Informative Precedent and Intrajudicial Communication

ETHAN BUENO DE MESQUITA AND MATTHEW STEPHENSON
Harvard University

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 new 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.

Political scientists have long recognized the importance of courts as political actors. However, while an extensive literature examines the judiciary's strategic interaction with the other branches of government (e.g., Ferejohn and Weingast 1992; Gely and Spiller 1992), less attention has been paid to the effects of the institutional structure of the courts themselves on patterns of judicial decision-making. Yet models of the judiciary's unique institutional dynamics are essential to understanding the courts, just as analyses of congressional committees (e.g., Gilligan and Krehbiel 1990; Shepsle and Weingast 1987) or the bureaucracy (e.g., Moe 1982) are essential to understanding the legislative and executive branches. We develop a formal model that demonstrates how the problem of communication between different levels of a hierarchical court system, such as the one in the United States, can, under some conditions, give judges incentives to defer to precedents established in prior cases and how this constraining effect influences the policy decisions made by the courts.

This approach sheds light on one of the most perplexing and controversial problems in the social-scientific study of judicial behavior: 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.[1] 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 reelection. Whatever the reason, the policy-oriented judge cares about actual judicial "outputs" more than any particular method of arriving at those outputs. Others, however, argue that judges are concerned with "legalism," that is, with correctly following the rules and norms of proper judicial reasoning. A legalist judge maximizes 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 subject of heated debate (e.g., Knight and Epstein 1996; Segal and Spaeth 1996a, b; Songer and Lindquist 1996).

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' concerns 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 rulings 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 the 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.

Ethan Bueno de Mesquita is a Ph.D. candidate, Department of Government, Harvard University, Littauer Center—North Yard, Cambridge, MA 02138 (mesquita@fas.harvard.edu). Matthew Stephenson is a Ph.D. candidate, Department of Government, Harvard University, and J.D. candidate, Harvard Law School, Cambridge, MA 02138 (mstephen@fas.harvard.edu).

We are indebted to Scott Ashworth, Bruce Bueno de Mesquita, Charles Cohen, Eric Dickson, Gilles Serra, Ken Shepsle, Judge Stephen Williams, several anonymous referees, and the participants in the Harvard Rational Choice Lunch Group, the Harvard Political Economy Research Workshop, the Texas A&M Conference on Institutional Games and the U.S. Supreme Court, and the William H. Riker Conference on Constitutions, Voting, and Democracy at Washington University for valuable comments. Stephenson gratefully acknowledges support from the John M. Olin Center for Law, Economics, and Business. Bueno de Mesquita gratefully acknowledges support from the Center for Basic Research in the Social Sciences.

[1] 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.

https://doi.org/10.1017/S0003055402000436 Published online by Cambridge University Press

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 U.S. 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 clear rules versus flexible standards, and patterns of long-term legal evolution.

COMPETING THEORIES OF JUDICIAL DECISION-MAKING

We focus on the principle of *stare decisis*,[2] 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 (e.g., Schauer 1987). Judges are not, in ordinary circumstances, supposed to overturn "settled law" (Nelson 2001).[3] 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 the legalist models of judicial decision-making because these models seem 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 that judge would have decided the precedent-setting case differently.[4]

In 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 ruling if the precedent did not exist. Thus, a policy-oriented judge would be expected to attach little importance to established precedent when making 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 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 given judge to sacrifice other policy goals for the sake of stability, that judge must believe that other judges will also value stability sufficiently highly that they will not overturn precedent. But if a judge believes that other judges do place a high value on stability, that judge 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 to a judge is considerably reduced 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."[5] While there may be some of this tit-for-tat in the real world, this

[2] *Stare decisis et non quieta movere*—"Stand by the thing decided and do not disturb the calm."

[3] 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*).

[4] 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 it is safe to assume that legalist judges in a common law system would attach substantial weight to precedent.

[5] If there is not specific, targeted retaliation against particular judges for failing to follow precedent, 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).

https://doi.org/10.1017/S0003055402000436 Published online by Cambridge University Press

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 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 (e.g., Hyde 1983; Nelson 2001).

Finally, several studies point out that deference to precedent may be valuable even to policy-oriented judges because of the informational function that judicial precedents serve (e.g., Rehnquist 1986; Shapiro 1972). 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 (e.g., 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 that 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," or "compelling interest"; 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 10 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 underinclusive.[6] 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 that 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.

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 the decisions in these cases, the appellate judge can announce how the law ought to be interpreted. The trial judge decides the vast majority of cases. We make the simplifying assumption that when making decisions the trial judge attempts to apply established law, without reference to the trial judge's own personal policy preferences. There are several possible substantive defenses of this assumption that could ultimately be modeled explicitly. These include considerations such as promotion being dependent on strict adherence to the law established by appellate decisions, a desire not to be overturned on appeal, and other factors. For the purposes of this model, it suffices to assume that some mechanism exists that leads the trial judge to behave as a faithful agent of the appellate judge. While we assume a nonstrategic trial judge, we do address the conditions under which a strategic trial judge might behave in this way later in the paper.

The trial judge's understanding of appellate rulings is imperfect, and the trial judge often does not decide cases exactly as the appellate judge would have. 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 (e.g., U.S. Supreme Court review of circuit court decisions) or because many cases are not appealed despite the existence of an automatic right to an appeal (e.g., federal circuit court review of district court decisions). Thus, the appellate court judge, to influence the application of the law to the majority of cases successfully, needs to communicate a preferred interpretation of the law to the trial court as accurately as possible, subject to time and resource constraints.

[6] A bright line rule is a rule that minimizes ambiguity by setting well-defined and simple categories of prohibited and permissible behavior.

https://doi.org/10.1017/S0003055402000436 Published online by Cambridge University Press

The appellate judge has preferences over the legal rule defined on a unidimensional continuum.[7] Denote the appellate judge's ideal point $j \in \mathbb{R}$. The policy continuum 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, the trial judge tries to implement whatever the appellate judge has declared to be the proper legal rule.

When deciding a case, the appellate judge issues a ruling with two components. First, the appellate judge 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 $r \in \mathbb{R}$. Second, the appellate judge declares whether this substantive holding is consistent with existing precedent or whether the appellate judge 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. Prior decisions are not assumed all to have been made by the current appellate judge. Rather, one can think of the game beginning with an existing line of precedent that may have been established before the current appellate judge took office. If the appellate judge issues a holding and declares that it is consistent with precedent, then the case is added to the relevant line of cases, and the trial judge will interpret the appellate judge's substantive holding in the context of the other substantive holdings in the line to ascertain the legal rule the appellate judge wishes to enact. If, however, the appellate judge declares a break 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.[8] The number of cases in the line of precedent, including the decision being made in the current round, is denoted t. Thus a line of precedent is a series of substantive rulings (r_1, $r_2, \ldots, r_{t-1}$), where r_{t-1} is the most recent ruling and the current appellate judge issues ruling r_t.

The declaration that the appellate judge 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 the trial judge, when attempting to understand an appellate ruling, observes t signals, each drawn from a normal distribution with variance σ^2. The means of these distributions are the rs associated with the substantive holdings of the different cases in the line of precedent. The trial court judge averages these signals to estimate the legal rule.[9] Thus, the trial judge's estimate of the legal rule is a sum of a series of normally distributed random variables. Such a sum is itself a normally distributed random variable.[10] Consequently, the trial court's estimate can be treated *as if* it were a single signal, denoted x, drawn from a normal distribution with mean $\mu_t = (\sum_{i=1}^{t} r_t)/t$ and variance σ^2/t (DeGroot 1970).

It is important to stress that the trial judge does not actually believe that all of the observed signals are drawn from a single distribution. If the trial judge did believe this, signals far from the estimated mean would make the trial judge less confident in this estimate. However, the trial judge recognizes that the various observed signals are drawn from different distributions with different means, reflecting earlier decisions, potentially made by different appellate judges. Given that the trial judge aggregates these signals, each additional signal, no matter how far it is from the mean, decreases the variance of the trial judge's estimate of the legal rule to be implemented. This variance is the variance of the trial judge's estimate, not the trial judge's estimate of the variance of a single underlying distribution.

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 σ^2. The mean of the distribution

[7] We refer to the judge's preferences over this continuum as policy preferences; however, we use policy broadly, to reflect public-policy preferences, normative judgments regarding fairness or justice, or a weighted combination of various factors.

[8] The discrete binary choice between maintaining and breaking with precedent is a simplifying assumption. Clearly, in the real world, multiple lines of precedent may be in play in any given decision and the judge may decide to maintain parts of existing precedent. Framing the issue in the manner that we do captures in starkest form the same essential decision problem that would arise in more complex adjudicative choices.

[9] The trial judge could, in theory, aggregate these signals in some other way (for example, by taking a weighted average), and this would in turn change the optimal decision of the appellate judge. We focus attention on a simple average for the following reasons. First, including the possibility of a weighted average would not alter the appellate judge's basic decision problem—the trade-off between greater precision (if the appellate judge's holding is aggregated with those of other judges) and the opportunity to announce the judge's most preferred legal rule (if the trial judge looks only at the appellate judge's ruling). The unweighted average is thus the simplest among a class of similar aggregation rules, any of which might be reasonable depending on other background institutional assumptions. Second, calculating the optimal weighted average is itself a complex problem that depends on the appellate judge's deriving the optimal weights as a function of the appellate judge's preferences and existing precedent. Third, not only is this calculation complex for the appellate judge, but the proper weights, once calculated, would have to be communicated to the trial judge. This communication, like the communication of the substantive holding, would be inherently noisy. Thus, the calculation and transmission of appropriate weights—essentially, a more refined and sophisticated principle of *stare decisis*—introduces another level of complexity that, though interesting, is beyond the scope of this paper and is therefore left to future research.

[10] In particular, if $X_1, \ldots, X_n$ are independent random variables and $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, \ldots, n$ and $a_1, \ldots, a_n$ are constants such that $a_i \neq 0$ for at least one i, then the random variable $a_1 X_1 + \cdots + a_n X_n$ is normally distributed with mean $a_1\mu_1 + \cdots + a_n\mu_n$ and variance $a_1^2\sigma_1^2 + \cdots + a_n^2\sigma_n^2$. (DeGroot 1970, 38). In our case, the trial judge is taking an average, $a_i = (1/t)$, $\mu_i = r_i$, and $\sigma_i^2 = \sigma^2$ for $i = 1, \ldots, t$. Consequently the average of the signals is itself a random variable with mean $\mu_t = (\sum_{i=1}^{t} r_t)/t$ and variance σ^2/t.

https://doi.org/10.1017/S0003055402000436 Published online by Cambridge University Press

of the signal in the period prior to the current appellate judge's decision is denoted μ_{t-1}, and we refer to this value as "existing precedent." Because x (the value of the draw from the distribution of the trial court's estimate) is the only information that the trial judge has, it is the trial judge's best guess as to what the appellate judge wants. Consequently the trial judge will treat x as the controlling legal rule when deciding cases.

The appellate judge's utility has two components. First, the appellate judge would like the decisions of the trial court judges to be as close as possible to the appellate judge's own ideal point. That is, the appellate judge 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 in which the policy dimension represents the level of care taken by a tort 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—that is, the appellate judge wants to minimize the expected value of the square of the distance between x and j, which is equal to $((\mu_t - j)^2 + (\sigma^2/t))$.

Second, if the appellate judge does not announce a break with precedent, it is costly to offer a ruling, r_t, that is substantively different from the existing precedent, μ_{t-1}. We are agnostic as to the relative importance of these two factors in the appellate judge's utility calculation; we weight the importance of the latter by the parameter $\alpha \geq 0$ on which we perform comparative static analysis.[11] 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 a decision was consistent with precedent without explaining how the substance of the 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. This aspect of the model captures the intuition that trial judges would have difficulty simply averaging signals that are quite distant from one another. Trial judges are able to aggregate such signals only if the appellate judge has invested substantially in explaining how to do so. This effort is costly to the appellate judge and becomes more costly the greater the distance between the appellate holding and existing precedent, reflecting the increasing difficulty of the trial judge's task. In our model the intuition that it would be difficult for the trial judge to aggregate a sharply divergent holding into existing precedent is captured by the costliness to the appellate judge of making such a holding.

One might object to the assumption that judges bear a cost when maintaining precedent but do not bear a cost when breaking with precedent. However, the assumption is benign with respect to the conclusions of our model. 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. There is a second, more conceptual justification for making this assumption. Any cost from breaking with precedent, beyond the loss of information already modeled, would arise from legalist values. We omit legalist values from the judicial utility function to see whether the norm of *stare decisis* is consistent even with purely policy-oriented judges and, indeed, whether what appear to be legalist values may emerge as a result of policy motivations.

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

ANALYSIS

The Appellate Judge's Decision

The appellate judge has two choices to make. The appellate judge must select a substantive holding (r_t) and must also decide whether to declare a break with precedent. These two decisions are interrelated. To determine whether or not to break with precedent, the appellate judge must first know the expected outcomes, which are a function of the different substantive rulings that the appellate judge would issue depending on whether or not precedent is maintained. Thus, we work backward, first determining the appellate judge's optimal substantive holding contingent on whether the judge has broken with or maintained precedent.

If the appellate judge breaks with precedent, the second term in the judge's expected utility function is zero, because $\tilde{\alpha} = 0$. Further, because the appellate judge has broken with existing precedent, t reverts to 1. Therefore, the line of cases considered by the trial judge

[11] 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 $0 \leq \alpha < 1$, the closeness of the trial court decision to the appellate judge's ideal point is weighted more heavily. If $\alpha = 0$, adherence to precedent is costless no matter how much the legal rule is changed substantively; in this special case, the appellate judge would always claim to adhere to precedent, because doing so imposes no constraints whatsoever on the distance the appellate judge can move the legal rule.

https://doi.org/10.1017/S0003055402000436 Published online by Cambridge University Press

contains only the 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 an appellate judge breaks with precedent, the substantive holding will be the judge's ideal legal rule, and the judge's expected utility in this case, $EU(\text{break})$, is $-\sigma^2$.

If the appellate judge does not break with precedent, the decision is more complicated. The appellate judge would like to issue a ruling that moves precedent toward the judge's own ideal point but is constrained by the cost of deviating too much from the line of precedent of which this decision becomes a part. The appellate judge thus chooses r_t by solving the following maximization problem:

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

Note that μ_t, the mean of the distribution of the trial judge's estimate of the legal rule, is a function of r_t. In particular, $\mu_t = [(t-1)\mu_{t-1} + r_t]/t$. Thus, we can rewrite the maximization problem

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

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 an estimate of the proper legal rule. This new mean, μ_t^*, is given by

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

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

$$EU(\text{maintain}) = -\left[\left(\frac{j + \alpha t^2 \mu_{t-1}}{1 + \alpha t^2} - j \right)^2 + \frac{\sigma^2}{t} \right] - \alpha \left[\mu_{t-1} - \left(\mu_{t-1} + \frac{t(j - \mu_{t-1})}{1 + \alpha t^2} \right) \right]^2,$$

which simplifies to

$$EU(\text{maintain}) = -\frac{\alpha t^2}{1 + \alpha t^2}(j - \mu_{t-1})^2 - \frac{\sigma^2}{t}. \quad \textbf{(4)}$$

In our model the central benefit to appellate judges of maintaining precedent is informational. The more cases the trial judges have to refer to, the more accurately they will understand the legal rule for which that line of cases stands. Thus, the appellate judge wishes to maintain precedent because it makes communication with trial courts less noisy. However, the use of precedent comes at a price. In particular, appellate judges bear a cost for deviating substantively too far from the line of precedent that they claim to uphold. This constrains judges who are maintaining precedent from implementing a legal rule that matches their personal ideal, as can be seen in the model. When precedent is maintained, $\mu_t^* = (j + \alpha t^2 \mu_{t-1})/(1 + \alpha t^2) \neq j$ (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 (μ_{t-1}). Further, it is interesting to note that the appellate judge, when maintaining precedent, chooses the holding $r_t^* = \mu_{t-1} + [t(j - \mu_{t-1})/(1 + \alpha t^2)]$, which is not equal to existing precedent (μ_{t-1}) or the appellate judge's ideal point (j) unless, by chance, $j = \mu_{t-1}$ or $t = 1 + \alpha t^2$. This ruling is always in the direction of the appellate judge's ideal point, but may be more or less extreme, relative to existing precedent. If $t > 1 + \alpha t^2$, the appellate judge will issue a ruling farther from existing precedent than the appellate judge's ideal point. Otherwise, the appellate judge will issue a ruling between the ideal point and existing precedent.

While the appellate judge can make sure, by announcing a break with precedent, that the mean of the trial court's signal is equal to the appellate judge's ideal point, the variance of this signal (that is, the noisiness of communication) will be higher. Again, this can be seen clearly in the model, since $\sigma^2 > \sigma^2/t$ as long as $t > 1$. The appellate judge thus faces a trade-off between the accuracy with which the legal rule is communicated to the trial courts and the proximity of that rule to the appellate judge's ideal. The appellate judge determines whether or not to break with precedent by comparing the expected utilities associated with each choice, breaking with precedent if and only if

$$\begin{aligned} &EU(\text{break}) - EU(\text{maintain}) \\ &= \sigma^2 \frac{1 - t}{t} + \frac{\alpha t^2}{1 + \alpha t^2}(j - \mu_{t-1})^2 > 0 \quad \textbf{(5)} \\ &\Leftrightarrow \frac{\alpha t^3}{(t-1)(1 + \alpha t^2)}(j - \mu_{t-1})^2 - \sigma^2 > 0. \end{aligned}$$

The Trial Judge's Decision

The trial judge in this model is not a strategic actor. Rather, the trial judge simply follows the instructions of the appellate judge, aggregating all existing holdings in the line if precedent has been maintained but considering only the appellate judge's most recent ruling if precedent has been broken. Because the focus of our model is the appellate judge's decision, taking the trial judge's aggregation rule as given, we do not formally analyze the ramifications of strategic behavior by the trial judge. Nonetheless, because we justified this simplifying assumption in part by claiming that trial judges want to please their superiors on the appellate

https://doi.org/10.1017/S0003055402000436 Published online by Cambridge University Press

bench or avoid being overturned on appeal, one might wonder if the trial judge would be better off attempting to deduce and implement the current appellate judge's ideal point, rather than following existing precedent. This turns out, however, not to be the case.

The trial judge could, of course, form an unbiased approximation of the appellate judge's ideal point by inverting the maximization problem that the appellate judge solved in equation (2), conditioning on the trial judge's estimate of preexisting precedent, μ_{t-1}.[12] That is, the trial judge could observe the signals associated with the first $t-1$ rulings in the line of precedent and use this information to estimate the precedent the current appellate judge faced. Then, conditioning on this preexisting precedent, the trial judge could invert the appellate judge's maximization problem to estimate what value of j would have caused the current appellate judge to issue the ruling the trial judge observed. This estimate is unbiased because the expected value of the trial judge's observation of any holding, r_i, in the line of precedent is r_i.

Notice, though, that the information regarding preexisting precedent can be used only to help the trial judge figure out what the current appellate judge's maximization problem was. Having done this, the only information the trial judge has regarding the appellate judge's ideal point is the signal that the trial judge observes, which is drawn from a normal distribution with mean equal to the current appellate judge's ruling, r_t^*. The variance of this distribution is σ^2, not σ^2/t. Consequently, following this inversion procedure cannot lead the trial judge to a better estimate of the appellate judge's ideal point than would be achieved if the appellate judge had simply broken with precedent and issued a ruling exactly at the appellate judge's ideal point. Indeed, because the trial judge approximates preexisting precedent with error, the estimate of j deduced from the inversion procedure is less precise than the estimate the trial judge would have formed if the appellate judge had broken with precedent.

The appellate judge maintains precedent only when the ruling of a trial judge who follows that precedent (that is, averages all the signals) will, in expectation, be closer to the appellate judge's ideal point than it would have been had the appellate judge broken with precedent. Since the trial judge's estimate of the appellate judge's ideal point formed by the inversion procedure is even worse than that produced by a break with precedent, the trial judge is more likely to please the appellate judge (and less likely to be overturned) by following precedent than by trying to deduce the appellate judge's ideal point. If this were not the case, the appellate judge would have broken with precedent. Thus, the trial judge would be better off acting as a faithful agent of the appellate judge, deciding cases according to precedent when instructed to do so.[13]

[12] This is possible only if the trial judge knows α and σ^2. If not, the trial judge is unable even to entertain the possibility of trying to deduce the appellate judge's ideal point from the observed signals.

[13] This is not to say that an aggregation rule other than averaging might not lead to superior outcomes in some circumstances. As discussed earlier, the optimal aggregation rule is a function of a host of complex institutional factors beyond the scope of this model. Nonetheless, the basic logic—that the appellate judge's choice to maintain precedent implies that the trial judge will do better by following precedent than by attempting to deduce the appellate judge's ideal point—holds under any aggregation rule.

RESULTS

Comparative Statics

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

It is clear from equation (5) that the desirability of breaking with precedent decreases as σ^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. Equation (5) also 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 much the appellate judge can move the expected trial court decision, the appellate judge is less willing to trade off control over the substantive rule for increased accuracy of transmission.

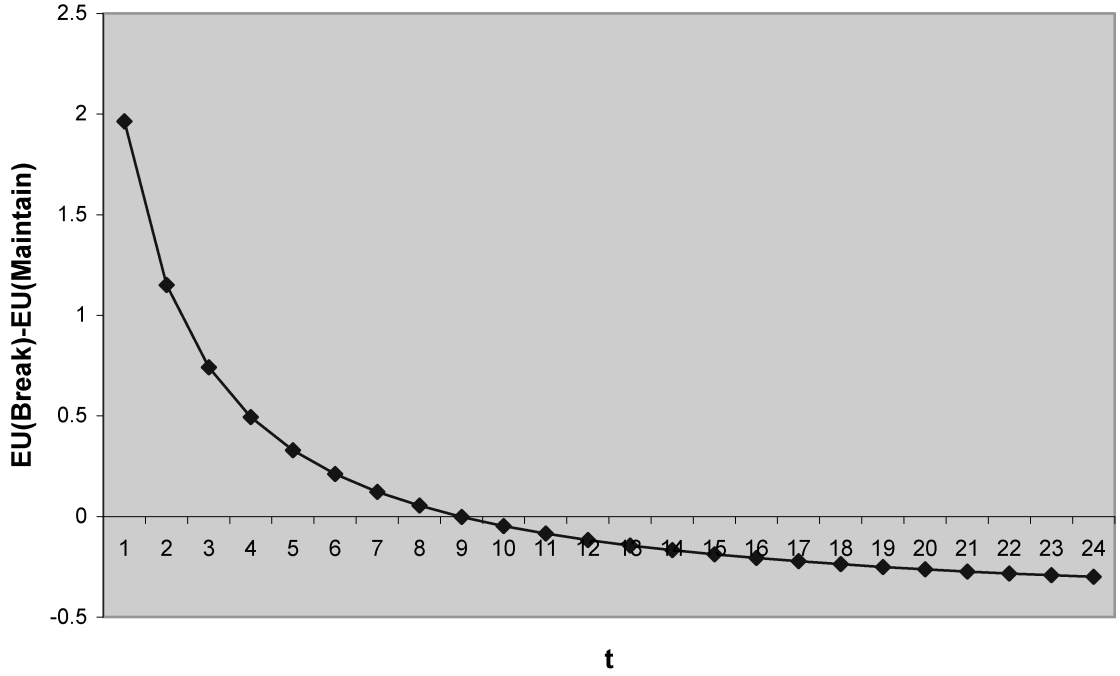
Equation (5) additionally shows that breaking with precedent becomes more attractive as α, the marginal cost of making a substantively divergent decision within an existing line of precedent, increases. The reason for this is that, when α is close to zero, the appellate judge can move μ_t very close to j, even while maintaining precedent. Thus, the informational benefit of situating the current decision in a long line of cases comes at very little cost in terms of substance. However, as α grows, the appellate judge's ability to move the legal rule close to j 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 $|\mu_t^* - \mu_{t-1}| = |(j-\mu_{t-1})/(1+\alpha t^2)|$, which is clearly decreasing in α.

Deriving the comparative statics on t, the number of cases in the line of precedent, is more complicated. Because t is a discrete variable, to calculate the change in the attractiveness of breaking with precedent as the age of the precedent increases, we take the first differences of equation (5) with respect to t:

$$[EU_{t+1}(\text{break}) - EU_{t+1}(\text{maintain})] - [EU_t(\text{break}) - EU_t(\text{maintain})] \equiv FD(t) = -(j-\mu_{t-1})^2 \times \alpha \frac{\alpha t^4 + 2\alpha t^3 + \alpha t^2 - 2t^3 + 2t + 1}{(1+\alpha t^2 + 2\alpha t + \alpha)(1+\alpha t^2)t(t-1)}. \quad (6)$$

https://doi.org/10.1017/S0003055402000436 Published online by Cambridge University Press

FIGURE 1. ***EU*(Break) − *EU*(Maintain) with $\alpha > \bar{\alpha}(2)$**



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

$$FD(t) < 0 \quad \text{if} \quad \frac{2t^3 - 2t - 1}{t^4 + 2t^3 + t^2} \equiv \bar{\alpha}(t) < \alpha,$$
$$FD(t) > 0 \quad \text{if} \quad \frac{2t^3 - 2t - 1}{t^4 + 2t^3 + t^2} \equiv \bar{\alpha}(t) > \alpha. \tag{7}$$

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, $\bar{\alpha}(t) = (2t^3 - 2t - 1)/(t^4 + 2t^3 + t^2)$, 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 decreases as t increases. That is, for a sufficiently large α, 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 α is below the threshold value of $\bar{\alpha}(t)$. But as t increases, this threshold value decreases, meaning that α 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.

To understand the intuition behind these comparative statics, recall that an increase in t has two effects on the desirability to the appellate 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 α is sufficiently high, the information effect always overwhelms the constraining effect. The reason for this is that when α is high, the appellate 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 α 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 α is, when t is sufficiently high the appellate judge is already so constrained that the marginal constraining effect of an increase in t is again negligible.[14]

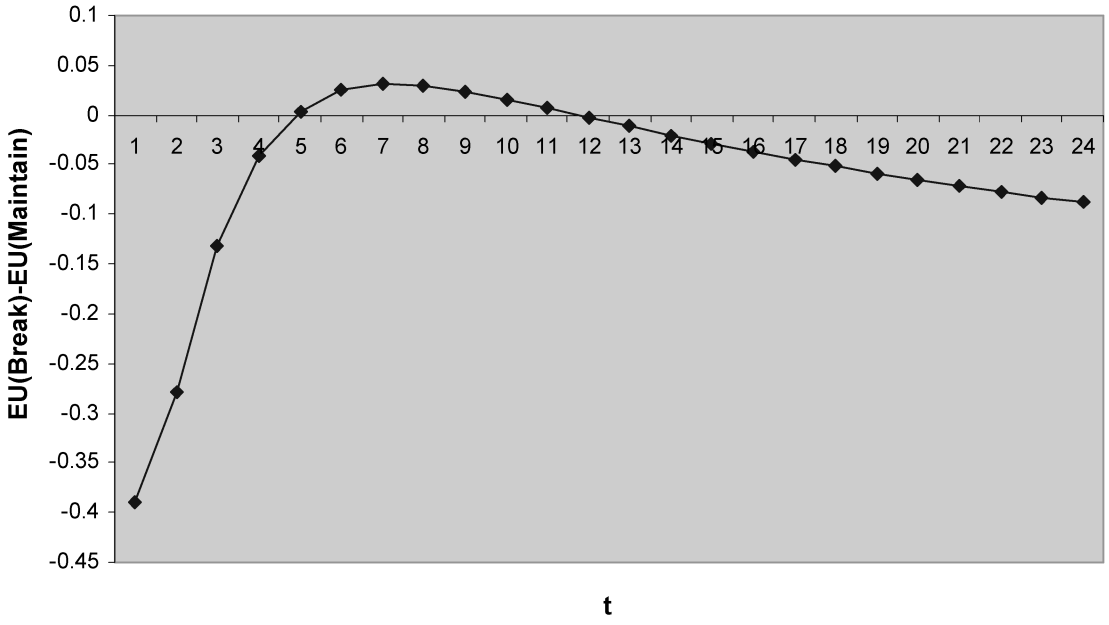
Limits of Legal Change

An interesting question in the context of our model is, How far on the substance dimension can a legal

[14] Note also that as t goes to infinity, the appellate judge is completely constrained by the old precedent. In this case, the value of α does not matter, and the payoff of an appellate judge who maintains precedent converges to $-(j - \mu_{t-1})^2$.

https://doi.org/10.1017/S0003055402000436 Published online by Cambridge University Press

FIGURE 2. ***EU*(Break) − *EU*(Maintain) with $\alpha < \bar{\alpha}(2)$**



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 that judge was exactly indifferent between moving existing precedent and breaking, leading these judges to move μ_t as far from μ_{t-1} as is possible—always in the same direction—without declaring a break from existing precedent? Returning to the tort liability example discussed earlier, imagine that over time appellate judges want a progressively stricter standard of care. We want to know whether it is possible for the standard of care to become infinitely more strict than the original standard announced by the first judge to consider the question, without any judge ever openly breaking with established precedent.

Of course, the idea that there are multiple appellate judges deciding cases in sequence implies a dynamic model, beyond the scope of this project. Nonetheless, our model does provide a framework that allows us to gain some theoretical leverage on this problem. If an appellate judge does not know what the ideal points of future judges will be, it is reasonable for the current judge to assume that the probability of a future judge having preferences any given distance to the left of existing precedent is the same as the probability of that judge having preferences the same distance to the right. Under such a symmetry assumption, the expected ideal point of a future judge lies at existing precedent. Consequently, the appellate judge's best guess of future precedent is current precedent, and so the decision problem in the dynamic model reduces to the one-shot problem analyzed above.[15]

[15] Of course, this is also the case if each appellate judge is myopic, caring only about the payoff in the current round.

Therefore, 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:

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

$$= \sum_{t=2}^{\infty}\left\{\frac{j-\mu_{t-1}}{1+\alpha t^2}\right\}, \qquad \textbf{(8)}$$

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

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

$$\frac{\sigma}{\sqrt{\alpha}}\sum_{t=2}^{\infty}\frac{1}{t}\frac{\sqrt{t-1}}{\sqrt{t}\sqrt{1+\alpha t^2}}. \qquad \textbf{(9)}$$

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

$$\frac{\sigma}{\sqrt{\alpha}}\sum_{t=2}^{\infty}\frac{1}{t}\frac{\sqrt{t-1}}{\sqrt{t}\sqrt{1+\alpha t^2}} < \frac{\sigma}{\sqrt{\alpha}}\sum_{t=2}^{\infty}\frac{1}{t}\frac{1}{\sqrt{\alpha t^2}}$$

$$= \frac{\sigma}{\alpha}\sum_{t=2}^{\infty}\frac{1}{t^2} = \left(\frac{\pi^2}{6} - 1\right)\frac{\sigma}{\alpha}. \qquad \textbf{(10)}$$

Thus, the distance that can be moved in one direction without a break from precedent is finite, even over

https://doi.org/10.1017/S0003055402000436 Published online by Cambridge University Press

an infinite number of turns. Specifically, the legal rule cannot move a distance greater than $((\pi^2/6) - 1)(\sigma/\alpha)$ from the decision in the first case in the line of precedent without at least one judge deciding to break. We can find comparative statics on the actual point of convergence by referring directly to equation (9). Notice that if α 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 α gets large the distance shrinks. Similarly, as σ 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.

DISCUSSION

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 (Knight and Epstein 1996; Phelps and Gates 1991). 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 to be a waste of resources. But if precedent is really influencing justices' decisions, then the persistent patterns of continuous dissent from landmark decisions is difficult to explain.

Our model accounts straightforwardly 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 an appellate judge decides that the additional accuracy is not worth the sacrifice in terms of substantive policy, then the judge will vote to break with precedent. If a given judge dissents in a landmark case, therefore, that judge 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 that judge's ideal simply because that rule had been established in a prior case. An appellate judge will vote to adhere to precedent only if, in doing so, the legal rule can be moved sufficiently closer to the judge's ideal point that the gain in terms of accuracy is worth the cost in terms of substance.

A second apparent 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's (1897) 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." 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's 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.

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 σ, which measures the inherent difficulty in communication between appellate and trial judges. The model demonstrated that as σ 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

https://doi.org/10.1017/S0003055402000436 Published online by Cambridge University Press

change of legal rules within the precedent. However, when the legal issue is 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 (e.g., Kaplow 1992; Schlag 1985; Sullivan 1992). 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, rulelike holdings might correspond to relatively higher values of α, whereas standards might correspond to lower values of α (recall that α 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 α 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 that should emerge in the long-term evolution and development of law in an Anglo-American common law system. Recall that when α 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 personal policy preferences, making a break with precedent unnecessary. And, while a very old precedent strongly constrains a 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 α 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.

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 we believe that 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 modeled 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 model yields 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 have 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

Cox, Archibald. 1976. *The Role of the Supreme Court in American Government.* New York: Oxford University Press.

DeGroot, Morris H. 1970. *Optimal Statistical Decisions.* New York: McGraw–Hill.

Ferejohn, John A., and Barry R. Weingast. 1992. "A Positive Theory of Statutory Interpretation." *International Review of Law and Economics* 12 (September): 263–79.

Gely, Rafael, and Pablo T. Spiller. 1992. "The Political Economy of Supreme Court Constitutional Decisions: The Case of Roosevelt's Court-Packing Plan." *International Review of Law and Economics* 12 (January): 45–67.

Gilligan, Thomas, and Keith Krehbiel. 1990. "Organization of Informative Committees by a Rational Legislature." *American Journal of Political Science* 34 (2): 531–64.

Holmes, Oliver Wendell. 1897. "The Path of the Law." *Harvard Law Review* 10: 457–78.

https://doi.org/10.1017/S0003055402000436 Published online by Cambridge University Press

Hyde, Alan. 1983. "The Concept of Legitimation in the Sociology of Law." *Wisconsin Law Review* 1983 (March–April): 379–426.
Kaplow, Louis. 1992. "Rules Versus Standards: An Economic Analysis." *Duke Law Journal* 42 (December): 557–629.
Knight, Jack, and Lee Epstein. 1996. "The Norm of *Stare Decisis*." *American Journal of Political Science* 40 (November): 1018–35.
Knight, Jack, and Lee Epstein. 1998. *The Choices Justices Make.* Washington, DC: CQ Press.
Kornhauser, Lewis A. 1989. "An Economic Perspective on *Stare Decisis*." *Chicago-Kent Law Review* 65 (1): 63–92.
Landes, William, and Richard A. Posner. 1976. "Legal Precedent: A Theoretical and Empirical Analysis." *Journal of Law and Economics* 19 (September): 249–307.
Macey, Jonathan R. 1989. "The Internal and External Costs and Benefits of *Stare Decisis*." *Chicago-Kent Law Review* 65 (1): 93–113.
Maltz, Earl. 1980. "Some Thoughts on the Death of *Stare Decisis* in Constitutional Law." *Wisconsin Law Review* 1980: 467.
Moe, Terry M. 1982. "Regulatory Performance and Presidential Administration." *American Journal of Political Science* 26 (May): 197–224.
Nelson, Caleb. 2001. "*Stare Decisis* and Demonstrably Erroneous Precedents." *Virginia Law Review* 87 (March): 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 (3): 567–96.
Rasmusen, Eric. 1994. "Judicial Legitimacy as a Repeated Game." *Journal of Law, Economics, and Organization* 10 (April): 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 (March): 345–76.
Schauer, Frederick. 1987. "Precedent." *Stanford Law Review* 39 (February): 571–605.
Schlag, Pierre J. 1985. "Rules and Standards." *UCLA Law Review* 33 (December): 379–430.
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. 1996a. "The Influence of *Stare Decisis* on the Votes of United States Supreme Court Justices." *American Journal of Political Science* 40 (November): 971–1003.
Segal, Jeffrey A., and Harold J. Spaeth. 1996b. "Norms, Dragons, and *Stare Decisis*: A Response." *American Journal of Political Science* 40 (November): 1064–82.
Shapiro, Martin. 1972. "Toward a Theory of *Stare Decisis*." *Journal of Legal Studies* 1 (1): 125–34.
Shepsle, Kenneth A., and Barry R. Weingast. 1987. "The Institutional Foundations of Committee Power." *American Political Science Review* 81 (March): 85–104.
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 (November): 1049–63.
Sullivan, Kathleen M. 1992. "The Justices of Rules and Standards." *Harvard Law Review* 106 (November): 22–97.
Talley, Eric. 1999. "Precedential Cascades: An Appraisal." *Southern California Law Review* 73 (November) 87–137.

https://doi.org/10.1017/S0003055402000436 Published online by Cambridge University Press