American Political Science Review
http://journals.cambridge.org/PSR

Rule Creation in a Political Hierarchy

CLIFFORD J. CARRUBBA and TOM S. CLARK

American Political Science Review / Volume 106 / Issue 03 / August 2012, pp 622 - 643
DOI: 10.1017/S000305541200024X, Published online:

Link to this article: http://journals.cambridge.org/abstract_S000305541200024X

How to cite this article:
CLIFFORD J. CARRUBBA and TOM S. CLARK (2012). Rule Creation in a Political Hierarchy. American Political Science Review,106, pp 622-643 doi:10.1017/S000305541200024X

doi:10.1017/S000305541200024X

Rule Creation in a Political Hierarchy

CLIFFORD J. CARRUBBA and TOM S. CLARK *Emory University*

Principal-agent relations are replete in politics; politicians are agents of electorates, bureaucrats are agents of executives, lower courts are agents of upper courts, and much more. Commonly, principals are modeled as the rule-making body and agents as the rule-implementing body. However, principals often delegate the authority to make the rules themselves to their agents. The relationship between the lower federal courts and the Supreme Court is one such example; a considerable portion of the law (rules) is made in the lower federal courts with the Supreme Court serving primarily as the overseer of those lower courts' decisions. In this article, we develop and test a principal-agent model of law (rule) creation in a judicial hierarchy. The model yields new insights about the relationship among various features of the judicial hierarchy that run against many existing perceptions. For example, we find a non-monotonic relationship between the divergence in upper and lower court preferences over rules and the likelihood of review and reversal by the Supreme Court. The empirical evidence supports these derived relationships. Wider implications for the principal-agent literature are also discussed.

Politics is replete with principal-agent relationships. Research often conceptualizes politicians, bureaucrats, and courts as agents of their own principals and models the principal as the rule-making body and the agent as the implementing body. Electorates elect politicians based on a policy mandate that they are expected to implement once in office; executives appoint bureaucrats to implement the laws of the land; lower courts implement the rules dictated by upper courts. Yet, principals often delegate to their agents the authority to make the rules themselves. For example, in the United States, bureaucracies such as the Food and Drug Administration and the Environmental Protection Agency (EPA) are often charged with rule-making responsibilities, but are subject to oversight by Congress and the courts.

The judicial hierarchy is a particularly important example of this phenomenon. Judge-made law constitutes a large portion of the rules that affect the daily lives of most people in the United States. Courts make law on such important constitutional matters as whether public education may be segregated along racial lines and on more routine matters such as the conditions that constitute wrongful termination. Because of its position at the apex of the judicial hierarchy, the Supreme Court is usually the focus of the study of judge-made law. That focus, however, misses how a substantial amount of judge-made law is created in the United States. The Supreme Court hears only a very small fraction of the cases filed in federal courts (about 0.02%); most law is both made and applied in the lower courts, especially the federal courts of appeals. Indeed, the Supreme Court is often guided by law that has previously been created by lower courts.[1]

To the degree that lower courts appear in theories of judicial lawmaking, they are usually treated as subordinate institutions responsible only for implementing rules generated by the Supreme Court. Scholars examine issues such as whether and to what degree lower courts deviate from applying the rules generated by the Supreme Court and under what conditions the Supreme Court reviews and reverses lower court decisions (e.g., Cameron, Segal and Songer 2000).

In this article we develop and test a theory of rule-creation and application in which the lower court is explicitly allowed to formulate its own rules, subject to possible review by an upper court. Introducing the ability of lower courts to choose rules as well as to decide cases allows us to answer three unexplored questions about law creation in a judicial hierarchy. First, how do characteristics of a case (i.e., the facts of the case, its importance, and the degree to which the lower and upper court have differing preferences over what rules should govern the relevant body of law) affect the rules the lower court will choose? Second, under what conditions is the upper court more likely to review the lower court's decision in light of the rule the lower court declares? And third, how is the upper court likely to treat the lower court's decision—the way a lower court decides a case and the rule used to justify that decision?

Several our theoretical findings represent surprising and normatively important relationships in answer to these questions. For example, the facts of a case can significantly influence the rule the lower court declares,

Clifford J. Carrubba is Associate Professor, Department of Political Science, Emory University, 327 Tarbutton Hall, 1555 Dickey Drive, Atlanta, GA 30322 (ccarrub@emory.edu).

Tom S. Clark is Associate Professor, Department of Political Science, Emory University, 327 Tarbutton Hall, 1555 Dickey Drive, Atlanta, GA 30322 (tom.clark@emory.edu).

We thank *APSR* Co-editor Greg Caldeira, Chuck Cameron, Justin Esarey, Micheal Giles, John Kastellec, Jeffrey Lax, Drew Linzer, Andrew Martin, Matt Stephenson, Craig Volden, and seminar participants at the Stanford University Graduate School of Business Political Economy Seminar, the University of California - Berkeley, and the University of California - San Diego for their helpful comments and suggestions. We thank Josh Strayhorn for research assistance. An earlier version of this article was presented at the 2010 Annual Meeting of the Midwest Political Science Association.

[1] The interest among students of the Supreme Court in intercircuit conflict, along with the Supreme Court's attention to intercircuit conflict, demonstrates the importance of law creation in the lower courts.

and therefore the merits of the case can potentially substantially influence the future direction and content of judge-made law. Further, surprisingly, at times the upper court intentionally allows the lower court to establish rules that are relatively distant from the upper court's preferred rule and (seemingly ironically) potentially reviews and reverses decisions that appear, on the surface, far more consonant with the upper court's preferences. This theoretical finding has important implications for the extent to which Supreme Court preferences over law determine how U.S. law is actually made and applied. Finally, we find non-monotonic relationships between the ideological divergence of the lower court from the upper court and the likelihood of both review and reversal by the upper court. This finding is novel and has substantial implications for the Supreme Court's influence on the application of law. We find empirical support for this last pair of theoretical predictions.

PREVIOUS APPROACHES

Past research in this area has focused most directly on how or to what extent judges higher in the hierarchy—e.g., the Supreme Court—can control or influence the decisions made by judges lower in the hierarchy. Rather than examining the rule-making function, it has focused on error correction (for a discussion, see Friedman 2006). Research has demonstrated that, when lower courts disregard their superiors' preferences, they are more likely to be reviewed (e.g., Cameron, Segal, and Songer 2000; Clark 2009). The operative assumption is that higher courts seek strategically to audit lower courts in an effort to "catch" those cases that it finds most displeasing. Recent studies of the rule-creation function of superior courts have examined the review strategies that they use to exercise control over the large bureaucracy of lower courts that interprets and applies those rules (e.g., Jacobi and Tiller 2007; Lax 2009; Staton and Vanberg 2008). This research characterizes rule making in a political hierarchy as a classical moral hazard problem (see also McNollgast 1995).

We approach the principal-agent problem from the bottom up, focusing on the role lower courts play in rule making. Although much of the research on lower courts has centered on the votes that lower court judges cast (e.g., Hettinger, Lindquist, and Martinek 2006), scholars lately have been paying greater attention to the role judges play in the construction of law.[2] In the most detailed study, Klein (2002) examines the creation of legal rules by circuit courts in the absence of guidance from the Supreme Court. Klein considers cases where the Supreme Court has not offered an authoritative legal rule to govern a situation (cases of first impression). Although he does not develop an explicit theory of intercourt interactions, Klein argues that the extent to which a judge adopts an existing rule depends critically on whether the judge is in sympathy with that rule. Importantly, however, Klein argues that judges on the court of appeals are *not* influenced by the anticipated preferences of the Supreme Court when deciding a case of first impression. From a different, purely theoretical perspective, Hammond, Bonneau, and Sheehan (2006) develop a trio of formal models they use to investigate how the Supreme Court's internal bargaining process influences the policies that the lower courts will adopt, as well as the conditions under which the lower courts can exercise greater or lesser influence over the law. Their analysis centers on the policy choices that lower courts will make under the shadow of oversight by the Supreme Court.

Because (a) most law is created in the courts of appeals and (b) cases enter the system in the lower courts, which have the first opportunity to resolve them, we approach the problem of rule creation in a hierarchy from the perspective of the lower courts. In the next section, we advance a series of formal models of law creation in a judicial hierarchy characterized by an informational asymmetry (i.e., lower courts are better informed about cases than are higher courts, which have not reviewed them). We then discuss the implications of our model and present empirical support for two of the predictions.

THE MODELS

Our models build from canonical principal-agent models. In particular, consistent with Cameron, Segal, and Songer (2000), we assume that a lower court hears a case with specific case facts and must determine the case's outcome. The upper court observes the lower court's decision, but does not observe the case facts unless it chooses to review the lower court's decision. Reviewing the lower court's decision requires paying some cost in time and resources but, in exchange, allows the upper court to learn the case facts and dictate its preferred outcome. Unlike Cameron, Segal, and Songer, we allow the lower court both to declare a rule and dispose of the case.[3] The rule is observable to the upper court before it decides to grant or deny certiorari, and therefore it potentially provides information that the upper court can use when deciding whether to hear a case. On reviewing the lower court's decision, the upper court can change the rule used in the case, thereby potentially changing the case's disposition (i.e., reversing the lower court). A final key feature of our model is that, whereas most models of judicial hierarchy assume that there is some exogenous cost associated with the displeasure of being reversed

[2] This focus in the research applies nearly equally to traditional research on the Supreme Court. However, with the rise of the case-space model as a powerful analytic tool for analyzing rule making in the courts, students of judicial politics have developed increasingly sophisticated models of rules and rule making (for a review, see Lax 2011).

[3] We also note that this formulation is distinct from the model developed in Tiller and Spiller (1999), who examine the way in which a lower court or administrative agency might strategically base its decisions on either adjudication or rule making. By contrast, our model assumes that lower courts are always engaged in both rule making (rule application) and adjudication.

(beyond the policy implications; e.g., Kastellec 2007), we do not assume such a cost. Rather, any reduction in the lower court's utility from review arises as a function of the degree of ideological divergence between the upper and lower courts and the actions that the two courts take.

The primary source of incomplete information between the courts in our model is the contextual nuances of an individual case (the case facts), rather than the courts' preferences or the actions they have taken (rules and dispositions). We focus on uncertainty about case facts for three reasons. First, judges have a close, long-lived working relationship, both within and across levels of the hierarchy (e.g., Hammond, Bonneau, and Sheehan 2006, 137). Thus, they are likely to have an appreciation of each others' preferences. Second, their actions are public and the subject of written documentation. Before deciding whether to review a case, the upper court justices are likely to have a fairly clear idea of how a case was decided and why. Third, by contrast, many cases come to the courts that present similar issues and problems, even though the specifics of the cases are always unique. Thus, the nuances of a given case are going to be comparatively hard to assess until it is actually heard. Our assumption of uncertainty about case facts does not, however, imply that the upper court is completely unaware of the case's context or issues. Rather, it implies that, until the upper court agrees to review a case, the lower court has an informational advantage on the particulars of the case and which disposition a given rule might induce.

Our baseline model is of a nonstrategic lower court, one that does not care about subsequent moves made by the Supreme Court and therefore is not influenced by the institutional structure of the hierarchy; this model serves as a benchmark against which we can compare behavior in subsequent extensions. We then introduce a strategic element—we allow the lower court to have preferences over the decisions made by its superior court. This court is similar to those in most models of judicial auditing (e.g., Cameron, Segal, and Songer 2000; Clark 2009).[4] Finally, we extend our model by introducing an additional informational asymmetry, assuming that the lower court is uncertain about the cost to the higher court of reviewing its decision. One can interpret this uncertainty as the cost of reviewing a case in terms of resources or interest in the issue (e.g., Clark and Strauss 2010). Proofs and supplemental results are presented in the Appendix.

Model 1: A Nonstrategic Lower Court

Players and Sequence of Play. We model law creation in the judicial hierarchy as a game between two players, the Upper Court (UC) and the Lower Court (LC), with commonly known "ideal rules" on the real number line, U and L, respectively. We assume, without loss of generality, that $U = 0$ and $L > 0$. First, Nature draws the case facts from a unidimensional fact-space. Formally, the case facts are represented by a point on the real number line, $f \in \mathbb{R}$. LC observes f, but UC does not. Second, LC chooses a rule, r, which maps the fact-space into a dichotomous disposition-space.[5] Third, UC observes r and the resulting disposition and must decide whether or not to audit LC's decision. If UC audits the decision, it observes f (resolving the information asymmetry) and is free to select a new rule it chooses.

Strategies. A strategy for the Lower Court is a rule, $r \in \mathbb{R}$; a strategy for the Upper Court is a pairing of an audit probability and a rule to select on auditing, $\langle a(r), r_{UC} \rangle$. Because the Upper Court always has a dominant strategy to select its ideal rule ($r_{UC} = U = 0$) on review, we collapse the Upper Court's strategy into a single choice: the probability of auditing the Lower Court, $a(r)$. Throughout this article, we use search-and-seizure cases as a stylized example. The fact-space represents the "intrusiveness" of a search; the disposition-space is the decision either to admit or exclude the evidence from the search. Specifically, a rule divides the fact-space into dispositions so that all cases with $f \geq r$ are placed into the "exclude" disposition, and all cases with $f < r$ are placed into the "admit" disposition.[6] One can think of the intrusiveness of the search as the extent to which the police invade one's privacy. A very intrusive search is one in which the police knock down the door to one's home and turn the house upside down looking for something without a warrant or probable cause. A less intrusive search might be one in which a police officer approaches a car that has been stopped for a traffic violation and asks the driver if he may inspect the trunk of the car and the driver consents. A rule is a point along the continuum from least intrusive to most intrusive searches at which the court decides that evidence obtained from searches that are less intrusive than that point is admissible and evidence from more intrusive searches is not admissible. The fact- and disposition-spaces are shown visually in Figure 1. Many rules may lead to the same disposition of any given case. For example, in a intrusive search (one with f very far to the right along $\mathbb{R}$), there are many rules to the left of f, which would all yield the "exclude" disposition.

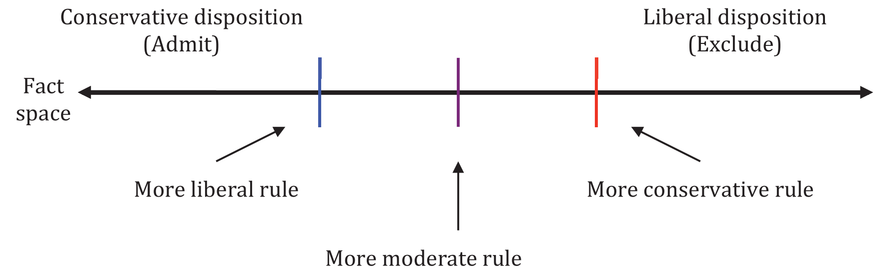
To understand the relationships among rules, facts, and dispositions, consider the following example. In *U.S. v. Katz* (1967), the Supreme Court was asked to consider whether warrantless wiretapping constitutes an unconstitutional search. The Court had previously chosen a rule in *Olmstead v. U.S.* (1928), which held that electronic eavesdropping without a warrant *did not*

[4] In a companion paper, we extend this version of the model to incorporate a non policy dimension into opinion writing (Clark and Carrubba 2012).

[5] Because appellate courts, such as the lower court in our model, that create legal rules generally accept the case facts as found by the trial court, and because entering a disposition that is inconsistent with the rule announced in the opinion as applied to those facts would constitute a serious dereliction of judicial duty, we assume that the disposition must be consistent with the announced rule—that judges cannot "lie" about case facts. The relaxation of this assumption is, in our view, a promising avenue for future research.

[6] We define $f = r$ to induce an "exclude" disposition for analytic convenience (i.e., it avoids the need for asymptotic language).

FIGURE 1. Visual Representation of the Case-space Model



Notes: Facts are represented by the unidimensional line; rules are represented by cutpoints in that line. Cases with case facts greater than the threshold receive the liberal disposition (in our example, exclude); cases with facts less than the threshold receive the conservative disposition (in our example, admit).

constitute an impermissible search. In *Katz*, however, the Court decided that warrantless wiretapping *does* constitute an unconstitutional search. Thus, the Court in *Katz* chose a rule that was further to the left than the rule in *Olmstead*—the *Katz* rule implied that a greater range of searches were too invasive to have their evidence placed into the "admit" category. In other words, the *Katz* rule is one to the left of the *Olmstead* rule, taking cases that would have been mapped into the "admit" disposition and now mapping them into the "exclude" disposition. In this example, the warrantless wiretap constitutes the facts of the case (i.e., the level of intrusiveness), though the two cases do differ in the precise location of the case facts (the exact nature of the search). The *Olmstead* and *Katz* rules constitute the rules that decide which disposition should be given to that set of facts.

Beliefs. The source of uncertainty is the facts of the case, f. Formally, UC's uncertainty about f is characterized by a belief that f is distributed across the fact-space according to an unbounded distribution $g(\cdot)$. For any rule selected by the Lower Court, this uncertainty induces a potential uncertainty about whether the disposition produced by the Lower Court's rule corresponds to the disposition that would be produced by the Upper Court's ideal rule. Let $b(r)$ represent UC's belief that the disposition announced by LC and, induced by rule r, corresponds to UC's preferred disposition.

Utilities. The two players receive utility from the rules they select and whether the resulting disposition results in his or her preferred disposition. A player's ideal rule defines her preferred disposition. Specifically, a player with ideal rule i receives $-(i-r)^2$ from selecting rule r; each player also receives positive utility, $\phi > 0$, if the disposition induced by her selected rule accords with her preferred outcome. In other words, we assume each player receives additive utility from the two components of a decision—the rule and the case disposition. We have assumed a nonstrategic lower court: LC's utility is not affected by any change either to the rule or disposition by UC or any other consequence of being reversed, such as the professional cost of being rebuked by the Supreme Court. Finally, we also assume that UC must pay a cost, k, to audit LC's decision.[7] This cost represents an opportunity or effort associated with auditing the Lower Court; it can also capture other nonspatial features of a case, such as its political or legal salience. The more important (salient legally or politically) the case, the smaller k. We are implicitly assigning a weight of 1 to the rule. The values of ϕ, k, and the ratio of those two parameters therefore tell us the relative weights of the rule, disposition, and the costs of review for the courts. The smaller ϕ and/or k, the greater relative importance of the legal rule to the courts.[8] Each court's utility is an additive combination of the two (Lower Court) or three (Upper Court) relevant components.

Equilibrium. The solution is a unique perfect Bayesian equilibrium in which the Lower Court always chooses its ideal rule and the Upper Court reviews the decision if it finds the rule sufficiently unattractive (i.e., far away from its own ideal rule). What constitutes a sufficiently far away rule is a function of (a) the cost of reviewing the Lower Court, k, and (b) the Upper Court's belief that it dislikes the disposition (it may still be uncertain whether it agrees with the disposition; i.e., whether f falls between LC's rule and UC's ideal rule). The two courts' optimal behavior is driven entirely by our assumption that the Lower Court is nonstrategic and therefore does not care about any move the Upper Court might make.

Proposition 1 (Nonstrategic Lower Court Equilibrium). *The Lower Court's strictly dominant strategy is always to set the rule at its ideal point. The Upper Court reviews the case if the rule is sufficiently far away from the Upper Court's ideal rule to justify the cost of review.*

[7] Without a cost associated with reviewing the Lower Court, it would of course be a weakly dominant strategy for the Upper Court to review all Lower Court decisions.

[8] Alternatively, one might directly model the relative weight of the disposition to the rule by weighting the payoff of the rule by $(1-\phi)$. Therefore as ϕ approaches 1, the disposition becomes relatively more important; as ϕ approaches 0, the rule becomes more important relative to the disposition. Where relevant in the description of the comparative statics, we footnote the effect that this weighting would have.

Comparative Statics. The nonstrategic model yields a series of comparative statics. First, the Upper Court is (weakly) more likely to review a decision when the costs of review (k) are smaller,[9] the Upper Court values the disposition (ϕ) more, or the preference divergence between the Upper and Lower Court is larger (as L increases). Second, the Upper Court is also more likely to reverse a decision when the preference divergence between the courts is larger. As L increases, the range of cases for which the courts disagree about the disposition increases, and therefore a review is more likely to yield a reversal.

Although these predictions are highly intuitive and, in some cases, essentially guaranteed by the strong assumptions we have made about the lower court's preferences, they serve two important functions. First, they establish a benchmark against which we can compare the predictions of our subsequent models. Second, as will become apparent later, they reveal that some of the most common empirical tests in the literature may not be discriminating among competing theories of judicial hierarchy.

Model 2: A Strategic Lower Court

We now consider the possibility of a strategic Lower Court. The only distinction between the strategic lower court model (SLCM) and the nonstrategic lower court model (NSLCM) is that we now assume that *the Lower Court cares about the final disposition and legal rule ultimately applied in its case*. That is, the Lower Court's utility is now determined by both its decision and its interaction with the Upper Court.

Players, Sequence of Play, and Beliefs. The players, sequence of play, and information environment are identical to the baseline model.

Utilities. UC's sources of utility and its utility function are identical to the baseline model. LC's utility, however, now depends on UC's moves. Rather than deriving utility from the rule it selects, LC receives utility from the rule and disposition standing *at the end of the game*. That is, LC receives $-(L-x)^2$, where x is either the rule selected by LC, if UC does not audit the decision, or the rule selected by UC if UC audits the decision. LC receives ϕ if and only if the disposition at the end of the game corresponds to LC's preferred disposition. All components of the utility functions are additive.

Equilibrium. To solve the model, we identify the set of isomorphic pure-strategy perfect Bayesian equilibria—there is a set of equilibria in which the Upper Court and Lower Court strategies are essentially identical functions of the case facts, but the precise rule selected depends on the precise facts of the case.[10] The set of equilibria is summarized next and illustrated in the left-hand panel of Figure 2. In Figure 2a and the statement of Proposition 2, we restrict attention to the case where $L > \sqrt{k}$, because this condition gives rise to the fullest range of possible equilibrium behavior. In the Appendix, we derive the full equilibrium and behavior without restriction on L. All equilibrium behavior discussed later holds independent of the relative location of L, and where relevant, we discuss the implications of $L < \sqrt{k}$.

Depending on the case facts, one of three outcomes obtains. When a search is not intrusive (i.e., case facts are far to the left in Figure 2a, $f < r_a$) the lower court sets a rule at $r^* = r_a$. For moderately intrusive searches ($r_a \le f < \sqrt{k}$), the Lower Court sets the rule at the case facts ($r^* = f$), and when the search is sufficiently intrusive ($f \ge \sqrt{k}$), the Lower Court sets the rule at $r^* = \sqrt{k}$. (As we demonstrate in the Appendix, if $L < r^*$ then the Lower Court sets the rule at L.) The disposition is a direct consequence of the rule, so the disposition is *Admit* if $r^* = r_a$ and *Exclude* otherwise. The Upper Court never reviews the Lower Court. This solution characterizes a set of equilibria because r_a can take on any value between $\underline{r}$ and $\bar{r}$ (we define these boundaries later). That is, the Lower Court and the Upper Court can arbitrarily coordinate on any $r_a \in [\underline{r}, \bar{r}]$.

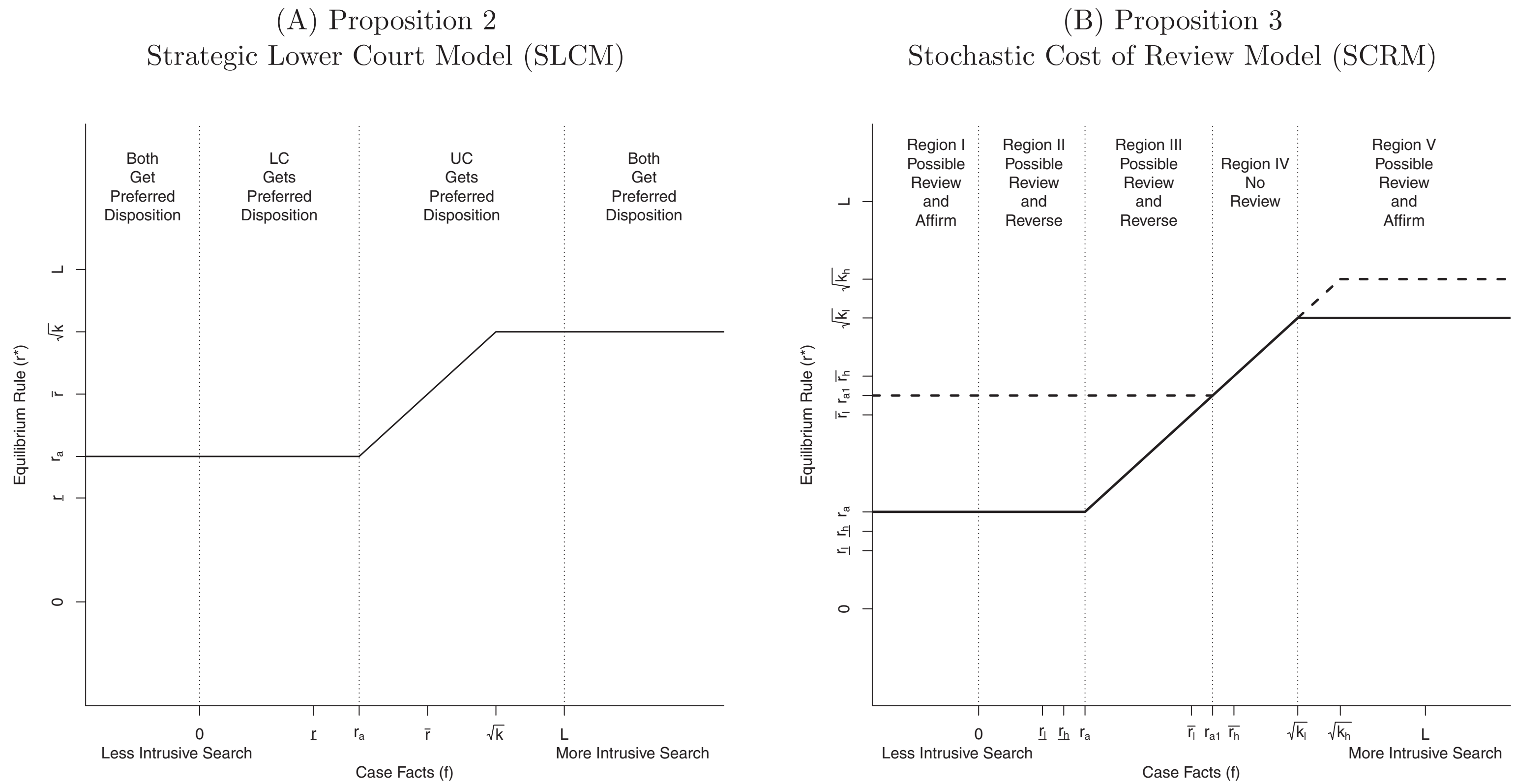
To provide intuition for this equilibrium, we first discuss the choice of rule and then the Upper Court's decision not to review. All else equal, a Lower Court prefers a rule closer to its ideal point than further away, and it prefers a rule that yields its preferred disposition over one that does not. The further the Lower Court's ideal rule from the Upper Court's ideal rule, however, the greater the chance that the rule does not produce the Upper Court's preferred disposition (i.e., there is a larger range of case facts for which the two courts disagree about the disposition). Consequently, the Upper Court has a stronger incentive to review the case. If the Upper Court reviews the case, the Upper Court's dominant strategy is to set the rule at its ideal point, which can cost the Lower Court its preferred disposition if the case facts lie between the courts' ideal rules. This result illustrates the following strategic tension: The Lower Court prefers to avoid review, because review always leaves the Lower Court worse off. Avoiding review requires one of two strategies. One is for the Lower Court to offer a rule that the Upper Court prefers not to review even though the Upper Court is unsure if it is getting its preferred disposition. The alternative strategy is for the Lower Court to reassure the Upper Court that it is getting its preferred disposition so that the Lower Court can pull the rule closer to its ideal point, which means selecting the best rule it can and still yield the Upper Court's preferred disposition.

Now return to Figure 2a. Lower Courts with case facts $f \le 0$ want the same disposition as the Upper

[9] The relationship is a weak one because once the costs of review are sufficiently low, further decreases in the cost of review (holding constant the divergence between L and U) will not further increase the incentive for review.

[10] We focus on pure-strategy equilibria, but conjecture that some mixed-strategy equilibria may exist under knife-edge conditions. We identify in the Appendix what we conjecture those mixed strategies to look like. See footnotes 24 and 25 for the necessary conditions for mixed strategy to exist.

FIGURE 2. Equilibrium Rules in Models 2 (Strategic Lower Court Model) and 3 (Stochastic Cost of Review Model) as a Function of Case Facts



Notes: The x-axis shows the case facts; the y-axis shows the equilibrium rule. A, Strategic Lower Court Model: In the two outside regions, both courts receive their preferred disposition in equilibrium; in the left-center region, the Lower Court receives its preferred disposition; in the right-center region, the Upper Court receives its preferred disposition. B, Stochastic Cost of Review Model: The solid line represents the risk-avoiding strategy; the dotted line represents the risk-seeking strategy (the two are equivalent in Region IV); review happens in Regions I–III and V when the Lower Court plays the risk-seeking strategy and the Upper Court is a low-cost type.

Court, Admit. These Lower Courts ideally would like to take advantage of this situation by signaling that the case disposition corresponds to the Upper Court's preferences and selecting a comparatively favorable rule (from the Lower Court's perspective). An optimal rule under this strategy would be one the Upper Court would have no incentive to review. However, any Lower Court with case facts between the Upper Court's ideal point and the declared rule would want to mimic this rule if the Upper Court is not going to review it. Doing so would again produce a disposition of Admit, but this time only the Lower Court wants this disposition. As a result, the Upper Court is going to be unsure if it agrees with the disposition induced by the Lower Court's rule. Lower Courts with case facts $f \leq r_a$ will pool on a rule in the set $\underline{r}$ to $\bar{r}$. Specifically, let $\underline{r}$ be the rule the Upper Court does not review even if it knows it is getting the undesired disposition and $\bar{r}$ be the rule the Upper Court is indifferent over reviewing, given all Lower Courts with case facts less than $\bar{r}$ are pooling on that rule.[11] There exists an equilibrium for any $r_a \in [\underline{r}, \bar{r}]$. In equilibrium, when the Upper Court observes that the Lower Court plays $r = r_a$ and the disposition Admit, the Upper Court has beliefs that make it weakly prefer to not review the case.

This strategy is not feasible for $f > r_a$. If the Upper Court observes $r > r_a$ and Admit, the Upper Court would prefer to review the case. Thus, Lower Courts with case facts greater than r_a make a different choice. If these courts choose the rule r_a they will produce the disposition Exclude. Because the Upper Court knows that the Lower Court wants a more permissive standard (i.e., a rule that leads to more Admit dispositions), the Upper Court knows it wants Exclude whenever the Lower Court makes that determination. Thus, the Lower Court is in a position to choose a more favorable rule than r_a. In fact, the Lower Court can choose any rule that produces the Exclude disposition up to one that the Upper Court would review solely to bring the rule back into line, $r = \sqrt{k}$. (The Upper Court prefers to pay the cost of review, k, to move any rule $r > \sqrt{k}$ to its own ideal rule, regardless of the disposition.) If the case facts are less than $\sqrt{k}$, the Lower Court sets the rule at f (any rule to the right of f would produce the Admit disposition and get reviewed). If $f > \sqrt{k}$, the best the Lower Court can do is to set the rule at $\sqrt{k}$; any rule greater than $\sqrt{k}$ will always be reviewed. In each case, however, if the Lower Court can select its preferred rule and reach an Exclude disposition (i.e., if $L \leq f$), then the Lower Court will select its preferred rule. As the Lower Court's ideal point, L, approaches the Upper Court's ideal point, 0, the range of case facts for which the Lower Court can select its ideal point expands and overtakes the regions where the Lower Court must concede somewhat on the legal rule.

To develop a substantive intuition for how the Lower Court might choose to balance a trade-off between the rule used and the case disposition, consider the question of whether to allow a warrantless search of a mobile home. This issue arose in a series of court of appeals decisions during the 1980s. The primary issue at stake was to determine which mobile homes should be treated like homes and which should be treated like cars for the purposes of a search. In terms of our model, the question is where in the fact-space to divide cases into those in which a warrantless search is permissible ("admit" disposition) and those in which a warrantless search is impermissible ("exclude" disposition). The Sixth Circuit Court of Appeals had an opportunity to weigh in on this issue in *U.S. v. Markham* (1988). In that case, the Court had to decide whether to allow a warrantless search of a mobile home. The mobile home was from out of state, had license plates on it, and had no utilities connected. The court of appeals allowed the search (it chose the more conservative "admit" disposition), but in doing so it adopted a more liberal rule than was necessary to reach the conclusion. In particular, the court held that if obtaining a warrant is not an unreasonable burden, then the law enforcement officials should do so. However, in this case, the facts were sufficiently inoffensive—the mobile home was clearly more of a car than a home—that a warrant was not necessary. The rule it used to reach that disposition was a more liberal rule than it needed to be, because it stated that under different circumstances, where the mobile home was more of a home than a car, law enforcement officials should seek to obtain a search warrant if possible. The court of appeals could have opted for a more conservative rule that simply held that searches of mobile homes fall under the "automobile exception" and do not require a warrant and still reached the "admit" disposition. This is notable, because the court of appeals panel was composed of three Reagan appointees who may reasonably have believed the Supreme Court was more liberal than the panel. In fact, in its opinion, the court of appeals explicitly noted the more liberal views of the Supreme Court, citing Justice Stevens' dissent in a previous mobile home search case. Thus, this example illustrates how the more conservative lower court adopted a relatively liberal rule while reaching a conservative disposition, thereby trading off the disposition and rule components of its decision. The Supreme Court denied a petition for certiorari in this case.

Notice that the Upper Court never reviews in equilibrium. When case facts are to the left of r_a, the Lower Court pools on a rule that the Upper Court prefers not to review, even though the Upper Court is unsure whether it agrees with the disposition induced by the Lower Court's rule. When case facts are to the right, the Lower Court and Upper Court know that they agree on the disposition. The Lower Court uses this slack to declare a rule that is as close to its ideal point as possible, but that the Upper Court still prefers not to review. Thus, the threat of review shapes the rule the Lower Court will apply. Because the Lower Court can perfectly anticipate the Supreme Court's decision, it can always choose a rule that is immune from review.

[11] It is important to note our model does not yield any predictions about which rule from the set will be selected as r_a. Rather, all we predict is that the Lower Court will coordinate on a single r_a from the set $[\underline{r}, \bar{r}]$.

Proposition 2 (Strategic Lower Court Equilibrium). *In any equilibrium to the SLCM, the rule selected by the Lower Court depends on the facts of the case. For case facts far enough to the left, the Lower Court pools on a rule that yields the "admit" disposition. For middle-range case facts, the Lower Court chooses a unique rule for each value of the case facts, where the rule yields the "exclude" disposition. For case facts far enough to the right, the Lower Court pools on the best rule the Upper Court will not review, which yields the "exclude" disposition. In equilibrium, the Upper Court never reviews the Lower Court decision.*

Model 3: Uncertainty about Upper Court Preferences

Although no review happens in equilibrium in the SLCM, we might easily imagine that review could happen if the Lower Court were uncertain about the Upper Court's preferences. Indeed, a standard feature of principal-agent models is that increasing uncertainty about the principal's preferences increases the chance of auditing. Simply assuming that observed review is a function of such uncertainty, however, is insufficient. Specifically, we desire to know two things. First, in what ways does review vary with such uncertainty? Second, and more important for our purposes, how might such uncertainty affect the law-creation choices made by the Lower Court? To answer these questions, we allow for uncertainty about Upper Court preferences by assuming that the Upper Court's cost to review a case is unknown to the Lower Court.[12]

As we demonstrate, Lower Court behavior in this model is almost identical to Lower Court behavior in the SLCM. The purchase from the stochastic cost of review model (SCRM) concerns Upper Court behavior. Whereas in the SLCM review never occurs because the Lower Court can anticipate the Upper Court's preferences perfectly and so can avoid review, the SCRM allows auditing to take place in equilibrium. As we describe later, we derive not just the possibility of review but also a variety of counterintuitive equilibrium relationships in the use of review.

Players, Sequence of Play, and Beliefs. The players and sequence of play in this model are essentially the same as in the preceding model, except that we now assume that the model begins with Nature drawing a "type" for the Upper Court. Specifically, Nature selects $i \in \{l, h\}$, which is private information to the Upper Court. The Upper Court's type determines the Upper Court's cost to review a Lower Court decision. Substantively, an Upper Court with low review costs can be thought of as a Court that cares a great deal about a case below or finds that reviewing a case is less taxing of its resources (not that the disposition or rule is more valuable but simply that the effort associated with reviewing the case is less).[13] The Upper Court's uncertainty about the facts of the case is characterized by the same beliefs as described earlier. The Lower Court's uncertainty about the Upper Court's type, i, is characterized by a belief $p = \Pr(i = l)$.

Utilities. The Lower Court's utility function is the same as in the SLCM. The Upper Court's utility function is the same except that the cost of reviewing a Lower Court decision is now indexed by the Upper Court's type, with $k_l < k_h$.

Equilibrium. To solve the model, we again identify the set of isomorphic pure-strategy perfect Bayesian equilibria.[14] The set of equilibria is characterized in Figure 2b.[15] As in the previous model, we again restrict attention to the case where $L > \sqrt{k}$ in the formal statement of the equilibrium and Figure 2b. We also assume a particular ordering of the various cutpoints; these restrictions are for illustrative purposes and do not affect the equilibria or their substantive interpretation.

As Figure 2b reveals, the strategies are isomorphic to the equilibrium strategy from the SLCM. The critical difference is that, if the Lower Court believes it is sufficiently likely that the Upper Court is a high-cost type (that is, p is sufficiently low), then the Lower Court will play the risk-seeking strategy that involves choosing a rule that an Upper Court with low costs would review. The intuition behind this result is straightforward: If the Lower Court believes the Upper Court is a high-cost type, then it will be willing to play a strategy that the high-cost type would find acceptable, but the low-cost type is willing to pay the cost to review.

Consider first Regions I and II in Figure 2b. In these regions, the Lower Court always pools on a rule—which rule it pools on is determined by the Lower Court's decision to pursue either the risk-seeking or risk-avoiding strategy. If it pursues the risk-seeking strategy (the dotted line), review occurs with positive probability. Sometimes the Upper Court will reverse the Lower Court (Region II), but sometimes it will affirm the Lower Court (Region I). The logic here is that the Upper Court is strategically auditing these cases—it prefers to review the case when it believes it is sufficiently

[12] We allowed for uncertainty over the cost of review rather than the Upper Court's preferred rule for reasons related to why we chose to allow for uncertainty over case facts. Judicial preferences are fairly stable and well known among the judiciary compared to the idiosyncratic factors that can influence treatment of any given case. For example, between the time a lower court issues a ruling and the Supreme Court is deciding whether to hear the case, the Supreme Court's docket might fill up, or not, with other pressing issues. Similarly, social, economic, or political factors can intervene and raise or lower the stakes associated with hearing that case. Thus, we assume the locus of uncertainty lies with the cost of reviewing the case rather than with justices' preferred rules.

[13] This latter case may occur, for example, when the Upper Court can consolidate a number of cases from below, effectively dividing the cost of review by the number of cases collected together.

[14] We again focus on pure-strategy equilibria, but conjecture that some mixed-strategy equilibria may exist under knife-edge conditions. In the Appendix we identify what we conjecture those mixed strategies to look like.

[15] The ordering of the cutpoints $\underline{r_h}$ and $\overline{r_l}$ is arbitrary. Specifically, $\underline{r_h} \leq \overline{r_l}$ whenever $b \geq \frac{k_h - k_l}{\phi}$. Because b can vary arbitrarily (for example, as a function of $g(\cdot)$), the specific ordering can vary arbitrarily.

likely that it has not received its preferred disposition. As in all situations of auditing under incomplete information, there will be instances in which the Upper Court reviews and learns that it would have preferred not to review (i.e., it prefers not to reverse the Lower Court). If the Upper Court had known it was receiving its preferred disposition (i.e., $f < 0$), then it would have been willing to accept the Lower Court's rule.

In Region III, a Lower Court playing the risk-avoiding strategy will switch from the pooling strategy to the separating strategy, giving the Upper Court its preferred disposition while choosing the best rule it can (from its own perspective) to reach that disposition. Meanwhile, a Lower Court playing the risk-seeking strategy will continue to pool on a comparably favorable rule that leads to its preferred disposition. A low-cost Upper Court will find this combination of an unfavorable rule and uncertainty about whether the disposition accords with its preferences worth the cost of review. If the Upper Court is a low-cost type then it reviews and reverses the Lower Court.

In Region IV, both the risk-seeking and the risk-avoiding strategies prescribe the same equilibrium rule: the separating strategy that induces the Upper Court's preferred disposition using the best rule from the Lower Court's perspective. In this region, review never takes place.

Finally, in Region V, if the Lower Court pursues the risk-seeking strategy (i.e., it continues the separating strategy through $\sqrt{k_h}$ and then pools on $\sqrt{k_h}$), then review happens with positive probability (whenever the Upper Court is a low-cost type). A remarkable feature of the Upper Court's strategy in this case is that it is reviewing cases it *intends* to affirm. This is because, in contrast with standard principal-agent models of the judiciary (e.g., Cameron, Segal, and Songer 2000; Clark 2009; Haire, Songer, and Lindquist 2003), in our model the Supreme Court cares not just about the case disposition but also about the rule used to reach that disposition. As a consequence, the Court will sometimes review cases where it believes it agrees with the disposition but is willing to spend the effort to change the rule used by the Lower Court. That is, in equilibrium, affirmances are not simply the product of cases where the Upper Court would have preferred ex post not to have reviewed the Lower Court; rather they can occur intentionally. Of course, this dynamic is distinct from models that consider multiple courts and the possibility of conflict among lower courts, where the Supreme Court may take a case it intends to affirm in order to resolve conflict among multiple lower courts. Such dynamics cannot arise in our model, because we have assumed a single lower court. Thus, perhaps surprisingly, our model predicts this behavior *independent* of those dynamics.

That final feature of the equilibrium is particularly striking—consider a final substantive illustration. In *Delaware v. Prouse* (1979), the Supreme Court was asked to decide whether a police officer violated the Fourth Amendment when he stopped a car simply to check the driver's registration and then observed marijuana on the floor of the car, seized the marijuana, and arrested the driver. The lower court, the Delaware Supreme Court, applied a rule that included a relatively limited view of the Fourth Amendment, resting its decision to exclude the evidence on the Delaware Constitution's protections. The Supreme Court chose to review the case to make clear that the Federal Constitution's Fourth Amendment indeed prohibited the search in this case. The Supreme Court affirmed the Delaware Supreme Court's decision to exclude the evidence, but applied a rule more protective of privacy. Thus, the Supreme Court intervened to review a case, the disposition with which it agreed, in order to move the rule closer to its own ideal point (further to the left), thereby offering a wider range of privacy protection than the rule applied by the lower court would have.

So long as the Lower Court plays the risk-avoiding strategy, it can always avoid review. Although it may sometimes concede more than it needs to in order to avoid review (it will choose a rule acceptable to the low-cost type when it faces a high-cost type), this strategy will never be reviewed by either type. When it plays the risk-seeking strategy, however, it runs the chance that it faces a low cost type and therefore will be reviewed when case facts are far to the left or right (when the two strategies do not prescribe the same equilibrium rule).

Proposition 3 (Stochastic Cost of Review Equilibrium). *In any equilibrium to the SCRM behavior is characterized as follows. The Lower Court's strategy is isomorphic to its strategy in the SLCM; however, the Lower Court must choose between a risk-averse and risk-seeking strategy. The only difference between the risk-seeking and risk-avoiding strategies is that the risk-seeking strategy prescribes a rule that is (weakly) closer to the Lower Court's ideal rule than the risk-avoiding strategy. The Lower Court chooses the risk-seeking strategy if it believes, with sufficiently high probability, that the Upper Court is a high-cost type. The Upper Court reviews the Lower Court's decision if it is a low cost type, the Lower Court has played a risk-seeking strategy, and the case facts are either sufficiently far to the left or sufficiently far to the right.*

IMPLICATIONS

The models developed in the preceding section yield a variety of implications about lawmaking in a judicial hierarchy—some of which are very natural and intuitive, others of which represent previously unappreciated insights—as well as implications for other institutions. First, they provide implications concerning the degree to which lower courts can influence the content of judge-made law. Second, they provide novel implications for the conditions under which and the ways in which superior courts oversee their inferior court's decision making. Third, our models provide insights that inform the microfoundations of some common assumptions used in theoretical models of strategic auditing. Finally, they provide implications for the kinds of evidence scholars can use to empirically evaluate theories of lawmaking in a judicial hierarchy.

Lower Court Influence on Legal Rules

Previous studies of the judicial hierarchy generally have not focused explicitly on the choice of rules by lower courts (but see Klein 2002). Instead, most theories focus on whether case dispositions are influenced by Supreme Court preferences (e.g., Cameron, Segal, and Songer 2000; Hettinger, Lindquist, and Martinek 2006). Two of our results about the influence that lower courts exercise over the legal rules they apply bear discussion. We first note a completely natural result: As the cost of reviewing a case increases (for example, as the Supreme Court becomes less interested in a particular area of the law), the Lower Court is able to apply rules that are closer to its own ideal rules.[16] The intuition is straightforward: The more costly the review, the more dissatisfied the Upper Court has to be with the Lower Court's decision for it to hear an appeal, and therefore the more slack the Lower Court has in setting a rule. This result, of course, is quite intuitive.

Second, our analysis shows that the Lower Court's proposed rule depends on the facts in the case. In particular, the Lower Court will be able to choose to apply legal rules that are more consistent with its own ideal rule when the context of the particular case (in our models, the facts of the case) allows it to do so while inducing a case disposition that is consistent with the Supreme Court's preferences.[17] Thus, in our search-and-seizure example, a liberal lower court is going to be able to issue a more desirable rule (from its perspective) when the police engage in an extremely intrusive search than when they engage in a less intrusive one. We note that this finding is consistent with the model developed in Cameron, Segal, and Songer (2000), but our model yields insights regarding the rule the Lower Court will apply and makes this relationship explicit.

A related but counterintuitive finding is that the more the Upper Court cares about the disposition relative to the rule being produced, the *less* influence the Lower Court has on the rule. So, if the Lower Court is making a decision in a case where the rule has broad implications relative to the significance of the particular case's outcome, the Lower Court actually has more freedom to sculpt a desirable rule (from its own perspective) than if the case is primarily important for the consequences of its immediate outcome. For example, how the courts weight the significance of the decision to invalidate a major piece of legislation (i.e., will the Social Security Act stand or fall) relative to the significance of the rule used to reach that decision (what types of federal taxes are permissible) can have an unexpected and counterintuitive effect on the courts' influence over the rule. To see why, notice first that the Upper Court's intensity of preferences over the disposition (ϕ) has no impact on the rule when the case facts are sufficiently far right ($f > r_a$—in our search-and-seizure example, when a search is sufficiently intrusive). In this case the Upper Court knows it is getting its preferred disposition, and the Lower Court is simply pulling the rule toward its ideal point as much as possible. How much it can do so solely depends on the costs of review. The disposition comes into play when the case facts are sufficiently close to the Upper Court's most preferred rule ($f \leq r_a$—in our search-and-seizure example, when a search is sufficiently not intrusive). Here, the more the Upper Court cares about the disposition, the stronger the incentive for the Upper Court to review a case where it is unsure whether the disposition is in line with its own preferences (i.e., whether it agrees with the disposition induced by the rule applied in the Lower Court). As a consequence, the Lower Court must apply a rule that is sufficiently attractive to the Upper Court to offset that uncertainty. As the Upper Court cares more about the disposition, that uncertainty becomes increasingly unattractive for the Upper Court, requiring an increasingly attractive rule to offset the risk of having a "bad disposition." Thus, the more the Upper Court cares about the disposition relative to the rule, the (weakly) less influence the Lower Court has on the law.[18]

Supreme Court Oversight

Of course, the models' implications for lower court influence over the law are just one side of the equation. The SCRM model also provides an account of how a strategic upper court chooses to review rules produced by strategic lower courts. These results are particularly instructive because review and reversal of lower courts are central behaviors at the core of studies of compliance and judicial hierarchy. On this point, our analysis of the SCRM reveals a series of novel implications that bear discussion. First, we uncover two non-monotonic relationships concerning Supreme Court review and reversal of lower court decisions. These findings contrast with existing literature, which generally predicts (weakly) monotonic relationships between upper-lower court preference divergence and the rate of both review and reversal (e.g., Cameron, Segal, and Songer 2000; Clark 2009; Lindquist, Haire, and Songer 2007).

First, we find a non-monotonic relationship regarding the Upper Court's decision to review the Lower Court. To see this relationship, note that as the Lower Court's ideal rule moves from the Upper Court's ideal rule, the various regions depicted in Figure 2b come into existence. As the Lower Court moves right from the Upper Court's ideal rule, it enters Regions II and then III. In these regions there is a constant rate of probabilistic review. Further along, however, the

[16] As k increases, $\underline{r}$, $\bar{r}$, and $\sqrt{k}$ all increase.

[17] To see why, recall from Figure 2 that, as the case facts move to the right, they allow the Lower Court to apply legal rules that are more consistent with the Lower Court's ideal rule while still yielding the Upper Court's preferred rule. Note that this result can only hold because we allow the courts to care about both rules and dispositions.

[18] As noted in footnote 8, one might alternatively weight the rule by $(1-\phi)$. In that case, $\underline{r} = \sqrt{\frac{k-\phi}{1-\phi}}$, and $\bar{r} = \sqrt{\frac{k-(1-b)\phi}{1-\phi}}$. For k sufficiently small, (e.g., $k < 1$ for $\underline{r}$), the same comparative statics hold. Therefore, only if the costs of review are relatively large would this alternative specification lead to different expectations.

Lower Court enters Region IV, where review does not occur. Then, as the Lower Court continues to diverge, it enters Region V, at which point the Upper Court again begins to review the Lower Court (probabilistically). Thus, the effect of increasing the divergence between the Lower Court and the Upper Court is first to decrease and then to increase the rate of review. In other words, because review takes place in Regions I–III and V (Figure 2b), but not in region IV, there is a non-monotonic relationship between Lower Court-Upper Court preference alignment and the probability that the Upper Court reviews a Lower Court decision.

We also find a comparable non-monotonic relationship concerning the Upper Court's decision to reverse a Lower Court decision. Again, as the Lower Court's ideal rule moves away from the Upper Court's ideal rule through Regions II and III, the probability that the Lower Court declares a disposition with which the Upper Court disagrees increases. As a result, conditional on reviewing the Lower Court, the probability that the Upper Court will reverse the disposition induced by the Lower Court's decision increases. Once the Lower Court's ideal rule enters Region V, however, the Upper Court is reviewing not because it is suspicious of the disposition, but rather because it finds the applied rule sufficiently distasteful (see Figure 2b). Thus, our model yields a second non-monotonic relationship: As the Lower Court diverges from the Upper Court, the rate of reversal should initially increase and then begin to decrease.

In addition to these two relationships, our model also demonstrates that the value of the case disposition affects patterns of review and reversal in unexpected ways. In particular, as the value the Upper Court places on the case disposition increases, so too does the ratio of affirmances to reversals. One might suppose that this relationship is the result of a simple expected utility calculation: The more the Upper Court cares about the disposition, the more cases it is willing to review to catch ones it wants to reverse. However, such behavior cannot arise in equilibrium. The rules it reviews to catch "bad" dispositions are those that yield an "admit" (i.e., rules that are relatively attractive but leave the Upper Court uncertain about whether it agrees with the disposition). Therefore, the Upper Court could review more cases overall, but it cannot affect the rate of affirmances to reversals. Rather, this result arises in our model for a different, less intuitive, reason. Consider again Figure 2b; as the value of the case disposition (ϕ) increases, so too does the lower boundary on Region IV. Thus, increasing the value of the disposition increases the range of case facts for which the Lower Court plays a strategy that is not reviewed by either type of Upper Court. (As a consequence, the range of cases that will be probabilistically reversed decreases, whereas the range of case facts that will be probabilistically affirmed remains constant.) The intuition behind this result is straightforward: Following from our previous discussion, as the Upper Court cares more about the disposition, the Lower Court is constrained in which rules it can use to reach its own preferred disposition. In particular, the Lower Court can only leave the Upper Court uncertain about whether it agrees with the disposition if it is applying a sufficiently attractive rule (from the Upper Court's perspective). This constraint causes the Lower Court to apply rules that induce the Upper Court's preferred disposition for a greater range of case facts. The Upper Court is therefore able to allocate its effort to cases in which it wants to affirm the Lower Court and simply revise the rule the Lower Court has applied.

Microfoundations for Common Theoretical Assumptions

Models of strategic auditing commonly assume an exogenous cost to the lower courts associated with being reversed. This cost is variably framed as a reputation cost or a cost associated with losing the case (e.g., Cameron, Segal, and Songer 2000; Clark 2009; Kastellec 2007). However, past research has criticized this modeling assumption (e.g., Cross 2005; Klein and Hume 2003). Our model assumes no such cost; instead the cost of review arises endogenously from our model. Here, lower courts seek to avoid being reviewed in order to exercise control over the law; review is costly because it takes their influence and control away from them. Because judges surely have an interest in exercising influence over the law (Posner 1995), our model helps provide deeper microfoundations for models of compliance that make use of a cost associated with being reversed.

Implications for Empirical Analyses

A final set of implications that derives from our theoretical models concerns the empirical analysis of compliance in the judicial hierarchy. The first, and perhaps most surprising, finding is that the Upper Court reviews cases in equilibrium that it *intends* to affirm. This is because, in contrast with standard principal-agent models of the judiciary (e.g., Cameron, Segal, and Songer 2000; Clark 2009; Haire, Songer, and Lindquist 2003), in our model the Supreme Court cares not just about the case disposition but also about the rule used to reach that disposition. As a consequence, the Court will sometimes review cases where it believes it agrees with the disposition, but is willing to spend the effort to change the rule used by the Lower Court. That is, in equilibrium, affirmances are not simply the product of cases where the Upper Court would have preferred ex post not to have reviewed the Lower Court (of course, such affirmances do occur in our model, as in Region I of Figure 2). Instead, affirmances sometimes occur intentionally. Our model is the first, to our knowledge, that predicts such behavior.[19] This result is particularly instructive, given that some research has suggested that relying on affirmances to understand

[19] Of course, this is not to claim that our model provides a full explanation for why Supreme Court cases lower court decisions. For example, one might imagine that the Supreme Court affirms lower court decisions to entrench a given proposition as a national precedent.

the Supreme Court's motivations may be misleading, because affirmances occur primarily because of incomplete information on the part of the Supreme Court (McGuire et al. 2009).

Second, our analysis highlights the need to account for the dual components of judicial decision making—dispositions and legal rules—when evaluating compliance. Our analysis demonstrates that the relationship between the legal rule and the disposition is complicated. Although scholars have been widely interested in legal rules, at least theoretically (e.g., Lax 2007; Staton and Vanberg 2008), they have paid less attention to them empirically (but see, Clark and Lauderdale 2010; Klein 2002). Instead, as a proxy, studies of compliance have often used case dispositions (e.g., Cameron, Segal, and Songer 2000; Clark 2009). In the context of judicial hierarchy, our models demonstrate that case disposition can be quite misleading. In fact, as the Lower Court moves the rule further away from the Upper Court's preferred rule, it must compensate by doing so in cases that allow it to give the Upper Court its preferred disposition. Thus, the disposition can actually be *inversely* correlated with the legal rule used by the Lower Court. Moreover, if we compare the nonstrategic lower court model with the SLCM and SCRM, we find that, in each model, the rule applied by the Lower Court is monotonically moving away from the Upper Court's preferred rule as a function of preference divergence between the Lower and Upper Court. Thus, if one were to examine the relationship between the rules adopted by the lower courts and their underlying preferences (e.g., Klein 2002), it might appear that the lower courts are not sensitive to Supreme Court preferences. However, because we find such a correlation in both the nonstrategic and strategic versions of our model, this relationship is insufficient to establish a claim that lower courts do not strategically anticipate the Supreme Court's intervention. In other words, finding, for example, that a lower court judge is more likely to make a liberal decision as he or she becomes more liberal relative to the Supreme Court is observationally equivalent across the nonstrategic and strategic versions of our model. This implies that very natural empirical tests of strategic behavior cannot necessarily discriminate among those different sets of premises about judicial preferences.

Implications beyond the Supreme Court

Although our argument has been cast in the context of the U.S. Supreme Court, the formal models and their insights should be applicable to a variety of other institutional settings. Most obviously, the basic argument closely maps to any judicial system with a hierarchical judicial structure in which the lower courts decide not just the decision in the case but also the underlying rule (rationale) justifying that decision. In virtually all advanced industrial democracies with judicial review, there is a judicial hierarchy in which the bulk of cases are handled by lower courts that have to reach dispositions and provide rationales (rules) for their decisions and are subject to review by the supreme tribunal. We also observe this type of structure in some international settings, including the European Union (since 1989 when the Court of First Instance was created to carry some of the European Court of Justice's caseload). More generally, these features can be found in settings outside of the judicial hierarchy. In particular, frequently administrative agencies issue and apply rules under the shadow of possible review and reversal by one or more principals, including presidents, legislatures, and even judiciaries. For example, when a legislature observes the implementation of a rule by the EPA, such as the standard it chooses to decide whether to issue a permit to build on a wetlands, that overseeing legislature observes the rule used and the outcome (was the permit issued), but does not know with certainty if it, too, would do the same thing. It can then choose whether to invest the time and energy to hold hearings, investigate the decision, and undertake the effort of rewriting law or otherwise working to change the agency's decision. Indeed, similar dynamics are at work in a wide array of administrative agencies engaged in rule making. Thus, our model suggests broader implications for the study of rule making in a political hierarchy beyond the particular setting within which we have developed our argument.

In sum, our models yield a series of implications concerning lower court influence over rule creation and application, Supreme Court oversight of lower court decision making, standard modeling approaches to judicial compliance, and empirical strategies for evaluating theories of compliance. Some of these implications conform to well-documented empirical regularities and theoretical results, whereas others offer novel insights that have not been previously appreciated.

EMPIRICAL ANALYSIS

In this section, we provide an empirical analysis of the third model developed earlier. Specifically, we consider two of the substantively most intriguing and counter intuitive relationships in the model: the relationships between the ideological divergence of the lower court and the likelihood of (i) review and (ii) reversal. First, as noted in the preceding section, our model predicts a non-monotonic relationship between lower court-Supreme Court divergence and the probability that a decision is reviewed.

Hypothesis 1 (Non-Monotonic Review). *Ideological distance between the Supreme Court and the lower court has a non-monotonic effect on the probability of review. The likelihood of review should initially decrease and then increase in ideological distance.*

We note that this prediction contrasts with the nonstrategic lower court model, which predicts a monotonically increasing rate of review in ideological divergence. Second, as discussed earlier, our model also predicts a non-monotonic relationship between lower court-Supreme Court divergence and the probability that a decision is reviewed.

Hypothesis 2 (Non-Monotonic Reversal). *Ideological distance between the Supreme Court and the lower court has a non-monotonic effect on the probability of reversal. The likelihood of reversal should initially increase and then decrease in ideological distance.*

We note that this prediction also contrasts with the non-strategic lower court model, which predicts a monotonic relationship between ideological divergence and the probability of reversal.

Data

To test these two predictions, we require data on cases decided by a rule-making/rule-applying lower court and decisions by a superior court to review those decisions (as well as whether to reverse them). We compiled a dataset from the U.S. Courts of Appeals (the mid-level court in the federal judiciary) and the U.S. Supreme Court. These data for our analysis came from several sources. First, to create a dataset of cases, we combined the Phase I and Phase II of the U.S. Courts of Appeals Database . Phase I is a dataset containing a random sample of cases decided by the court of appeals between 1925 and 1996; Phase II is a dataset containing every court of appeals decision subsequently reviewed by the Supreme Court. By combining these two databases, we had a set of cases, some of which were reviewed by the Supreme Court, providing the variation needed to test Hypothesis 1. However, we also needed to know whether the Supreme Court ultimately affirmed or reversed the Court of Appeals decision, and this information is not provided in the Phase II database. It is provided, however, in the U.S. Supreme Court Judicial Database. Unfortunately, there is no link between these databases. Thus, we traced every case in the Phase II database to its Supreme Court decision and linked the Supreme Court database with our combined Court of Appeals database.[20] We then created two variables, $reviewed_i$ and $reversed_i$, which equal 0 if the Supreme Court did not review or did not reverse case i, and 1 if the Supreme Court did review or did reverse case i (conditional on review).[21] The resulting dataset included 9,485 cases; the Supreme Court reviewed 1,854 (just under 20%) of the cases and reversed the lower court 921 times (just under 50%).

The key predictor of review and reversal decisions is the degree of ideological preference divergence between the two courts. To measure preference divergence between the circuit court panel that decided the case and the Supreme Court, we employed the Judicial Common Space scores (Epstein et al. 2007) of the panel median and the Supreme Court median during the year the panel made its decision. These scores are transformed versions of the Martin and Quinn (2002) ideal point estimates of Supreme Court justices and the Giles, Hettinger, and Peppers (2002) estimates of courts of appeals judges' preferences (in the Common Space dimension). We created a variable $distance_i$, which equals the absolute difference between the panel median and Supreme Court median for case i. The observed range of $distance_i$ is from 0 to 0.83; the average distance between a lower court panel and the Supreme Court is 0.27 (sd = 0.18).

The Raw Data

We begin by considering the raw data. Figure 3 shows a nonlinear scatterplot smoother (loess) for each of the two relationships. Specifically, it shows the relationship between the ideological distance from the lower court panel to the Supreme Court and the probability of review and of reversal. A striking pattern emerges: The probability of review initially *decreases* in ideological distance and then *increases*. By contrast, but again as predicted, the probability of reversal initially *increases* in ideological distance and then *decreases*. These patterns in the raw data strongly support the two non-monotonic predictions described by Hypotheses 1 and 2. Indeed, it is striking that the nonparametric fit to these data conforms so closely to the non-monotonic relationships predicted by the models of strategic lower courts. Moreover, these relationships are inconsistent with the model of a nonstrategic lower court.

Estimation

One feature of these data bears attention. Our data constitute a choice-based sample—we have the universe of cases reviewed by the Supreme Court, but only a sample of the cases not reviewed. As a consequence, we worried that bias in the sample would lead to bias in the estimated relationship among variables in the data; to correct for this, we must weight each observations appropriately (Greene 2003, 673). Specifically, the cases heard by the Supreme Court were overrepresented by a factor of 5.1; the cases not reviewed were underrepresented by a factor of 0.8. To account for the nature of our sample, we employed a quasi-likelihood empirical model in which each observation is weighted according to its under- or over-representation.

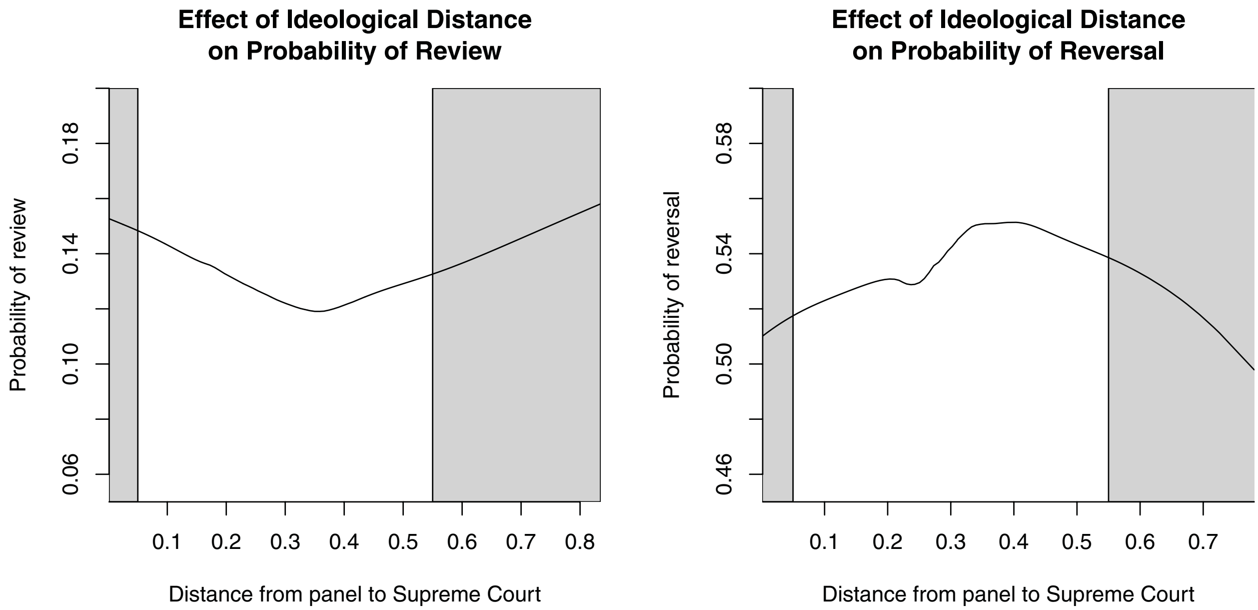
We specified an empirical model that considers the relationship between the lower court's ideological divergence from the Supreme Court and the decisions to review and reverse the lower court.[22] Specifically, we empirically modeled the Supreme Court's two related

[20] We discarded from our data all court of appeals cases not decided by three-judge panels.

[21] To create the reversed variable, we referenced the dis variable in the Spaeth database. We coded our variable as 0 if dis=affirmed and 1 if dis=reversed or reversed and remanded. We coded *reversed* as missing if dis is assigned any other value, because those values are more ambiguous in their substantive interpretation in the context of our model.

[22] A structural estimation of the model would require us to fully operationalize all of the model's parameters and behaviors to assess from which region in Figure 2 each observation is made; that poses many measurement challenges beyond the scope of this article. For example, we have no easy way of measuring the value of the disposition or the location of the lower court's rule. Thus, we focused on modeling the implications captured by Hypotheses 1 and 2.

FIGURE 3. Patterns in the Raw Data



Notes: The left-hand panel shows the relationship between the probability of review and the degree of ideological distance between the lower court and the Supreme Court. The right-hand panel shows the relationship between the probability of reversal and the degree of ideological distance between the lower panel and the Supreme Court. In both panels, the line is a non-monotonic scatterplot smoother (loess). Because of the case-control sampling strategy, it is important not to draw inferences about the scale of the level of review. Gray regions show bottom and top 10% of distribution of observed distance_i.

decisions. First, it must decide whether to review a case and, then if so, whether to affirm or reverse the case. To do so, we specified a likelihood function customized to the Supreme Court's choice set.

Following our hypotheses, we modeled the probability that the Supreme Court reviews a case as a function of the ideological distance between the Supreme Court and the panel that decided the case. To capture the hypothesized non-monotonic relationship between ideological distance and the Supreme Court's decisions, we included both a measure of the absolute distance between the Supreme Court and the lower court and the square of that measure. (There are, of course, many ways of modeling the non-monotonicity, but the relationships uncovered in Figure 3 suggest that a parabolic relationship is a good approximation.) We modeled the probability a case is reviewed as a logistic function of these covariates as follows:

$$\begin{aligned} f(\mathbf{X_i}) &= \Pr(review_i = 1 | distance_i) \\ &= \text{logit}^{-1}\left(\beta_0 + \beta_1 distance_i + \beta_2 distance_i^2\right) \end{aligned}$$

Similarly, we modeled the probability that a case is reversed, on review, as a function of the ideological distance between the lower court and the Supreme Court. Formally, we modeled the probability of reversal as a logistic function of these covariates as follows:

$$\begin{aligned} g(\mathbf{Z_i}) &= \Pr(reversed_i = 1 | distance_i) \\ &= \text{logit}^{-1}\left(\gamma_0 + \gamma_1 distance_i + \gamma_2 distance_i^2\right) \end{aligned}$$

Again, we included both distance_i and $\text{distance}_\text{i}^2$ as predictors of review to capture the nonmonotonic relationship predicted by Hypothesis 2. Given these representations, the likelihood model for the data is given by

$$\begin{aligned} \mathcal{L}(\mathbf{y}|\mathbf{X_i}, \mathbf{Z_i}) = \prod_{i=1}^{n} & w_i \cdot (1 - f(\mathbf{X_i}))^{(1-reviewed_i)} \\ & \times \left[(f(\mathbf{X_i}) \cdot g(\mathbf{Z_i}))^{reversed_i} \cdot (f(\mathbf{X_i}) \right. \\ & \left. \cdot (1 - g(\mathbf{Z_i})))^{1-reversed_i} \right]^{reviewed_i}, \end{aligned} \quad \textbf{(1)}$$

where w_i is the weight (capturing over- or underrepresentation in the sample) for observation i. We programmed and maximized the log-likelihood in R.

One might also suppose that strategic whistleblowing by potential dissenting lower court judges may be correlated both with the lower court's ideological alignment with the Supreme Court and the Supreme Court's decision to review and/or reverse the lower court's

TABLE 1. Estimates from the Empirical Model

	Model 1	Model 2
Review parameters		
β_0 (intercept)	−3.5	−1.9
	(0.1)	(0.1)
β_1 (distance)	−1.6	−1.4
	(0.7)	(0.7)
β_2 (distance squared)	2.0	1.6
	(1.1)	(1.1)
β_3 (lower court dissent)		1.5
		(0.1)
Reversal parameters		
γ_0 (intercept)	−0.1	−0.2
	(0.2)	(0.2)
γ_1 (distance)	1.8	1.9
	(1.4)	(1.3)
γ_2 (distance squared)	−2.7	−2.8
	(2.3)	(2.1)
γ_3 (lower court dissent)		(0.3)
		(0.1)
N	7166	7166
Log-likelihood	−4121.4	−3314.9
AIC	8255	6644

Notes: Estimates from maximizing the log of equation (1); point estimates with standard errors in parentheses; review parameters capture effect of covariates on the decision to review a case; reversal parameters capture effect of covariates on the decision to reverse a case once reviewed.

decision (e.g., Hettinger, Lindquist, and Martinek 2004; Kastellec 2007). Thus, we also estimated our empirical model including this control variable.[23]

RESULTS

Our results are reported in Table 1. Notice first that the point estimates and variance for each of the parameters are essentially identical across the two specifications (with and without the dissent control). Because Model 2 is the best fit to the data (according to the AIC), we focus on it for our substantive interpretations. All of the substantive conclusions we draw, however, hold for both models.

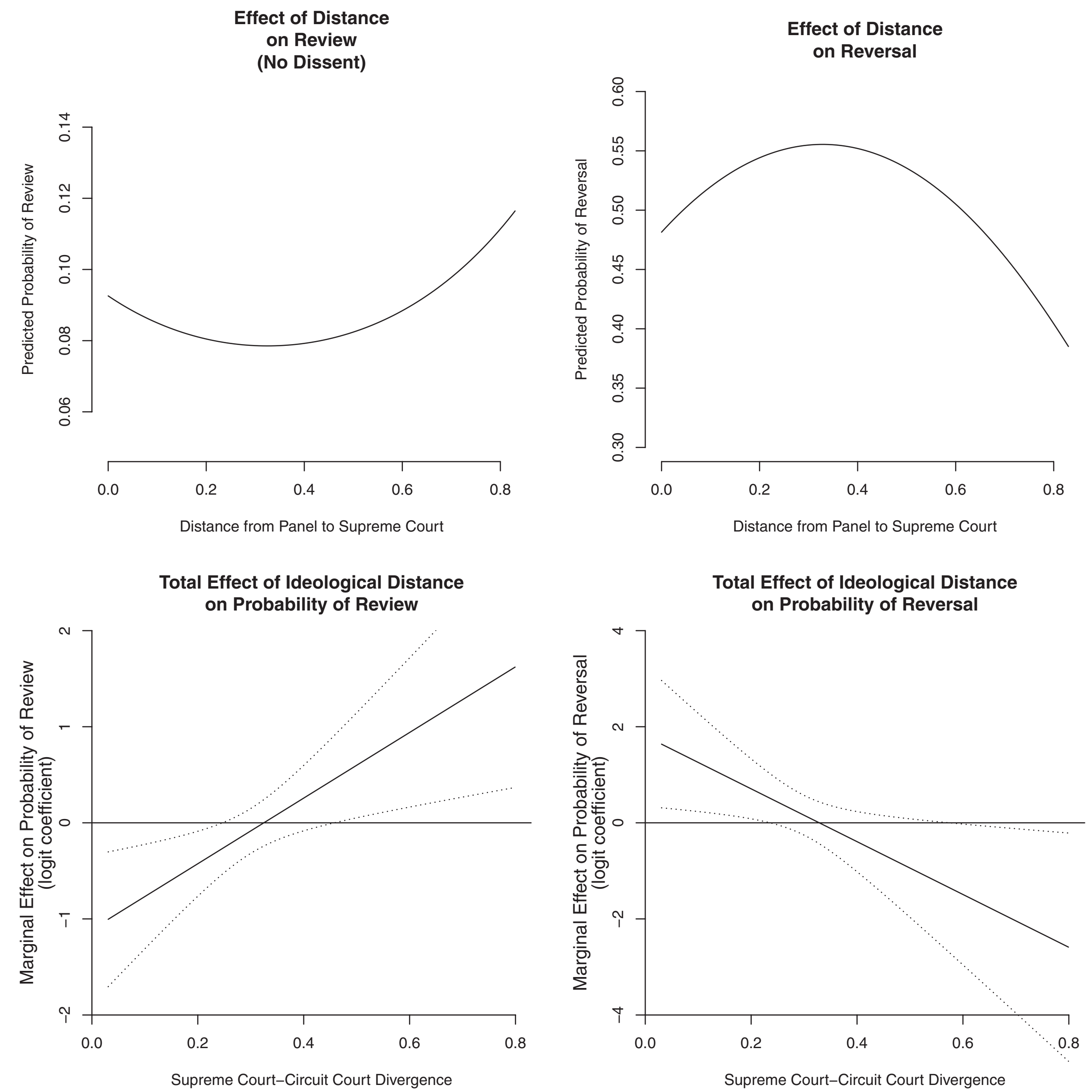
Our empirical analysis results are consistent with our theoretical predictions. Consider the review parameters (the β's). The negative estimate of β_1 and the positive estimate of β_2 indicate that the probability of review is a non-monotonic function of distance$_i$; as expected, the probability of review initially decreases and then increases as the courts become more ideologically divergent. In fact, the substantive magnitude of this effect is considerable, as is demonstrated in the top-left panel in Figure 4. As a lower court becomes more ideologically divergent, the probability of review initially decreases, but then begins to increase. The bottom-left panel in Figure 4, moreover, shows the marginal effect of ideological distance on the probability of review. As predicted, the marginal effect of preference divergence changes from a negative effect at low levels of divergence to an increasingly positive effect at high levels of divergence. This differential effect of ideological distance on the probability of review is precisely what was predicted in Hypothesis 1.

Consider next our estimates of the reversal parameters (the γ's). Again, we find support for our hypothesis: The positive estimated coefficient associated with distance$_i$ and the negative estimated coefficient associated with distance$_i^2$ together indicate a pattern whereby the probability of reversal initially increases as the courts become more divergent and then decreases. This relationship is shown in the top-right panel in Figure 4: The probability of reversing a decision that has been reviewed initially increases as the lower court becomes more divergent from about 40% to nearly 55%; however, as the lower court becomes increasingly distant, the probability of reversal *drops* to about 35%. The bottom-right panel shows the marginal effect of ideological divergence across the range of observed divergence. As predicted, the marginal effect of preference divergence changes from a positive effect at low levels of divergence to an increasingly negative effect at high levels of divergence. This empirical pattern comports with the expectation outlined in Hypothesis 2, providing additional support for the (counterintuitive) predictions derived from the theoretical model.

We briefly note in conclusion the estimates associated with our control variable. A nonunanimous lower court decision (one with a dissenting judge) is much more likely to be reviewed by the Supreme Court than a unanimous decision; the point estimates associated with this variable in the review part of the model are substantively very large. The predicted probability that a case from a lower court of average distance will be reviewed is 2% without a dissent, but 9% with a dissent. The presence of a dissent in the lower court has a comparable effect on the probability that the Supreme Court will reverse the lower court. In the absence of a dissent, the predicted probability that the Supreme Court will reverse a lower court decision is 52% (average distance, using estimates from Model 2). When there is a dissenting opinion, the predicted probability rises to 60%. This finding is largely in line with the literatures on certiorari and whistleblowing, which suggests an incentive for lower court judges to

[23] One might also suspect that there are circuit-specific factors that increase or decrease the Supreme Court's interest in a case (i.e., the cost of review and therefore the likelihood of review), and there is certainly correlation between which circuit a case comes from and the panel's ideological relationship to the Supreme Court (some circuits are more liberal, whereas others are more conservative). We tried evaluating our empirical model including fixed effects for the circuit from which a case came. Although the parameter point estimates from that model were nearly identical to those reported later, because of multicollinearity the model's Hessian was singular and therefore not invertible. Thus, we chose not to report those point estimates.

FIGURE 4. Substantive Effect of Ideological Distance on Probability of Review and Reversal



Notes: The top two panels show the substantive magnitude of the estimated effect of ideological divergence between the Supreme Court and the lower court on the probability of review and reversal; estimates come from empirical Model 1, using Model 2 specification. The bottom two panels show the marginal effect of ideological distance on the probability of review and reversal across the observed values of ideological distance; dotted lines show 95% confidence intervals.

dissent in an effort to signal to the Supreme Court that it should hear the case (e.g., Beim and Kastellec 2011; Caldeira, Wright, and Zorn 1999; Perry 1991). However, as noted earlier, the inclusion of this control variable does not considerably change the estimated relationship between ideological alignment and the Supreme Court's decisions.

The implications of these empirical findings are considerable. As noted earlier, many of the standard tests used in the literature to assess whether the lower courts have strategic or myopic preferences cannot discriminate among various accounts of judicial behavior. In particular, by considering dispositions or rules alone, those studies risk missing important dynamics. Our theoretical model, by incorporating those dual considerations, reveals a series of discriminating predictions, and the analysis here reveals striking support for those hypotheses.

CONCLUSION

A major area of inquiry in the study of American politics concerns rule making in a hierarchical institution. Perhaps chief among the substantively compelling institutions are the courts. With the judicial hierarchy's broad powers for lawmaking, the ways in which it shapes or constrains judicial decisions have broad implications for understanding much of the law. As noted at the outset of this article, the courts of appeals are often the final voice on important legal questions. We have developed a model of Supreme Court-lower court interactions that goes beyond existing theory. By combining a theoretical analysis of rule creation with a bottom-up perspective on the judicial hierarchy, we have uncovered a variety of striking features about the way in which law is created by lower courts. Our analysis yields several equilibrium predictions about judicial behavior that intuition predicts should exist yet no previous theory can explain. For example, our model predicts that the Supreme Court will intentionally review and affirm lower courts to alter the legal rule used in their decisions. In addition, our model highlights several difficulties with previous empirical strategies used to assess compliance and strategic behavior in the judicial hierarchy. Moreover, several of the results and comparative statics that derive from our analysis run counter to modal perceptions and intuitions and lead naturally to new empirical tests that future research can and should contemplate.

An empirical analysis of two of the SCRM model's most striking predictions has yielded results consistent with two of its counterintuitive predictions. Specifically, we have shown that there is a non-monotonic relationship between ideological divergence and both the likelihood of review and the likelihood of reversal. As lower courts become more divergent, they become at first *less* likely to be reviewed and then *more* likely to be reviewed. In contrast, as lower courts become more divergent, they become first *more* likely, but then *less* likely, to be reversed. Although we investigated only a limited set of the model's predictions, the evidence is striking.

Nevertheless, there is work still to be done. The findings reported here are but one step forward. The results of this study can and should be incorporated into the broader literatures on the development of law (e.g., Kornhauser 1992; Lax 2007), judicial compliance and top-down rule making (e.g., Cameron, Segal, and Songer 2000; Clark 2009), and the structure of legal rules (e.g., Lax 2009; Staton and Vanberg 2008). The formal models we develop have broader implications for rule making in a hierarchy (e.g., other judicial hierarchies, administrative bureaucracies). Our results may be exported to those settings for further interpretation and empirical analysis. Short of that future research, our study yields a series of important lessons about the judicial hierarchy. Perhaps most important, the evidence reveals that strategic considerations by the lower courts lead to more complicated nuanced incentives and patterns of behavior than standard models would predict.

APPENDIX: PROOFS AND SUPPLEMENTAL RESULTS

Proof. *Proof of Proposition 1.* The following strategy profile constitutes the unique perfect Bayesian equilibrium to the model:

$$r^* = L$$

$$a^*(r) = \begin{cases} 1 & \begin{cases} L \geq \sqrt{k} & \text{and disposition is Exclude} \\ k \leq L^2 + \phi \int_{-\infty}^{0} g(f)\,df & \text{and disposition is Admit} \end{cases} \\ 0 & \text{otherwise} \end{cases}$$

$$b^*(r) = \begin{cases} 1 & \text{if disposition is Exclude} \\ \int_{-\infty}^{0} g(f)\,df & \text{if disposition is Admit} \end{cases}$$

First notice that because UC's decision cannot affect LC's utility, it is always optimal for LC to select $r = L$. This gives LC the best possible rule and always results in LC's preferred disposition. Thus, $U_{LC}(r = L) = \phi$; this is the best possible utility for LC. It is optimal for UC to audit LC's decision only when $U_{UC}(audit) \geq EU_{UC}(\neg audit)$. By the utility function defined earlier, $U_{SC}(audit) = \phi - k$. UC's expected utility from not auditing depends on the rule selected by LC and whether the realized case facts and disposition implied by r correspond to UC's preferred disposition. When UC observes the rule $r = L$ and the disposition "exclude" (E), then UC knows that $f > L$. Given $f > L$, UC prefers the disposition E. Thus, $EU_{UC}(\neg audit|r = L, E) = \phi - L^2$. Thus, on observing E, UC prefers to audit only when $k \leq L^2$. When UC observes $r = L$ and the disposition "admit" (A), then UC's expected utility from not auditing depends on the realized case facts and the rule selected by LC. Specifically, $EU_{UC}(\neg audit|r = L, A) = -L^2 + \phi \int_{-\infty}^{0} g(f)\,df$. Thus, on observing A, UC prefers to audit only when $k \leq L^2 + \phi \int_{-\infty}^{0} g(f)\,df$. ■

Proof. *Proof of Proposition 2.* Note, Proposition 2 characterizes a continuum of isomorphic equilibria where there are three cases. The cases are determined by a cutpoint in the fact-space, $r_A \in [\underline{r}, \bar{r}]$. To characterize equilibrium behavior, we first establish a series of intermediate results.

Lemma 1. *The Upper Court reviews a case whenever $r \geq \sqrt{k - (1 - b)\phi}$.*

Proof. $U_{UC}(review|r) = \phi - k$ and $EU_{UC}(\neg review|r) = -r^2 + b\phi$, where b is the Upper Court's beliefs over the probability that the Lower Court has declared a rule that yields the Upper Court's preferred disposition. The parameter b is a decreasing function of r; by Brouwer's fixed point theorem a solution to this equality exists. The Upper Court reviews a case whenever $EU_{UC}(review|r) > EU_{UC}(\neg review|r)$. ■

Corollary 1. *If the Upper Court observes ($r \geq 0$, Exclude), the Upper Court reviews if $r > \sqrt{k}$ and does not review if $0 \leq r \leq \sqrt{k}$.*

Proof. If the Upper Court observes ($r \geq 0$, *Exclude*), $b = 1$ because the Upper Court always wants Exclude for any $f > 0$, and the Upper Court knows that $f > r$ on observing ($r \geq 0$, *Exclude*). The Upper Court is indifferent over changing the rule when it knows it is getting its preferred disposition when $r = \sqrt{k}$. ■

Corollary 2. *The Upper Court does not review if* $r \in [0, \underline{r}]$ *for any* $b \in [0, 1]$*, where* $\underline{r} = \sqrt{k - \phi}$.

Proof. The Upper Court prefers not to review for $b \geq 0$ if $EU_{UC}(review|r) = \phi - k \leq EU_{UC}(\neg review|r) = -r^2$. ■

Lemma 2. *The Lower Court never prefers to be reviewed.*

Proof. This proof is constructed in three steps. First, the Lower Court never offers $r < 0$. If the Lower Court offers $r = 0$, the Upper Court never reviews ($U_{UC}(\neg review|r = 0) = \phi > U_{UC}(review|r = 0) = \phi - k$). If the Lower Court offers $r < 0$ and is reviewed, the Lower Court gets the same outcome as if it set $r = 0$. If the Lower Court offers $r < 0$ and is not reviewed, the Lower Court is strictly worse off. If $f < r$, $U_{LC}(r < 0|\neg review) = -(r - L)^2 + \phi < U_{LC}(r = 0|\neg review) = -L^2 + \phi$. If $f \in [r, 0]$, $U_{LC}(r < 0|\neg review) = -(r - L)^2 < U_{LC}(r = 0|review) = -L^2 + \phi$.

Second, the Lower Court never offers $r > L$. If the Lower Court offers $r = L$ and is not reviewed, the Lower Court gets its optimal outcome. If the Lower Court offers $r = L$ and gets reviewed, the Lower Court is also reviewed for any $r > L$. Because $\frac{\partial(EU_{UC}(\neg review|r))}{\partial r} = -2r + \frac{\partial b}{\partial r}\phi$, and we know $\frac{\partial b}{\partial r} < 0$ for $r > 0$, therefore EU_{UC} is decreasing in r.

Because the Lower Court never has an incentive to offer a rule outside of the Pareto set, the Lower Court never prefers to be reviewed. If $f < 0$, $U_{LC}(\neg review) = -(r - L)^2 + \phi > U_{LC}(review) = -L^2 + \phi$. If $f \in [0, L]$, $U_{LC}(\neg review) = -(r - L)^2 + \phi$ or $-(r - L)^2$. Either way, $U_{LC}(review) = -L^2 < U_{LC}(\neg review|f \in [0, L])$. And if $f > L$, $U_{LC}(\neg review) = -(r - L)^2 + \phi > U_{LC}(review) = -L^2 + \phi$. ■

Corollary 3. *By proof of Lemma 2, the Lower Court never makes an offer outside the Pareto set,* $r \in [0, L]$*, for any* $b \in [0, 1]$.

Lemma 3. *For* $f \geq \sqrt{k}$*, the Lower Court's sequentially rational move is to offer* $r = \sqrt{k}$.

Proof. Consider first $f > L$. By Corollary 1, the Upper Court reviews if and only if $r \geq \sqrt{k}$. By Lemma 2, the Lower Court never prefers to be reviewed. The rule closest to the Lower Court's most preferred rule that will not be reviewed is $r = \sqrt{k}$. Consider next $\sqrt{k} \leq f \leq L$. By Corollary 1, the largest rule that the Upper Court will not review is $r = \sqrt{k}$. By Lemma 2, the Lower Court never prefers to be reviewed. Thus, a Lower Court with case facts $\sqrt{k} \leq f \leq L$ sets $r = \sqrt{k}$. ■

Lemma 4. *All equilibria*[24] *must involve pooling by Lower Courts with case facts* $f < 0$ *and the subset of Lower Courts with case facts* $0 < f < r_A$ *over a rule in the set* $r \in [\underline{r}, \bar{r}]$*, where* $\underline{r} = \sqrt{k - \phi}$ *and* $\bar{r} = \sqrt{k - (1 - b)\phi}$.

Proof. If the Lower Court offers $r > \sqrt{k}$, the Upper Court reviews by Lemma 1. By Corollary 2, the Upper Court never reviews any rule where $0 \leq r \leq \underline{r}$. Because the Lower Court never prefers to be reviewed (Lemma 2), and the Lower Court wants a rule as close to its ideal rule as possible, we know $r \in [\underline{r}, \sqrt{k}]$ must hold.

Suppose $r' \in (\underline{r}, \sqrt{k}]$ is offered by a Lower Court with case facts $f < 0$, and it separates from all Lower Courts with case facts $0 < f < L$. The Upper Court does not review by Corollary 1. A Lower Court with case facts $0 < f < L$ strictly prefers to offer $r = r'$ over offering $r \leq \underline{r}$ because r' will not be reviewed. Further, if a Lower Court with case facts $0 < f < \underline{r}$ chooses any other $r \in (\underline{r}, \sqrt{k}]$ in equilibrium, the Upper Court knows $f > 0$, $b = 0$, and by Lemma 1 the Upper Court reviews. Because by Lemma 2 any lower Court with case facts $0 < f \leq r'$ prefers to pool on r' ratherthan be reviewed, we have a contradiction and separation is unsustainable.

Finally, because the equilibrium must entail pooling behavior, and the Upper Court cannot review the Lower Court with certainty on observing the equilibrium rule, r^* (by Lemma 2, the Lower Courts with case facts $f \leq \underline{r}$ would deviate to $\underline{r}$), the Upper Court at most must be indifferent over reviewing the rule. Because from above all Lower Courts with case facts $f \leq r'$ will pool on r', by Lemma 1 the Upper Court is indifferent over reviewing a rule when $\bar{r} = \sqrt{k - (1 - b)\phi}$. Thus $r^* = r_A \in [\underline{r}, \bar{r}]$ must hold. ■

Lemma 5. *Given the possible equilibrium behavior defined in Lemma 4, the beliefs that can support this behavior are as follows. On equilibrium path, on observing* $(r^*, Admit)$, $b^* = \frac{\int_{-\infty}^{0} g(f)df}{\int_{-\infty}^{r^*} g(f)df}$. *Off equilibrium path, on observing* $(0 \leq r < r^*, Admit)$*, b can cover the entire support. Off equilibrium path, on observing* $(r^* < r < \bar{r}, Admit)$, $b^* \leq \frac{r^2 - k + \phi}{\phi}$ *must hold.*

Proof. For $(r^*, Admit)$, r^* is on equilibrium path, and so beliefs must be updated by Bayes rule. Thus, $b^* = \frac{\int_{-\infty}^{0} g(f)df}{\int_{-\infty}^{r^*} g(f)df}$. For $(0 \leq r < r^*, Admit)$, the Lower Court always prefers offering r^* to any deviation to the left. Thus, Upper Court moves, and beliefs are unconstrained in this range. Finally, for $(r^* < r' < \bar{r}, Admit)$, the Upper Court must prefer to review the rule in equilibrium, otherwise the Lower Court would deviate to r'. The Upper Court prefers to review whenever it believes the likelihood of a reversal on disposition is sufficiently high. By Lemma 1, this condition holds whenever $b^* \leq \frac{r^2 - k + \phi}{\phi} = \hat{b}$ holds. ■

Lemma 6. *For a Lower Court with case facts* $r^* < f < \sqrt{k}$*, the optimal r is the one closest to L for which the disposition is* Exclude, $r = f$.

Proof. If a Lower Court with case facts $r^* < f < \sqrt{k}$ sets the rule greater than f, the disposition would be Admit. As proven in Lemma 5, on observing $(r > r^*, Admit)$ the Upper Court reviews and holds off equilibrium path beliefs $b \leq \frac{r^2 - k + \phi}{\phi}$. Thus, the Lower Court sets the rule at the largest value that would yield a disposition of Exclude and not be reviewed, $r = f$. ■

With these intermediate results established, we now demonstrate equilibrium behavior in each of the three cases. Formally, the equilibria characterized in Proposition 2 are given as follows:

$$r^*(f) = \begin{cases} \sqrt{k} & \text{if } f > \sqrt{k} \\ f & \text{if } f \in \left(r_a, \sqrt{k}\right] \\ r_a & \text{if } f \leq r_a \end{cases}$$

[24] This result excludes the possibility of a mixed-strategy equilibrium. We conjecture that mixed-strategy equilibria may exist for the knife-edge condition where $r^* = \bar{r}$.

$$a^*(r) = \begin{cases} 1 & r \geq \sqrt{k} \\ [0,1] & r \geq r_a \quad \text{and disposition is admit} \\ 0 & \text{otherwise} \end{cases}$$

$$b^*(r) = \begin{cases} \hat{b} & \text{if } r = r_a \\ [0,1] & \text{if } \begin{cases} r \leq 0 & \text{and disposition is Exclude} \\ r < r_a & \text{and disposition is Admit} \\ r > \sqrt{k} & \end{cases} \\ 1 & \text{if } r > 0 \quad \text{and if disposition is Exclude} \\ [0, \hat{b}) & \text{if } r \in \left(r_a, \sqrt{k}\right) \quad \text{and disposition is Admit} \end{cases}$$

where $\hat{b}$ is defined by Bayes rule and makes the Upper Court indifferent about auditing when $r = r_a$.

Case 1 **($f \geq \sqrt{k}$):** By Lemma 3, $r^* = \sqrt{k}$, which leads to the disposition Exclude. By Corollary 1, UC does not review the rule. To see that LC never has an incentive to deviate, consider first a deviation to any $r' < r^*$. For any beliefs and any move by UC, LC is strictly worse off. Consider next a deviation to any $r'' > r^*$. By Corollary 1, UC reviews this rule. By Lemma 2, LC never prefers to be reviewed and therefore has no incentive to deviate to $r'' > r^*$.

Case 2 **($f < r_A \in [\underline{r}, \overline{r}]$):** By Lemma 4, LC picks $r^* = r_A$, which leads to the disposition Admit. By Lemma 5, on observing $(r^*, Admit)$, $b^* = \frac{\int_{-\infty}^{0} g(f)df}{\int_{-\infty}^{r^*} g(f)df}$. By Lemma 4, $b^* > \frac{r^2 - k + \phi}{\phi}$ and therefore UC does not review.

Suppose a deviation to $r' < r^*$. For any beliefs and any move by UC, LC is strictly worse off. Now, consider a deviation to r' where $L > r' > r^*$. By Lemma 5, $b \in [0, \hat{b}]$, and UC reviews. By Lemma 2, LC never prefers to be reviewed. Therefore, there is no incentive to deviate. Finally, suppose a deviation to $r'' > L$. By Corollary 3, LC will never deviate to any $r > L$. UC's beliefs and moves are unconstrained.

Case 3 **($r_A \leq f < \sqrt{k}$):** In this case, $r^* = f$. $r^* = f$ gives a disposition of Exclude. $r^* = f$, *Exclude* implies $f \geq r^*$, and therefore $b = 1$. By Lemma 6, LC has no incentive to deviate from r^* in either direction. ■

Lemma 7. *Given Proposition 2, if $L \geq r^*(f)$, the behavior characterized by Proposition 2 holds in equilibrium; if $L < r^*(f)$, then the following strategy characterizes the Lower Court's strategy in all perfect Bayesian equilibria:*

$$r^*(f) = \begin{cases} L & \text{if } \begin{cases} f > r_A \\ f < r_A \text{ and } b \geq b^* \end{cases} \\ r_A & \text{otherwise} \end{cases}$$

Proof. Suppose first $L \geq r_A$. Playing $r^* = L$ leads to a disposition of Exclude because, by assumption, $L < r^*(f)$. By Corollary 1, UC does not review. LC can do no better because it is choosing its optimal rule. Now, suppose $L < r_A$. If on observing $r < r_A$, $b \geq b^*$, then by Lemma 1, UC does not review. Again, LC is choosing its optimal rule and can do no better. Conversely, if on observing $r < r_A$, $b < b^*$, then by Lemma 1, UC does review. By Lemma 2, LC's best reply remains r_A. ■

Proof. *Proof of Proposition 3.* Note, as in Proposition 2, we characterize a continuum of isomorphic equilibria, where there are multiple cases. To characterize equilibrium behavior in any one of these isomorphic equilibria, we first establish a series of intermediate results.

Lemma 8. *The dominant strategies characterized by Lemmas 1 and 2, as well as their corollaries, continue to hold in this model by subscripting the k parameter with k_i, $i = l, h$.*

Proof. The Upper Court's objective function has not changed, except that the cost parameter, k is now subscripted by the Upper Court's type. Substituting in a subscripted k parameter does not alter the Upper Court's dominant strategies described by Lemma 1. ■

Lemma 9. *All equilibria*[25] *must involve pooling by Lower Courts with case facts $f < 0$ and the subset of Lower Courts with case facts $0 < f < r_A$ over a rule in the set $[\underline{r}_l, \overline{r}_h]$.*

Proof. Define $\underline{r}_i = \sqrt{k_i - \phi}$ and $\overline{r}_i = \sqrt{k_i - (1-b)\phi}$, where $i = l, h$. If a Lower Court with case facts $f \leq \overline{r}_h$ offers $r \geq \overline{r}_h$, the Upper Court always reviews (Lemma 1 and $k_l < k_h$). The Upper Court never reviews any rule where $0 \leq r \leq \underline{r}_l$ (Corollary 2, Lemma 8, and $k_l < k_h$). Because the Lower Court never prefers to be reviewed (Lemma 2 and 8) and the Lower Court wants a rule as close to its ideal rule as possible, we therefore know $r \in [\underline{r}_l, \overline{r}_h]$ must hold. Whether $\underline{r}_h$ or $\underline{r}_l$ is the lower boundary on this range depends on whether $EU_{LC}(\underline{r}_h | f \leq 0) = p(-L^2 + \phi) + (1-p)(-(L - \underline{r}_h)^2 + \phi) \geq EU_{LC}(\underline{r}_l | f \leq 0) = -(L - \underline{r}_l)^2 + \phi)$ holds. Substituting in for $\underline{r}_h$ and $\underline{r}_l$, and solving for p yield $p \leq \frac{(\sqrt{k_l} - L)^2 - (\sqrt{k_h} - L)^2}{L^2 - (\sqrt{k_h} - L)^2}$.

Now, to see that only pooling behavior can be supported, assume that a Lower Court with case facts $f \leq 0$ separates from all Lower Courts with case facts $f \in (0, L)$. Now suppose first the Lower Court with case facts $f \leq 0$ is playing $r' \in [\underline{r}_l, \overline{r}_l]$. A Lower Court with case facts $0 < f < r'$ strictly prefers to offer $r = r'$ over offering $r \leq \underline{r}_H$ because r' will not be reviewed and the alternatives guarantee the Lower Court a worse rule and possibly a worse disposition as well. If a Lower Court with case facts $f \in (0, r']$ offers any other $r'' \neq r' \in (\underline{r}_h, \overline{r}_h)$ the Upper Court knows $f > 0$, $b = 0$, and by Lemmas 1 and 8 either type of Upper Court reviews. Because the Lower Court never prefers to be reviewed (Lemmas 2 and 8), we have a contradiction.

Suppose, on the other hand, that the Lower Court with case facts $f \leq 0$ is playing $r' \in (\overline{r}_l, \overline{r}_h]$. A Lower Court with case facts $0 < f < r'$ strictly prefers to offer $r = r'$ over offering $r \leq \underline{r}_h$ because $EU_{LC}(r') = p(-(r' - L)^2 + \phi) + (1-p)(-L^2 + \phi) \geq U_{LC}(\underline{r}_l) = -(\underline{r}_l - L)^2 + \phi$ and $EU_{LC}(r') \geq U_{LC}(\underline{r}_h) = p(-L^2 + \phi) + (1-p)(-(\underline{r}_h - L)^2 \phi)$ must hold (otherwise the Lower Court with case facts $f \leq 0$ would deviate and play $\underline{r}_l$ or $\underline{r}_h$ as well). If a Lower Court with case facts $f \in (0, r']$

[25] As earlier, this result excludes the possibility of a mixed-strategy equilibrium. We conjecture that mixed-strategy equilibria may exist for the knife-edge conditions where $r^* = \overline{r}_l$ or where the Lower Court is indifferent between playing a risk-averse and a risk-seeking strategy. For the latter knife-edge condition, this possibility introduces no new observable behavior, because the Lower Court would simply be choosing among already identified equilibrium strategies.

offers any other $r'' \neq r' \in (\underline{r}_h, \overline{r}_h)$, the Upper Court knows $f > 0$, $b = 0$, and by Lemmas 1 and 8 either type of Upper Court reviews. Because the Lower Court never prefers to be reviewed (Lemmas 2 and 8), we once again have a contradiction. ■

Definition 1. Let $\hat{p} = \frac{(\sqrt{k_l}-L)^2-(\sqrt{k_h}-L)^2}{L^2-(\sqrt{k_h}-L)^2}$.

Lemma 10. *Given the possible equilibrium behavior defined in Lemma 9, beliefs that can support this behavior are as follows. On equilibrium path, on observing $(r^*, Admit)$, $\hat{b} = \frac{\int_{-\infty}^{0} g(f)df}{\int_{-\infty}^{r^*} g(f)df}$. If $r^* \leq \overline{r}_l$, on observing the off-equilibrium path play $(0 \leq r' < r^*, Admit)$, b can cover the entire support. If $r^* \in (\overline{r}_l, \overline{r}_h]$, on observing the off-equilibrium path play $(0 \leq r' < r^*, Admit)$, b can cover $\hat{b} \leq \frac{r'^2-k_l+\phi}{\phi}$. Off equilibrium path, on observing $(r^* < r' < \overline{r}_h, Admit)$, b can cover $\hat{b} \leq \frac{r'^2-k_h+\phi}{\phi}$.*

Proof. For $(r^*, Admit)$, r^* is on equilibrium path, and so beliefs must be updated by Bayes rule. When $r^* \leq \overline{r}_l$, for $(0 \leq r' < r^*, Admit)$ the Lower Court always prefers offering r^* to any deviation to the left. Thus, Upper Court moves, and beliefs are unconstrained in this range. When $r^* \in (\overline{r}_l, \overline{r}_h]$, for $(0 \leq r' < r^*, Admit)$ the Lower Court prefers r^* to any deviation to the left as long as an Upper Court with $k_i = k_l$ prefers to review the deviation. By Lemmas 1 and 8, this condition holds when $b \leq \frac{r'^2-k_l+\phi}{\phi}$ holds. Finally, for $(r^* < r' < \overline{r}_h, Admit)$, the Upper Court with $k_i = k_h$ must prefer to review the rule in equilibrium, otherwise the Lower Court would deviate to r'. The Upper Court prefers to review whenever it believes the likelihood of a reversal on disposition is sufficiently high. By Lemmas 1 and 8, this condition holds whenever $b^* \leq \frac{r^2-k_h+\phi}{\phi} = \hat{b}$ holds. Note that the beliefs described here are sufficient, but not necessarily necessary for the behavior in Lemma 9 to be sequentially rational. ■

Lemma 11. *There are multiple types of equilibria that may exist, but we focus on those where courts with $f < 0$ always pool on a single r. (a) When there are multiple $r' < \overline{r}_l$, the Lower Court always prefers to choose the largest. (b) When there are multiple $r' \in (\overline{r}_l, \overline{r}_h)$, the Lower Court always prefers to choose the largest. (c) There may be a case where $EU_{LC}(r' < \overline{r}_l | f < 0) = EU_{LC}(r \in (\overline{r}_l, \overline{r}_h) | f < 0)$. This, however, is a knife-edge condition. (d) We rule out arbitrary mappings from model parameters to rules.*

Lemma 12. *For a Lower Court with case facts $r_{A_i} < f < \sqrt{k_i}$, where $i = l$ if $p \geq \hat{p}$ and $i = h$; otherwise, the optimal r is the largest one for which the disposition is* Exclude, *$r = f$.*

Proof. If a Lower Court with case facts $r_A < f < \sqrt{k_i}$ sets the rule greater than f, the disposition would be Admit. As proven in Lemma 10, on observing $(r' > r_A, Admit)$, the Upper Court reviews and holds off-equilibrium path beliefs $b \leq \frac{r'^2-k_h+\phi}{\phi}$. Thus, the Lower Court sets the rule at the largest value that would yield a disposition of Exclude and not be reviewed, $r = f$. ■

Lemma 13. *A Lower Court with case facts $\sqrt{k_l} \leq f < \sqrt{k_h}$ sets $r = \sqrt{k_l}$ if $p \geq \frac{(\sqrt{k_l}-L)^2-(f-L)^2}{L^2-(f-L)^2}$ and $r = f$* otherwise.

Proof. By Corollary 1, the Upper Court reviews if and only if $r > \sqrt{k_i}$. Thus, if the Lower Court offers $r \leq \sqrt{k_l}$, the Lower Court receives $-(l-r)^2$. If the Lower Court offers $r \in [\sqrt{k_l}, f]$, the Upper Court reviews iff $k = k_l$, and the Lower Court receives $-pL^2 - (1-p)(L-r)^2$. Thus, the Lower Court sets the rule closest to its most preferred rule that ensures neither type of Upper Court reviews the case ($r = \sqrt{k_l}$) when $p \geq \frac{(\sqrt{k_l}-L)^2-(f-L)^2}{L^2-(f-L)^2}$, and $r = f$ otherwise. By Lemmas 2 and 8, the Lower Court never prefers to choose a rule that is reviewed with certainty, so no other behavior obtains. ■

Lemma 14. *A Lower Court with case facts $\sqrt{k_h} \leq f < L$ sets $r = \sqrt{k_i}$ where $i = l$ if $p \geq \hat{p}$ and $i = h$ otherwise.*

Proof. This proof is the same as Lemma 13, except for the fact that the Lower Court gets reviewed with certainty for any $r > \overline{r}_h$, and so the best risky proposal the Lower Court can make is $r = \overline{r}_h$. ■

Lemma 15. *For $f > L$, the Lower Court's sequentially rational move is to offer $r = \sqrt{k_i}$ where $i = l$ if $p \geq \hat{p}$ and $i = h$ otherwise.*

Proof. By Corollary 1, the Upper Court reviews if and only if $r > \sqrt{k_i}$. By Lemmas 2 and 8, the Lower Court never prefers to be reviewed when k is known. Thus, the Lower Court either sets the rule closest to its most preferred rule that ensures neither type of Upper Court reviews the case ($r = \sqrt{k_l}$) or that only the Upper Court with low costs, k_l, reviews the case ($r = \sqrt{k_h}$). The Lower Court offers $r = \sqrt{k_l}$ when $U_{LC}(r = \sqrt{k_l}) = -(\sqrt{k_l} - L)^2 + \phi \geq EU_{LC}(r = \sqrt{k_h}) = p(-L^2 + \phi) + (1-p)(-(\sqrt{k_h} - L)^2 + \phi)$. Solving for p yields $p \geq \hat{p}$. ■

With these intermediate results established, we demonstrate equilibrium behavior. The following characterizes all pure-strategy perfect Bayesian equilibria to the model.

The Lower Court's equilibrium strategy is given by

$$r^*(f; p \leq \hat{p}) = \begin{cases} r_{a_h} \in \left[\underline{r}_l, \overline{r}_h\right], \text{Admit} & \text{if } f \leq r_{a_h} \\ f, \text{Exclude} & \text{if } f \in \left(r_{a_h}, \sqrt{r_h}\right] \\ \sqrt{k_h}, \text{Exclude} & \text{if } f > \sqrt{k_h} \end{cases}$$

$$r^*(f; p > \hat{p}) = \begin{cases} r_{a_l} \in \left[\underline{r}_l, \overline{r}_l\right], \text{Admit} & \text{if } f \leq r_{a_l}. \\ f, \text{Exclude} & \text{if } f \in \left(r_{a_l}, \sqrt{r_l}\right]. \\ \sqrt{k_l}, \text{Exclude} & \text{if } f > \sqrt{k_l} \end{cases}$$

The Upper Court's equilibrium strategy is given by

$$a^*(r; i = l) = \begin{cases} 1 & \text{if } \begin{cases} r \geq \sqrt{k_l} \\ r \in \left(r_{a_l}, \sqrt{k_l}\right] \text{ and disposition is Admit} \\ r \in (\overline{r}_l, \overline{r}_h] \text{ and disposition is Admit} \end{cases} \\ [0,1] & \text{if } r \in \left(\underline{r}_l, r_{a_l}\right) \\ 0 & \text{otherwise.} \end{cases}$$

$$a^*(r; i = h) = \begin{cases} 1 & \text{if } \begin{cases} r \geq \sqrt{k_h} \\ r \in \left(r_{a_h}, \sqrt{k_h}\right] \text{ and disposition is Admit} \end{cases} \\ [0,1] & r \in \left(\underline{r}_l, r_{a_h}\right) \\ 0 & \text{otherwise} \end{cases}$$

The Upper Court's beliefs are given by

$$b^*(r;i) = \begin{cases} \hat{b} & \text{if } r = r_{a_i} \text{ and disposition is Admit} \\ [0,1] & \textit{if} \begin{cases} r \leq 0 \text{ and disposition is Exclude} \\ r < r_{a_i} \text{ and disposition is Admit,} \\ r > \sqrt{k_i} \end{cases} \\ 1 & \text{if } r > 0 \text{ and if disposition is Exclude} \\ [0, \hat{b}) & \text{if } r \in (r_{a_i}, \sqrt{k_i}) \text{ and disposition is Admit} \end{cases}$$

where $\hat{b}$ is defined by Bayes' rule and makes the Upper Court with type i indifferent about auditing when $r = r_{a_i}$.

Now, consider five cases. **Case 1** ($f < r_{A_l}$): By Lemma 9, the Lower Court pools on $r_{A_l} \in [\underline{r}_l, \overline{r}_l]$ if $p > \hat{p}$, which leads to a disposition of Admit. If, however, $p \leq \hat{p}$, the Lower Court pools on $r_{A_h} \in [\underline{r}_l, \overline{r}_h]$, which leads to a disposition of Admit. By Lemmas 1 and 8, the Upper Court reviews if $i = l$ and $r_{A_h} \in (\overline{r}_l, \overline{r}_h]$. The Upper Court's beliefs are given by Lemma 10.

Case 2 **($r_{A_l} \leq f < r_{A_h}$):** By Lemma 9, the Lower Court pools on $r_{A_h} \in [\underline{r}_h, \overline{r}_h]$ if $p \leq \hat{p}$, leading to a disposition of Admit. However, if $p > \hat{p}$, then by Lemma 12 the Lower Court plays a separating strategy and plays $r = f$, leading to a disposition of Exclude. By Lemmas 1 and 8, the Upper Court reviews if $i = l$ and the Lower Court plays $r_{A_h} \in (\overline{r}_l, \overline{r}_h]$. In the event the Lower Court plays the pooling strategy, and the Upper Court's beliefs are given by Lemma 10. When the Lower Court plays a separating strategy, then the Upper Court believes it has received its preferred disposition with certainty, because any $r > 0$ and Exclude implies $f > 0$.

Case 3 **($r_{A_h} \leq f < \sqrt{k_l}$):** By Lemma 12, the Lower Court plays a separating strategy and plays $r = f$, leading to a disposition of Exclude. By Lemmas 1 and 8, the Upper Court does not review. Observing $r > 0$ and Exclude implies $f > 0$, so $b = 1$.

Case 4 **($\sqrt{k_l} \leq f < \sqrt{k_h}$):** By Lemma 13, the Lower Court will play $r = \sqrt{k_h}$ if $p \geq \frac{(\sqrt{k_l}-L)^2-(f-L)^2}{L^2-(f-L)^2}$. Otherwise, the Lower Court will continue to play its separating strategy and offer $r = f$. Both strategies lead to a disposition of Exclude. By Lemmas 1 and 8, the Upper Court does not review $r = \sqrt{k_l}$ if $i = l$, but will review a rule $r = f > \sqrt{k_l}$ if $i = l$. By Lemmas 1 and 8, the Upper Court reviews neither rule if $i = h$. In any case, observing $r > 0$ and Exclude implies $f > 0$, which implies $b = 1$.

Case 5 **($f \geq \sqrt{k_h}$):** By Lemma 15, the Lower Court offers $r = \sqrt{k_l}$ if $p > \hat{p}$ and $r = \sqrt{k_h}$ otherwise. Any of these rules leads to a disposition of Exclude. By Lemmas 1 and 8, the Upper Court does not review either rule if $i = h$ and reviews any rule $r > \sqrt{k_l}$ if $i = l$. Observing $r > 0$ and Exclude implies $f > 0$, which implies $b = 1$. ■

Lemma 16. *Given Proposition 3, if $L \geq r^*(f;p)$, the behavior characterized by Proposition 3 holds in equilibrium; if $L < r^*(f;p)$, then the following strategy characterizes the Lower Court's strategy in all perfect Bayesian equilibria:*

$$r^*(f;p) = \begin{cases} L & \text{if} \quad \begin{cases} f > r_A \\ f < r_A \text{ and } b \geq b^* \end{cases} \\ r_A & \text{otherwise} \end{cases}$$

Proof. The proof follows immediately from the proof of Lemma 7 demonstrated above. ■

REFERENCES

Beim, Deborah, and Jonathan P. Kastellec. 2011. "Strategic Dissents in the Judicial Hierarchy: Theory and Evidence from Death Penalty Cases." Presented at the Annual Meeting of the American Political Science Association, Seattle.

Caldeira, Greg A., John R. Wright, and Christopher J. W. Zorn. 1999. "Sophisticated Voting and Gate-keeping in the Supreme Court." *Journal of Law, Economics and Organization* 15: 549–72.

Cameron, Charles M., Jeffrey A. Segal, and Donald R. Songer. 2000. "Strategic Auditing in a Political Hierarchy: An Informational Model of the Supreme Court's Certiorari Decisions." *American Political Science Review* 94: 101–16.

Clark, Tom S. 2009. "A Principal-Agent Model of en Banc Review." *Journal of Law, Economics, and Organization* 25 (1): 55–79.

Clark, Tom S., and Clifford J. Carrubba. 2012. "A Theory of Opinion Writing in the Judicial Hierarchy." *Journal of Politics* 74 (2): 584–603.

Clark, Tom S., and Benjamin Lauderdale. 2010. "Locating Supreme Court Opinions in Doctrine Space." *American Journal of Political Science* 54 (4): 871–90.

Clark, Tom S., and Aaron B. Strauss. 2010. "The Implications of High Court Docket Control for Resource Allocation and Legal Efficiency." *Journal of Theoretical Politics* 22 (2): 247–68.

Cross, Frank B. 2005. "Appellate Court Adherence to Precedent." *Journal of Empirical Legal Studies* 2: 369–405.

Epstein, Lee, Andrew D. Martin, Jeffrey A. Segal, and Chad Westerland. 2007. "The Judicial Common Space." *Journal of Law, Economics and Organization* 23 (2): 303–25.

Friedman, Barry. 2006. "Taking Law Seriously." *Perspectives on Politics* 4 (2): 261–76.

Giles, Michael W., Virginia A. Hettinger, and Todd Peppers. 2002. *Measuring the Preferences of Federal Judges: Alternatives to Party of the Appointing President.* Emory University. Working paper.

Greene, William H. 2003. *Econometric Analysis*. Upper Saddle River, NJ: Prentice Hall.

Haire, Susan B., Donald R. Songer, and Stefanie A. Lindquist. 2003. "Appellate Court Supervision in the Federal Judiciary: A Hierarchical Perspective." *Law and Society Review* 37 (1): 143–68.

Hammond, Thomas H., Chris W. Bonneau, and Reginald S. Sheehan. 2006. *Institutional Games and the U.S. Supreme Court*. Charlottesville: University of Virginia Press.

Hettinger, Virginia A., Stephanie A. Lindquist, and Wendy L. Martinek. 2004. "Comparing Attitudinal and Strategic Accounts of Dissenting Behavior on the U.S. Courts of Appeals." *American Journal of Political Science* 48: 123–37.

Hettinger, Virginia A., Virginia A. Lindquist, and Wendy L. Martinek. 2006. *Judging on a Collegial Court: Influences on Federal Appellate Decision Making*. Charlottesville: University of Virginia Press.

Jacobi, Tonja, and Emerson H. Tiller. 2007. "Legal Doctrine and Political Control." *Journal of Law, Economics, and Organization* 23 (2): 326–45.

Kastellec, Jonathan P. 2007. "Panel Composition and Judicial Compliance on the United States Courts of Appeals." *Journal of Law, Economics and Organization* 23 (2): 421–41.

Klein, David E. 2002. *Making Law in the United States Courts of Appeal*. Cambridge: Cambridge University Press.

Klein, David E., and Robert J. Hume. 2003. "Fear of Reversal as an Explanation of Lower Court Compliance." *Law and Society Review* 37 (3): 579–606.

Kornhauser, Lewis A. 1992. "Modeling Collegial Courts *II*: Legal Doctrine." *Journal of Law Economics and Organization* 8: 441–70.

Lax, Jeffrey R. 2007. "Constructing Legal Rules on Appellate Courts." *American Political Science Review* 101 (3): 591–604.

Lax, Jeffrey R. 2009. *Political Constraints on Legal Doctrine: How Hierarchy Shapes the Law.* Columbia University Working paper.

Lax, Jeffrey R. 2011. "The New Judicial Politics of Legal Doctrine." *Annual Review of Political Science* 14: 131–57.

Lindquist, Stefanie A., Susan B. Haire, and Donald R. Songer. 2007. "Supreme Court Auditing of the U.S. Courts of Appeals: An Organizational Perspective." *Journal of Public Administration Research Theory* 17 (4): 607–24.

Martin, Andrew D., and Kevin M. Quinn. 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953–1999." *Political Analysis* 10 (2): 134–53.

McGuire, Kevin T., Georg Vanberg, Charles E. Smith Jr., and Gregory A. Caldeira. 2009. "Measuring Policy Content on the U.S. Supreme Court." *Journal of Politics* 71 (4): 1305–21.

McNollgast. 1995. "Politics and the Courts: A Positive Theory of Judicial Doctrine and the Rule of Law." *Southern California Law Review* 68: 1631–89.

Perry, H.W., Jr. 1991. *Deciding to Decide: Agenda Setting in the United States Supreme Court*. Cambridge, MA: Harvard University Press.

Posner, Richard A. 1995. *Overcoming Law*. Cambridge, MA: Harvard University Press.

Staton, Jeffrey K., and Georg Vanberg. 2008. "The Value of Vagueness: Delegation, Defiance, and Judicial Opinions." *American Journal of Political Science* 52 (3): 504–19.

Tiller, Emerson H., and Pablo T. Spiller. 1999. "Strategic Instruments: Legal Structue and Political Games in Administrative Law." *Journal of Law, Economics and Organization* 15 (2): 349–77.