the desirability of breaking with precedent increases. Thus, our model calls into question the conventional view that rules are more stable and predictable than standards. We find a trade-off between two types of stability. Rules will be associated with periods of little substantive change punctuated, more frequently, by sudden breaks. Standards, on the other hand, will be characterized by more constant, but gradual, substantive change but will be overturned outright, less often. While we do not have a normative position regarding this trade-off, our model demonstrates that the question of whether rules or standards foster greater stability and predictability is more complex than is commonly appreciated.

Finally, our model suggests patterns the should emerge in the long-term evolution and development of law in an Anglo-American common law system. Recall that

when  $\alpha$  is sufficiently low, legal rules are less vulnerable to being overruled when they are very young and when they are very old. A judge can easily adapt the substance of a young precedent to reflect her policy preferences, making a break with precedent unnecessary. And, while a very old precedent strongly constrain judge's ability to influence the substance of the law, it provides tremendous informational value. Middle-aged precedents are more vulnerable to being overruled. Thus, when  $\alpha$  is sufficiently low, we might expect the following pattern of legal development. When courts initially confront a new legal issue, the law will likely be characterized by a number of false starts. The legal rule specified by the first judge to confront the issue will be refined by a number of subsequent judges, but as the rule solidifies, it becomes more vulnerable to being overruled. This occurs because as the rule develops it may begin to exert a significant constraint on the decisions of judges before it can provide sufficient informational benefits to compensate for this constraining effect. Thus, we expect several rules to be proposed, refined, and ultimately rejected. However, once a legal rule survives the precarious intermediate stage of its development it will become increasingly entrenched, overturned only if it is confronted by a judge with substantially divergent preferences.

## 7 Conclusion

We have developed an informational model of judicial decision-making in which deference to precedent is useful to outcome-oriented appellate judges because it improves the accuracy with which they can communicate legal rules to trial judges. Although our model makes a significant contribution to understanding judicial decision-making it is important to highlight that much work remains to be done in developing informational theories of judicial behavior. For example, while we have discussed a series of judicial decisions made over time, we have only modelled judges as one-shot decision-makers. A formal treatment of repeated decision making might offer important insights, especially regarding the potentially moderating effect of the belief that one has multiple opportunities to effect changes in a substantive legal rule. Further, several possibilities exist for game-theoretic extensions of our simple decision-theoretic model. A particularly interesting avenue to explore would be a game with multiple appellate judges with diverse policy preferences. Another possibility is to model trial court judges as strategic actors with policy preferences of their own.

Nonetheless, our simple model yielded a host of interesting implications and hypotheses regarding conditions under which judges will maintain or break with precedent, the constraining effect that precedent has on judicial decision-making, the voting behavior of Supreme Court Justices, the relationship between a precedent's age and

its authority, the effect of legal complexity on the level of deference to precedent, the relative stability of rules and standards, and long-term patterns of legal evolution. Perhaps most importantly, we demonstrated that “legalist” features of judicial decision-making are consistent with an assumption of policy-oriented judges. Thus, the informational approach to the study of judicial behavior can generate new insights and help to reconcile long-standing debates in the literature.

## References

- Brisbin, Richard A. 1996. “Slaying the Dragon: Segal, Spaeth and the Function of Law in Supreme Court Decision Making.” *American Journal of Political Science* 40:1004-1017.
- Cross, Frank B. and Emerson H. Tiller. 1998. “Judicial Partisanship and Obedience to Legal Doctrine: Whistleblowing on the Federal Courts of Appeals.” *Yale Law Journal* 107:2155.
- Cox, Archibald. 1976. *The Role of the Supreme Court in American Government* (New York: Oxford University Press).
- Edwards, Harry T. 1998. “Collegiality and Decision making on the D.C. Circuit.” *Virginia Law Review* 84:1335-1370.

Holmes, Oliver Wendell. 1897. "The Path of the Law." *Harvard Law Review* 10:457.

Hyde, Alan. 1983. "The Concept of Legitimation in the Sociology of Law." *Wisconsin Law Review* 1983:379.

Kaplow, Lewis. 1992. "Rules Versus Standards: An Economic Analysis." *Duke Law Journal* 42:557.

Knight, Jack and Lee Epstein. 1996. "The Norm of Stare Decisis." *American Journal of Political Science* 40:1018-1035.

Knight, Jack and Lee Epstein. 1998. *The Choices Justices Make* (Washington, D.C.: CQ Press).

Kornhauser, Lewis A. 1989. "An Economic Perspective on Stare Decisis." *Chicago-Kent Law Review* 65:63-92.

Korobkin, Russell B. 2000. "Behavioral Analysis and Legal Form: Rules vs. Standards Revisited." *Oregon Law Review* 79:23.

Landes, William and Richard A. Posner. 1976. "Legal Precedent: A Theoretical and Empirical Analysis." *Journal of Law and Economics* 19:249.

Macey, Jonathan R. 1989. "The Internal and External Costs and Benefits of Stare Decisis." *Chicago-Kent Law Review* 65:93-113.

Maltz, Earl. 1980. "Some Thoughts on the Death of Stare Decisis in Constitutional Law." *Wisconsin Law Review* 1980:467.

Nelson, Caleb. 2001. "Stare Decisis and Demonstrably Erroneous Precedents." *Virginia Law Review* 87:1-84.

Phelps, Glenn A. and John B. Gates. 1991. "The Myth of Jurisprudence: Interpretive Theory in the Constitutional Opinions of Justices Rehnquist and Brennan." *Santa Clara Law Review* 31:567-596.

Posner, Richard A. 1993. "What do Judges and Justices Maximize? (The Same Thing Everybody Else Does)." *Supreme Court Economic Review* 30(0):1-41.

Rasmusen, Eric. 1994. "Judicial Legitimacy as a Repeated Game." *Journal of Law, Economics, and Organization* 10:63-83.

Rehnquist, James C. 1986. "The Power that Shall be Vested in a Precedent: Stare Decisis, the Constitution and the Supreme Court." *Boston University Law Review* 66:345-376.

Revesz, Richard L. 1997. "Environmental Regulation, Ideology, and the D.C. Circuit." *Virginia Law Review* 83:1717-1772.

Revesz, Richard L. 1999. "Ideology, Collegiality, and the D.C. Circuit: A Reply to Chief Judge Harry T. Edwards." *Virginia Law Review* 85:805-851.

Schlag, Pierre J. 1985. "Rules and Standards." *UCLA Law Review* 33:379.

Segal, Jeffrey A. and Harold Spaeth. 1993. *The Supreme Court and the Attitudinal Model* (New York: Cambridge University Press).

Segal, Jeffrey A. and Harold J. Spaeth. 1996(a). "The Influence of Stare Decisis on the Votes of United States Supreme Court Justices." *American Journal of Political Science* 40:971-1003.

Segal, Jeffrey A. and Harold J. Spaeth. 1996(b). "Norms, Dragons, and Stare Decisis: A Response." *American Journal of Political Science* 40:1064-1082.

Shapiro, Martin. 1972. "Toward a Theory of Stare Decisis." *Journal of Legal Studies* 1:125-134.

Songer, Donald R. and Stefanie A. Lindquist. 1996. "Not the Whole Story: The Impact of Justices' Values on Supreme Court Decision Making." *American Journal of Political Science* 40:1049-1063.

Sullivan, Kathleen M. 1992. "The Justices of Rules and Standards." *Harvard Law Review* 106:22.

Talley, Eric. 1999. "Precedential Cascades; An Appraisal." *Southern California Law Review* 73:87-137.