

C. Project Description

1 Introduction

In its landmark decision in *Regents of the University of California v. Bakke* (1978), the U.S. Supreme Court ruled that, under certain circumstances, universities may take race and other factors into account when they make admissions decisions. Since that time the justices have never overruled *Bakke* and, as such, it should stand as the law of the land, a decision that all courts must follow. But it does not. In *Hopwood v. Texas* (1996, 963), the U.S. Court of Appeals for the Fifth Circuit 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, to combat the perceived effects of a hostile environment at the law school, to alleviate the law school’s poor reputation in the minority community, or to eliminate any present effects of past discrimination by actors other than the law school.” With these words, the judges of the 5th Circuit, at least according to their colleagues in dissent, took the dramatic step of defying precedent established by the top of the judicial hierarchy, the Supreme Court of the United States.

Scholars and journalists alike have spilt much ink over *Hopwood*, as well as over the Fourth Circuit’s recent decision, in *United States v. Dickerson* (1999), holding that states under its supervision need not follow *Miranda v. Arizona* (1966). And, yet, these decisions are merely the most striking instances of a more general phenomenon, lower court deviation from precedents set by a higher court— a phenomenon that can take far subtler forms (*e.g.*, distinguishing, limiting, or avoiding precedents). As one observer noted 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, not to mention the scholarship we touch on below, raises a question that, depending on one’s perspective, may be posed two different ways: *Why do lower courts defy higher courts*, or, given the minute percentage of lower court cases that are heard and reversed (these days, under 1%), *why do lower courts comply with higher courts?*

Scholarly attempts to address these questions take several forms.¹ One is a line of inquiry aimed at identifying 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 1999). Another has focused on socialization and conformity to legal culture as the critical causal mechanism. Robert Cover’s

¹This literature stresses, as we do, explanations for why lower courts defy higher courts. But there also are many, albeit “not comprehensive” (Songer 1991, 43), studies seeking to describe or assess the extent of defiant (or compliant) behavior among lower federal court judges (see, *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, 1987; Songer and Reid, 1989; Songer and Sheehan, 1990; Vines 1969). These studies have reached mixed conclusions or, as Songer (1991, 43) notes, “the overall extent and frequency of noncompliance and evasion is unclear.”

(1975) noted study of the enforcement of the Fugitive Slave Act by abolitionist judges, for example, emphasizes the moral quandary posed by the judges' twin commitments to abolition and rule of law (see also Howard 1981).

More recently, scholarly efforts, conducted both by social scientists and legal scholars, 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 has swept the social sciences in the last decade. But scholars 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 four distinct and, to some extent, competing models. The first, recently suggested by prominent legal scholars, is a model based on the theory of teams. The remaining three are “principal-agent models” that assume heterogeneous policy preferences among judges and examine the incentives and opportunities created by various institutional features of the modern judicial hierarchy. The first agency account emphasizes litigant policing, the second stresses strategic auditing, and the third highlights implicit “tournaments” among lower courts.

The goal of this research is to advance the new judicial institutionalism, empirically and theoretically, by invoking the various models to answer our primary research question— why do lower courts defy (comply with) higher courts? In what follows we lay our plans for accomplishing this end. Sections 2 and 3—the first devoted to the theory of teams and the second to the principal-agent approach—provide overviews of the different models and of the various hypotheses they generate.² Section 4 turns to matters of research design, measurement, and data, explaining the procedures we will use to test the hypotheses. We conclude with a discussion of the importance and implications of our proposed endeavor.

2 Team Theory: A Legal Approach

Within the legal academy, analytically-minded scholars who theorize about the logic of the judicial hierarchy have converged on an approach that draws on the theory of teams. A branch of economics, this theory is concerned with the efficient organization of individuals who share a common goal but control different decision variables and base their decisions on (possibly) different information (Radner 1987). As applied to the judiciary, the essential idea is to assume that judges have no value conflicts whatsoever; that is, they all share the same utility function. Hence, the object of judging is to decide as many cases “correctly” as possible. The point of the analysis is to see whether notable features of the judiciary (*e.g.*, a hierarchy, *stare decisis*, and collegial courts) emerge endogenously in such a setting, as a logical consequence of the effort by the “judicial team” to organize itself effectively (*e.g.*, Caminker 1994, 1995; Daughy and Reinganum 1998; Dorf 1995; Kornhauser 1989, 1995, 1997; Rogers 1995; Shavell 1995). The flavor of the analysis is well-conveyed by Kornhauser’s summary of his 1995 study:

I have argued that one may derive several features of the federal judicial system in the United States from a relatively simple model. When one regards the judiciary as a resource-constrained team that seeks to maximize the expected number of correct decisions, one would expect, under plausible assumptions about the “technology of legal deliberation,” a system with distinct trial and appellate courts. Moreover, one

²As part of the proposed research we will elaborate the empirically testable implications of each model more rigorously and completely than we do in the text and than scholars have done to date. But, for present purposes, the models are sufficiently well developed to generate a wealth of hypotheses.

would expect a pattern of precedent similar to the one we observe: strict vertical precedent, no horizontal precedent at the trial level, and strict horizontal precedent at the appellate level. In this model, then, even if the judiciary has no substantive reason to promote certainty or uniformity, these qualities will emerge as collateral consequences of the organizational aim of getting decisions right.

Some of the results in team models are driven by resource scarcity. Vertical *stare decisis*, for example, is seen partly as a device for economizing on judicial time and effort, driving litigants to settle early and avoid pointless appeals. Other results are driven by information asymmetries. For example, when rule-making in hard cases is delegated to a resource-rich group of specialists (*e.g.*, the U.S. Supreme Court), vertical *stare decisis* becomes an efficient way for less well-endowed and hence less knowledgeable judges in lower courts to make decisions cheaply and, on average, correctly. The best they can do is to follow the lead of thoughtful specialists who supply the team with a public good in the form of a legal rule or doctrine.

Within this class of models, the phenomenon of “defiance” cannot arise, but a rather similar phenomenon, *underruling*, does. Underruling (in contrast with overruling) occurs when a hierarchically-subordinate court rejects a rule supplied by a hierarchically-superior court (Caminker 1994; Kornhauser 1995). How can this occur in a team model? In two ways. First, the case before the lower court may appear to be covered by the rule, but in fact is sufficiently different that a new and different rule is required. Second (and more interesting), circumstances may have changed so much since the original rule was announced that it is now antiquated and needs to be replaced by a new rule, or, additional evidence acquired after the articulation of the initial rule shows that the rule was made in error and now requires revisions. The second rationale might be seen as applicable to the *Bakke* “underruling.”

2.1 A TEAM MODEL: STARE DECISIS AS AN INFORMATION CASCADE. Team models of the judicial hierarchy are so new that formalization is rare and systematic empirical evaluation, even rarer.³ We hope to remedy this situation by proposing a model that is in line with the existing conceptualization but capable of generating testable propositions. This model assumes *stare decisis* emerges endogenously in the judicial hierarchy as a result of information asymmetries and resource constraints. The key mechanism is an “information cascade,” which occurs when it is optimal for an individual, having observed the actions of those ahead of her, to follow the behavior of the preceding individual without regard to her own information. Since publication of the seminal paper on cascades (Bikhchandani et al. 1992), models of this kind have been used to study many social phenomena, from democratic transitions in Eastern Europe to trends in fashion. And while they have yet to be invoked to study *stare decisis*, the application seems quite natural. The model specifies the conditions in which *stare decisis*—following the previous action of the Supreme Court—emerges as a result of an information

³Even so, several well-known studies by social scientists and legal scholars are suggestive, if not definitive. Research by Songer and Reid (1989) shows that decisions produced by the Courts of Appeals and District Courts in economic cases became more liberal after the Supreme Court’s “switch in time that saved nine”—and not as a result of the influx of Roosevelt appointees to the federal bench between 1933 and 1936. Songer and Sheehan (1990) provide convincing evidence that the appointment of Democratic judges to the lower federal courts had no appreciable effect on their rulings in most areas of the law; rather, they tended to follow the Supreme Court’s decisional patterns. Finally, Eisenberg and Johnson (1991) and Ashenfelter, Eisenberg, and Schwab (1995) find that at the district and circuit court levels putatively liberal and putatively conservative judges decide cases the same way. This held even in ideologically-charged areas such as civil liberties and race discrimination.

cascade. It also indicates how a cascade can end, such that lower courts, under certain circumstances, will deviate from precedent.

We sketch a simple version of the model. Assume n lower courts. Let there be two possible states of the world, $\Omega = \{1, 2\}$. Assume there are two possible legal rules or doctrines, $D = \{1, 2\}$, one of which must be employed by a judge hearing a case of the given variety. For each judge, utility $U_i = 1$ if $\omega = d$ (that is, the doctrine matches the state of the world), but 0 otherwise. (Recall that all judges have the same utility function in a team model. The subscripts index the judges.) In other words, if the true state of the world is 1, each member of the judicial team wishes to implement $d = 1$. But if the state of the world is 2, each will prefer $d = 2$. The prior probabilities of each state, $q(1)$ and $q(2) = 1 - q(1)$, are given positive numbers and common knowledge. Judges hear cases sequentially. Upon hearing a case, a judge receives a private, noisy signal about the state of the world, $s \in \mathcal{S} = \{1, 2\}$. All the signals are conditionally independent. Given a state of the

world ω , with probability $p_i > \frac{1}{2}$ $s = \omega$ and with probability $1 - p_i$ $s \neq \omega$. In other words, if

$\omega = 1$ then with probability $p_i > \frac{1}{2}$ $s = 1$; but if $\omega = 2$ then $s = 2$ with $p_i > \frac{1}{2}$. Thus, having heard

a signal, the expected utility of adopting Rule 1 is $\gamma \cdot 1 + (1 - \gamma) \cdot 0 = \gamma$, where γ is the Bayesian posterior belief that $\omega = 1$. Judges cannot see the signals observed by earlier judges, for to do so would require knowing all the material that emerged in the earlier proceeding. But the instant judge can see the rule used by the earlier judges, *and infer what their signals must have been if they ruled the way they did*. The Bayesian posterior belief of the instant judge then reflects not only her private signal, but what she can infer about the private signals of judges who earlier had heard a similar case. The critical feature of this model is that, quite rapidly, the information received by the prior judges “overwhelms” the value of the private signal heard by the instant judge. At that point, the instant judge will rule in a given way (reflecting the earlier signals) regardless of her private signal—an “information cascade” begins. And, from this point forward, no additional information will be revealed by the judges’ actions, since subsequent judges no longer can learn anything from looking at the actions of judges who ruled after the cascade began. Hence, they will all follow precedent. Bikhchandani et al. also show that more precise signals lead more readily to a cascade. They further demonstrate that publicly observable information—a common signal—can end a cascade.

2.2 HYPOTHESES. This sketch model deserves deeper analytic treatment,⁴ and we fully intend to undertake that task. Nonetheless, the results in Bikhchandani et al. (1992) allow us to generate

⁴To create a genuine model of stare decisis, one must incorporate important and distinctive elements of judicial organization. Four seem critical. The first is collegial courts at the appellate level. How do collegial courts aggregate the signals received by the individual judges on the court, how does this affect the precision of the (pooled) signals heard by the collegial court, and what effect does the information aggregation have on subsequent courts? Adding this feature will require incorporating the new results on the Condorcet jury theorem into a model of cascades (Austen-Smith and Banks 1996). A second is the pyramid structure of the judiciary—with single-judge trial courts, three-judge appeals courts, and the nine-justice Supreme Court. How does the sequential process of hearing by progressively higher precision courts affect cascades? Third, the Supreme Court can intervene in the sequential process of hearing cases when it wants to, by granting certiorari to a case. In other words, the sequence of hearing cases is not exogenous but partly determined by the Supreme Court itself. When will the Supreme Court intervene, and what effect does its intervention have in initiating a cascade? Fourth, how do the courts respond to social change, that is, new evidence that the state of the world has changed? This is likely to be critical in ending a cascade. By investigating these topics, we develop the sketch model into a genuine, “legal model” of stare decisis.

several clear predictions—all of which center on the notion of signals. Specifically, on this account, *lower courts will be more likely to deviate from (underrule) Supreme Court precedent when:*

H1.1: There are complex issues in the case such that the signals it generates may be imprecise.

H1.2: There are signals, observed by most judges, indicating that important social change has occurred in the area governed by the precedent (so the precedent may need revision).

H1.3: There are signals, observed by most judges, indicating that the precedent was mistaken or in error (so the precedent may need revision even absent social change).

H1.4: There are conflicting signals from the Supreme Court over the doctrine in the governing case.

3 Agency Theory: Strategic Models

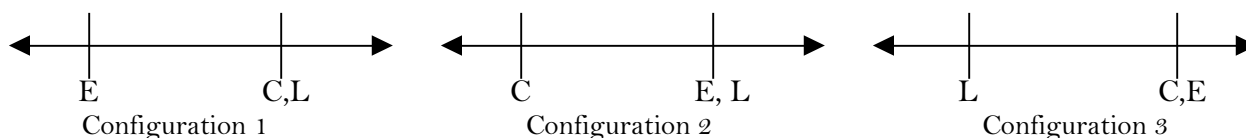
While the legal approach encapsulated in the team model holds sway with many law school professors, it is agency theory that has attracted the attention of social scientists (and, in particular, political scientists) (*e.g.*, Moe 1984; Songer, Segal, and Cameron 1994). The fundamental assumption of agency theory is that value conflicts are pervasive in organizations. The creation and application of rules—generic problems in bureaucracies—are occasions for political struggle. The outcomes of these struggles reflect the power of the bureaucratic contestants. In turn, the details of organizational design and operating procedures (“the rules of the game”) determine power. Theoretical investigations proceed by introducing agents (with value conflicts) into a game-theoretic rendering of an organizational procedure. Within the context of the model, the researchers deduce the consequences for organizational performance.

When applied to the federal judiciary, the agency perspective refuses to privilege one set of rulings as “correct” and another as “incorrect.” Rather, legal doctrine within the federal judiciary emerges from an unrelenting struggle between the few—the hierarchical superiors—and the many—the hierarchical subordinates. In turn, the power of the contestants is determined by the details of the organizational design and the rules of legal procedure. But which details are important in determining power? In broad terms, features that allow *lower* courts to take hidden actions, acquire private information, or otherwise evade sanctions will enhance their discretionary power; features that reveal hidden actions or information, or increase the sanctioning capacity of the *higher* courts will enhance their power.

With respect to sanctioning capacity, the U.S. federal judiciary is remarkable as a hierarchical organization to the extent that nominal superiors have almost no “high powered” incentives at their disposal. The justices of the U.S. Supreme Court cannot hire, fire, promote, demote, financially reward or penalize members of the Courts of Appeals or District Courts. Accordingly, agency analysis has focused on the one clear power the Supreme Court does have over lower courts; namely, the power to reverse their decisions. It is often posited that reversal itself is a sanction among federal judges. (On the empirical plausibility of this assumption, see Caminker 1994.)

Although agency analysis too remains in its infancy, it has generated three strategic models. The first two emphasize *litigant policing* by affected parties and *strategic auditing* by the Supreme Court in settings of “adverse selection,” while the third highlights implicit “*tournaments*” to avoid review among lower court judges in a setting of “moral hazard.” That is, the first two models emphasize the importance of hidden information and the latter, hidden actions.

While there are other differences among these models (reviewed below), all three draw a distinction between the preferences of the *enacting* Court (embodied in the existing legal doctrine), the preferences of the *current* or *contemporary* Supreme Court (the preferences of the Court at the time the lower court is considering the case), and the preferences of the *lower court* hearing the case. Accordingly, in all principal-agent models of the judiciary, three configurations of the players in ideological space (depicted below) are of particular interest. In Configuration 1, the lower court undertakes doctrinal deviation if it pursues its own preferences. But, by so doing, it engages in what we might call “hierarchical conformity,” as the lower court’s action actually conforms to the preferences of the current Supreme Court. In Configuration 2, the lower court—if it pursues its own preferences—engages in doctrinal conformity but hierarchical “deviation” to the extent that it will reach a decision distant from the preferences of the contemporary Supreme Court. In Configuration 3, a lower court that pursues its own preferences engages in both doctrinal and hierarchical deviation.



E=Enacting Court; C=Contemporary (or Current) Court; L=Lower Court.

Based on these configurations, the following hypotheses are common to all three principal-agent models. *Lower courts will be more likely to engage in doctrinal deviation when:*

H 2.1: The spatial location of preferences conforms to Configuration One rather than to Configurations Two or Three. That is, a lower court L is most likely to deviate from E when it is allied with the contemporary Court against the enacting Court.

H 2.2. The distance from E and the C increases, controlling for configuration.⁵ Thus, even when the lower court prefers the enacting Court’s doctrine, increasing the distance from the enacting Court to the contemporary Court leads reversal-sensitive judges to shy away from the enacting Court’s doctrine.

H 2.3. The distance from E and L increases, controlling for configuration (see note 5).

H 2.4. The distance from C and L increases, controlling for configuration (see note 5).

Beyond these, the models (in the main) offer distinct hypotheses, as our discussion of each (see Sections 3.1, 3.2, and 3.3 below) clarifies.

⁵This distance is large in Configurations One and Two, and small in Configuration Three. Thus, if one does not control for configuration, the hypothesis will not hold.

3.1 LITIGANT POLICING AND JUDICIAL FIRE ALARMS. The notion of “fire alarms” has become standard in the literature on congressional oversight of the bureaucracy (McCubbins and Schwartz 1984; Lupia and McCubbins 1994). The basic idea is the following. Congress delegates power to an agency whose interests differ somewhat from its own. Because monitoring the agency’s behavior is costly, Congress allocates its efforts elsewhere. Hence, the agency can take actions unseen by Congress (the model is thus one of “moral hazard”). But a directly affected party, an interest group, observes the actions perfectly. The interest group can notify Congress (sound the “fire alarm”) if the agency takes actions that Congress would oppose. If notified, Congress can intervene, reverse the action, and sanction the agency. In essence, Congress delegates implementation to the agency but supervision to an interest group. This results, allegedly, in a high degree of congressional control of the agency’s actions even in the absence of direct monitoring or supervision by Congress.

We extend these ideas to the federal judiciary but shift from hidden actions to hidden information. Consider again the judicial signaling model introduced in Section 2; however, expand the players so that they now consist of two litigants, a trial judge, three Court of Appeals judges, and nine Supreme Court justices and allow the possibility that some lower judges do not share the doctrinal preferences of the Supreme Court. In this revised game, each player involved in each legal proceeding (trial, first appeal, and second appeal) receives independent, private signals of the state of the world. Critically, however, an appeal by the losing litigant is voluntary and costly. Consequently, a losing litigant will appeal only if she has a reasonable expectation of ultimately prevailing via that appeal. Otherwise, she will avoid the additional expense of futile litigation. A litigant will have the expectation that she will prevail despite an adverse judgment under two circumstances. First, she can believe in the prospect of ultimate victory if her private information about the state of the world is so convincing to her that she sees the lower court’s adverse judgment as a “mistake” due to its receipt of a very unlikely and unrepresentative signal. In some cases, of course, a litigant who holds this belief will be mistaken: it was her signals, rather than those of others, that were misleading. Second, she can believe in the prospect of success if she sees the court’s adverse judgment as a deliberate deviation from the doctrine that will ultimately prevail should she pursue the appeal. In these circumstances, an appeal is a kind of “fire alarm” to a higher court with a discretionary docket, signaling a doctrinal deviation by a lower court. Importantly, though, this fire alarm cannot induce perfect compliance by lower courts because mistaken appeals from litigants with unrepresentative private signals introduce noise in the signal. This noise, in turn, opens the door for deliberate, strategic deviation by lower courts with preferences different from those of the Supreme Court.

3.1.1 Hypotheses. This is but a brief sketch of a model of imperfect policing of lower courts through litigant appeals. Though we intend to explore its empirical implications in careful detail, the underlying logic of the model points clearly to the following hypotheses. *Lower courts will be more likely to deviate from Supreme Court precedent when:*

H3.1: The litigant has few legal resources, leading the lower court to believe that an appeal to the Supreme Court is less likely in the face of a deviant decision.

H3.2: The issue is of low salience to the Supreme Court, leading the lower court to believe that the High Court is less likely to grant certiorari even in the face of a deviant decision.

H3.3: The trial court judge and the appellate court panel share similar ideological dispositions, enabling them implicitly to collude against a litigant (that is, to keep information from the Supreme Court).

H3.4: All members of the appellate court panel share similar ideologies, so that dissents—which may provide the Supreme Court with information—are less likely.⁶

3.2 STRATEGIC AUDITING. Theorists of principal-agency relations place great weight on the importance of strategic auditing by hierarchical superiors (Andeoni, Erard, and Feinstein 1998; Banks 1989). Cameron, Segal, and Songer (2000) extend the logic of strategic auditing to judicial politics in a game-theoretic model of the Supreme Court's certiorari decisions. In this model, a lower court and a higher court differ over the proper treatment of cases. A case presents a fact pattern, a location in an "attitudinal" space. A legal doctrine indicates how to group cases in the fact space by appropriate legal treatment (*e.g.*, all cases below a cut point should be treated this way; all above the cut point another way.) The lower court, having heard the case, knows its precise location in the attitudinal space. The higher court can see the action taken by the lower court (*e.g.*, in a search and seizure case, it can see whether the evidence was admitted or excluded). But the high court cannot know whether the lower court followed the high court's preferred doctrine, because it cannot know the precise location of the case without hearing the case. At best, easily observed facts indicate only the approximate location of the case. The high court can pay an "audit cost" by granting certiorari and hearing the case, thus learning its exact location of the case. If the lower court has cheated, the high court can reverse, thus enforcing its preferences. But, because hearing cases is costly, the high court would rather not grant certiorari if it can avoid it. In the model, the high court uses the observed ideology of the lower court, the observable facts in the case, and the action of the lower court as cues in picking cases to review.

3.2.1 Hypotheses. Cameron and his colleagues (2000) derive and test hypotheses about Supreme Court decisions only, not about lower court behavior. But, because their model treats lower and higher court behavior as strategically interactive, we can develop it to yield such predictions (see also H2.1-H2.4). Specifically (and as a starting place for empirical analysis), we posit that *lower courts will be more likely to deviate from Supreme Court precedent when:*

H4.1: The number of amicus briefs in the case is low (thereby increasing the costs of review for the high court, making an audit less attractive).

H4.2: The ideological diversity on the lower court panel is low (thereby decreasing the likelihood of dissent and, thus, of an audit).

H4.3: The issues in the case are complex (so the Supreme Court will have difficulty using reliable cues in selecting deviant cases for review).

3.3 TOURNAMENTS. Theorists of principal-agency relations also have emphasized the advantages of placing subordinates into competitive relationships with one another (Lazear 1995, 1997; Lazear and Rosen 1981,). This idea is extended to the judiciary in Cameron (1993)

⁶See also Cross and Tiller 1998.

and McNollgast (1994), which explore how the structure of the judiciary puts lower courts into an implicit tournament with one another, to avoid review and reversal.

The basic tournament model takes the following form (see Kornhauser 1997). There is one high court and n lower courts. Each court has spatial preferences over a one-dimensional policy space. There is complete information so that each lower court knows the high court's ideal point and the high court knows the location in policy space of the decision of each lower court. Each lower court seeks to maximize its utility, which depends on the location of the final decision in the case it decides. In addition, if reviewed and reversed, a lower court suffers an ε loss of utility. The higher court wants to maximize its utility, which is a function of all the decisions in *all* the cases. The game has two stages. In stage one, each lower court issues a decision. In stage two, the higher court selects at most one case for review (this reflects, in stark form, the severe resource limitations of the high court). Cameron (1993) shows that there is a sub-game perfect equilibrium to this game in which each lower court decides its case by announcing the ideal policy point of the high court as its decision and the high court reviews no case, because it has already achieved its optimum outcome. In other words, the high court establishes perfect control over the lower courts, regardless of the number of lower courts, and despite the fact that the high court can review at most one of them.

The intuition behind this theoretical result is well captured by the Japanese proverb, "The nail that sticks out will be driven in." The high court will review the most deviant lower court, so no court wishes to be the most deviant. Thus, the lower courts are pitted against one another, implicitly competing to avoid review (hence, a tournament, in which the loser is sanctioned). The competition drives the lower courts to perfect acceptance of a precedent they do not prefer. If the lower courts could coordinate with one another, they could in effect "riot" and overwhelm the high court's review capacity. But each lower court has an incentive to free ride on the efforts of the others in this coordination game. The free-riding kills the "riot" equilibrium.

3.3.1 Hypotheses. Again, this sketch version of the model is only the starting place for a more thorough theoretical investigation.⁷ But the logic of the tournament approach suggests that *lower courts will be more likely to deviate from precedent when:*

H5.1: The issues presented in the case are of low saliency to the Supreme Court (since the Supreme Court will direct most of its fire at less frequent but critical high saliency deviations).

H5.2: The number of lower court cases in an issue area increases (thereby increasing the lower court's belief that deviations worse than the one it commits exist).

H5.3: The judges on the lower court panel have been reversed by the Supreme Court relatively frequently in prior cases (that is, judges who are "repeat offenders" or "reversal-insensitive" and, thus, more likely to deviate in future cases.)

⁷We see "population uncertainty" as central to a more satisfactory model. There are some judges who care little about reversal. If a sufficient number of such judges hear cases in a given area of law, they can provide "cover" for other judges who wish to deviate but fear reversal. The number of reversal-insensitive judges hearing cases in a given area must be a matter of uncertainty for a non-reversal-insensitive judge hearing a case in a given area. Recent advances in game theory allow careful consideration of situations of this type (see Myerson, 1997a, 1997b, 1998).

H5.4: The population of repeat offenders—that is, the proportion of those judges who are frequently reversed by the Supreme Court—in the lower courts rises (since a larger pool of “reversal-insensitive” judges will encourage deviation by their non-reversal-sensitive counterparts).

H5.5: The ideological distance between the population of judges on the Courts of Appeals and the Supreme Court increases.

4 Research Design

With the four models now noted, we turn to our scheme for assessing the various hypotheses. In general terms, our plan is as follows: (1) generate a random sample of U.S. Supreme Court decisions; (2) track the responses of lower courts to these decisions (the dependent variable for all the hypotheses); (3) collect data necessary to animate the independent variables; and (4) implement statistical models, testing for the influence of theoretically-critical variables. The legal model and the fire-alarm models are nested so they can be distinguished via straightforward t-tests. We employ non-nested hypothesis tests (J-tests and non-nested F-tests) to compare the three agency models.

4.1 DEPENDENT VARIABLE. Our dependent variables, for all the hypotheses, are the responses of the U.S. Courts of Appeals to U.S. Supreme Court decisions. To operationalize these responses, we draw a 10% random sample (from Spaeth’s U.S. Supreme Court Judicial Data Base) of all orally-argued Supreme Court cases decided between the 1946 and 1990 terms (or 10% of $5,640=564$ cases).⁸

After drawing the sample, we follow the lead of many other judicial specialists (*e.g.*, Johnson 1979; Segal and Spaeth 1996; Hurwitz and Reddick 1996; Spriggs and Hansford 1998) and rely on *Shepard’s Citations* to determine how courts of appeals responded to each U.S. Supreme Court case. *Shepard’s* (which we access via LEXIS) identifies every decision by a U.S. Court of Appeals that “treated” the Supreme Court case. It also, in most instances, specifies the nature of circuit court’s treatment, that is, whether the court “overruled,” “questioned,” “limited,” “criticized,” “followed,” “distinguished,” “explained,” or “harmonized” the precedent.⁹ Thus, we simply code each circuit court case on the basis of *Shepard’s* categories, which themselves seem to run along an ordinal scale (according to how “negative” the citing case interpreted the cited case). At the very least, *Shepard’s* treats the categories as an ordinal scale and trains its coders to do so.

Given this ordinal characterization (from “overruled” through “harmonized”) we can model lower court reaction to Supreme Court decisions using the ordered probit model (McKelvey and Zavoina 1976). If the data demonstrate a lack of ordinality, the relationships can alternatively be tested by multinomial logit. This will make interpretation a bit more complicated, as there are a matrix of coefficients to interpret rather than a vector, but should pose no additional problems.

⁸We end with the 1990 term in order to provide a sufficient time horizon for lower court response.

⁹*Shepard’s* only allows each headnote to be assigned one of these codes. If multiple codes could be applied, *Shepard’s* adheres to the “strongest letter” rule to determine which code to use.

Since the average U.S. Supreme Court case generates about 14.3 circuit court responses that fall into one of *Shepard's* response categories, our overall data set will consist of approximately 8,065 court of appeals cases—a more-than-ample number for empirical analysis. Moreover, the measure resulting from this procedure—that is, the scale of circuit court responses to Supreme Court decisions—provides a highly suitable and satisfactory way for us to capture egregious defiance of precedent, subtler evasions or modifications, as well as compliance. In other words, it is a *valid* indicator of the concepts we seek to tap.

At the same time, we, like others who have invoked *Shepard's*, recognize that it does not supply a perfect rendering of judicial treatment of Supreme Court cases (see, *e.g.*, Johnson 1979; Spaeth and Segal 1999). But, also like our predecessors, we know of no other method for classifying responses to U.S. Supreme Court cases that is, at the very least, as *reproducible* and *efficient* as *Shepard's*—especially when dealing with large numbers of cases. Even more the point, recent work by Spriggs and Hansford (1998) demonstrates a surprising degree of *reliability* in the case treatment analyses conducted by *Shepard's*, thus countering the chief criticism of it. The following table, reproduced from the Spriggs and Hansford study, makes this clear:

Reliability of *Shepard's* Citations' Legal Treatment Codes, 1946-1995

<i>Treatment Code</i>	<i>Intercoder Agreement</i>
Overruled	99.5%
Questioned	96.4
Limited	99.0
Criticized	94.5
Followed	81.9
Distinguished	87.9
Explained	79.1
Harmonized	94.7

4.2 INDEPENDENT VARIABLES. The table below lists the key independent variables (in italics) for each of the working hypotheses discussed above, the operationalizations, and the sources of the data to be used. Since many of the operationalizations either are entirely transparent (*e.g.*, the number of dissents and special concurrences in the enacting case as an indicator of “conflicting signals from the Supreme Court”) or conventional in the field (*e.g.*, the Songer and Sheehan 1992 [see also Galanter 1974; Wheeler et al. 1987] categorization of litigant status), we describe below only the more novel or difficult ones.¹⁰

¹⁰Of course, we recognize that some of the models produce a handful of the same predictions. We handle this problem of behavioral equivalence in the standard ways; namely, we distinguish the models via their unique predictions and through non-nested hypothesis testing.

<u>Model</u>	
<u>Hypothesis (in abbreviated form)</u>	<u>Measurement of Independent Variable/Data Source</u>
<i>Lower courts will be more likely to deviate from precedent when:</i>	
<u>Team Theory</u>	
H1.1: There are <i>complex issues</i> in the case.	Issue complexity, following Maltzman, Spriggs and Wahlbeck (2000), is the number of issues and laws involved in the case, as coded in Spaeth. ¹¹
H1.2: There are <i>signals, observed by most judges, indicating that important social change</i> has occurred in the area governed by the precedent.	A publicly-observable signal of social change is the average number of articles on the issue in question (<i>e.g.</i> , crime, civil rights, the economy), as reported in <i>Reader's Guide</i> in the three years preceding the circuit court's decision compared to the average number of articles in that area over the time span of the study (Flemming et al. 1997; Rosenberg 1991).
H1.3: There are <i>signals, observed by most judges, indicating that the precedent was mistaken or in error.</i>	A publicly-observable signal of legal error is the number of law review articles (as identified in LEXIS and in the <i>Guide to Legal Periodicals</i>)—published in the three years preceding the circuit court's decision—that criticized the Supreme Court case under analysis.
H1.4: There are <i>conflicting signals</i> from the Supreme Court.	Conflicting signals are the number of dissents and special concurrences filed by Supreme Court justices in the case under analysis, as coded in Spaeth.
<u>All Principal-Agent Models</u>	
H 2.1: The <i>lower court is allied with the contemporary Court against the enacting Court.</i>	See Sections 4.2.1-4.2.3 of the text.
H 2.2. The <i>distance from the enacting Court and the current Court</i> increases, controlling for configuration.	See Section 4.2.1 of the text.
H 2.3. The <i>distance from the enacting Court and lower court</i> increases, controlling for configuration	See Section 4.2.2 of the text.
H 2.4. The <i>distance from the current Court and lower court</i> increases, controlling for configuration.	See Section 4.2.3 of the text.
<u>Litigant Policing</u>	
H3.1: The <i>litigant has few legal resources.</i>	Status of the litigant (identified from LEXIS) follows Songer and Sheehan's 1992 coding scheme (which, in turn, followed Galanter 1974; Wheeler et al. 1987).
H3.2: The <i>issue is of low salience</i> to the Supreme Court.	Issue salience is the number of cases—in the three years preceding the circuit court's decision—that the Supreme Court decided in the same issue area as the case under analysis, as coded in Spaeth.
H3.3: The <i>trial court judge and the appellate court panel share similar ideological dispositions.</i>	See Section 4.2.4 of the text.
H3.4: <i>All members of the appellate court panel share similar ideologies.</i>	See H3.3 (simply take the standard deviation of the ideology scores).
<u>Strategic Auditing</u>	
H4.1: The <i>number of amicus briefs</i> in the Court of Appeals case is low.	Amicus curiae brief counts come from Epstein and Caldeira's Amicus Curiae Data Base.
H4.2: <i>The ideological diversity</i> on the lower court (if a Court of Appeals) is low.	See H3.3

¹¹Here and throughout the table, Spaeth=Harold J. Spaeth's U.S. Supreme Court Judicial Data Base.

H4.3. <i>The issues in the case are complex.</i>	See H1.1
<u>Tournaments</u>	
H 5.1: <i>The issues presented in the case are of low saliency to the Supreme Court.</i>	See H3.2.
H 5.2: <i>The number of lower court cases in an issue area increases.</i>	The number of lower court cases is the weighted absolute number of cases in the broadly defined issue, as computed from the Songer Court of Appeals database.
H 5.3: <i>The judges on the lower court have been reversed by the Supreme Court relatively frequently in prior cases.</i>	The variable will be the absolute number of reversals in the past five years. From Spaeth, we will identify every case from the Court of Appeals that results in a reversal, and then, using LEXIS, locate the judges who were reversed.
H 5.4: <i>The overall population of “repeat offenders” rises.</i>	Same data as H5.3 (simply aggregated over all Courts of Appeals).
H 5.5: <i>The ideological distance between the population of judges on the Courts of Appeals and the Supreme Court increases.</i>	The ideological distance is the absolute difference between the percentage of Supreme Court justices who are Democrats and the percentage of Court of Appeals judges who are Democrat (as identified in Zuk et al. 1997.)

4.2.1 The distance between the contemporary Supreme Court’s preferences and the enacting Supreme Court’s doctrine (H2.2). For the ideological location of the original decision (that is, the decision of the enacting Court), we follow the lead of Hausegger and Baum (1999) and take the median issue-specific liberalism scores for all members of the majority coalition at the time the Court reached its decision in the case, which can be computed from the Spaeth data set (see Epstein and Mershon 1996 for an examination of this type of measure). Then, for the ideology of the Supreme Court (at the time of the lower court response), we use the median issue-specific liberalism score for all members on the Court, again using Spaeth. In both cases, “issue-specific” refers to the twelve value categories that Spaeth created. (Using the hundreds of narrowly-defined issue categories will result in many issues with far too few cases.) Because we measure both the original precedent and the current Court on the same scale, we can simply take the absolute value of the difference between them as our measure of ideological distance.

4.2.2 The distance between the enacting Court and the appellate court panel (H2.3). To assess this hypothesis, we require a measure of the enacting Court doctrine and of the appellate court preferences, and a method of scaling them. We already have a measure of the ideology of the enacting Court (see 4.2.1). We next need a measure of the ideology of individual judges on the Courts of Appeals who were in the majority in the case under analysis. We create this in accord with the methodology established by Giles, Hettinger, and Pepper (1998), who devised a weighted average of the preferences of the nominating President and the home state senators and successfully validated that measure against other, more common approaches (*e.g.*, the party of the appointing president). All in all, their measure accounts for about 10% of the variance in voting, which is well in line with others and which is not particularly low given that many factors other than ideology affect the voting behavior of Court of Appeals judges, including precedent (*e.g.*, Songer, Segal, and Cameron 1994; Van Winkle 1996) and ambition (*e.g.*, Segal and Spaeth 1993). In other words, even if one had a perfect measure of a lower appellate court judge’s ideology, it may be unable to explain more than 10% of the variance in voting.

In accord with the Giles et al. approach, we estimate, for each broad issue area in Songer’s (1996) U.S. Court of Appeals Data Base,¹² the following:

$$Y_{ij} = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_2 X_4 + \beta_5 X_3 X_5$$

Where: Y_{ij} = the percentage of liberal votes by judge j in issue i ; X_1 = the appointing President’s ideology; X_2 = the ideology of the senior home-state senator; X_3 = the ideology of the junior home-state senator; X_4 = whether the senior home-state senator is of the President’s party; and X_5 = whether the junior home-state senator is of the President’s party.

We use the predictions from these regressions to impute each judge’s ideology in each issue area.¹³

To scale the resulting measures, we can take advantage of the fact that many of the justices on the U.S. Supreme Court also have served on the U.S. Court of Appeals. Thus, for those justices who served on the Court of Appeals we run an issue-specific regression of the form

$\mathcal{Y}_{ij} = \alpha + \beta_i x_{ij}$ where y_{ij} represents the percentage of liberal votes by justice j in issue i , and x_{ij} represents the percentage of liberal votes by judge j in issue i .¹⁴

4.2.3 The distance between the current Court and the appellate court panel (H2.4). We measure the current Court via the procedures described in 4.2.; the appellate court as in 4.2.2, and scale them also as in 4.2.2.

4.2.4. The distance between the trial court judge and the appellate court panel (H3.3). We measure the appellate court panel as described in Section 4.2.2. We then propose to create imputed liberalism scores for district court judges, using the regression results obtained above. These predicted values will not be the predicted liberalism scores of district court judges on district court cases, but rather the predicted liberalism scores of district court judges on Court of Appeals cases. Fortunately, that is precisely what we require.¹⁵

4.3 TIME HORIZONS. Data collection and analysis will proceed as indicated on the Budget Justification page.

¹²The issue categories in the Songer data set are the same as those in Spaeth’s U.S. Supreme Court Judicial Data Base.

¹³We will use Songer’s data base to create Y_{ij} ; the Segal, Timpone and Howard (1999) ideology scores for Presidents serving since Franklin Roosevelt; and NOMINATE scores for Senators. Since NOMINATE scores overwhelmingly load on the first dimension, we rely those scores for all issue categories.

¹⁴Almost needless to say, the appropriate scaling parameters may change over time: 60% liberal Supreme Court decisions may correspond to 60% liberal circuit court decisions in the 1960s, but 65% in the 1970s. Within the constraints imposed by the need for parsimony, we can test for plausible patterns of change by adding time-dependent parameters to the model.

¹⁵If we had district court liberalism scores, we could follow the same process to impute the ideology of district court judges. But such information is not necessary, nor is it necessarily appropriate. A trial court who decides 60% of her civil rights cases liberally may be more, less, or just as liberal as a Court of Appeals judge who decides 60% of his cases liberally. We cannot assume a 1-to-1 scale of district court decisions to Court of Appeals decisions.

5 Conclusion

The American legal system rests on the firmly-entrenched assumption that lower court judges will, in fact, adhere to precedent established by higher courts (see, *e.g.*, Dworkin 1978; George 1998; Howard 1981; Llewelyn 1931). Without the uniformity imposed by this principle of vertical stare decisis, law in the United States would be a jumbled patch work, taking on different, probably inconsistent meanings in every circuit and state. The grand ideal of Western jurisprudence, “Equal Justice Under Law,” would have no foundation in practice and the rule of law—one law, consistently applied without fear or favor—would have little real meaning in American courts. The resulting chaos would be a disaster, politically, economically, and morally.

On this much—the stakes at play in the judicial hierarchy—virtually all judicial specialists agree. Where they depart is over these questions: How prevalent is lower court deviation from precedent? Under what circumstances does it occur? What causal mechanisms are at work when precedent holds and when it fails? And, among these causal mechanisms, which are the most important in explaining the successes and failures of precedent?

Our research, which explicitly focuses on the prevalence and causes of doctrinal deviation (conformity), attempts to advance the discussion surrounding each of these questions. We believe the empirical findings will deepen our factual knowledge and prove interesting to scholars of the judiciary. Perhaps even more important, by showing, both theoretically and empirically, how specific features of the design of the federal judiciary advance or retard vertical stare decisis, we will be able to uncover, at least partially, the institutional foundations of the rule of law in American courts. While these results too should be of interest to students of the American judiciary, we also believe they will hold value for all social scientists concerned with the effect of various structural incentives created by the design and operation of organizations.