

3-2000

Strategic Auditing in a Political Hierarchy: An Informational Model of the Supreme Court's Certiorari Decisions

Charles M. Cameron
Columbia University

Jeffrey A. Segal
SUNY Stony Brook

Donald Songer
University of South Carolina - Columbia, dsonger@sc.edu

Follow this and additional works at: https://scholarcommons.sc.edu/poli_facpub



Part of the [Political Science Commons](#)

Publication Info

Published in *American Political Science Review*, Volume 94, Issue 1, 2000, pages 101-116.

<http://journals.cambridge.org/action/displayJournal?jid=PSR>

© 2000 by Cambridge University Press

This Article is brought to you by the Political Science, Department of at Scholar Commons. It has been accepted for inclusion in Faculty Publications by an authorized administrator of Scholar Commons. For more information, please contact digres@mailbox.sc.edu.

Strategic Auditing in a Political Hierarchy: An Informational Model of the Supreme Court's Certiorari Decisions

CHARLES M. CAMERON *Columbia University*

JEFFREY A. SEGAL *SUNY at Stony Brook*

DONALD SONGER *University of South Carolina*

We examine how the Supreme Court uses signals and indices from lower courts to determine which cases to review. In our game theoretic model, a higher court cues from publicly observable case facts, the known preferences of a lower court, and its decision. The lower court attempts to enforce its own preferences, exploiting ambiguity in cases' fact patterns. In equilibrium, a conservative higher court declines to review conservative decisions from lower courts regardless of the facts of the case or the relative ideology of the judges. But a conservative higher court probabilistically reviews liberal decisions, with the "audit rate" tied to observable facts and the ideology of the lower court judge. We derive comparative static results and test them with a random sample of search-and-seizure cases appealed to the Burger Court between 1972 and 1986. The evidence broadly supports the model.

Hierarchical control of organizations is problematic throughout the realm of politics. Congress and presidents attempt to control agencies, upper levels of bureaucracies attempt to control lower levels, and higher courts strive to control lower courts. With incomplete information about their subordinates' decisions and knowledge, superiors in rule-based hierarchies often employ some form of auditing. In this article we study how the Supreme Court uses signals and indices from lower courts to pluck a relative handful of cases from a plethora of potential candidates for review. Our point of departure is the role of review in enforcing the doctrinal preferences of the Supreme Court within the judicial hierarchy.

We begin by presenting a game-theoretic model of the Court's certiorari process. The model goes much farther than earlier efforts to incorporate concepts commonly employed by judicial scholars. For example, we include the legal concepts of case facts, doctrine, and holding; formalize the attitudinal model; and include an explicit role for judicial culture. Most important, the model puts asymmetric information at its analytic center, following the hints in recent empirical

studies based on interviews with justices of the Supreme Court and their clerks (Perry 1991).

The model explains several well-established patterns in certiorari, such as the importance of amicus briefs by the Solicitor General. It also generates several entirely new and nonobvious predictions. We test the predictions on a random sample of 274 decisions in the area of search-and-seizure law heard in the U.S. Courts of Appeals and appealed to the Supreme Court during the Burger years. The data strongly display the patterns predicted by the model. Thus, in addition to contributing to the certiorari literature, we hope to advance the burgeoning literature on the strategic behavior of courts (e.g., Caldeira and Wright 1988; Epstein and Knight 1998; Hall 1992).

The article is organized as follows. We first discuss the role of information in the certiorari process and the place of certiorari decisions in the incentive system of the judicial hierarchy. We then present the formal model, followed by the empirical analysis. We next discuss the implications of the findings and offer conclusions in the final section. All proofs are gathered in the Appendix.

INFORMATION AND INCENTIVES IN THE CERTIORARI PROCESS

The Supreme Court's decision to hear a case—in legal parlance, to grant certiorari—is perhaps the best studied area of the high court's decision making. A full review of the literature is outside the scope of this article and is unnecessary here, but it is helpful to review several major themes relevant to these decisions.¹

¹ The Supreme Court uses the Rule of Four to grant certiorari, but a majority is necessary to prevail on the merits. Presumably, forward-thinking justices take the latter into account when voting to grant certiorari. Absent a careful analysis of the strategic properties of the Rule of Four (to the best of our knowledge, no formal model of this procedure has appeared in print), we treat the Court as a unitary actor with respect to the lower courts.

Charles M. Cameron is Associate Professor of Political Science, Columbia University, New York, NY 10027 (Cmc1@columbia.edu). Jeffrey A. Segal is Professor of Political Science, State University of New York at Stony Brook, Stony Brook, NY 11794 (jeffrey.segal@sunysb.edu). Donald R. Songer is Professor of Government and International Studies, University of South Carolina, Columbia, SC 29208 (dsonger@sc.edu).

We acknowledge the support of National Science Foundation Grant SES-9112755 and thank seminar participants at the 1995 annual meeting of the Midwest Political Science Association, the 1997 annual meeting of the Law and Society Association, Berkeley Business School, New York University Law School, the University of Iowa, Stanford Graduate School of Business, the University of Southern California Business School, Washington University, SUNY Binghamton, and George Washington University. We also thank Greg Caldeira, Alessandra Casella, Lee Epstein, John Ferejohn, Andrew Gelman, Lewis Kornhauser, Nolan McCarty, John Matsusaka, Richard Revez, David Z. Seide, Charles Shipan, Pablo Spiller, Matt Spitzer, Oliver Williamson, the anonymous referees, and especially Jean-Pierre Benoit for their extremely helpful comments. Cameron gratefully acknowledges additional support from the Hoover Institution, Stanford University.

Certiorari and Incentives in the Judicial Hierarchy

Certiorari and Interactive Incentives. The certiorari system creates incentives for upper and lower courts alike. One cannot get very far with a deductive theory of certiorari without thinking about those incentives, but a strange feature of the contemporary field of judicial politics is that the literatures on certiorari decisions and judicial compliance have developed largely independently. Most of the empirical studies of certiorari decisions ignore doctrinal deviations by lower courts (we note some exceptions below). In turn, empirical studies of the responsiveness of lower courts to Supreme Court precedent typically ignore review probabilities. In other words, there has been scant appreciation of the fact that the certiorari process creates an *interactive* incentive system between higher and lower courts, with the behavior of appellate courts affecting the behavior of the Supreme Court and vice versa.

A partial exception is the empirical work of Ulmer (1984), who used conflict with Supreme Court precedent as a predictor of certiorari, although he did not recognize that the probability of certiorari is apt to affect conflict with precedent. Ulmer noted that counsel for appellants seem to believe conflict with precedent galvanizes the Court into granting certiorari because such conflict is claimed in more than half of all petitions (six or more conflicts were claimed in more than 10% of the petitions). Attorneys may try to manipulate this cue by padding their claims about conflicts, so Ulmer tried to classify his sample into cases that do and do not contain "real" conflict with precedent. Using this distinction, he found that actual conflict is the single most important variable for understanding grants of certiorari; it explains a higher proportion of the variance than any of the other cues noted by Tanenhaus et al. (1963) and other scholars.

Ulmer's subjective coding of conflict is not amenable to replication, however. Moreover, his analysis puts dated precedents and more recent ones on a par although lower courts almost certainly defer more conscientiously to the preferences of the contemporaneous Court. Finally, Ulmer's dichotomization of real conflict leaves little room for nuance. Many scholars recognize that conflict involves subtle gradations of noncompliance (Feeney 1975). Nonetheless, one cannot read Ulmer's analysis without becoming aware of certiorari as a tool for policing the doctrinal decisions of lower courts.

Even if one accepts that certiorari creates an interactive incentive system, formalizing a model of it requires answers to two more questions. What does doctrinal compliance really mean? What sanctions and rewards are available to the Supreme Court if it detects noncompliance?

Doctrinal Compliance. The concept of a legal doctrine is central to the process of legal reasoning. What is a legal doctrine, and what does it mean to comply with a doctrine? We take as fundamental the conception implicit in classic descriptions of legal reasoning (e.g.,

Levi 1948): A legal doctrine indicates which fact situations are to be grouped together and treated similarly. In other words, it creates a set of equivalence classes in a fact or case space (Kornhauser 1992). When a lower court conforms to a legal doctrine, it accepts the equivalence classes defined by the higher court and acts in accordance with them. Nonconformity means that the lower court treats the case inappropriately for the class to which the case belongs. This concept of doctrine can be seen as an example of the broader notion of rule-governed behavior, a phenomenon general to bureaucracies. We formalize (part of) these conceptions of doctrine and doctrinal compliance below and employ them in the empirical analysis.

Sticks, Carrots, and Culture. What sanctions or rewards can a higher court bring to bear on lower courts that do not comply with its doctrine? In the American system of jurisprudence, the formal answer is "very few." Higher courts cannot promote, demote, or fire; they cannot cut salaries, give bonuses, or offer stock options. Thus, the Supreme Court possesses none of the motivational tools typically employed by hierarchical superiors. This is one of the most striking features of the federal judiciary considered as a hierarchical organization.

A focus on traditional motivational tools is too constricted a view of judges, however. In fact, two sources of control are available to higher courts. First, lower courts care about the disposition of cases: They wish to see justice done, at least as they conceive it. If a higher court reverses the decision of a lower court, the latter may well view the ultimate disposition of the case as much less attractive than if its judgment had stood. Hence, reversal itself can be a kind of sanction, at least for judges who care about the disposition of cases. Second, informal sanctions supplement the formal rules. "Judicial culture" famously includes a desire to avoid reversals. Frequent reversals bring the derision of colleagues and a decline in professional status. Higher courts are well aware of this sanctioning power. For example, Perry (1991, 267) notes that Supreme Court clerks "frequently talked about the need to 'slap the wrist' of a judge below." The importance of judicial culture should not be surprising. Federal judges belong to a very special and relatively close-knit society, and their informal culture is apt to affect their decisions.

Enforcing Doctrine Versus Creating Doctrine. Our model stresses the role of certiorari in enforcing doctrine, but this is a very partial view of the process. Perhaps equally important is the selection of cases as vehicles for creating *new* doctrine, for example, in novel fact situations. These often occur in the context of intercircuit conflict, when different circuits take different positions on new issues, which the Court must then answer. Indeed, the incremental, fact-soaked creation of new rules is one of the most interesting and distinctive elements of judicial politics. The justices' emphasis on finding "good" cases and "well-percolated" cases underscores this important part of the certiorari process (Perry 1991, chap. 8). In addition, the model ignores aggressive grants, that is, situations in

which the Supreme Court grants certiorari in order to affirm a lower court decision and make it binding on the entire nation.²

We defend the approach we take by noting that law creation is an important part of certiorari, but modeling it involves even more complex theoretical issues than the enforcement of doctrine, which is our focus. The development and testing of a rigorous theory of doctrinal enforcement is an important step toward a more general theory of the judicial hierarchy.

Structure of the Hierarchy. Our analysis simplifies the judicial structure into upper and lower courts. In reality, the federal judicial hierarchy has three levels, and courts at the first level play an important role in fact-finding. Our simplified structure enables us to explore the essential principal-agent dilemma. Extensions to this model could examine additional features of the judicial hierarchy, such as the special role of trial courts.

Information and Certiorari

From “Cue Theory” to Strategic Manipulation. Political scientists have long recognized the central importance of information in the Supreme Court’s certiorari process. Tanenhaus et al. (1963) argue that the Court must economize in its search for information, since each year it is confronted with a blizzard of cases and tens of thousands of pages of documents. In particular, they argue, the Court is apt to rely heavily on easily distinguishable but highly informative “cues,” such as splits in the lower court, to winnow out the potential cases of interest.

Tanenhaus’s cue theory, as it came to be known, provoked a flurry of empirical work that continues to this day. Studies tried to determine which cues the Court uses (if any), whether cues are dispositive or merely suggestive, and whether the Court’s use of cues changes over time (Armstrong and Johnson 1982; Teger and Kosinski 1980). Development of the underlying theory of information acquisition received little attention, however. Consequently, scholars largely ignored the strategic manipulation of cues, although they soon recognized the importance of reputation in the behavior of the Solicitor General (Caldeira and Wright 1988; Tanenhaus et al. 1963; Ulmer 1984).

The neglect of the strategic dimension of cue theory began to change when Caldeira and Wright (1988) demonstrated the effect of amicus briefs on the probability of granting certiorari.³ When coupled with new findings about the litigation strategies of interest groups, Caldeira and Wright’s finding strongly suggests that at least one cue, the presence of an amicus brief,

is deliberately manipulated by interested parties in order to affect the behavior of the Court (Epstein 1991). This move to greater theoretical sophistication continued in Perry’s (1991) landmark study of the certiorari process. Perry adopted concepts from Jervis’s (1970) classic work on signaling in international relations, including the distinction between *signals*, which are manipulable, and *indices*, which are not. Nonetheless, Perry did not address many of the issues that occupy center stage in game-theoretic accounts of signaling.

Meaning and credibility in certiorari signaling. Contemporary signaling theory revolves around two issues: meaning and credibility (Banks 1991; Farrell 1993). With respect to meaning, the key issues are the identity of the signal sender, the nature of his or her private information, and the interpretation of the message by the signal receiver. In the context of certiorari, the obvious questions are: Who is the signaler? What does the signaler know that the Supreme Court does not? What signals might reflect this private information? How can the Court understand or interpret the signal? With respect to credibility, the key issue is sometimes called “the Jervis paradox”: If sending a particular message benefits a signaler with private information, then why do not others without the information do the same thing? If they do, then why does this not destroy the credibility of the message? In the context of certiorari, if a lower court that follows Supreme Court doctrine can signal its conformance by sending a particular message, then why do not courts who flout doctrine send the same signal? If they do, how can the signal from the first court credibly convey the intended message: “There is no need to review me since I am conforming to your doctrine”? The contemporary theory of signaling games provides powerful tools for answering these questions.

A JUDICIAL SIGNALING GAME

The Model

Our model shares some similarities with other auditing models in the social sciences, such as the importance of mixed strategy equilibria, but its structure differs from most others in three ways: It includes a publicly observable index, the message space for the signaler is restricted to two actions, and both actors agree about the disposition of some cases even when they differ strongly about others. Consequently, the equilibria in our model also differ in some regards from most others (for a review of those models and their characteristics, see Andreoni, Erard, and Feinstein 1998).

The players are higher court H and lower court L . Play of the game determines the legal disposition of a search-and-seizure case. A “case” is a set of facts (i.e., a point in a fact space). Each point in the fact space corresponds to a degree of intrusiveness of the search. That is, points in the fact space map into the real line X , where each point x indicates a degree of intrusiveness. Each case thus corresponds to a point x in X .

² Overall, 23.3% of the search-and-seizure cases were affirmed by the Burger Court (U.S. Supreme Court database, using orally argued citation and split vote as the unit of analysis). Some of these affirmations may be consistent with our model. That is, because of differences between the observed and actual intrusiveness of the search, the Court may take a case that seems suspect only to find on further review that it was correctly decided.

³ Cue theory aside, strategic analyses of certiorari date to Schubert 1959; also see Brenner 1979.

Some of the facts in a case are publicly observable, others are not. For example, it is easily determined whether a search took place in a car or home, or whether a warrant was issued; whether the police had "probable cause" for the search is a more difficult determination that can only be reached after careful review of evidence and testimony.

Let $x = \hat{x} + t$, where \hat{x} is the intrusiveness of the case based on the publicly observed facts, and t is an additional increment or decrement of intrusiveness based on the nonpublicly observable facts. We assume $t \in T = X$ and is distributed according to distribution F , with common knowledge density f assumed everywhere continuous, differentiable, and nonzero. We assume throughout that F is log concave (displays the monotone hazard rate property). This condition is met by most common probability distributions, including the normal, logistic, chi-squared, exponential, and Laplace (Bagnoli and Bergstrom 1989). We further assume the density $f(t)$ is independent of the realization of \hat{x} .

Sequence of Play, Information Structure, and Strategies.

The lower court hears the case, learns all the facts in the matter (i.e., it learns \hat{x} and t and hence x), and either admits or excludes the evidence. The higher court sees only the publicly observable facts (i.e., it learns \hat{x} but not t) and the decision of the lower court. H may deny certiorari (i.e., decline to review the case), and L 's decision stands. Or, at cost k , H may review the case, learn the private information t , and affirm or reverse L .

Given this information structure, the action of L becomes a "signal" about its private information, exactly as noted by Songer (1979), while the publicly observable degree of intrusiveness becomes an "index" in Perry's terminology.

A strategy for L is a function

$$s: T \times X \rightarrow \Delta(M),$$

where $\Delta(\cdot)$ denotes the set of probability distributions over a finite set, and $M = \{m_1, m_2\}$, with $m_1 =$ exclude and $m_2 =$ admit. Thus, $s(t; \hat{x})$ gives the probability of excluding the evidence given the private information t and the public information \hat{x} .

If H hears the case it will affirm or reverse L in such a way that the evidence is admitted or excluded according to its own preferred doctrine. Hence, we can take H 's strategy as simply a reviewing or auditing strategy:

$$r: M \times X \rightarrow \Delta(A),$$

where $A = \{a_1, a_2\}$, with $a_1 =$ grant certiorari and $a_2 =$ deny certiorari. Thus, $r(m_i; \hat{x})$ is the probability of granting certiorari given action m_i by L and the observable degree of intrusiveness \hat{x} .

Preferences. Both courts have preferences defined partly in relation to the ultimate disposition of the case. Each wishes the case to be adjudicated correctly, given its own notions of justice. More specifically, each has a preferred legal rule, which uses a cut-point x_i in X , $i =$

H, L . Court i believes the evidence should be excluded in cases that are more intrusive than x_i and admitted in cases less intrusive than x_i . This captures the notion of a legal rule or doctrine and the "correct" versus "incorrect" dispositions of cases, concepts of critical importance in the judicial setting. Each court wishes to decide cases correctly in view of its preferred legal rule. Note, however, that the two courts differ in the value of the ideal doctrinal cut-point. The reason courts prefer different legal rules is not modeled here, but it could result, for example, from different perceptions of the social costs and benefits of curbing the police.

The unidimensional space X is exactly the "attitude" space of the "attitudinal model" assumed in many empirical studies of judicial decision making (Segal and Spaeth 1993). The ideal cut-point partitions the space in precisely the fashion noted in attitudinal studies. This partitioning, combined with the mapping from the n -dimensional fact space into the unidimensional attitude space, in turn partitions the fact space in exactly the way required by the concept of a judicial doctrine. Conveniently, however, most of the action of the model takes place in the simple attitude space.

To complete the specification of the doctrinal component of preferences, we normalize each player's payoffs. If a case is decided correctly from its perspective, then the court receives a payoff of 0; if incorrectly, the payoff is -1 . This stylization reflects the different value of a correct versus an incorrect decision to the courts while simplifying the development and exposition of the model. In the interest of brevity and clarity in presentation, we take $x_L < x_H$ so that H is more conservative than L ; results concerning the other, completely symmetrical, case are noted where appropriate.

In addition to preferences about outcomes, L prefers not to be reversed. If reversed, it suffers an $\epsilon > 0$ utility loss; this loss reflects the influence of judicial culture. We do not try to account for the origin or maintenance of this loss, although one can view it as sustained through repeated play within the legal community. Finally, the higher court loses $k \in (0, 1)$ if it hears the case.⁴ This auditing cost reflects the time and effort of hearing the case, which involves a cost in other cases unheard and in leisure and other activities forgone. Given the normalization of the payoffs, ϵ and k can be seen as the utility losses from reversal and auditing relative to the utility loss of an incorrectly decided case.

The utility function for the higher court is thus:

$$u_H = \begin{cases} 0 & \text{if } \begin{cases} x < x_H \text{ and } m = \text{admit and } a = \text{deny} \\ x \geq x_H \text{ and } m = \text{exclude and } a = \text{deny} \end{cases} \\ -1 & \text{if } \begin{cases} x < x_H \text{ and } m = \text{exclude and } a = \text{deny} \\ x \geq x_H \text{ and } m = \text{admit and } a = \text{deny} \end{cases} \\ -k & \text{if } a = \text{grant} \end{cases}$$

⁴ We bound the auditing cost by 1; otherwise, the higher court would never find it profitable to review cases (since the most it can gain as a result, exclusive of the auditing cost, is 1).

and for the lower court:

$$u_L = \begin{cases} 0 & \text{if } \begin{cases} x < x_L \text{ and } m = \text{admit} \\ x_L \leq x < x_H \text{ and } m = \text{exclude and } a = \text{deny} \\ x \geq x_H \text{ and } m = \text{exclude} \end{cases} \\ -1 & \text{if } \begin{cases} x < x_L \text{ and } m = \text{exclude and } a = \text{deny} \\ x_L \leq x < x_H \text{ and } m = \text{admit} \\ x \geq x_H \text{ and } m = \text{admit and } a = \text{deny} \end{cases} \\ -1 - \varepsilon & \text{if } x_L \leq x < x_H \text{ and } m = \text{exclude and } a = \text{grant} \\ -\varepsilon & \text{if } \begin{cases} x < x_L \text{ and } m = \text{exclude and } a = \text{grant} \\ x \geq x_H \text{ and } m = \text{admit and } a = \text{grant} \end{cases} \end{cases}$$

Expected Utilities and Best Responses. The expected utility for the lower court from strategy $s(t; \hat{x})$ follows straightforwardly:

$$u_L(s(t; \hat{x})) = \begin{cases} s(t; \hat{x})[r(m_1; \hat{x})(1 - \varepsilon) - 1] & \text{if } x < x_L \\ s(t; \hat{x})[1 - r(m_1; \hat{x})(1 + \varepsilon)] - 1 & \text{if } x_L \leq x < x_H \\ (1 - s(t; \hat{x}))[r(m_2; \hat{x})(1 - \varepsilon) - 1] & \text{if } x \geq x_H. \end{cases}$$

L 's best-response correspondence is immediate:

$$BR_L(r(m; \hat{x}), t; \hat{x}) = \begin{cases} s(t; \hat{x}) = 0 & \text{if } \begin{cases} x < x_L \\ x_L \leq x < x_H \text{ and } r(m_1; \hat{x}) > \rho \end{cases} \\ s(t; \hat{x}) \in [0, 1] & \text{if } x_L \leq x < x_H \text{ and } r(m_1; \hat{x}) = \rho \\ s(t; \hat{x}) = 1 & \text{if } \begin{cases} x \geq x_H \\ x_L \leq x < x_H \text{ and } r(m_1; \hat{x}) < \rho, \end{cases} \end{cases}$$

where $\rho \equiv 1/(1 + \varepsilon)$.

L 's best response correspondence indicates there are three strategically distinct regions in X : the region below x_L , the region between x_L (inclusive) and x_H , and the region above x_H (inclusive). If x falls into the first region, then L always admits the evidence; if x falls into the third region, then L always excludes the evidence; if x falls into the second region, the "conflict region," then L admits or excludes the evidence depending on whether the probability of certiorari (given exclusion and the observable facts) falls above or below a critical value ρ . If x falls into the conflict region and the probability of certiorari equals ρ , then L is indifferent between admitting and excluding the evidence and may randomize between the two pure strategies in any fashion.

This best-response correspondence is quite intuitive. The interests of the two players correspond when x is very high or very low (i.e., at or above H 's ideal cut-point or below L 's). Therefore, when x falls into those regions L can pursue her ideal policy without fear of reversal, yielding L a best response that is invariant to H 's strategy. When x falls into the conflict region, the interests of the players are opposed. In this

case, a sufficiently imposing chance of reversal will keep L from pursuing her ideal policy. For cases that fall into the conflict region, the role of reversal costs is critical; as the cost of reversal falls to zero, the probability of an audit must go to one if the lower court is to be dissuaded from implementing its own preferred doctrine.

Finally, note that for any set of observable facts \hat{x} it is possible for L to have the best response "admit" or "exclude" (since by assumption $F(t_a) > 0$ and $1 - F(t_b) > 0$, where $t_a \equiv x_L - \hat{x}$ and $t_b \equiv x_H - \hat{x}$). This has an important implication: All information sets can be reached in equilibrium, so the need to determine beliefs at unreached information sets does not arise.

Let $\mu(x; m_i, \hat{x})$ denote H 's beliefs about x given the observable facts \hat{x} , L 's action m_i , and L 's strategy. For example, $\mu(x < x_H; \text{admit}, \hat{x}')$ indicates the probability H puts on x lying below x_H after observing "admit" in a case with observable obtrusiveness \hat{x}' . The following fact is important: Given L 's best-response correspondence, H 's beliefs about x must be concentrated on the region below x_H after having observed "admit," regardless of the value of \hat{x} . Therefore, after observing "admit," the expected utility to H of review strategy $r(\text{admit}, \hat{x})$ must be $-r(\cdot)k$. This expression is maximized at $r(\text{admit}, \hat{x}) = 0$ for all \hat{x} . Accordingly, in any equilibrium, $r(\text{admit}; \hat{x}) = 0$ for all \hat{x} .

In addition, given L 's best-response correspondence, H 's belief about x must be concentrated above x_L following the observation of "exclude," regardless of the value of \hat{x} . Recall that the payoff to H from denying certiorari given exclusion is zero if $x \geq x_H$ but -1 if $x < x_H$. Therefore, following the observation of "exclude," the expected utility to H of review strategy $r(\text{exclude}, \hat{x})$ is for all \hat{x}

$$-r(\text{exclude}, \hat{x})k - (1 - r(\text{exclude}, \hat{x}))\mu(x_L \leq x < x_H, \text{exclude}, \hat{x}), \quad (1)$$

where, from Bayes's rule,

$$\begin{aligned} \mu(x_L \leq x < x_H, \text{exclude}, \hat{x}) &= \frac{\int_{t_a}^{t_b} s(\hat{x} + t)f(t)dt}{1 - F(t_b) + \int_{t_a}^{t_b} s(\hat{x} + t)f(t)dt} \end{aligned}$$

(recall that $x_L = \hat{x} + t_a$ and $x_H = \hat{x} + t_b$). The solution to maximizing expression 1 with respect to $r(\text{exclude}, \hat{x})$ depends on the relative magnitudes of k and $\mu(x_L \leq x < x_H, \text{exclude}, \hat{x})$. Thus, in any equilibrium, for all \hat{x}

$$r(\text{exclude}; \hat{x}) = \begin{cases} 1 & \text{if } \mu(x_L \leq x < x_H) > k \\ \alpha \in [0, 1] & \text{if } \mu(x_L \leq x < x_H) = k \\ 0 & \text{if } \mu(x_L \leq x < x_H) < k. \end{cases} \quad (2)$$

Equilibria

We seek perfect Bayesian equilibria (PBE) to the judicial signaling game. In doing so it is convenient to consider a level of \hat{x} , \hat{x}^* , such that H 's posterior belief about the true location of the case, given observed \hat{x} and "exclude" by L , leave him just indifferent between granting and denying certiorari, even if L definitely excludes the evidence for all cases truly lying in the conflict region (i.e., even if $s(t, \hat{x}) = 1 \forall x \in [x_L, x_H]$). Using the middle part of equation 2 and Bayes's rule, \hat{x}^* is implicitly defined as the level of \hat{x} such that

$$\frac{F(t_b) - F(t_a)}{1 - F(t_a)} = k. \quad (3)$$

Note that \hat{x}^* may not exist if k is sufficiently large, that is, if the left-hand quantity in equation 3 is less than k for all values of \hat{x} . If \hat{x}^* does exist, then the assumption of log concavity of $F(\cdot)$ assures that it will be unique.

PROPOSITION 1.

1. If \hat{x}^* does not exist, then the following is a unique PBE:

$$s(t, \hat{x}) = \begin{cases} 0 & \text{if } x < x_L \\ 1 & \text{otherwise} \end{cases}$$

$$r(m, \hat{x}) = 0 \forall m \text{ and } \forall \hat{x}$$

and beliefs are determined everywhere by Bayes's rule.

2. If \hat{x}^* exists, then the PBE are characterized by:

$$s^*(t, \hat{x})$$

$$= \begin{cases} 1 & \text{if } \begin{cases} x \geq x_H \\ x \geq x_L \text{ and } \hat{x} \geq \hat{x}^* \end{cases} \\ \beta(t, \hat{x}) \in [0, 1] \text{ such that } \mu(x_L \leq x < x_H, \text{ exclude, } \\ \hat{x}) = k \text{ if } x_L \leq x < x_H \text{ and } \hat{x} < \hat{x}^* \\ 0 & \text{if } x < x_L \end{cases}$$

$$r^*(m_i, \hat{x}) = \begin{cases} \frac{1}{1 + \epsilon} & \text{if } \hat{x} < \hat{x}^* \text{ and } m_i = \text{exclude} \\ 0 & \text{if } \begin{cases} \hat{x} \geq \hat{x}^* \text{ and } m_i = \text{exclude} \\ m_i = \text{admit} \end{cases} \end{cases}$$

and beliefs are determined everywhere by Bayes's rule.⁵

Part 2 of the proposition allows a great deal of latitude in admissible strategies for L . The latitude arises for cases whose observable intrusiveness is below \hat{x}^* and whose true location lies within the conflict region. Hence, part 2 describes many equilibria. In all these equilibria, however, H 's posterior belief that x lies in the conflict region, after observing "exclude" and \hat{x} , must exactly equal k . In this sense, all of L 's strategies characterized in part 2 can be said to be

"belief equivalent." All would be observationally equivalent to an outside observer not privy to t .

Example. Perhaps the best way to illustrate the review equilibria in part 2 of the proposition is through an example. Suppose $x_L = 0$, $x_H = 1$, $k = 1/2$, $\epsilon = 1/2$, and f is a normal distribution with mean zero and standard deviation 1. In this example, if $\hat{x} < \hat{x}^*$, then $s(t, \hat{x}) = s(t', \hat{x}) \forall x \in [x_L, x_H]$. In other words, if a case lies in the conflict region and $\hat{x} \leq \hat{x}^*$, then L keys the probability of excluding the evidence only on the case's observed intrusiveness.

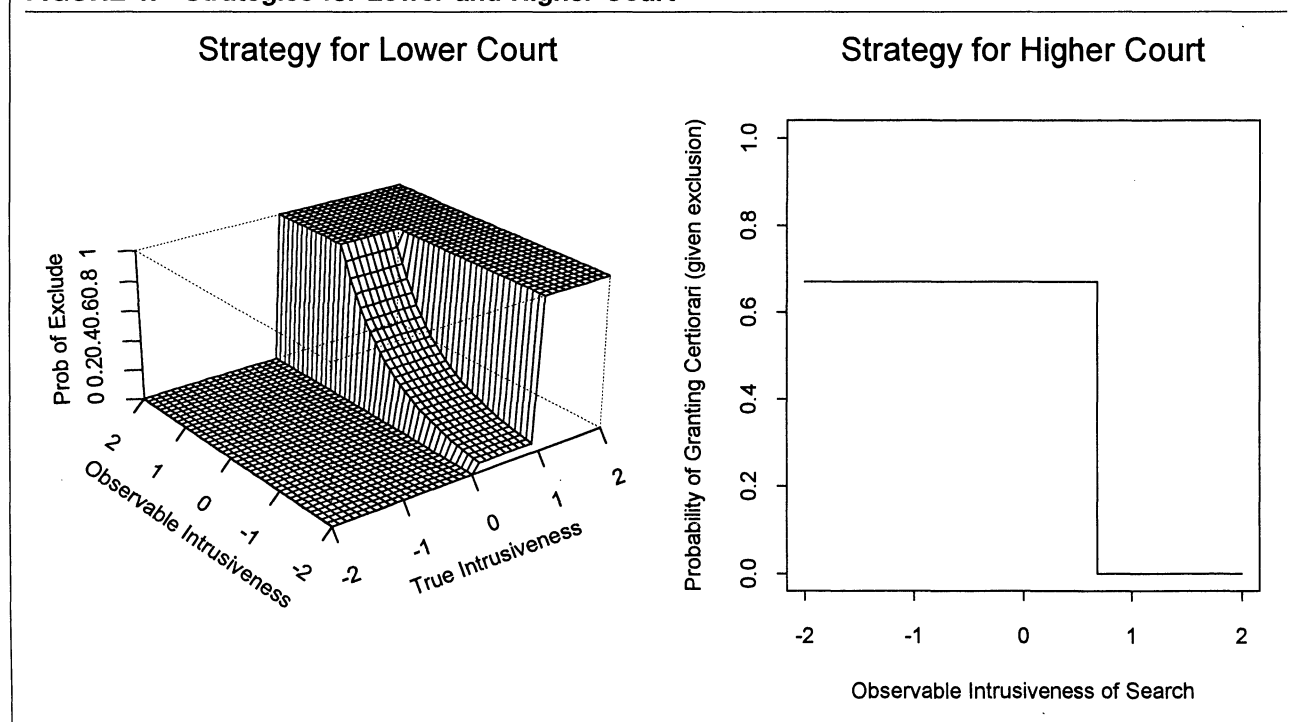
Consider the strategy of the lower court, shown in the left-hand panel of Figure 1. If $x < x_L$ or $x \geq x_H$, then the lower court has obvious actions regardless of the value of \hat{x} . When $x < x_L$, the lower court admits the evidence; when $x \geq x_H$, the lower court definitely excludes the evidence. It can do so because even if audited it will not be reversed. If x falls into the conflict region, then L 's strategy depends on the value of \hat{x} , the publicly observable level of intrusiveness. When \hat{x} is sufficiently high, at or above the critical level \hat{x}^* (which can be calculated to be about .69 in this example), the high court will not review the case even given exclusion of the evidence, for reasons explained below. Given this, the lower court can implement its preferred doctrine and exclude the evidence. When \hat{x} is lower than \hat{x}^* , the lower court uses a probabilistic strategy keyed to the observable facts. For instance, when $\hat{x} = .3$, the probability of excluding the evidence is about .64.

Now consider the strategy of the higher court. If it observes the admission of the evidence, then it accepts the lower court's judgment. This follows from the fact that L is more liberal than H , so any evidence it finds admissible will surely be found admissible by H as well. The higher court has a more difficult decision if it observes the exclusion of the evidence (the right-hand panel of Figure 1). If the true situation is $x \geq x_H$, then H would like L 's judgment to stand and would prefer not to review the case. But if $x < x_H$, then H would like to review the case and reverses L , even though granting certiorari entails cost k . Unfortunately for the higher court, it can only observe \hat{x} and the fact of exclusion. So, given the observable degree of intrusiveness, the two ideal doctrinal cut-points, the cost of auditing, and the probability distribution $F(t)$, the higher court must decide whether it is worthwhile to check the lower court after observing "exclude." When \hat{x} is large enough (equal to or greater than \hat{x}^*), it will not be worthwhile; that is, the probability that x truly falls into the conflict region between the two doctrinal cut-points will not be large enough to warrant the audit. When $\hat{x} < \hat{x}^*$, the higher court's calculation goes the other way. When H observes "exclude" from a liberal lower court, it employs a probabilistic auditing strategy keyed to the value of ϵ . In this case, the probability of granting certiorari after observing "exclude" is $2/3$.

These strategies of the two players together determine outcomes. For example, if the liberal lower court admits the evidence, then the higher court does not

⁵ Recall that the proposition assumes the higher court is more conservative than the lower court. There is a "mirror" proposition when the higher court is more liberal than the lower court.

FIGURE 1. Strategies for Lower and Higher Court



grant certiorari, and the lower court's decision stands. If the lower court excludes the evidence and $\hat{x} < .69$, then the higher court grants certiorari with a $\frac{2}{3}$ chance. In the event of certiorari, if x actually falls in the conflict region, then H reverses L ; if it does not fall in that region, then H affirms. (Note that the model predicts some grants and affirms as well as grants and reversals.) For instance, when x falls into the conflict region and $\hat{x} = .2$, L admits the evidence with a 42% chance and excludes it with a 58% chance; if it excludes, then H denies certiorari with a 33% chance, so the judgment stands, and grants certiorari with a 67% chance; if it so grants, then H reverses L . Finally, when $\hat{x} \geq .69$, L decides the case according to its own preferred doctrine, and H always denies certiorari, even if L excluded the evidence.

Meaning and Credibility. How do meaning and credibility play out in the model? The signals "admit" or "exclude" are inherently ambiguous, since x lies on a continuum, but the messages are dichotomous (i.e., the most informative equilibrium can only be a partial pooling equilibrium). Because L is liberal, however, the meaning of the signal "admit" (i.e., $x < x_H$) is sufficiently clear to make H 's decision easy, provided the message is credible; indeed, the message is perfectly credible when L is more liberal than H . The signal "exclude" from L is also ambiguous, with more serious ramifications for H . A liberal lower court always (weakly) prefers the higher court to interpret "exclude" as meaning $x \geq x_H$, but this interpretation cannot be inherently credible because of the potential tension between the two courts. Instead, H uses an index, the observable degree of intrusiveness, to judge

the message credibility: If the observable facts indicate intrusiveness above the critical threshold \hat{x}^* , then the interpretation $x > x_H$ is sufficiently credible to warrant accepting L 's decision. If not, then H will occasionally audit L . The audit rate will be keyed to the lower court's sensitivity to reversals, and sensitive courts will be audited less frequently than insensitive ones.

Hypotheses about Review

Because the model treats certiorari as part of an interactive incentive system, it generates hypotheses not only about the probability of review but also about the exclusion of evidence in the lower court. Both types of hypotheses can be used to test the model; because our primary interest is the certiorari decision, we focus on the former.

The first two hypotheses follow straightforwardly from the propositions but make explicit the consequences of a lower court that is more conservative than the higher court.

HYPOTHESIS 1 (*liberal admission or conservative exclusion*). The probability the higher court grants certiorari is zero if $x_L < x_H$ and the lower court admits the evidence, or $x_L > x_H$ and the lower court excludes the evidence.

HYPOTHESIS 2 (*liberal exclusion or conservative admission*). (a) If $x_L < x_H$ and the lower court excludes the evidence, then the probability the higher court grants certiorari is zero only if the search is sufficiently intrusive ($\hat{x} \geq \hat{x}^*$); (b) if $x_L > x_H$ and the lower court admits the evidence, then the probability the higher court

grants certiorari is zero only if the search is sufficiently unintrusive ($\hat{x} < \hat{x}^$).*

A striking feature of these hypotheses, which are new to the literature, is the way the higher court's behavior depends on the lower court preferences *relative* to the higher court. The asymmetry of H 's response to, say, "admit" arises because the meaning of the signal depends on the nature of the sender (a liberal or conservative lower court).⁶ These propositions—particularly the first, which might be called the Nixon goes to China proposition—are strongly reminiscent of Calvert's (1985) work on the value of information from a biased source ("if a cold warrior like Nixon goes to China it must be time for a change in American policy").

Now consider two circuit courts, both more liberal than the Supreme Court but one less liberal than the other. How does the critical value of the index \hat{x}^* , the level of intrusiveness above which the Supreme Court does not review even following "exclude," differ for the two courts? This question is answered by the third hypothesis.

HYPOTHESIS 3 (proximity of lower court). *Consider two lower courts. (a) If $x_L < x_H$ for both lower courts, then (i) the critical value \hat{x}^* is lower for the less liberal lower court, and (ii) the probability of review is (weakly) lower for that court given the same observable facts and actions. (b) If $x_L > x_H$ for both lower courts, then (i) the critical value \hat{x}^* is lower for the more conservative lower court, and (ii) the probability of review is (weakly) greater for that court given the same observable facts and actions.*

Hypothesis 3 indicates that moving a lower court closer to the higher court expands the region in which cases escape review, a sensible result. Hypothesis 3 also indicates that if one estimates the probability of review as a function of observable facts and the ideological distance between the higher and lower court, then the coefficient on the distance between the two courts' ideologies should be positive. This hypothesis is new to the literature.

HYPOTHESIS 4 (change in high court's preferences). *(a) As the court becomes more conservative or more liberal, \hat{x}^* moves in the same direction (e.g., if x_H increases, then \hat{x}^* increases). Consequently, (b) as the higher court becomes more conservative, cases from liberal lower courts become (weakly) more likely to be reviewed (controlling for observable facts), and cases from conservative lower courts become (weakly) less likely to be reviewed.*

⁶ Referring to Figure 1, it may seem plausible to strengthen hypothesis 2 so that the probability of review, if nonzero, is a constant with respect to observed facts. Yet, such a strengthening requires the additional assumption that ϵ is not only uncorrelated with the distance between the courts but also identical across all lower courts. Whether the probability of review, when positive, varies with respect to observable facts can be seen as a test of this rather strong assumption.

Hypothesis 4 is closely related to hypothesis 3 and also seems quite sensible, although it is new to the literature.

One of the most established facts about certiorari is that amicus briefs and the involvement of the Solicitor General greatly increase the probability the Supreme Court will accept a case. A complete analysis of those two factors would require modeling third-party signaling and is beyond our scope here. In a somewhat similar model, however, Banks and Weingast (1992) note that third-party signals can be seen as reducing auditing costs. The following hypothesis considers the effects of changes in the cost of review.

HYPOTHESIS 5 (change in the cost of review). *A decrease in the cost of review (a) increases \hat{x}^* if $x_L < x_H$ and decreases \hat{x}^* if $x_L > x_H$. Consequently, (b) the probability of review (weakly) increases, controlling for observable levels of intrusiveness.*

This hypothesis offers one reason amicus briefs or involvement of the Solicitor General may increase the probability of certiorari: They reduce information costs for the Supreme Court.

Hypothesis 5 and the no-review equilibrium in part 1 of proposition 1 have an interesting implication. When the Supreme Court has little interest in a technically demanding area (such as admiralty or patent law), the value of k will be large, and the Court will largely or entirely abandon that area. For instance, during the last 50 years the Court has only rarely heard admiralty cases, once a substantial portion of its caseload. Conversely, when the Court has a substantial interest in a less technically demanding area (e.g., civil liberties), hypothesis 5 suggests it is likely to hear many such cases. The topic seems to act as a "cue" for the Court, though in fact the causal mechanism is not so much signaling as the magnitude of information costs relative to the possible gains from review. The importance of this type of cue is a recurrent finding in the literature.

EMPIRICAL ANALYSIS

Data

The data for this analysis were assembled as follows. As part of a larger project on circuit court decision making, we drew a random sample of U.S. Courts of Appeals search-and-seizure opinions (including per curiams) published in the *Federal Reporter* from 1961 through 1990. We determined the universe of these cases by searching Westlaw (www.westlaw.com) for all "searches and seizures," all Fourth Amendment cases, and criminal law topics 219, 226, 364, 394, and 207, plus relevant combinations of keywords. From this universe we then drew a random sample stratified by year, consisting of 40 cases per year. Any that were not in fact search-and-seizure cases were discarded and replaced with the next case from the universe. For years with less than 40 cases, all published cases were included in the sample.

The original sample of more than 1,100 published appeals court decisions yielded only 18 cases reviewed

by the Supreme Court. This is not surprising, given the rarity of certiorari. Therefore, the next step was to add all other appeals court search-and-seizure cases reviewed by the high court. Such a sample, known as a choice-based sample, is commonly used when the event of interest is rare. Special statistical techniques are required for such samples, to which we return shortly.

Each case in the final sample was coded in various ways: whether the appeals court upheld or struck down the search, whether the case was appealed to the Supreme Court, whether certiorari was granted, and so on. We also coded fact patterns for each case using the procedures employed by Segal (1984). The identity and votes of each judge on the appellate court panel were also recorded. Using the identity information, appeals court judges were matched with data on their personal background and characteristics (see Songer, Segal, and Cameron 1994). Characteristics were unavailable for some, since often district court judges and retired judges sit as the third member of appellate court panels. Personal characteristics were available for the lead judge in most of the cases, however.

Our model indicates five key variables. The dependent variable is whether certiorari was granted (*Grant*). The independent variables are the signal, that is, whether the lower court admitted the evidence (*Admit*); an index, the publicly observable intrusiveness of the search (*Intrude*); another index, the ideology of the lower court (*JI*); and the ideology of the Supreme Court.

Using the case facts, we constructed the measure of the publicly observable intrusiveness of the search in the following way. Segal (1984) estimates the likelihood the Supreme Court will uphold a search, based on the fact pattern in cases and a proxy for the ideology of the Court. In particular, Segal estimates weights on each variable using a logit regression. Since the Court's decision to uphold or strike down search-and-seizure cases depends partly on the intrusiveness involved, the logit weights on case facts provide a convenient way to measure intrusiveness.⁷ More specifically, we construct the variable *intrude* as follows:

$$\begin{aligned} \text{Int} = & -3.256 \times \text{Incident} - 1.049 \times \text{Afterlaw} + .06 \\ & \times \text{Unlawful} - 1.928 \times \text{Warrant} + 3.25 \times \text{Home} \\ & + 2.054 \times \text{Person} + 2.733 \times \text{Business} + 2.243 \\ & \times \text{Car} - 1.411 \times \text{Except} + 1.766 \times \text{Extent}. \end{aligned}$$

The definitions for the case fact variables are identical to those of Segal (1984), but we reverse the sign on the variables, as we wish to measure intrusiveness rather than lack of it. We also do not employ the constant or Segal's proxy for Supreme Court ideology, since we wish to measure intrusiveness rather than probability of upholding or reversing. Furthermore, we do not

employ Segal's variable *Probcaus*, which indicates whether the lower court found probable cause for the search. We argue that, for the most part, findings about probable cause do not reflect publicly observable facts but the private information the lower court obtained through its scrutiny of the case. It is exactly in its determination of probable cause that a lower court can try to exploit its informational advantage to alter the outcome of a search-and-seizure case in its preferred direction. It is worth noting that lower court determinations of probable cause do not affect the Supreme Court's decision on the merits, according to Segal's analysis (i.e., the coefficient on this variable was not statistically different from zero). This makes perfect sense from an informational standpoint: Once the Supreme Court has heard a case, it can make its ruling based on its own knowledge of previously unobservable facts, not what the lower court says about them.

We created a measure of the doctrinal preferences of the lower court in a two-step fashion, following Songer, Segal, and Cameron (1994).⁸ First, we regressed the votes of the appellate judges on their background characteristics (e.g., whether they were appointed by a Republican president or were a prosecutor before becoming a judge). The logit weights from this regression were used to form an index of conservatism for each judge for whom background information was available (low numbers on the index indicate a presumptively liberal judge; higher numbers a more conservative one).⁹ This is obviously far from perfect as a measure of doctrinal preferences. Nonetheless, presumptively conservative judges, based on simple background characteristics, tend to vote more often to uphold searches relative to presumptively liberal judges (Songer, Segal, and Cameron 1994). Hence, the index has a degree of face validity. Thus, the first stage yields a measure of ideology for each judge on the three-judge appellate panel.

The formal model assumes a single measure of ideology for a lower court, an abstraction adopted to simplify an already complex model. In the empirical analysis, however, we have to consider how the Supreme Court may determine the ideology of the lower court judges. Several possibilities present themselves. Most obviously, the Court may cue off the ideology of the opinion writer (measured by our variable *JI*); we explore this possibility below. Or the Court may try to judge the general ideological tenor of the lower court, perhaps best reflected in the mean ideology of the judges in the majority. Alternatively, the Court may use its limited information very efficiently, cueing off the

⁸ See Brace, Hall, and Langer 1998 and Giles, Hettinger, and Peppers 1998 for similar uses of regression-based techniques to impute judicial ideologies.

⁹ The logit is $.353 \text{ region} - .311 \text{ appointing president} + .381 \text{ religion} + .189 \text{ prosecutor} + .152 \text{ judicial experience}$, defined as: judge's region (1 = South, 0 = non-South), the appointing president's ideology (1 = liberal, 0 = moderate, -1 = conservative [Tate and Handberg 1991]), the judge's religion (1 = Catholic, 0 = other), prosecutorial experience (1 = yes, 0 = no), and judicial experience (1 = yes, 0 = no). A simple index based on the number of presumptively conservative traits works almost as well as the multivariate regression in predicting votes of appellate judges.

⁷ The following equation provides the mapping from fact space to attitude space discussed in the previous section. Whether this simple linear mapping is adequate to capture important doctrinal subtleties is an interesting question, but Segal's work demonstrates its empirical power.

most liberal judge in the majority in conservative decisions and the most conservative judge in the majority in liberal decisions. Information about a dissenting judge, if any, also may be useful. Data limitations prevent a detailed analysis of these possibilities.¹⁰

We control for the doctrinal preferences of the Supreme Court in a simple way: We restrict our attention to the last three natural courts (courts in which there are no personnel changes) of the Burger years (Burger Courts 3–5 in Table 5-2 of Epstein et al. 1997). That is, we examine cases from January 7, 1972, to September 25, 1986—from the addition to the Court of William Rehnquist up to the addition of Antonin Scalia—which yields 273 usable appealed cases. During this period, in the arena of criminal justice the dominant coalition on the Court was relatively stable and quite conservative. Moreover, the bulk of the appellate court panels that the high court faced were almost certainly more liberal than the Supreme Court, at least with respect to search and seizure.¹¹ This makes the detection of relevant patterns in the data much easier.

Method

Hypotheses. The data allow us to address hypotheses 1–3 in the signaling model. Since the Burger Court was almost certainly as conservative, if not more so, in search-and-seizure cases as most of the lower court panels it faced (Blasi 1983), the model predicts a series of patterns that are neatly captured in the conditioning plot shown in Figure 2. Such plots are powerful devices for detecting interactions or (as is relevant in this case) patterns across statistical regimes (Cleveland 1993). The plot is read in the following way. Each panel shows the relationship between the probability of granting certiorari and the intrusiveness of the search, under a specific condition. In the top two panels, the lower court admitted the evidence. In the bottom two panels, the lower court excluded the evidence. In the left-hand panels (top and bottom), the judge who wrote the opinion for the lower court was quite liberal. In the right-hand panels, the judge who wrote the opinion for the lower court was conservative (but, recall, less conservative than the Supreme Court).

The signaling model predicts several patterns when the data are arranged in this fashion (Figure 2). The

top two panels should show a zero probability of review at all levels of intrusiveness: If the lower court took the conservative action, then the Burger Court should not review it (from hypothesis 1). The prediction for the lower panels is more complicated. Refer to the step function in the right-hand panel of Figure 1. The signaling model predicts such a step function in each of the lower panels of Figure 2, but it should be located farther to the right in the left-hand panel than in the right-hand panel (hypothesis 3.a.i.), and it may be beyond any level of intrusiveness actually observed. So, at any value of intrusiveness, the probability of review should be no greater in the lower right-hand panel than the lower left-hand panel (hypothesis 3.a.2). Finally, in the lower panels, if the probability of granting certiorari approaches zero, it should do so only for high levels of intrusiveness (hypothesis 2).

Choice-Based Sampling. The estimated models need to reflect the fact that the sample is choice based. This is easily done by weighting the observations (see Manski and Lerman 1977; Greene 1991, Section 36.5, provides a simple example). In the initial sample, the proportion of appealed cases actually heard by the Supreme Court was 5%. These were subsequently oversampled, so that they account for about 27% of the final sample. Consequently, cases granted appeal are overrepresented by $.27/.05 = 5.4$, and cases not granted review are underrepresented by $.73/.95 = .77$. To account for this, the granted cases need to be weighted by the factor $.05/.27 = .19$, and the rejected cases need to be weighted by the factor $.95/.73 = 1.3$. In all the models presented below, including the scatter plot smoothers, this procedure was followed.

Data Display

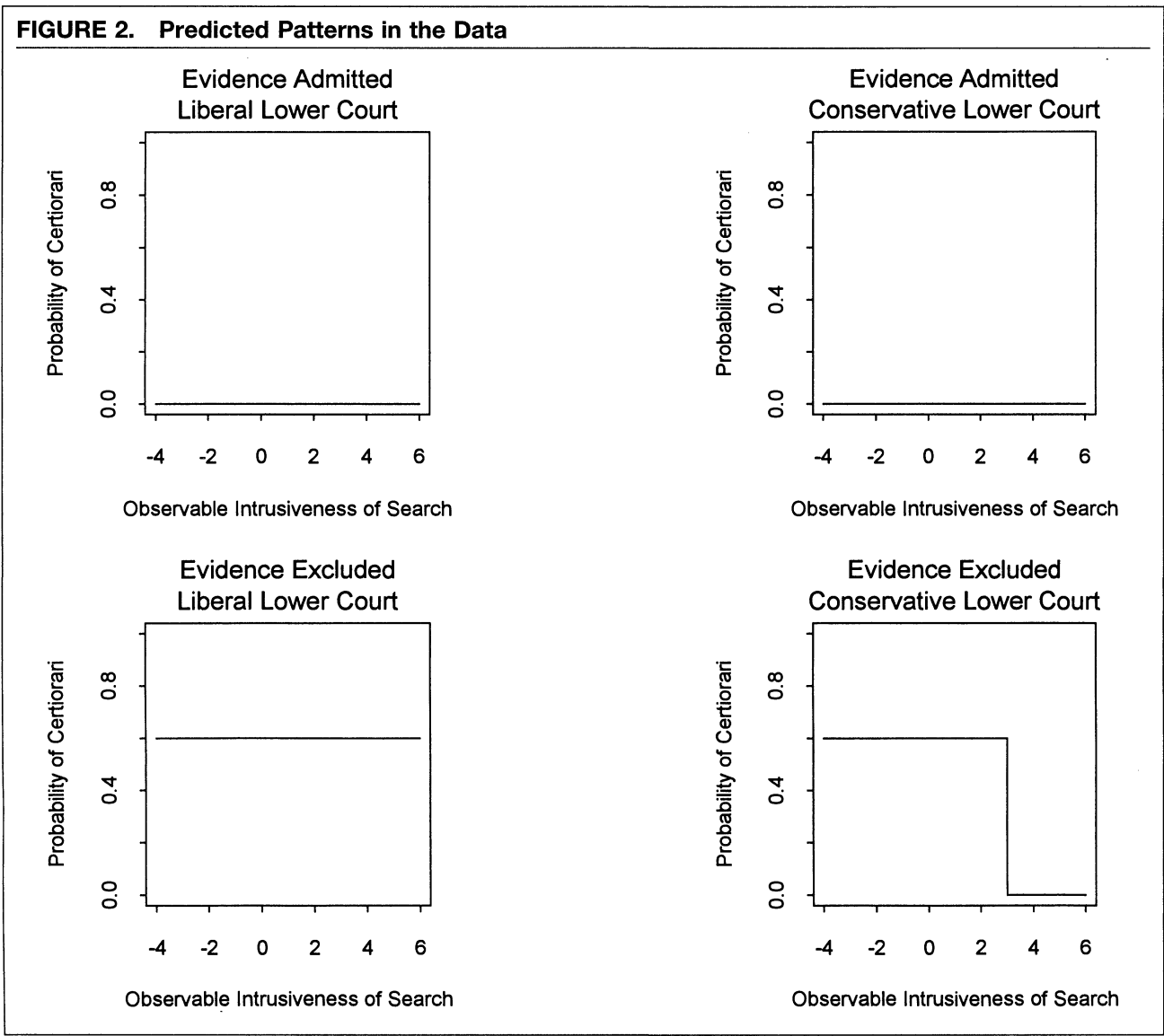
Figure 3 displays all the data in the form of the critical conditioning plot. Liberal lower courts are defined as those whose ideology score (*JI*) is at or below the median value (.152). To help uncover the systematic variation in the data without imposing any pattern *ex ante*, we include the fit from a nonparametric scatter plot smoother.¹² The outstanding patterns in the data are clear, namely, the radical difference between cases in which the evidence was admitted (the top row) and those in which the evidence was excluded (the bottom row). In the former, the probability of review is zero, except at the very highest levels of intrusiveness, where it appears to increase slightly. In the latter, the probability of review is never near zero, except for very intrusive searches struck down by conservative courts.

Figure 3 uncovers another predicted pattern: The probability of granting certiorari as a function of intrusiveness is flat (or slightly increasing at high levels of intrusiveness) when the lower court admits the evidence; flat when liberal courts exclude the evidence;

¹⁰ Because data on district court and retired judges sitting by designation were not available, more than one-third of our cases would have to be excluded if we were to use a composite measure of panel ideology. By focusing on the opinion writer, we cut missing data to below 10%. Moreover, it is not unrealistic to imagine that the Supreme Court cues off the opinion writer, especially as the justices' certiorari memos typically highlight his or her name.

¹¹ To test whether the Court's decisions remained stable under Chief Justice Burger, we added to the model dummy variables for each chief justice from Warren through Rehnquist (minus one for the excluded dummy) and an interaction for the presence of that chief justice and each passing term of the Court from 1962 through 1991. This, as we have previously shown, is the best measure of changing search-and-seizure doctrine between Courts. Nevertheless, if we look only within the Burger Court, we find that the coefficient for its interaction is 0.06, with a standard error of 0.05. Thus, we are reasonably confident in our assertion that under Burger, the Court's preferences remained stable.

¹² In each panel the smoothing curve is a locally weighted (loess) regression (span = 1) incorporating the Manski-Lerman weights. Other smoothers yield similar patterns, and the results are not particularly sensitive to the span of the regression. On these techniques, see Beck and Jackman 1998.



and strongly decreasing when conservative courts strike down a search. Finally, in the lower panels the estimated probabilities of granting certiorari—ranging from about 10% to more than 80%—are very high, in the context of known patterns about certiorari. For example, the overall probability of granting certiorari for search-and-seizure cases appealed to the Burger Court was 7%.¹³

Parametric Fitting: The Signaling Hypotheses

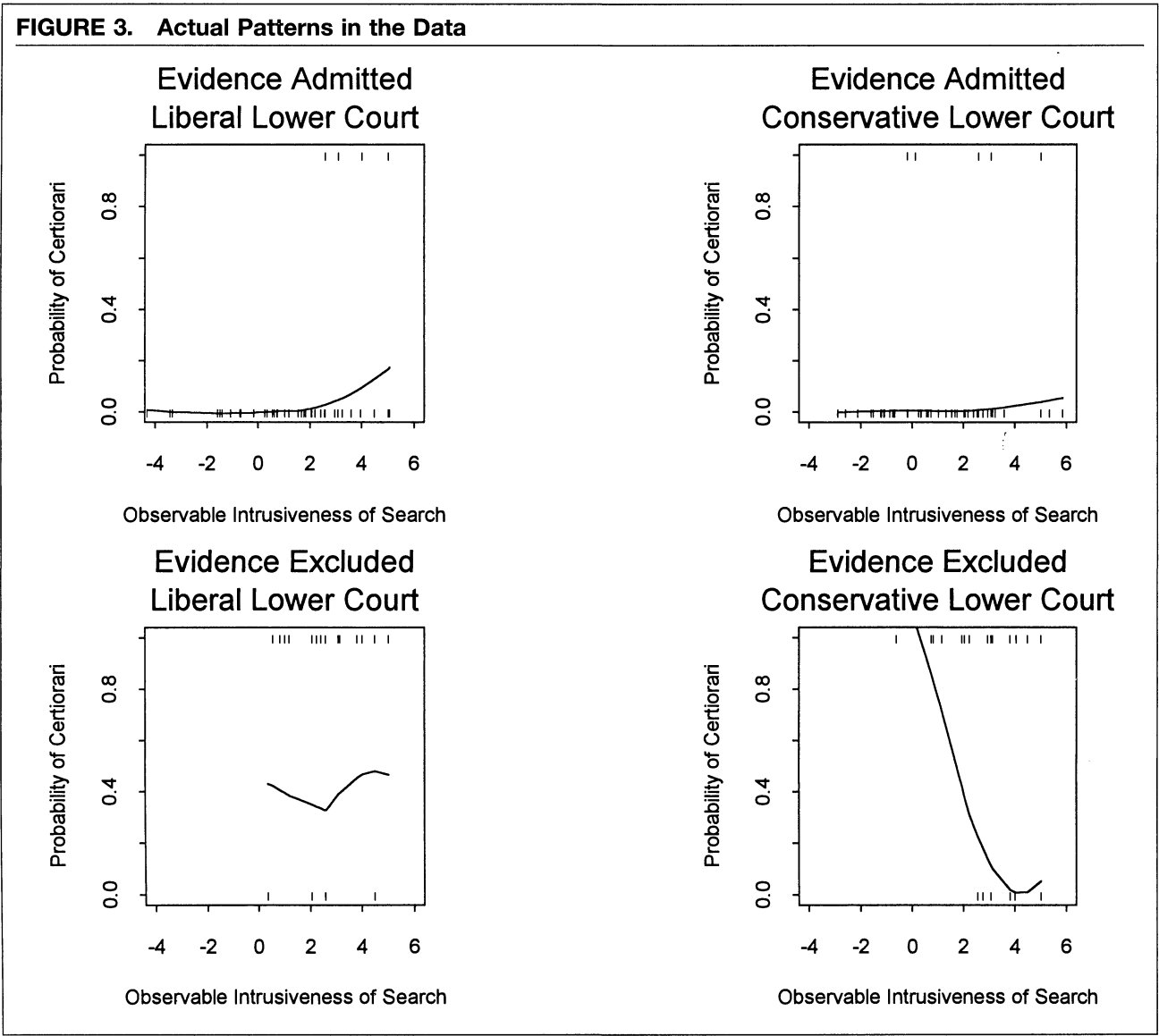
Bearing in mind that we model only one aspect of the certiorari process, Table 1 presents empirical tests of the theoretical model. As required by the theory, the statistical models are switching regime regressions (Goldfeld and Quandt 1973), which we estimate as logistic regressions. The regimes are the same shown in the critical conditioning plot in Figure 2. In models 1

and 2, the slope in the lower right-hand panel is estimated as a linear function of observed intrusiveness; in models 3 and 4 it is estimated as a step function.

In all four models, the coefficient on the variable admit is the intercept for cases in which the lower court admitted the evidence. Admit:intrude is the slope coefficient on the level of intrusiveness for the admitted cases. Across all four models, these variables have stable, statistically significant coefficients. It is straightforward to calculate the implied probabilities of granting certiorari, and these are shown in Figure 4 (using model 4). In the top two panels of that figure, the probability of the Supreme Court granting certiorari if the lower court admitted the evidence is essentially zero, except at the highest levels of intrusiveness, where it increases to about 15%.

The patterns predicted when the lower court excluded the evidence are more complex, and the models explore these patterns in some detail. In all four

¹³ Data compiled by the authors from *U.S. Law Week*.



models, the variable exclude is the intercept for cases in which the lower court excluded the evidence. In no model is this term statistically significantly different from zero, which indicates a baseline probability of granting certiorari for these cases of approximately 50%. Model 1 allows the slope on intrusiveness to vary linearly depending on the ideology of the lower court (as measured by *JI*, the ideology of the opinion writer). For liberal judges ($JI \leq .1$) the coefficient may take one value; for conservative ones ($JI > .1$) it may take another. As shown in Table 1, if the lower court was liberal, then the probability the Supreme Court would grant certiorari was not affected by the level of intrusiveness. But if the lower court was conservative, then the probability fell with the level of intrusiveness.¹⁴

¹⁴ Although our theoretical model explicitly calls for cut-points and regime changes, estimating the model with the original interval-level measurements for lower court ideology and the intrusiveness of the search does not change the results. Rerunning model 1 yields the following equation: $.046 (1.08) - 6.56 (1.68) \times Upheld - 0.19$

Model 2 is identical to model 1 but drops the statistically insignificant variable that allowed intrusiveness to affect the probability of certiorari for cases with evidence excluded by a liberal court. (The insignificance of this variable is compatible with the theory [see the lower left-hand panel in Figure 2].) Excluding this variable has only a very small effect on the overall fit of the model, as shown in Table 1 by the residual deviance. In model 2, the predicted behavior of the Supreme Court varies dramatically depending on the ideology of the lower court, given exclusion of the evidence. If the lower court was liberal ($JI \leq .1$), then the predicted probability of granting certiorari was slightly more than 50% regardless of how intrusive the search. If the lower court was conservative ($JI > .1$), then the probability declined dramatically with intru-

$(0.36) \times Intrude + 2.01 (2.23) \times Ideology + 1.13 (0.50) \times (Upheld \times Intrude) - 1.09 (0.71) \times (Ideology \times Intrude)$, with standard errors in parentheses.

TABLE 1. Probability of Granting Certiorari by Lower Court Decision and Intrusiveness of Search (Logit Models)				
Variable	Model 1	Model 2	Model 3	Model 4
Admit	−5.96 (1.27) (−4.7)	−5.96 (1.27) (−4.7)	−5.96 (1.27) (−4.7)	−5.96 (1.27) (−4.7)
Admit:intrude	0.73 (0.35) (2.1)	0.73 (0.35) (2.1)	0.73 (0.35) (2.1)	0.73 (0.35) (2.1)
Exclude	0.38 (1.03) (0.4)	0.08 (0.62) (0.1)	0.14 (1.06) (0.1)	−0.01 (0.62) (−0.02)
(Exclude & liberal):Intrude	−0.14 (0.38) (−0.4)		−0.07 (0.39) (−0.2)	
(Exclude & conserv.):intrude	−0.58 (0.36) (−1.6)	−0.49 (0.26) (−1.9)		
Step down (step at <i>intrude</i> = 1)			−1.72 (1.25) (−1.4)	−1.57 (0.92) (−1.7)
Residual deviance	61.8	61.9	62.9	62.9
Degrees of freedom	268	269	268	269

Note: In all cases, the dependent variable is the probability of granting certiorari. Standard error is given in first parenthesis, *t* value in second parenthesis.

siveness, falling to less than 10% for very intrusive searches.¹⁵

Models 3 and 4 more directly test the theoretical model by forcing the relationship in the lower right-hand panel of the conditioning plots in Figure 2 to take the form of a step function. We place the step at *intrude* = 1 (similar results obtain when the step is placed anywhere in the range 0–4). The step “down” from the intercept is given in Table 1. Model 3, like model 1, allows the probability of certiorari, given exclusion by a liberal lower court, to vary with the intrusiveness of the search. But as in model 1, this variable is not statistically significantly different from zero. Model 4 drops this term. Model 4 fits the data only negligibly worse than models 1 and 2 (residual deviance = 62.9).

The fit from model 4 is shown in Figure 4. As noted above, the predicted probability of granting certiorari for most upheld cases is zero and for liberal lower courts that excluded the evidence is about 50%. As shown in the lower left-hand panel of Figure 4, this constant probability can now be seen as the upper step in a step function. Among more conservative lower courts that excluded the evidence (the lower right-hand panel), the probability of review remains at about 50% for less intrusive searches (the upper step in the step function). For more intrusive searches, the probability falls to about 18% (the lower step in the step function).

¹⁵ The results are not sensitive to the exact location of the break between more liberal and less liberal lower courts, that is, the results are similar if the break falls anywhere between 0 and .3 on the ideology scale.

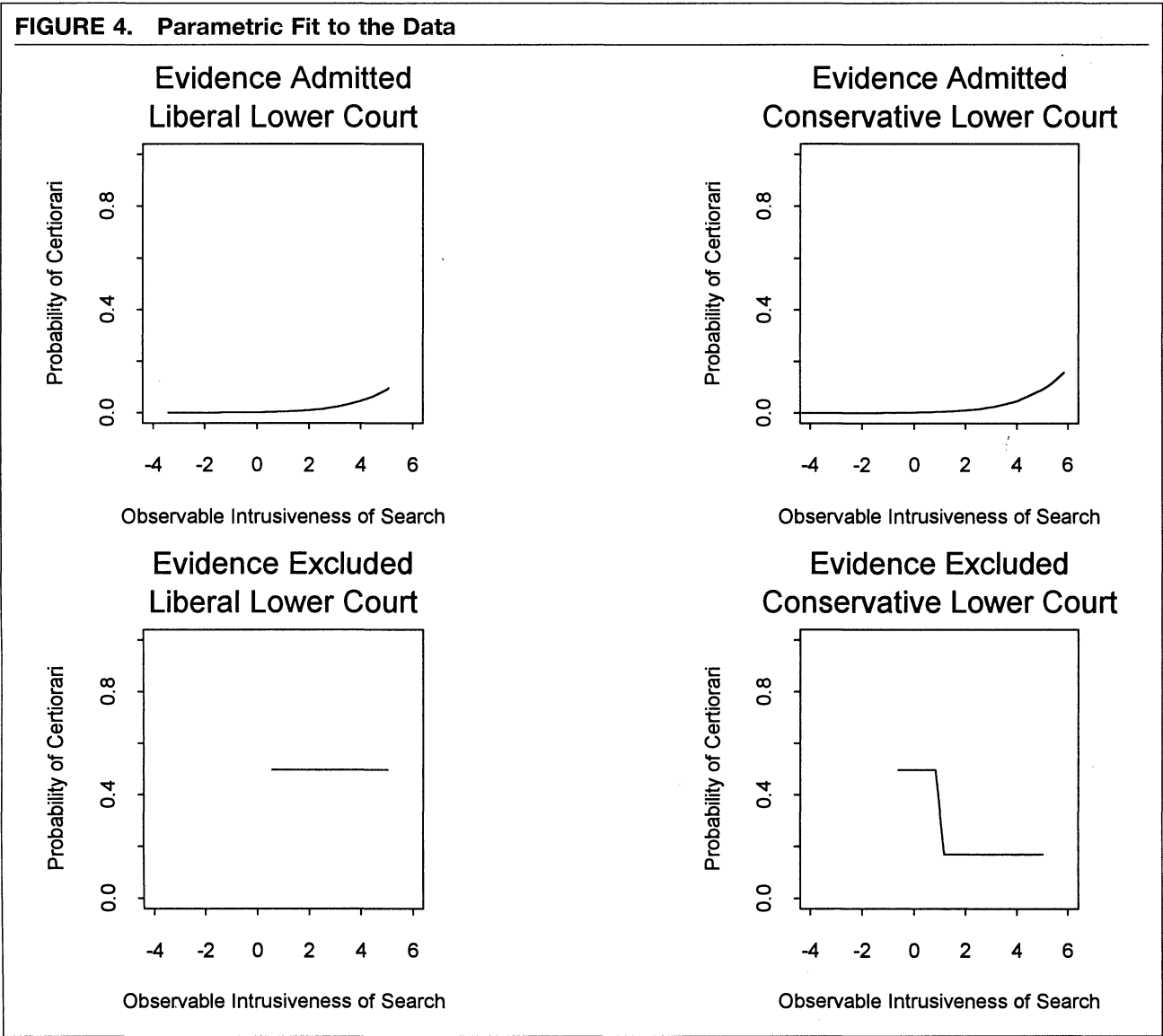
Discussion

The patterns revealed by the statistical analysis and shown in Figure 4 strongly resemble those predicted by the signaling model. The dramatic difference between the top and bottom rows in the figure confirms the “Nixon goes to China principle.” The more subtle predictions, involving the location of the step, are consistent with the lower panels.

Two discrepancies deserve brief discussion. First, among cases in which the lower court admitted the evidence, the probability of review rises slightly at the highest levels of intrusiveness. If some of the lower courts were more conservative than the Supreme Court, then the model predicts this should not completely escape review for very intrusive searches. The presence of some courts like this, or variables outside the model, could account for the slight increase in review probabilities at high levels of intrusiveness after admission of the evidence. Second, the lower step in the step function does not rest on zero but is located somewhat higher. If more liberal courts are mistakenly measured as less liberal, then this measurement error (as well as variables outside the model) could lead to an increase in the measured value for the lower step.

CONCLUSION

Our signaling theory of the Supreme Court’s certiorari decisions was tested against a random sample of cases heard in the federal appellate courts. The theory is an incomplete view of the process: It emphasizes the role of certiorari in enforcing the doctrinal preferences of



the Supreme Court throughout the judicial hierarchy but ignores its important role in the evolutionary creation of doctrine. The data are far from perfect. For example, we rely on fairly crude measures of judicial ideology and a relatively coarse coding of fact patterns in the cases. Despite these shortcomings, which ought to militate against finding patterns in the data, our analysis uncovers considerable support for the signaling theory as well as support for a variety of new propositions. Of course, we tested our model in only one area of decision making during the reign of one chief justice. Whether the data fit the model in other areas and during other regimes remains to be seen.

Our research makes three contributions. First, we use a formal model to advance our substantive understanding of certiorari. Most of the patterns uncovered here are new, despite the extensive research on the subject. We believe it unlikely that an analyst would uncover the patterns shown in figures 3 and 4 without a formal signaling model. Second, the research inte-

grates the study of the judicial hierarchy with the “new economics of organization” (Moe 1984). The application of formal principal-agent theory to the judicial hierarchy yields several benefits, both within the field of judicial politics and, more broadly, for the study of political hierarchies. Within the field of judicial politics, two phenomena, hitherto regarded as distinct, emerge as two components of a single underlying process that can be analyzed in a unified way. The two phenomena are judicial compliance (the responsiveness of lower courts to changes in the doctrinal directives of higher courts) and certiorari; both can be treated as interactive components in a political struggle over doctrine within the judicial hierarchy. From this perspective, the problem of doctrinal control in the judiciary is not sui generis but a particularly interesting case of an issue common to rule-governed hierarchies. To underscore the point, our model of strategic auditing, although tailored to the certiorari process, has broader applicability to other hierarchical settings in

which players have political preferences about rules. Third, the research provides empirical support for a formal model, which is still in relatively short supply.

APPENDIX

Proof of Proposition 1

Part 1 (\hat{x}^ does not exist).* Note that the case deals with $F(t_b) - F(t_a)/1 - F(t_a) < k \forall \hat{x}$. $F(t_b) - F(t_a)/1 - F(t_a) > k \forall \hat{x}$ cannot arise because $\lim_{\hat{x} \rightarrow \infty} F(t_b) - F(t_a)/1 - F(t_a) = 0$. This is true because $t_a, t_b \rightarrow -\infty$ as $\hat{x} \rightarrow \infty$. Thus, $F(t_a), F(t_b) \rightarrow 0$ and $1 - F(t_a) \rightarrow 1$ as $\hat{x} \rightarrow \infty$, and the quotient goes to zero. For any nonzero k there must exist a sufficiently large \hat{x} so that $F(t_b) - F(t_a)/1 - F(t_a) < k$. (A) Existence. If H does not review any cases, then the indicated strategy for L is clearly a best response. By assumption, $F(t_b) - F(t_a)/1 - F(t_a) < k \forall \hat{x}$ when $s(t, \hat{x}) = 1 \forall x \in [x_L, x_H]$. Then, from equation 2, $r(\text{exclude}, \hat{x}) = 0 \forall \hat{x}$. That $r(\text{admit}, \hat{x}) = 0 \forall \hat{x}$ is discussed in the text. (B) Uniqueness. There are three cases to consider, none of which can be an equilibrium. (1) The indicated strategy for L , but H reviews at least one case with probability greater than zero. By construction, H can profitably deviate to the indicated strategy. (2) The indicated strategy for H , but L deviates in some fashion from deciding cases according to its preferred rule. Then there are cases where L can profitably deviate to the indicated strategy. (3) H reviews some cases with probability greater than zero, and L deviates from deciding cases according to its preferred rule for at least one case in the conflict region (i.e., $\exists x \in [x_L, x_H]$ s.t. $s(t, \hat{x}) < 1$) (deviations from the indicated strategy for L are never profitable for cases below x_L or above x_H [inclusive] and need not be considered further). By construction, the Bayesian posterior $\mu(x_L \leq x \leq x_H, \text{exclude}, \hat{x}) < k$ when $s(t, \hat{x}) = 1 \forall x \in [x_L, x_H]$. But $\mu(x_L \leq x \leq x_H, \text{exclude}, \hat{x})$ decreases when $s(t, \hat{x}) < 1$ for some x in the conflict region. Thus, $\mu(x_L \leq x \leq x_H, \text{exclude}, \hat{x})$ remains less than k when $s(t, \hat{x}) < 1$ for some x in the conflict region, and therefore H can profitably deviate to the strategy indicated in the proposition.

Part 2 (\hat{x}^ exists).* By inspection, the strategies characterized in part 2 comprise best responses: $s^*(t, \hat{x})$ is compatible with $BR_L(r(m; \hat{x}), t; \hat{x})$, and $r^*(m_i, \hat{x})$ is compatible with equation 2 and the requirement that $r(m_2, \hat{x}) = 0 \forall \hat{x}$ (the latter is explained in the text). *Q.E.D.*

Proof of Hypotheses 1 and 2

Follows immediately from proposition 1. *Q.E.D.*

Proof of Hypothesis 3

(a) We consider the case in which $x_L < x_H$; via symmetry, similar arguments apply when $x_L > x_H$. Recall that \hat{x}^* is defined implicitly by equation 3. The partial derivative of equation 3 with respect to \hat{x}^* is

$$\frac{f(t_a)(1 - F(t_b)) - f(t_b)(1 - F(t_a))}{(F(t_a) - F(t_b))^2} \frac{k}{1 - k},$$

which will be nonzero provided

$$\frac{f(t_a)}{1 - F(t_a)} \neq \frac{f(t_b)}{1 - F(t_b)},$$

that is, the hazard rate at t_a is not equal to the hazard rate at t_b . A necessary condition for this is that $x_L \neq x_H$, which will also be sufficient provided F displays the monotone hazard

rate property (MHRP) (under MHRP, if $t' > t$, then the hazard rate at t' is greater than the hazard rate at t), as assumed. Then, via the implicit function theorem,

$$\frac{\partial \hat{x}^*}{\partial x_L} = - \frac{\partial s}{\partial x_L} / \frac{\partial s}{\partial \hat{x}^*} = \frac{f(t_a)[1 - F(t_b)]}{f(t_a)[1 - F(t_b)] - f(t_b)[1 - F(t_a)]}.$$

The numerator is positive. From the definitions of t_a and t_b , $x_H > x_L \Rightarrow t_b > t_a$, which, using MHRP, in turn implies the denominator is negative. (b) is obvious given (a). *Q.E.D.*

Proof of Hypothesis 4

The argument is similar to that of hypothesis 3. Again considering the case in which $x_L < x_H$, via the implicit function theorem,

$$\frac{\partial \hat{x}^*}{\partial x_H} = - \frac{\partial s}{\partial x_H} / \frac{\partial s}{\partial \hat{x}^*} = \frac{f(t_b)[1 - F(t_a)]}{f(t_b)[1 - F(t_a)] - f(t_a)[1 - F(t_b)]} > 0$$

using MHRP. The other case follows via symmetry. *Q.E.D.*

Proof of Hypothesis 5

As in hypotheses 4 and 5. For $x_L < x_H$

$$\begin{aligned} \frac{\partial \hat{x}^*}{\partial k} &= - \frac{\partial s}{\partial k} / \frac{\partial s}{\partial \hat{x}^*} \\ &= \frac{[1 - F(t_a)][F(t_b) - F(t_a)]}{-f(t_a)[1 - F(t_b)] - f(t_b)[k(1 - k) + F(t_a)]} < 0, \end{aligned}$$

as the denominator is negative (provided $0 < k \leq 1$), and the numerator is positive. The other case follows from symmetry. *Q.E.D.*

REFERENCES

- Armstrong, Virginia, and Charles A. Johnson. 1982. "Certiorari Decisions by the Warren and Burger Courts: Is Cue Theory Time Bound?" *Polity* 15 (1): 143–50.
- Andeoni, James, Brian Erard, and Jonathan Feinstein. 1998. "Tax Compliance." *Journal of Economic Literature* 36 (June): 818–60.
- Bagnoli, Mark, and Ted Bergstrom. 1989. "Log-Concave Probability and Its Applications." Working Paper 89-23 (October 1989). University of Michigan Center for Research on Economic and Social Theory.
- Banks, Jeffrey. 1991. *Signaling Games in Political Science*. Chur, Switzerland: Harwood Academic.
- Banks, Jeffrey, and Barry Weingast. 1992. "The Political Control of Bureaucracies under Asymmetric Information." *American Journal of Political Science* 36 (2): 509–24.
- Beck, Nathaniel, and Simon Jackman. 1998. "Beyond Linearity by Default: Generalized Additive Models." *American Journal of Political Science* 42 (2): 596–627.
- Brace, Paul, Melinda Gann Hall, and Laura Langer. 1998. "Measuring the Preferences of State Supreme Court Justices." Paper presented at the annual meeting of the Midwest Political Science Association, Chicago.
- Brenner, Saul. 1979. "The New Certiorari Game." *Journal of Politics* 41 (May): 649–55.
- Blasi, Victor. 1983. *The Burger Court: The Counter-Revolution That Wasn't*. New Haven, CT: Yale University Press.
- Caldeira, Gregory, and John Wright. 1988. "Organized Interests and Agenda Setting in the U.S. Supreme Court." *American Political Science Review* 82 (December): 1109–28.
- Calvert, Randall. 1985. "The Value of Biased Information: A Rational Choice Model of Political Advice." *Journal of Politics* 47 (May): 530–55.

- Cleveland, William S. 1993. *Visualizing Data*. Summit, NJ: Hobart.
- Epstein, Lee. 1991. "Courts and Interest Groups." In *The American Courts: A Critical Assessment*, ed. John B. Gates and Charles A. Johnson. Washington, DC: CQ Press. Pp. 335–72.
- Epstein, Lee, and Jack Knight. 1998. *The Choices Justices Make*. Washington, DC: CQ Press.
- Epstein, Lee, Jeffrey Segal, Harold Spaeth, and Thomas Walker. 1997. *The Supreme Court Compendium: Data, Decisions and Developments*. Washington, DC: CQ Press.
- Farrell, Joseph. 1993. "Meaning and Credibility in Cheap Talk Games." *Games and Economic Behavior* 5 (October): 514–31.
- Feeney, Floyd. 1975. "Conflicts Involving Federal Law: A Review of Cases Presented to the Supreme Court." In *Structures and Internal Procedures: Recommendations for Change*, ed. Commission of Revision of the Federal Court Appellate System. Washington, DC: Government Printing Office. Pp. 93–111.
- Giles, Micheal, Virginia Hettinger, and Todd C. Peppers. 1998. "Alternative Measures of Preferences for Judges of the Courts of Appeals." Paper presented at the annual meeting of the Midwest Political Science Association, Chicago.
- Goldfeld, Stephan M., and Richard E. Quandt. 1973. "The Estimation of Structural Shifts by Switching Regressions." *Annals of Economic and Social Measurement* 2 (October): 475–85.
- Greene, William H. 1991. *LIMDEP Version 6.0: User's Manual and Reference Guide*. Bellport, NY: Econometric Software.
- Hall, Melinda Gann. 1992. "Electoral Politics and Strategic Voting in State Supreme Courts." *Journal of Politics* 54 (May): 427–46.
- Jervis, Robert. 1970. *The Logic of Images in International Relations*. Princeton, NJ: Princeton University Press.
- Kornhauser, Lewis. 1992. "Modeling Collegial Courts II: Legal Doctrine." *Journal of Law, Economics, and Organization* 8 (October): 441–70.
- Levi, Edward H. 1948. *An Introduction to Legal Reasoning*. Chicago: University of Chicago Press.
- Manski, Charles F., and Steven R. Lerman. 1977. "The Estimation of Choice Probabilities from Choice Based Samples." *Econometrica* 45 (8): 1977–88.
- Moe, Terry. 1984. "The New Economics of Organization." *American Journal of Political Science* 28 (November): 739–77.
- Perry, H. W., Jr. 1991. *Deciding to Decide: Agenda Setting in the United States Supreme Court*. Cambridge, MA: Harvard University Press.
- Schubert, Glendon A. 1959. *Quantitative Analysis of Judicial Behavior*. Glencoe, IL: Free Press.
- Segal, Jeffrey A. 1984. "Predicting Supreme Court Decisions Probabilistically: The Search and Seizure Cases, 1962–1981." *American Political Science Review* 78 (December): 891–900.
- Segal, Jeffrey, and Harold Spaeth. 1993. *The Supreme Court and the Attitudinal Model*. New York: Cambridge University Press.
- Songer, Donald R. 1979. "Concern for Policy Outputs as a Cue for Supreme Court Decisions on Certiorari." *Journal of Politics* 41 (November): 1185–94.
- Songer, Donald, Jeffrey Segal, and Charles Cameron. 1994. "The Hierarchy of Justice: Testing a Principal-Agent Model of Supreme Court-Circuit Court Interactions." *American Journal of Political Science* 38 (August): 673–96.
- Tanenhause, Joseph, Marvin Schick, Matthew Muraskin, and Daniel Rosen. 1963. "The Supreme Court's Certiorari Jurisdiction: Cue Theory." In *Judicial Decision Making*, ed. Glendon Schubert. Glencoe, IL: Free Press. Pp. 111–32.
- Tate, C. Neal, and Roger Handberg. 1991. "Time Binding and Theory Building in Personal Attribute Models of Supreme Court Voting Behavior, 1916–88." *American Journal of Political Science* 35 (May): 460–80.
- Teger, Stuart, and Douglas Kosinski. 1980. "The Cue Theory of Supreme Court Certiorari Jurisdiction: A Reconsideration." *Journal of Politics* 42 (August): 834–46.
- Ulmer, Sidney S. 1984. "The Supreme Court's Certiorari Decisions: Conflict as a Predictive Variable." *American Political Science Review* 78 (December): 901–11.