

# Strategic Defiance and Compliance in the U.S. Courts of Appeals<sup>\*</sup>

Chad Westerland, Jeffrey Segal, Lee Epstein, Charles Cameron, and Scott Comparato<sup>\*\*</sup>

## Abstract

Why do lower courts treat Supreme Court precedents favorably or unfavorably? To address this question, we outline two distinct theoretical frameworks. The first, based on the theory of teams, assumes a shared conception of the judicial role, and emphasizes judicial learning by circuit judges. The second, drawing inspiration from principal-agent theory, assumes pervasive value conflicts among judges, and emphasizes power and control. We use the frameworks to structure an empirical analysis of a random sample of 500 Supreme Court cases, yielding over 10,000 subsequent treatments in the U.S. Courts of Appeals. We find support for both the team-theoretic and agency-theoretic perspectives. The results have important implications for understanding legal change, political control of the judiciary, and strategic litigation.

---

<sup>\*</sup> Chad Westerland is the corresponding author. Email: cwesterl@email.arizona.edu; Post: Department of Political Science, University of Arizona, Tucson AZ 85721-0027; Phone: 520.621.9780; Fax: 520.621.5051.

<sup>\*\*</sup> Chad Westerland is Assistant Professor of Political Science at the University of Arizona; Jeffrey A. Segal is Chair of Political Science and SUNY Distinguished Professor at Stony Brook University; Lee Epstein is the Beatrice Kuhn Professor of Law and Political Science at Northwestern University; Charles M. Cameron is Professor of Politics and Public Affairs, Princeton University; and Scott Comparato is Assistant Professor of Political Science, Southern Illinois University. We presented earlier versions of this paper at the 2006 annual meeting of the Midwest Political Science Association and the 2006 Empirical Legal Studies conference. We thank Raffaele Engle, Justina Geraci, Scott Graves, Nicole Liguori, and Vanessa Maldonado for research assistance. We also owe thanks to Donald Songer, Tom Hansford, Andrew D. Martin, Jim Spriggs and Nancy Staudt for supplying valuable insights and to the National Science Foundation for supporting our work on the hierarchy of justice. On the project's web site [<http://epstein.law.northwestern.edu/research/defiance.html>] is a full replication archive.

# 1. Introduction

In its landmark decision in *Bakke v. California* (1978), the U.S. Supreme Court ruled that universities may, under certain circumstances, take race and other factors into account when they make admissions decisions. But even before the justices had the opportunity to reconsider *Bakke* in *Grutter v. Bollinger* (2003), the U.S. Court of Appeals for the Fifth Circuit took matters into its own hands. In *Hopwood v. Texas* (1996, 963), it held “that the University of Texas School of Law may not use race as a factor in deciding which applicants to admit in order to achieve a diverse student body.” With these words, the judges of the Fifth Circuit, at least according to their colleagues in dissent, took the dramatic step of defying precedent established at the top of the judicial hierarchy, the Supreme Court of the United States.

Scholars and journalists alike spilt much ink over *Hopwood*, as well as decisions by other courts overturning well-established Supreme Court precedents – cases such as the Fourth Circuit’s *United States v. Dickerson* (1999), holding that states under its supervision need not follow *Miranda v. Arizona* (1966); and the Missouri Supreme Court’s overruling of *Stanford v. Kentucky* (1989) in *Simmons v. Roper* (2003). And, yet, these decisions are merely the most striking instances of a more general phenomenon, lower court deviation from earlier precedents set by a higher court – a phenomenon that can take far subtler forms (e.g., distinguishing or limiting precedents). Indeed, as one observer noted well over half a century ago, “[Many] precedents have been rejected through the stratagem of distinguishment; others have been the subject of conscious judicial oversight. As a consequence, judicial discretion among ‘inferior’

judges is not so confined and limited as legal theorists would have it” (Comment 1941, 1448-9; see also Canon and Johnson, 1998; Murphy 1959).

This observation raises a question that, depending on one's perspective, may be posed two different ways: *Why do lower courts defy higher court precedent*, or, given the minute percentage of lower court cases that are heard and reversed (currently well under 1%), *why do lower courts comply with higher court precedent?*

Scholarly attempts to address these questions take several forms.<sup>1</sup> One line of inquiry seeks to identify the circumstances that lead to deviations, subtle or overt. Baum (1978), for example, suggests that lower courts will be less responsive to the U.S. Supreme Court in controversial civil liberties cases, and that the clarity of the precedent, the perceived legitimacy of the Court's ruling, and perception by lower court judges of the chances of review also affect the likelihood of compliance (see also Canon and Johnson, 1998). Another has focused on socialization and conformity to legal culture as the critical causal mechanism. Robert Cover's (1975) noted study of the enforcement of the Fugitive Slave Act by abolitionist judges, for

---

<sup>1</sup> This literature stresses, as we do, explanations for why lower courts defy higher courts (see, e.g., Benesh and Reddick, 2002, who analyze lower court responses to precedent reversed by the Supreme Court). But there also are many, though somewhat narrower, studies seeking to describe or assess the extent of defiant or compliant behavior among lower federal court judges (e.g, Baum, 1980; Beatty, 1972; Beiser, 1968; Canon, 1973; Canon and Kolson, 1971; Gruhl, 1980; Johnson, 1987; Manwaring, 1968; Peltason, 1961; Reid, 1988; Tarr, 1977; Songer and Sheehan, 1990; Vines, 1969). These studies have reached mixed conclusions or, as Benesh and Reddick (2002, 535) note, “The literature on judicial ... compliance is voluminous, albeit somewhat contradictory.”

example, emphasizes the moral quandary posed by the judges' twin commitments to abolition and the rule of law (see also Howard, 1981).

More recently, scholarly efforts, conducted both by social scientists and legal academics, have shifted focus from individual socialization to structural incentives created by the design and operation of organizations. In broad terms, this move is part of the new institutionalism that swept the social sciences in the 1990s. But judicial specialists adopting a new institutional perspective have failed to converge on a singular model of lower court behavior in the hierarchy of justice. Quite the opposite: They have elaborated two distinct and, to some extent, competing accounts. The first, suggested by prominent law professors, is based on the theory of teams and assumes largely homogeneous policy preferences among judges. The second is inspired by principal-agent theory and so assumes heterogeneous policy preferences. Both examine the incentives and opportunities created by various institutional features of the modern judicial hierarchy.

The goal of this paper is to contribute to the burgeoning literature on new judicial institutionalism, empirically and theoretically, by deploying the two approaches to address our primary research question: why do lower courts defy (or, alternatively, comply with) high court precedent? As we explain momentarily, this (seemingly simple) question involves untangling complex interactions between the original decision of the Supreme Court, subsequent interpretations of that decision by the Supreme Court and the relevant circuit court, as well as the preferences of the contemporary lower court panel and the contemporary Supreme Court.

Following from this theoretical discussion, we use team theory and agency theory to

structure an empirical analysis of a random sample of 500 Supreme Court cases, yielding 10,244 subsequent treatments in the U.S. Courts of Appeals. We find that the actual practice of hierarchical interpretation in the federal judiciary affords a degree of support for both perspectives. Panels of lower court judges appear attentive to the preferences of the contemporary Supreme Court and the preferences of the median judge on the circuit in ways suggestive of agency perspectives. At the same time, they appear responsive to earlier decisions in their circuit in ways suggestive of team perspectives. In short, neither account alone appears adequate for understanding judicial behavior—a result with important implications for how the federal judiciary operates, including the dynamics of legal change, political control of the judiciary, and opportunities for strategic litigation.

## 2. Theoretical Frameworks

When politicians and other policy makers bandy about the term “precedent” in reference to trial or intermediate courts, they are typically referring to vertical stare decisis: when lower court judges base their decisions on precedents established by upper courts. But, as Figure 1 shows, a lower court's treatment high court precedent can be far more complex than that.

-- Insert Figure 1 about here --

To see why, first consider the boxes in Figure 1—each of which contain potential influences on the interpretation of an **Enacting Supreme Court** decision by a **Contemporary Lower Court**. Between the time of the decision by the Enacting Supreme Court and the interpretation by the Contemporary Lower Court, two other sets of actors also may interpret the

precedent: **Subsequent Supreme Courts** (those interpreting the precedent after the Enacting Supreme Court) and **Earlier Lower Courts** (those interpreting the precedent prior to the Contemporary Lower Court). Finally, at the time of the decision by the Contemporary Lower Court, it is hierarchically supervised by the Contemporary Supreme Court.

Now consider the arrows in Figure 1— each of which denotes a potential source of influence on the interpreting court. As depicted, four sources can directly influence the Contemporary Lower Court: certainly the Enacting Supreme Court’s original decision in the case (*vertical stare decisis*), but also the interpretations of that case by Subsequent Supreme Courts (again, *vertical stare decisis*); the Earlier Lower Court’s previous interpretations (*horizontal stare decisis*); and the lower court’s expectations about the *preferences* of the Contemporary Supreme Court on that case or, what we might call, *hierarchical conformity*.

The Enacting Supreme Court may also influence the Contemporary Lower Court *indirectly*, in three ways. First, Enacting Supreme Court decisions may affect the interpretations given by Subsequent Supreme Courts, which can directly influence the Contemporary Lower Court. Second, the Enacting Supreme Court can affect the interpretations given by Earlier Lower Courts (those that rule on the issue prior to the Contemporary Lower Court), which then affects the Contemporary Lower Court through horizontal stare decisis. Third, the decision by the enacting high court could influence the preferences of the Contemporary Supreme Court, which, as above, influences the decisions of the Contemporary Lower Court.

Judicial scholars have not missed these potential influences. Actually, both horizontal relations between coordinate courts and hierarchical relations between superior and inferior

courts have generated extensive literatures. (Llewellyn 1960; Levi 1948; Hansford and Spriggs 2006; Spaeth and Segal 1999; Cohen 2001; Klein 2001). Unfortunately, though, these literatures typically consider horizontal and vertical relations *in isolation from one another*. It is relatively novel to map the full range of relationships as they bear on lower court treatment of high court precedent.

As a result, we lack analyses, formal or otherwise, detailing how the combined influences of horizontal and vertical stare decisis interact to determine lower court behavior. For purposes of structuring an empirical investigation, however, we can draw on strands from the existing literature to fashion two quite distinct analytical frameworks that can structure an empirical investigation: *team-oriented* and *agency-based* approaches to judicial politics. The former usually assumes a shared conception of justice among judges, combined with intrinsic uncertainty about the world and pervasive resource scarcity. Attention to signals from earlier courts, both high and low, help team members more efficiently discover, administer, and modify legal rules on which all agree. In contrast, agency-theoretic approaches assume pervasive value conflict among judges. On this account, hierarchy forces the doctrinal preferences of superiors on resistant and rambunctious inferiors, while horizontal stare decisis is an implicit bargain struck across generations to give longevity to actors' policy decisions.

In what follows, we elaborate both approaches to hierarchical interpretation. We also draw out their apparent empirical implications with an eye toward assessing them independently and collectively.

## 2.1 Team-Theoretic Framework

Within the legal academy, analytically-minded scholars who theorize about relations among courts have converged on an approach that draws on the theory of teams. A branch of economics, this theory concerns itself with the efficient organization of individuals who share a common goal but control different decision variables and who base their decisions on possibly different information (Marshack and Radner 1972). As applied to the judiciary, the essential idea is to assume that judges have no value conflicts whatsoever; formally, they all share the same utility function. Hence, the object of judging is to decide “correctly” as many cases as possible in the face of an uncertain world and resource constraints (see e.g., Caminker 1994a, b; Daughety and Reinganum 2000; Kornhauser 1989, 1995; Rogers 1995; Shavell 1995).

To help fix ideas, it may be useful to consider a simplified set-up illustrating the thrust of recent work. In the model, a case presents itself to a judge, who must render a judgment,  $x$ . For the sake of exposition, we treat such decisions as dichotomous, e.g., “admit the evidence” versus “exclude the evidence” or “statute is constitutional” versus “statute is unconstitutional,” and so on. So,  $X = \{0,1\}$ . For the judge, the correct judgment depends on the “state of the world”  $\omega \in \Omega$ . The concept of the “state of the world” connotes all the facts in the case and all the social and other circumstances surrounding the case. For purposes of illustration, assume the state of the world may only be of two types,  $\Omega = \{0,1\}$ .

The utility to the judge of her judgment is

$$u(x; \omega) = \begin{cases} 1 & \text{if } x = \omega \\ 0 & \text{if } x \neq \omega \end{cases}$$

In other words, the judge would like to match the judgment to the state of the world – for all members of the judicial team, this is the “correct” judgment in the case.

Unfortunately for the judge, the true state of the world cannot be known with certainty: she is necessarily somewhat uncertain which judgment is best, given the circumstances in the case, social and political relations, technology, and so on. However, the judge has an initial belief about the probability  $\omega = 1$ ,  $\mu$ . In addition, the judge may observe one or more noisy signals about the state of the world,  $s \in S = \{0,1\}$ . Each signal depends probabilistically on the state of the world. For example, the probability of observing  $s = 0$  when  $\omega = 0$  is  $\tau(0)$  (and the probability of observing  $s = 1$  when  $\omega = 0$  is  $1 - \tau(0)$ ), while the probability of observing  $s = 1$  when  $\omega = 1$  is  $\tau(1)$  (and the probability of observing  $s = 0$  when  $\omega = 1$  is  $1 - \tau(1)$ ). Clearly, the closer  $\tau(0)$  and  $\tau(1)$  are to one, the more precise or informative the signals. When  $\tau(0)$  and  $\tau(1)$  are close to .5, the signals are uninformative.

Finally, we consider one more probability,  $\pi$ , the probability the state of world exogenously shifts from one state to the other. In other words, despite earlier evidence about the state of the world in a previous period, social circumstances may have changed so much that the alternative judgment is now the better choice for the case. Again, the judge may observe a signal,  $w$ , providing noisy information about the likelihood of such a social change. Having observed one or more signals about the state of the world (and hence the best legal doctrine), and one or

more signals about social change, the judge updates her subjective perception of the most probable state of the world, using Bayes's Rule. Thus, judicial learning about cases and doctrine are typically at the heart of team models. Call the resulting posterior beliefs  $\nu(\omega | s, w)$ . Then, using this posterior belief, she renders the judgment maximizing her expected utility.

Just to be clear, the "moving pieces" in the model are 1) prior beliefs about the state of the world (that is, what is the best judgment in the case), 2) the number of "s" signals about the likely state of the world (that is, signals about the best judgment in the case) 3) the precision of those signals (that is, the closeness of  $\tau(0)$  and  $\tau(1)$  to one), 4) the number of "w" signals about social change, 5) the precision of those signals, 6) the posterior belief about the state of the world ( $\nu(\omega | s, w)$ ), and 7) the judgment in the case.

Using models of this flavor, Cooter, Kornhauser, and Lane (1979), Blume and Rubinfeld (1982), and Bueno de Mesquita and Stephenson (2006) formally analyze inter-temporal relations between coordinate courts. In Cooter et al., a court (or succession of such courts – within a team setting, the two are virtually equivalent) learns about the consequences of a legal rule by hearing cases. In other words, the act of hearing cases generates "s" signals about the state of the world. It may also observe "w" signals about changes in the state of the world. The court then updates its appreciation of the rule's social consequences (it estimates its posterior belief  $\nu(\omega | s, w)$ ) and adjusts its prior rule incrementally. Thus, learning and altered social circumstances lead the court to update its posterior beliefs about the best legal doctrine and modify its prior rulings in light of those altered perceptions. Note as well that an implication of Bayesian updating is that signals

contradicting the court's prior belief will be more influential if the prior belief itself was weaker (e.g., if an initial signal from the Supreme Court helping to establish the earlier prior belief was itself imprecise or uncertain).

Blume and Rubinfeld (1982) and Bueno de Mesquita and Stephenson (2006) examine the role of transactions costs in slowing the modification of precedent. To grasp the thrust of their analyses, suppose the court's beliefs have become quite certain about the state of the world. But then the court observes "w" signals suggesting the state of the world has changed. If sufficiently compelling, the signals would lead the court to change its doctrine, as in Cooter et al. But if agents outside the court have incorporated expectations about the court's behavior into contracts or other forms of private behavior, the agents may face steep transition costs from the new – albeit superior – doctrine. Blume and Rubinfeld (1982) show that, if the court cares about those transition costs, it may somewhat delay its response to the new information (the "w" signals). In a somewhat similar vein, Bueno de Mesquita and Stephenson (2006) emphasize the chilling effect of possible mis-communication between a court and agents who respond to its decisions. In the model, the confusion and interpretive uncertainty of the agents in the face of a new rule may deter a court from modifying its precedent despite information (e.g., the "w" signals) that the altered doctrine would be superior. In neither model do transaction costs afford unlimited sticking power to precedents in the face of new circumstances. Rather, the transactions costs simply mute the court's responses to "w" signals.

These studies have straightforward implications for horizontal relations in the U.S. Courts of Appeals. First, a succession of "s" signals from similar earlier cases in the circuit

confirming an initial precedent in the circuit should tend to convince the contemporary court that the initial precedent was correct. Conversely, a succession of “s” signals questioning the precedent should do the opposite. Similarly, an absence of indications of social change (“w” signals) should increase adherence to earlier precedent in the circuit, while an abundance of signals suggesting social change may lead the contemporary court to deviate from earlier precedent in the circuit. A confused or weak initial signal from the Supreme Court will lead to weaker priors in the earlier courts, magnifying the impact of subsequent “s” and “w” signals. In this case, adherence to the initial “weak” precedent is apt to be lower.

Finally, Cameron and Kornhauser (2005) use a formal team-theoretic model to study hierarchical relations in the judiciary. An obvious point in team models of the judicial hierarchy is that cases heard by the Supreme Court are apt to generate many precise “s” signals to lower courts because the Supreme Court is composed of nine highly skilled judges with abundant resources (Kornhauser 1995; Caminker 1994a, b). So, lower courts are apt to follow the lead established by the high court – one or more favorable treatments of a precedent by the high court are likely to lead to more favorable treatments by the lower court. Conversely, repeated unfavorable treatments by the high court are likely to generate similar behavior in the lower court. Cameron and Kornhauser’s model of error correction provides an additional reason why a lower court may adhere to a high court precedent even when its application leads to an apparent error (that is, even though the lower court receives an “s” signal during the case suggesting the state of the world is different from that suggested by earlier high court cases). In the model, adherence to an apparently incorrect high court precedent ultimately assures more accurate

processing of cases because it encourages litigants in “hard cases” (that is, ones in which  $\omega$  is quite uncertain) to appeal. The appeal reveals additional information to a resource-rich higher court that is then likely to adjudicate the case correctly. The model thus underscores the sticking power of precedent in team models of the judicial hierarchy.

### **2.1.1. Empirical Expectations: Team-Theoretic Approach**

From the existing team-theoretic models of judicial politics, we derive the following empirical expectations about lower court treatment of Supreme Court precedent—each of which we summarize in terms of the relevant variables.

1. (*Weak initial signals from the Supreme Court*) When the original precedent sends mixed, imprecise or weak signals, it will have little effect on the beliefs of lower courts about the best doctrine. Under these circumstances, all forms of stare decisis are likely to be weaker: lower courts are more likely to rely on their own signals. Accordingly, we expect that when the initial opinion resulted from a fractured court and generated numerous dissents or concurring opinions and the case involved many complex issues, the prior beliefs of lower courts about proper behavior are apt to shift little. If so, Earlier and Contemporary Lower Courts are more likely to treat the high court precedent negatively. Conversely, precedents in simpler cases established by a unanimous court or that failed to generate many dissents are more likely to be treated favorably in subsequent lower court opinions.
2. (*Social change*). When the original precedent is perceived as antiquated due to rapid

social change, lower courts are more likely to treat negatively the precedent. Conversely, absent much social change, lower courts are more likely to treat positively high court precedent.

3. (*Subsequent Supreme Court treatments*) When a succession of high courts has indicated the original precedent needs revision due to new evidence or changed circumstances, the Contemporary Lower Court is more likely to treat the precedent negatively. Conversely, if a succession of high courts have continued to favor the original precedent, the Contemporary Lower Court is more likely to treat the precedent favorably.
4. (*Earlier Lower Court treatments*). When the lower court in earlier incarnations considered the original precedent and treated it negatively, the Contemporary Lower Court is more likely to treat the precedent negatively. Conversely, if the lower court in earlier incarnations considered the original precedent and treated it positively, the Contemporary Lower Court is more likely to treat the precedent positively.

## **2.2 Agency-Theoretic Framework**

While the team conception of judiciaries has gained momentum in law schools, it is agency theory that has attracted the attention of most political scientists interested in the operation of judiciaries (e.g., Moe 1984; Songer et al 1994; Brent 1999). The fundamental assumption of agency theory is that value conflicts are pervasive in organizations. Hence, the creation and application of legal rules are occasions for political struggle.

To fashion agency models of the judiciary it might seem natural to simply add conflict

over the “correct” judgment to the team framework: Judges with different preferences would still face doctrinal uncertainty and engage in learning over time. Existing agency models, however, typically suppress the kind of judicial uncertainty so prominent in team models in order to focus on the implications of value conflict. To see the point, reconsider the earlier illustrative model and assume the state of the world is fixed and known with certainty by all, such that inherent doctrinal uncertainty is removed from the table. But also suppose that judges have doctrinal preferences indexed by judge type  $t \in T = \{0,1\}$ . That is, assume there are two ideological varieties of judges. For both,

$$u(x;t) = \begin{cases} 1 & \text{if } x = t \\ 0 & \text{if } x \neq t \end{cases}$$

Thus, type “0” judges wish to render judgment  $x = 0$ , while type “1” judges wish to render  $x = 1$ . In this setting learning, based on the “ $s$ ” signals about the state of the world or “ $w$ ” signals about social change, plays no role whatever. Rather, the emphasis is on a struggle over outcomes by the two varieties of judges. Consequently, the key moving parts to the models typically are 1) the ideological distance between the enacting high court and the contemporary high court (whose effects will be transmitted to lower courts when vertical control devices are effective), 2) the ideological distance between the Contemporary Lower Court and the enacting high court (which will manifest itself when the lower court can assert a degree of autonomy from the higher court), 3) the ideological distance between the Contemporary Lower Court and the contemporary high court (which may weaken the effectiveness of vertical control devices).

Invoking a model of this general kind, Rasmusen (1994) formally analyzes horizontal

stare decisis in an agency setting. The model examines the behavior of a succession of courts over time that differ in preferences and that may overturn the judgments of earlier courts as well as render judgments in new cases themselves. Rasmusen's model identifies the fundamental dilemma of horizontal stare decisis in an agency framework: *each court would like to annihilate and re-write the opinions of all earlier courts but would like its own opinions to be immortal.* The extent of horizontal stare decisis reflects the efforts of successive courts to resolve this inter-generational tension.

One obvious outcome in such a setting is the total absence of horizontal stare decisis, so later courts simply overturn earlier precedent. But Rasmusen shows the possibility of other equilibria. In these equilibria, a future-minded court respects the opinions of earlier courts (at least to some degree), thereby upholding an implicit cross-generational deal. This deal is attractive to the contemporary judge because it affords her own opinions a degree of durability in the future, when future judges will similarly defer to her opinions in accord with the implicit deal. At least some deference to precedent is the cost for achieving durability for her own rulings.<sup>2</sup>

Two features of the "horizontal stare decisis" equilibrium stand out. First, adherence to precedent is less likely when the sitting court finds a precedent highly objectionable. This will be more likely if the policy preferences of the sitting court are distant from those of the enacting court. Second, adherence to precedent is less likely if the precedent is old. Essentially, the intergenerational log-roll involves a moving window: older precedents are discarded while younger ones are afforded respect. These two features seem likely to emerge in any agency-

---

<sup>2</sup> The model is similar to overlapping generations models and equilibria are constructed accordingly.

theoretic model of horizontal stare decisis.

Applying the insights of Rasmusen's analysis to hierarchical interpretation is relatively straightforward. In light of the model, one would expect a sitting Supreme Court to be more inclined to treat positively precedents established by an ideologically and temporally proximal predecessor, and less likely to treat positively those from a temporally and ideologically distant predecessor. As this prediction deals with Supreme Court treatment of precedent, we do not test it directly. To the extent a hierarchical power relation exists between the higher court and the lower courts (as discussed momentarily), however, *these effects should be transmitted to lower courts*. Consequently, lower courts will be more inclined to treat negatively precedents enacted by a Supreme Court ideologically distant from the Contemporary Supreme Court, especially if the precedent is an older one.

Turning to vertical relations, we can identify many more modeling exercises (Cameron et. al 2000, Doughety and Reinganum 2006, McNollgast 1995, Spitzer & Talley 2000). In these, a hierarchically inferior court attempts to deviate from the doctrine preferred by a hierarchically superior court with distinct preferences. And, the hierarchically superior court uses a variety of devices to force its doctrinal preferences on the resistant lower court.

As one would expect from the broader principal-agent literature (Bendor et al 2001, Dixit 2002), information asymmetries are critical in agency models of this kind. For example, the inferior court may be able to take an action that is effectively hidden from a distracted or overburdened superior court. Or, the lower court may have private access to facts about the case (by virtue of having heard the case) that could justify an apparently deviating decision. But the high

court is not defenseless in the face of these tactics. The Supreme Court can use the certiorari process to “audit” the lower court to determine if it is “cheating”; if so, the Supreme Court can reverse the lower court’s decision (Cameron et al. 2000, Spitzer and Talley 2000). Allies of the high court on the lower court may act as whistleblowers, alerting the high court to attempted deviations (Cross and Tiller 1998, Doughety and Reinganum 2006).

As in the certiorari process, many other features of the judicial hierarchy, such as litigant-based appeals, can be understood as devices for the Contemporary Supreme Court to control lower courts, despite information asymmetries. To the extent these devices are effective, the treatment of earlier precedent in the lower courts will track the preferences of the Contemporary Supreme Court and not those of the lower courts themselves. But, to the extent these devices are less than completely effective, the treatment of earlier precedent in the lower courts will also reflect the preferences of the lower court judges hearing the case.

### **2.2.1 Empirical Expectations: Agency Approach**

Because agency theory predict *patterns* in the signs of variables it is more convenient to discuss empirical expectations in terms of these patterns rather than each variable separately.

First, agency models of horizontal relations among courts identify two critical variables in the adherence of the sitting Supreme Court to its own precedents: 1) the difference in preferences between the sitting Supreme Court and the enacting Supreme Court, and 2) the age of the original precedent. Accordingly, *to the extent the Supreme Court’s devices for vertical control are effective*, an increase in the distance between the sitting Supreme Court’s preferences

and the enacting Supreme Court's preferences should increase the likelihood a lower court panel deviates from Supreme Court precedent. Similarly, an increase in the age of the Supreme Court precedent should increase the probability of a lower court deviation. In addition, to the extent vertical control devices are effective, increasing differences in preferences between the lower court panel and the enacting Supreme Court and increasing differences in preferences between the lower court panel and the sitting Supreme Court should *not* additionally increase the probability the panel deviates from Supreme Court precedent.

Conversely, *to the extent the Supreme Court's devices for vertical control are ineffective*, the lower courts ought display a degree of autonomy from the desires of the sitting Supreme Court. Accordingly, increasing differences between the sitting Supreme Court and the enacting Supreme Court should not increase the probability a lower court deviates from earlier Supreme Court precedent. Instead, we expect that increasing differences in preferences between the lower court panel and the enacting Supreme Court will increase the probability the panel deviates from earlier Supreme Court precedent. And this effect may be enhanced when the distance between the Contemporary Lower Court and the sitting Supreme Court is larger.

### **3. Data and Methods**

In order to explore the empirical implications of the two theoretical frameworks, we employ a four-step approach: (1) generate a random sample of U.S. Supreme Court cases; (2) track the responses of lower courts to the doctrine established in these cases (the dependent variable for all the hypotheses); (3) collect the data necessary to animate the independent variables; and (4)

implement statistical models, testing for the influence of the variables suggested by each framework. Our focus throughout is the response of panels in the U.S. Courts of Appeal to precedent established by the U.S. Supreme Court.

We drew a random sample (from Spaeth 2006) of 500 cases from all orally-argued Supreme Court cases decided between the 1953 and 1990 terms.<sup>3</sup> Next, following other studies, we employed *Shepard's Citations* to determine how circuit court panels responded to each U.S. Supreme Court case, through 2000 (Johnson 1987, Spaeth and Segal 1999, Spriggs and Hansford 2000, 2001). Shepard's (which we accessed via LEXIS) identifies every decision produced by a U.S. Court of Appeals that "treated" the Supreme Court cases. It also specifies the nature of the circuit court's treatment (e.g., "followed," "explained," "criticized," the precedent) – thereby enabling us to capture compliance with and deviations from extant rules in ways widely recognized as reliable, valid and accessible by legal practitioners and scholars.<sup>4</sup>

The random sample of 500 U.S. Supreme Court cases generated 10,244 citations distributed over the Shepard's treatment categories, or a mean of 20.5 circuit court citations per case. Following Spriggs and Hansford (2000), we collapsed the finely articulated Shepard's

---

<sup>3</sup> The random sample of size 500 represents about 10% of all orally argued cases (N=4879) during the time frame and using `dec_type=1, 6, or 7` and `analu=0` in Spaeth (2006). We end with the 1990 term in order to provide a sufficient time horizon for lower court responses.

<sup>4</sup> Spriggs and Hansford (2000) demonstrate a high degree of reliability in the case treatment conducted by *Shepard's*. While *Shepard's* is surely not perfect, no other method has been shown to have a similar degree of reliability and efficiency, especially when coding large numbers of cases.

categories into three broader categories: “Deviate,” “Neutral,” and “Comply.”<sup>5</sup> “Comply” was the modal response, though roughly one out of three responses fell into the “Deviate” category.<sup>6</sup> Of these, most resulted from the Shepard’s category “distinguished,”<sup>7</sup> but “criticized” and “overruled” cases occurred as well. We code “Deviate” as the baseline category of 1, “Neutral” as 2, and “Positive” as 3. Therefore, negative coefficients indicate an increasing likelihood of a negative treatment, and positive coefficients indicate an increasing likelihood of a positive treatment.

-- Insert Table 1 about here --

Table 1 lists the key independent variables suggested by each framework, as well as the measures and sources of data used.<sup>8</sup> Because many are entirely conventional, they require little elaboration. To measure social change, for example, we counted the number of articles appearing

---

<sup>5</sup> We experimented with a dichotomous classification, alternative tripartite ones, and a five-category dependent variable. Broadly speaking, the results we report are robust to these specifications.

<sup>6</sup> 5109 (49.87%) fell into the “Comply category; 2126 (21.11%) were “Neutral” (which included “explained” and “harmonized” treatments) and 2973 (29.02%) were “Deviate” (which included “distinguished,” “questioned,” “criticized,” “overruled,” “not followed,” and “limited” treatments).

<sup>7</sup> Some commentators have suggested to us that “distinguished” treatments are not always deviations. While we agree that they may not be as harsh as, say, overrulings, *Shepard's* characterizes them as deviant treatments—as do Hansford and Spriggs (2006) who conducted extensive reliability and validity checks on *Shepard's*. Our inspection of the treatments confirms the Hansford/Spriggs analysis (see also Spriggs and Hansford 2000).

<sup>8</sup> Table A-1, in the Appendix, provides descriptive statistics on the variables.

in the *New York Times* in the three years prior to the panel decision concerning the issue in question.<sup>9</sup> We assume that an increase in the *New York Times* coverage of an issue area suggests increasing amount of social change within that particular area. Following Maltzman, Spriggs, and Wahlbeck (2000), issue complexity is the number of legal provisions plus the number of legal issues present in the precedent as coded directly from Spaeth's U.S. Supreme Court Database. The database is also the source for the number of concurring and dissenting opinions in the precedent. Subsequent signals concerning the precedent that are generated by both the Supreme Court and the respective circuit are simply coded as the number of positive and negative treatments of the precedent by the Supreme Court and within the circuit prior to the panel's decision. We also include the age of the precedent at the time of the panel's decision.

Somewhat more novel are our measures for assessing ideological distances, which are at the center of the agency-theoretic framework. To enable these comparisons, we adopt the approach in Epstein et al. (2007). In brief, Epstein and her colleagues developed a measurement strategy designed to place Supreme Court justices and Court of Appeals judges into a policy space, the "Judicial Common Space" (JCS). To assess the preferences of justices, the authors relied on the Martin-Quinn scores (Martin and Quinn 2005); for circuit court judges, they invoked the inferential measure developed by Giles, Hettinger, and Peppers (2001), based upon the Poole and Rosenthal Common Space scores for the judges' senatorial sponsors. The final step taken by Epstein et al. was to transform the scale of the Martin-Quinn scores into that of the Common Space scores, thus resulting in a comparable metric for the ideology of Supreme Court justices and Court of Appeals judges.

---

<sup>9</sup> For ease of interpretation, we divide the number of articles by 1000.

We employ these scores to measure ideological distances for the actors identified as critical in the agency-theoretic framework (these measures play no role in the team-theoretic framework, of course). More specifically, we attribute to the Enacting Supreme Court, the JCS score of the median member of the majority coalition;<sup>10</sup> for the Contemporary Supreme Court and the contemporary panel, we employ the score of the median member of the respective court or panel.

-- Insert Figure 2 about here --

Figure 2 illustrates the use of the JCS scores, depicting the scores for a lower court deviation mentioned earlier: *Dickerson* (a departure from *Miranda*). As shown, the enacting (*Miranda*) Court was well to the left of the Fourth Circuit panel and the Supreme Court in 1990 (the year *Dickerson* was decided). Thus, from an agency perspective, it is not surprising that the Fourth engaged in doctrinal deviation in *Dickerson*: its behavior can be seen as an example of hierarchical conformity to the Contemporary Supreme Court's doctrinal preferences. Alternatively, from a team perspective, perhaps the lower court acted on signals in earlier years from other Fourth Circuit panels or the Supreme Court itself, indicating the original decision required modification in light of its consequences. Indeed, in the years leading up to *Dickerson*, the Supreme Court had not been entirely unwavering to its 1966 rule.<sup>11</sup>

---

<sup>10</sup> For variables relying on the enacting Court, we checked an alternative measure: the median member of the Court rather than the median member of the majority coalition. The results (as depicted in Table 2) do not change in any meaningful statistical or substantive way.

<sup>11</sup> *Shepard's* lists several instances of negative treatments, including *Illinois v. Perkins* (1990), *Oregon v. Elstad* (1985), and *Roberts v. United States* (1980). See also the Fourth Circuit's decision in *Dickerson*.

## 4. Empirical Analysis

This is but arm-chair speculation on our part. Only by systematically (and, eventually, simultaneously) analyzing the empirical implications of the team and agency frameworks can we reach conclusions of a higher and more certain quality.

To this end, we constructed three ordinal logit models (recall the dependent variable has three categories), as shown in Table 2.<sup>12</sup> The first contains the variables associated with the team framework, the second with variables drawn from the agency framework, while the third combines the two. We consider each separately.

-- Insert Table 2 about here --

### 4.1 Team Model

The theoretical discussion identified four hypotheses about a lower court's adherence to precedent: 1) weak initial signals by the Supreme Court increase the probability of deviation; 2) dramatic social change increases the probability of deviation; 3) negative signals (unfavorable treatments of the precedent) by earlier panels in the circuit increase the probability of deviation while supportive signals (favorable treatments of the precedent) increase the probability of

---

<sup>12</sup> The ordered logit model assumes that the effects of the estimated coefficients are constant across the choices (the proportional odds assumption). Straightforward Wald tests can be used to test this assumption. In the fully specified model, the number of dissenting opinions, the number of negative treatments by the circuit, and the age of the precedent demonstrate some evidence of failing the proportional odds assumption. We estimated a series of constrained and unconstrained ordinal logit models in which the necessary assumption is relaxed. Based on this diagnostic work and alternative model specifications, we find no need to alter the substantive conclusions presented here.

compliance; and 4) negative signals by earlier high courts increase the probability of deviation while supportive signals increase the probability of compliance. The variables in the left column of Table 2 test each of these hypotheses.

The data provide weak support for the second hypothesis but little or none for the first. Social change, as measured by number of *New York Times* stories in an issue area, does increase the probability of deviation from the precedent; the coefficient on the variable is statistically significantly different from zero but only at the .10 level. The three variables associated with a weak initial signal by the Supreme Court appear to have little if any impact of lower courts' subsequent behavior. From the perspective of team theory, one may interpret the result as suggesting that both unified and divided majority opinions from the Supreme Court function with equal force.

The four variables associated with the two subsequent-treatment hypotheses, on the other hand, operate exactly as predicted by the team framework. Favorable treatments of the precedent by earlier panels increase the probability of favorable treatment, while unfavorable treatments increase the probability of deviation. Similar results hold for treatments by the Supreme Court. Somewhat surprisingly, though, the magnitudes of the coefficients on the circuit treatments are much larger than those of the corresponding variables for Supreme Court treatments. (We return to the substantive significance of the “treatment” coefficients shortly).

#### **4.2 Agency Model**

The middle column in Table 2 presents a model incorporating the identified agency variables. As shown, the Enacting Supreme Court-Contemporary Supreme Court distance is

strongly statistically significant and signed correctly. But none of the other preference variables attain statistically significant levels.

Because the coefficient on Enacting Supreme Court-Contemporary Supreme Court distance is significantly different from zero, the results may be interpreted as supportive of an agency approach. From the perspective of agency theory, the importance of this variable coupled with the low impact of the two other distance measures suggests that vertical controls in the judiciary hierarchy are quite effective.<sup>13</sup>

### **4.3 Combined Model**

Thus far our discussion of the results has mirrored existing theory by explicating team and agency models separately. Not unexpectedly, though, it is a combination of the two perspectives that provides the greatest insight into the operation of the judiciary.

As depicted in Table 2, within the combined model several variables stand out as important. First, the age of precedent emerges as statistically significant, with the sign suggested by the discussion of horizontal stare decisis in agency models. Substantively, positive treatments are about 12% less likely for the oldest precedent in our sample than for a new precedent.<sup>14</sup>

---

<sup>13</sup> A variety of alternative specifications consistently yield similar results. For example, the panel-Contemporary Supreme Court distance interacted with the other distance measures was never statistically significant.

<sup>14</sup> As a robustness check, we operationalized the age of the precedent in a variety of different ways. To examine possible non-linearities in the impact of the age of precedent variable, we have included a squared term for the variable. However, the coefficient on the variable was not statistically different from zero.

Second, the coefficients on the two circuit level treatment variables not only remain statistically significant but are of substantive importance as well. Highlighting this point is Figure 3, which examines the substantive effects of the number of positive and negative treatments of the precedent within a circuit. As a circuit treats a precedent positively with increasing frequency, the likelihood of another positive treatment substantially increases. With no prior circuit positive treatments, the likelihood of a positive treatment is .42. Precedents with merely three positive treatments are at even odds to have a positive treatment. This increases to a predicted probability of .64 with 10 positive treatments and to over .90 with 30 positive treatments.

-- Insert Figure 3 about here --

The number of negative circuit court treatments has a similar effect. Again, holding all other values at their means (including the number of positive circuit treatments), a case with three negative treatments is still at even odds for a positive treatment. However, a case with 12 negative treatments becomes slightly more likely to receive a negative reception (.39) than a positive one (.38). By 30 negative treatments, the probability of a positive treatment is less than .20, while the likelihood of a negative treatment increases to .62. This suggests that intra-circuit development of Supreme Court precedent is an essential element of the team approach.

Finally, the distance from the Enacting Supreme Court to the Contemporary Supreme Court remains highly statistically significant, and too is substantively interesting, as Figure 4 illustrates. There, we show the impact of the distance between the Contemporary Supreme Court and the Enacting Supreme Court on the likelihood of compliance in a circuit.

-- Insert Figure 4 about here --

Setting all the other variables in the combined model at their means, we estimate the likelihood of compliance at .50 (the 95% confidence interval is [.49-.51]); that figure for the probability of defiance is .28 [.27-.29].<sup>15</sup> Note, though, how that probability increases (or decreases) depending on earlier treatments of the precedent in the circuit. When a circuit has not negatively treated a precedent and lends its support to a past precedent (that is, when we set the number of lower court positive treatments at 1 standard deviation above its mean, or 10 positive treatments and negative treatments at 0; the dashed line with open circles in Figure 4), and when the distance between the Enacting and Contemporary Supreme Courts is minimal (0), circuits are extremely unwilling to act in defiance of the Court: close to 70% of the time they will comply; in fewer than 15 percent will they deviate. Those percentages decrease (increase) to about 58% and 22% as the distance variable moves to maximum levels. Now, consider the willingness of the circuits to deviate from a precedent as negative treatments within a circuit accumulate with no positive treatments. Setting the number of negative circuit treatments at 1 standard deviation above the mean (8 negative treatments; the dashed line with diamonds in Figure 4), the probability of a positive treatment is only .39. Moving to the most extreme levels of distance between the Enacting and Contemporary Supreme Courts, the odds of defiance are slightly better than 50-50;<sup>16</sup> compliance falls to less than 30 percent (predicted probability of .28 [.24-.31]).

Finally, observe that even if the circuit has failed to send any positive or negative signals about an existing precedent (the solid line in Figure 4), the circuit court panels appear to attend to the ideological distance between the enacting and Contemporary Supreme Courts. Indeed, when

---

<sup>15</sup> The predicted probability for neutral treatment is .22 [.21-.23].

<sup>16</sup> Under these conditions, the probability of a negative treatment is .51.

they perceive that distance to be minimal, they are far more likely to comply than to deviate (a predicted probability of .49 versus .29). But as that distance increases, the odds of compliance decrease to the point, at the most extreme levels, where circuits are more likely deviate than to comply (.37 versus .41).

## **5. Discussion and Conclusion**

Broadly speaking, we find that circuit court panels are quite attentive to earlier legal developments within their circuit, in the fashion suggested by team theory. But we also find that circuit panels are quite sensitive to the shifting preferences of the Contemporary Supreme Court—so much so that earlier treatments of the precedent by the justices do not seem to matter much once one controls for the preferences of the Contemporary Supreme Court (see Table 2). This finding is strongly evocative of agency theory.

The picture that emerges from the data is, we think, broadly the following. A Supreme Court precedent initiates a period of learning within a circuit, as the judges hear cases, ponder them, and devise the most congenial ways to apply it. Over time, the accumulating experience of the judges weighs ever more heavily within the circuit. At the same time, however, the circuit judges remain acutely aware of the shifting ideological composition of the Supreme Court. Thus, as cases present themselves, the circuit judges anticipate high court treatment of their efforts, and continually nudge or bump the circuit legacy to reflect the likely desires of their hierarchical superiors. This anticipation of Supreme Court responses explains why the high court's earlier treatments seem to have little direct impact in the combined model: in response to ideological

change on the Supreme Court, the lower court judges have already adjusted their rulings.

This interpretation of the statistical results necessarily remains speculative, but it finds support in a variety of other studies employing very different methods and data. Klein (2001), for example, demonstrates the importance of intra-circuit considerations, such as creating good legal rules. Songer and Sheehan (1990) show that the appointment of liberal (Democratic) judges had no meaningful effect on their behavior in most areas of the law; instead, these judges tended to follow the Supreme Court's decisional patterns. Likewise, Brent (1999) finds that the lower courts, regardless of their ideological propensities, grew increasingly reluctant to rule in favor of free exercise claimants after the Supreme Court's rulings in *Smith* and *Boerne*. Caminker (1994a) provides evidence that federal courts have adopted a “predictive” approach when discerning state law pursuant to the Erie doctrine.<sup>17</sup> Other legal scholars have noted the anticipatory, predictive, or forward-looking behavior of lower courts as well (e.g., Bradford 1990 and Kniffin 1982).

If this interpretation of the data is correct, several important implications follow for how one views the operation of the federal judiciary. Particularly striking are the implications for *legal change*, *strategic litigation*, and *political control of the judiciary*.

With respect to legal change, the circuits certainly appear responsive to the Supreme Court's initiation of precedent, even for cases from fractured courts or with many dissents. But then the circuits shape doctrine themselves (albeit with an eye on the Contemporary Supreme Court). Repeated negative treatments within a circuit tend to beget further negative treatments. As a consequence, doctrinal development may well produce a degree of judicial erosion and

---

<sup>17</sup> On the other hand, Cross (2005) asserts that appellate courts are more likely to follow their own ideological inclinations and the preferences of the enacting Supreme Court when deciding whether to adhere to precedent.

perhaps, eventually, virtual eradication of the doctrine in question (see generally Hansford and Spriggs 2006; Spriggs and Hansford 2001). Moreover, the data suggest the mere presence of Supreme Court justices who are ideological distinct from an extant rule could induce circuit judges to deviate from it.

Because this process of legal change is apt to take place at different rates – or even perhaps even in different directions – in different circuits, somewhat distinctive doctrines are apt to be active within the circuits at any given time. This has broad implication for litigation strategies, particularly for forum shopping. For instance, some circuits are apt to be much more receptive than others to the efforts of interest groups striving to move the law in favorable directions. Thus, finding the right circuit would certainly seem to be a relevant consideration since, even if Supreme Court precedent does not favor a litigant, the presence of negative treatments within a circuit may be critical for the disposition of the case there.

Finally, the results underscore the importance of appointments to the Supreme Court for shaping doctrine, not simply on the high court itself, but throughout the judicial hierarchy (Epstein and Segal 2005). Because our findings indicate that lower court judges are so sensitive to even incremental alterations in the Court's ideology – typically alterations brought about by membership changes – presidents and senators who desire sweeping legal change may be able to achieve a considerable measure of success simply through manipulating the Supreme Court appointments process. As the composition of the high court changes, the doctrine in the lower courts will tend to track along.

We conclude with two obvious directions for future research. First, on the theoretical

side, the empirical results suggest that team and agency theorists should consider how the two perspectives can speak to one another more directly. For example, agency theorists might consider the implications of judicial uncertainty and the learning dynamics studied in team models. Team theorists might consider how hierarchical pressure can alter learning dynamics. Second, on the empirical side, a useful extension of the present research would examine circuit treatments of precedents created within the circuits.<sup>18</sup> If the data display similar patterns – attention to prior circuit treatments but responsiveness to changing Supreme Court preferences – the portrait of what we might call “dynamic hierarchical interpretation” offered here would be even more compelling.

## Appendix

-- Insert Table A-1 about here --

## References

- Baum, Lawrence. 1978. “Lower Court Response to Supreme Court Decisions: Reconsidering a Negative Picture.” *Justice System Journal* 3:208-219.
- Baum, Lawrence. 1980. “Response of Federal District Judges to Court of Appeals Policies: An Exploration.” *Western Political Quarterly* 33:217-224.
- Bendor, Jonathan, Amihai Glazer, and Thomas Hammond. 2001. “Theories of Delegation.” *Annual Review of Political Science* 4: 235-69.
- Benesh, Sara C. and Malia Reddick. 2002. “Overruled. An Event History Analysis of Lower Court Reaction to Supreme Court Alteration of Precedent.” *Journal of Politics* 64: 535-550.

---

<sup>18</sup> Other extensions include the relationship between state courts of last resort and the Supreme Court, and between the circuits and the federal trial courts. See, e.g., Haire, Lindquist, and Songer 2003.

- Blume, Lawrence E. and Daniel L. Rubinfeld. 1982. "The Dynamics of the Legal Process." *Journal of Legal Studies* 11: 405-419.
- Brent, James. 1999. An Agent and Two Principals: U.S. Courts of Appeals Responses to *Employment Division, Department of Human Resources v. Smith* and the Religious Freedom Restoration Act." *American Politics Quarterly* 27: 236-266.
- Bueno de Mesquita, Ethan, and Matthew Stephenson. 2006. "Informative Precedent and Intrajudicial Communications," pp. 205-229 in James Rogers, Roy Flemming and Jon Bond (eds.) *Institutional Games and the U.S. Supreme Court*. Charlottesville: University of Virginia Press.
- Cameron, Charles and Lewis Kornhauser. 2005. "Decision Rules in a Judicial Hierarchy." *Journal of Institutional and Theoretical Economics* 161: 264-292.
- Cameron, Charles, Jeffrey Segal, and Donald Songer. 2000. "Strategic Auditing in a Political Hierarchy: An Informational Model of the Supreme Court's Certiorari Decision.," *American Political Science Review* 94: 101-116.
- Caminker, Evan H. 1994a. "Precedent and Prediction: The Forward-Looking Aspects of Inferior Court Decisionmaking." *Texas Law Review* 73: 1-82.
- Caminker, Evan H. 1994b. "Why Must Inferior Courts Obey Superior Court Precedent?" *Stanford Law Review* 46: 817-873.
- Canon, Bradley C. 1973. "Reactions of State Supreme Courts to a U. S. Supreme Court Civil Liberties Decision." *Law and Society Review* 8:109-34.
- Canon, Bradley C., and Charles A. Johnson. 1998. *Judicial Policies: Impact and Implementation*.

Washington, D.C.: CQ Press.

Cooter, Robert, Lewis Kornhauser and David Lane. 1979. "Liability Rules, Limited Information, and the Role of Precedent." *Bell Journal of Economics* 10: 366-373.

Cohen, Jonathan. 2001. *Inside Appellate Courts: The Impact of Court Organization on Judicial Decision Making in the United States Courts of Appeal*. Ann Arbor: University of Michigan Press.

Comment. 1941. "The Attitude of Lower Courts to Changing Precedents." *Yale Law Journal* 50:1448-1459.

Cover, Robert. 1975. *Justice Accused*. New Haven: Yale University Press.

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

Cross, Frank and Emerson Tiller. 1998. "Judicial Partisanship and Obedience to Legal Doctrine: Whistleblowing on the Federal Courts of Appeal." *Yale Law Journal* 107: 2156-2176.

Dixit, Avinash. 2002. "Incentives and Organizations in the Public Sector: An Interpretive Review." *Journal of Human Resources* 37: 696-727.

Doughety, Andrew and Jennifer Reinganum. 2000. "Appealing Judgments." *Rand Journal of Economics* 31: 502-525.

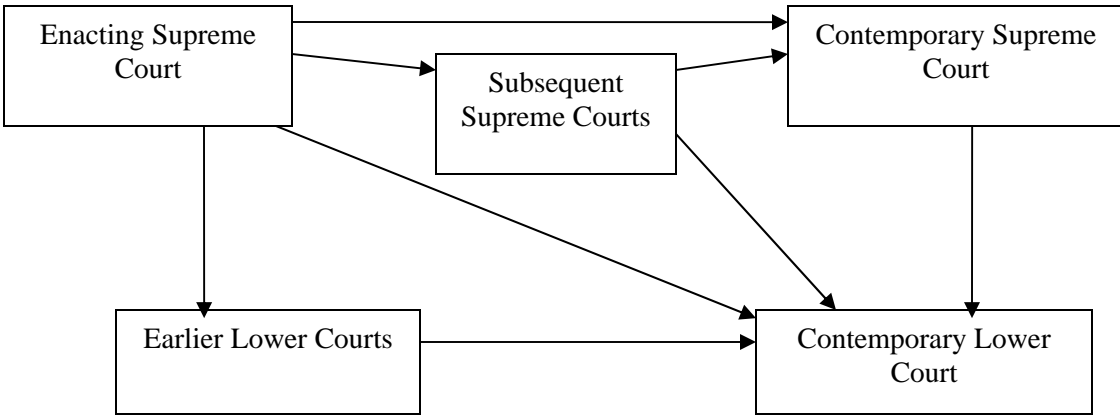
Doughety, Andrew and Jennifer Reinganum. 2006. "Speaking Up: A Model of Judicial Dissent and Discretionary Review." *Supreme Court Economic Review* 14:1-41.

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

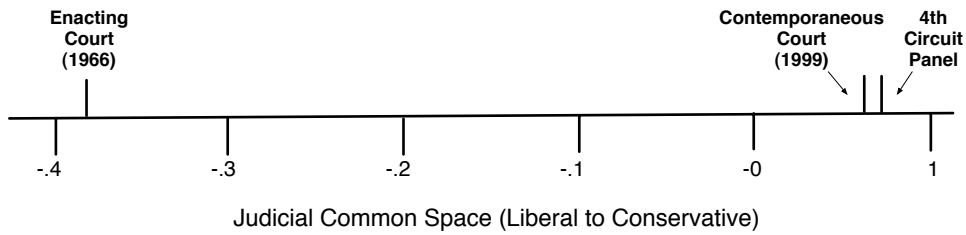
- Epstein, Lee, Andrew Martin, Jeffrey Segal, and Chad Westerland. 2007. "The Judicial Common Space." *Journal of Law, Economics, and Organization* 23: 303-325.
- Epstein, Lee, and Jeffrey A. Segal. 2005. *Advise and Consent: The Politics of Federal Judicial Appointments*. New York: Oxford University Press.
- Giles, Michael, Virginia Hettinger, and Todd Peppers. 2001. "Picking Federal Judges: A Note on Policy and Partisan Selection Agendas." *Political Research Quarterly* 54: 623-641.
- Gruhl, John. 1980. "The Supreme Court's Impact on the Law of Libel: Compliance by Lower Federal Courts." *Western Political Quarterly* 33:502-19.
- Haire, Susan B., Stefanie Lindquist, and Donald R. Songer. 2003. "Appellate Court Supervision in the Federal Judiciary: A Hierarchical Perspective." *Law and Society Review* 37: 143-167.
- Hansford, Thomas G. and James F. Spriggs II. 2006. *The Politics of Precedent on the U.S. Supreme Court*. Princeton: Princeton University Press.
- Howard, J. Woodford. 1981. *Courts of Appeals in the Federal Judicial System: A Study of the Second, Fifth, and District of Columbia Circuits*. Princeton, New Jersey: Princeton University Press.
- Johnson, Charles A. 1987. "Law, Politics, and Judicial Decision Making: Lower Federal Court Uses of Supreme Court Decisions." *Law and Society Review* 21:325-340.
- Klein, David. 2001. *Making Law in the US Courts of Appeals*. New York: Cambridge University Press.
- Kniffin, Margaret N. 1982. "Overruling Supreme Court Precedents: Anticipatory Action by the United States Courts of Appeals." *Fordham Law Review* 51: 53-89.

- Kornhauser, Lewis A. 1989. "An Economic Perspective on Stare Decisis." *Chicago-Kent Law Review* 65: 63-92.
- Kornhauser, Lewis A. 1995. "Adjudication by a Resource-Constrained Team: Hierarchy and Precedent in a Judicial System." *Southern California Law Review* 68: 1605-1629.
- Maltzman, Forest, James F. Spriggs II, and Paul Wahlbeck. 2000. *Crafting Law on the U.S. Supreme Court*. New York: Cambridge University Press.
- Marshack, Jacob and Roy Radnor. 1972. *Economic Theory of Teams*. New Haven: Cowles Foundation and Yale University Press.
- Martin, Andrew D. and Kevin Quinn. 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953-1999." *Political Analysis* 10: 134-153  
(scores updated at <http://mqscores.wustl.edu/press.php>)
- McNollgast. 1995. "Political Control of the Judiciary: A Positive Theory of Judicial Doctrine and the Rule of Law." *University of Southern California Law Review* 68: 1631-1683.
- Moe, Terry. 1984. "The New Economics of Organization," *American Journal of Political Science* 28: 739-777.
- Murphy, Walter F. 1959. "Lower Court Checks on Supreme Court Power." *American Political Science Review* 58: 1017-1031.
- Peltason, Jack. 1961. *Fifty Eight Lonely Men: Southern Judges and School Desegregation*. Urbana, IL: University of Illinois Press.
- Reid, Tracie V. 1988. "Judicial Policy-Making and Implementation: An Empirical Investigation." *Western Political Quarterly* 41: 509-527.

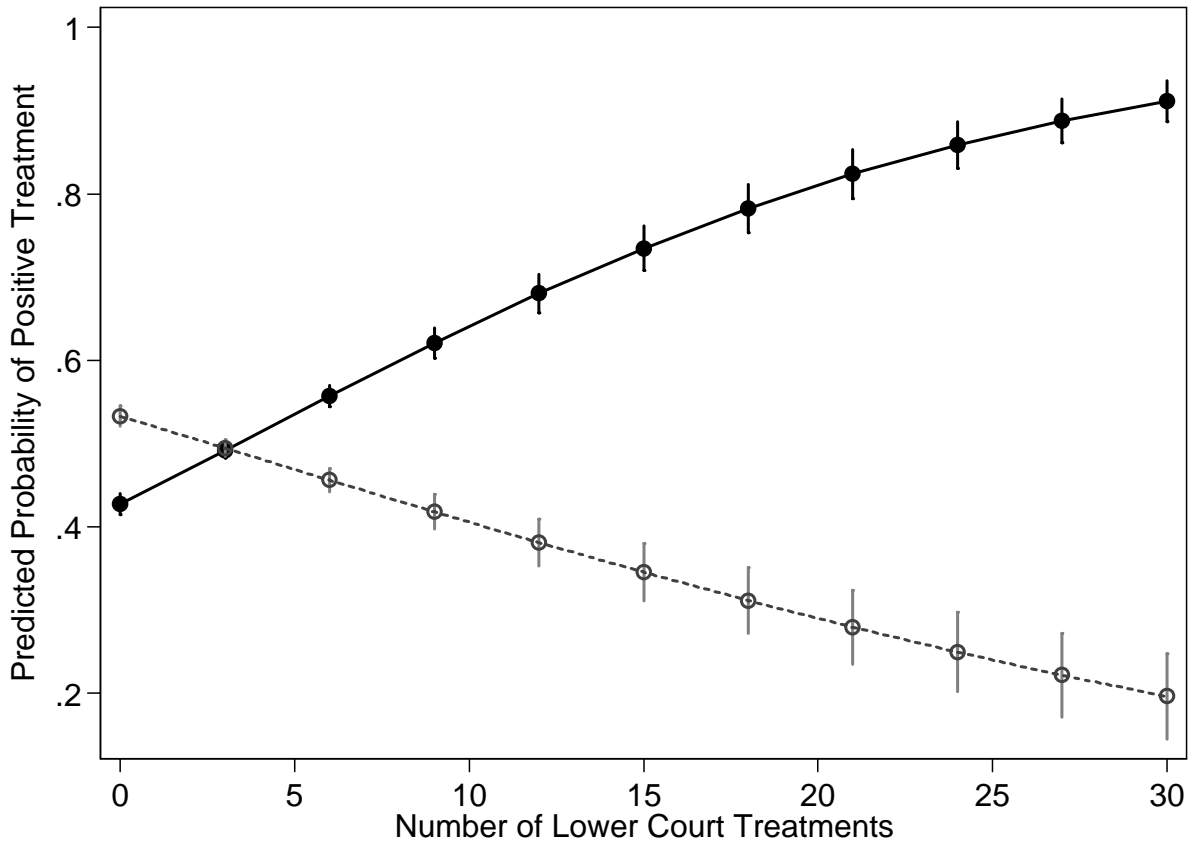
- Rogers, John. 1995. "Lower Court Application of the 'Overruling Law' of Higher Courts." *Legal Theory* 1: 179-204.
- Shavell, Steven. 1995. "The Appeals Process as a Means of Error Correction." *Journal of Legal Studies* 23: 379-426.
- Songer, Donald R., Jeffrey A. Segal, and Charles Cameron. 1994. "The Hierarchy of Justice: Testing a Principal-Agent Theory of Supreme Court-Circuit Court Interactions." *Journal of Politics* 38: 673-696.
- Songer, Donald R. and Reginald S. Sheehan. 1990. "Supreme Court Impact on Compliance and Outcomes: *Miranda* and *New York Times* in the United States Courts of Appeals." *Western Political Quarterly* 43:297-319.
- Spaeth, Harold J. 2006. "The Original United States Supreme Court Judicial Database, 1953-2003 Terms." Available at: <http://www.as.uky.edu/polisci/ulmerproject/sctdata.html>
- Spaeth, Harold J., and Jeffrey A. Segal. 1999. *Majority Rule or Minority Will: Adherence to Precedent on the U.S. Supreme Court*. New York: Cambridge University Press.
- Spitzer, Matt and Eric Talley. 2000. "Judicial Auditing," *Journal of Legal Studies* 29: 649-683.
- Spriggs, James F. II, and Thomas G. Hansford. 2000. "Measuring Legal Change." *Political Research Quarterly* 53: 327-341.
- Spriggs, James F. II, and Thomas G. Hansford. 2001. "Explaining the Overruling of U.S. Supreme Court Precedent." *Journal of Politics* 63: 1091-1111.
- Tarr, G. Alan. 1977. *Judicial Impact and State Supreme Courts*. Lexington, MA: Lexington.



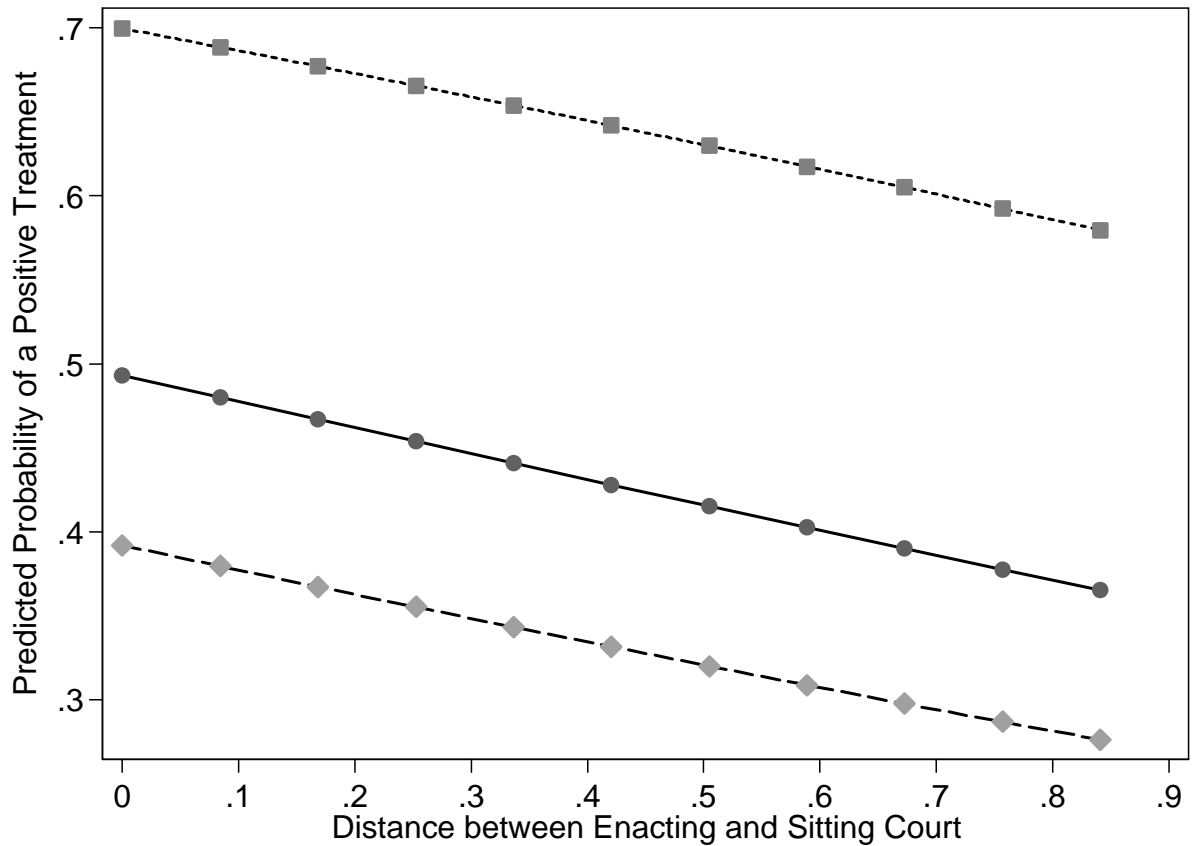
**Figure 1: Interpreting Precedent.** Four external relationships directly affect the Contemporary Lower Court's interpretation of precedent. the decisions of the Enacting Supreme Court (initial vertical stare decisis), the interpretations of that decision by Subsequent Supreme Courts (secondary vertical stare decisis), the interpretations of those decisions by Earlier Lower Courts (lower court horizontal stare decisis); and the Contemporary Lower Court's reading of the preferences of the Contemporary Supreme Court (hierarchical conformity).



**Figure 2: Enacting and contemporaneous judicial regimes in the Judicial Common Space for the Fourth Circuit in *U.S. v. Dickerson* (1999).** The Enacting Court is the median of the majority coalition in *Miranda* (1966). The Contemporaneous Court is the median of the Supreme Court at the time the circuit decided *Dickerson*.



**Figure 3: Predicted probability of a positive treatment in a circuit given treatments within the circuit prior to the panel decision.** The solid line represents the predicted probability of a positive treatment as the number of earlier positive circuit treatments increases, holding all other values at the mean. The dashed line represents the probability of a positive treatment given an increasing number of earlier negative treatments in the circuit, holding all other variables at the mean. The vertical lines represent 95% confidence intervals. To generate predicted probabilities, we used the spost package for Stata 9.0.



**Figure 4: Predicted probability of a positive treatment in the circuits over the range of the distance between the Enacting and Contemporary Supreme Courts.** The top dashed line with squares is the predicted probability with the number of positive circuit treatments at one standard deviation above the mean (10 positive treatments), negative treatments at 0, and the rest of the variables at the mean. The middle solid line sets all values at the mean and varies the Enacting-Contemporaneous distances. The bottom dashed line with diamonds sets negative treatments at one standard deviation above its mean (8 negative treatments), positive treatments at 0, and the rest at the mean.

<b>Concept</b>	<b>Measurement</b>
Signals of social change	Measured by <code>NYT Topic Coverage</code> , the number of New York Times stories in the topical area of the case, during the preceding three years, from Proquest.
Unclear initial signal from the Enacting Supreme Court	Measured by three variables, <code>Dissenting Opinions</code> and <code>Concurring Opinions</code> , the number of dissenting and concurring opinions in the original Supreme Court case, as coded by Spaeth (2006), and <code>Case Complexity</code> , the number of legal provisions plus the number of legal issues present in the original Supreme Court case, as coded by Spaeth (2006).
Earlier signals (treatments) by the Supreme Court	Number of treatments of the precedent by the Supreme Court before the lower court case. Coded as two variables, <code>Negative SC Treatments</code> and <code>Positive SC Treatments</code> .
Earlier signals (treatments) by the judges in the circuit	Similar to the earlier Supreme Court treatments, but the number of earlier positive and negative treatments of the precedent in the Circuit. Coded as two variables, <code>Negative LC Treatments</code> and <code>Positive LC Treatments</code> .
Age of Supreme Court precedent	<code>Age of SC Precedent</code> , age in years of the Supreme Court precedent
High court distances	<code>Contemporary-Enacting Distance</code> , the distance in the JCS between the Contemporary Supreme Court and the Enacting Supreme Court, with the Enacting Court measured by the location of the median member of the majority.
Contemporary Lower Court panel – Enacting Supreme Court distance	<code>Panel-Enacting Distance</code> , similar to <code>Contemporary-Enacting Distance</code> but measured from the median member of the contemporary panel to the Enacting Supreme Court
Contemporary Lower Court panel – Contemporary Supreme Court distance	<code>Panel-Contemporary Distance</code> , similar to <code>Panel-Enacting Distance</code> , but measured from the median member of the contemporary panel to the median member of the Contemporary Supreme Court.

**Table 1: Definitions of Variables.**

	Team-Theoretic Approach	Agency-Theoretic Approach	Conjoined Model
NYT Topic Coverage	-.086 (.047)	---	-.066 (.047)
Case Complexity	-.029 (.027)	---	-.006 (.028)
Concurring Opinions	.031 (.022)	---	.064* (.032)
Dissenting Opinions	-.018 (.029)	---	-.001 (.031)
Positive SC treatments	.016* (.007)	---	.013 (.007)
Negative SC treatments	-.028* (.011)	---	-.006 (.011)
Positive LC treatments	.083* (.005)	---	.087* (.006)
Negative LC treatments	-.052* (.006)	---	-.051* (.006)
Contemporary SC-Enacting SC Distance	---	-.669* (.130)	-.625* (.140)
Panel-Enacting SC Distance	---	-.042 (.097)	-.067 (.099)
Panel-Contemporary SC Distance	---	-.022 (.123)	.022 (.125)
Age of SC precedent	---	.002 (.002)	-.011* (.003)
Tau1	-.875 (.070)	-1.041 (.049)	-1.023 (.080)
Tau2	.055 (.069)	-.138 (.048)	-.089 (.080)
N	10210	10206	10173
AIC	2.029	2.067	2.025
BIC	-73463.168	-73070.921	-73167.140

**Table 2. Regression results.** The dependent variable is whether the lower court: complies with Supreme Court precedent, treats the precedent neutrally, or deviates from the precedent. Ordered logit estimates, robust standard errors in parenthesis. \*  $p < .05$ . LC= lower court; SC=Supreme Court. For more information about the variables, see Table 1.

	Mean	Standard Deviation	Minimum	Maximum
NYT Topic Coverage	.277	.390	0	2.893
Case Complexity	2.264	.736	1	11
Concurring Opinions	.394	.633	0	3
Dissenting Opinions	.873	.666	0	4
Positive SC treatments	2.324	3.594	0	27
Negative SC treatments	1.348	2.770	0	16
Positive LC treatments	3.439	6.333	0	50
Negative LC treatments	2.419	5.509	0	44
Contemporary-Enacting Distance	.225	.161	0	.841
Panel-Enacting Distance	.305	.221	0	1.246
Panel-Contemporary Distance	.244	.163	0	.861
Age of SC precedent	10.838	8.300	0	45

**Table A-1. Descriptive statistics.**