LOWER COURT DISCRETION

PAULINE T. KIM*

Empirical scholars typically model the judicial hierarchy in terms of a principal-agent relationship in which the Supreme Court, the principal, sets policy and the lower federal courts, as agents, must faithfully implement that policy. The law is a signal—the means by which the Court communicates its preferences. This Article argues instead for recognizing the law as an independent normative force. Empirical scholars fail to take seriously the role of law because they reject as implausible formalistic accounts of its operation. This Article advances a more nuanced account of how law shapes the decisionmaking environment of the lower federal courts, one that focuses on the presence of discretion. It explores how different types of discretion afford distinct types of power over lawmaking and case outcomes, and how that discretionary power is allocated between district and appellate courts. Paying attention to discretion suggests features of the judicial hierarchy that are commonly overlooked in principal-agent models. For example, judges' goals, and therefore their strategies, will vary depending upon whether they seek to influence law development or merely to shape case outcomes. The Article also questions the normative assumption, implicit in principal-agent models, that lower federal courts should decide cases in accordance with the policy preferences of the Supreme Court. Because judges inevitably have discretion when applying the law, a norm of compliance with superior court precedent does not necessarily require lower courts to follow the policy preferences of the Supreme Court. The reasons judicial discretion exists, such as allocating power within the judicial hierarchy, may argue against such a centralization of power in the Supreme Court.

 Professor of Law, Washington University School of Law, St. Louis. My thanks to Nancy Staudt, Sam Bagenstos, Margo Schlanger, Barry Friedman, Barbara Flagg, and Ron Levin, who provided helpful feedback on earlier drafts of this Article. I also benefited from the comments of participants at Washington University's Workshop on Empirical Research in the Law (WERL) and faculty workshops at Washington University School of Law and Duke Law School. Tom Clark and James Bloom provided invaluable research assistance.

INTRODUCTION

Court decisionmaking has long been a subject of study for both legal theorists and political scientists. Despite the common focus of their efforts, several quite distinct approaches have developed. The dominant approach in traditional legal scholarship has been normative, with scholars asking how judges *should* decide important questions of law. Even when this work has taken a positive turn by asking how judges *in fact* decide cases, the work has emphasized the importance of doctrine, focusing on court opinions as evidence of the reasons for their decisions. By contrast, social scientists who study the courts have usually emphasized positive explanations of court behavior and focused their attention on case outcomes rather than written opinions. One approach, dominant among political scientists for a time, relies on a social-psychological paradigm. Often referred to as the "attitudinal" model, it sees judicial decisionmaking as determined by the attitudes or preferences of individual judges, whose votes in particular cases reflect their sincere policy preferences largely unconstrained by legal precedent.[1] More prominent in recent years are positive political theories that emphasize strategic interactions among judges and between judges and other political actors.[2] This approach shares the attitudinalists' assumption that judges seek to advance their policy preferences; however, it posits that in doing so, they act strategically, taking account of the likely response of other

[1] *See generally* JEFFREY A. SEGAL & HAROLD J. SPAETH, THE SUPREME COURT AND THE ATTITUDINAL MODEL (1993) (elaborating and defending "attitudinal model," which holds that Supreme Court decisions reflect ideological attitudes and values of Justices).

[2] *See generally* Frank B. Cross & Emerson H. Tiller, *Judicial Partisanship and Obedience to Legal Doctrine: Whistleblowing on the Federal Courts of Appeals*, 107 YALE L.J. 2155 (1998) (arguing that policy-oriented court of appeals judges follow legal doctrine strategically in order to avoid reversal); William N. Eskridge, Jr., *Overriding Supreme Court Statutory Interpretation Decisions*, 101 YALE L.J. 331 (1991) (describing establishment of statutory policy as dynamic game among Court, Congress, and President in which each tries to impose its policy preferences in light of expected responses of other players); McNollgast, *Politics and the Courts: A Positive Theory of Judicial Doctrine and the Rule of Law*, 68 S. CAL. L. REV. 1631 (1995) (modeling judicial decisionmaking as product of strategic interactions between upper and lower courts); Donald R. Songer, Jeffrey A. Segal & Charles M. Cameron, *The Hierarchy of Justice: Testing a Principal-Agent Model of Supreme Court-Circuit Court Interactions*, 38 AM. J. POL. SCI. 673 (1994) (same).

actors and the institutional context in which they operate. Although this approach is more closely associated with political science, a number of legal scholars have also embraced strategic accounts of judicial decisionmaking.

Scholars have undertaken empirical studies to test these competing theories of decisionmaking.[3] With an emphasis on quantitative analysis, these efforts have produced a great deal of data about what judges do—or at least how they vote—in a wide variety of situations. What these large-scale studies of judicial decisionmaking generally lack, however, is a satisfactory account of the law. Typically, they measure judicial output in terms of ideological outcome, focusing on whether a case was decided in a liberal or conservative direction and largely ignoring the legal reasoning set forth in the written opinion. For scholars ascribing to social-psychological explanations, the lack of attention to law is not surprising, given that their theoretical approach presumes that law is irrelevant. Positive political theorists have been more willing to recognize law as operative, but they account for it in instrumental terms. For example, a recent study sought to demonstrate that judicial adherence to legal principles is "explained by . . . judges' concerns with the external policy effects of their rulings."[4] In other words, the theory goes, "judges care about precedent *because* they care about policy."[5] On this view, the law is merely a cipher—a stand-in for the *real* motivations behind a decision.

This Article challenges that perspective, arguing that the law has independent normative force that cannot be reduced to purely strategic explanations. In other words, it takes the position that law matters in its own right and that both theoretical and empirical efforts to understand how judges make decisions will be enhanced by paying more attention to legal doctrine and legal norms. Other scholars have

[3] *See generally* Gregory A. Caldeira, John R. Wright & Christopher J.W. Zorn, *Sophisticated Voting and Gate-Keeping in the Supreme Court*, 15 J.L. ECON. & ORG. 549 (1999) (empirically testing whether Supreme Court Justices engage in strategic voting in certiorari decisions); Forrest Maltzman & Paul J. Wahlbeck, *Strategic Policy Considerations and Voting Fluidity on the Burger Court*, 90 AM. POL. SCI. REV. 581 (1996) (empirically testing whether Supreme Court Justices act strategically in changing their votes between initial conference and final vote); Richard L. Revesz, *Environmental Regulation, Ideology, and the D.C. Circuit*, 83 VA. L. REV. 1717 (1997) (empirically testing whether D.C. Circuit judges are strategically ideological when reviewing environmental regulations); Songer et al., *supra* note 2 (empirically testing extent of courts of appeals' responsiveness to and congruence with Supreme Court policy); Paul J. Wahlbeck, James F. Spriggs II & Forrest Maltzman, *Marshalling the Court: Bargaining and Accommodation on the United States Supreme Court*, 42 AM. J. POL. SCI. 294 (1998) (examining Supreme Court draft opinions to test empirically whether they are written strategically). R

[4] Ethan Bueno de Mesquita & Matthew Stephenson, *Informative Precedent and Intrajudicial Communication*, 96 AM. POL. SCI. REV. 755, 755 (2002).

[5] *Id.*

critiqued the body of positive scholarship on court decisionmaking along similar lines. Barry Friedman recently argued that "positive scholars need to pay greater attention to the norms of law, i.e., how law and legal institutions operate."[6] Similarly, Tiller and Cross have asserted that "the nature and effect of legal doctrine has been woefully understudied"[7] and have called for "greater attention to the core elements of legal analysis."[8] These scholars have made various suggestions to improve research on court decisionmaking: for example, Friedman's argument that positive scholars should pay more attention to legal norms and institutions,[9] or Tiller and Cross's suggestion that scholars need to identify the "basic decision structures" judges use in their opinions.[10] My approach here is consistent with some of the suggestions made by these scholars; however, this Article goes beyond the general critique that law is undertheorized in positive accounts of the law. It draws on jurisprudential insights about the nature of law in order to develop a theory about how law shapes the decisionmaking environment of lower federal court judges[11]—one that I hope will be useful in shaping future empirical investigations.

In considering how law impacts lower courts, I take particular aim at the principal-agent model of the judicial hierarchy on which most of the existing theoretical and empirical work on lower courts is based.[12] Positive political theorists commonly accept that the role of lower courts is best understood in terms of a principal-agent relationship.[13] Such a model conceives of the Supreme Court as the principal,

---

[6] Barry Friedman, *Taking Law Seriously*, 4 PERSP. ON POL. 261, 262 (2006).

[7] Emerson H. Tiller & Frank B. Cross, *What is Legal Doctrine?*, 100 NW. U. L. REV. 517, 517 (2006).

[8] *Id.* at 532.

[9] *See* Friedman, *supra* note 6, at 266. R

[10] *See* Tiller & Cross, *supra* note 7, at 528. R

[11] Although the Supreme Court is obviously an important institution, the lower federal courts—by virtue of their numbers and the Supreme Court's limited resources for reviewing their decisions—are the most immediately relevant actors for the overwhelming majority of federal court litigants and deserve study in their own right. The relationships between state supreme courts and the U.S. Supreme Court, and between courts at various levels of a state court hierarchy, bring different institutional considerations to bear and are therefore beyond the scope of this Article.

[12] A recent study empirically tested both a principal-agent model and a "team theory" of the interaction between the Supreme Court and the federal courts of appeals. Chad Westerland, Jeffrey A. Segal, Lee Epstein, Scott Comparato & Charles M. Cameron, Lower Court Defiance of (Compliance with) the U.S. Supreme Court (Apr. 9, 2006) (unpublished manuscript, on file with the *New York University Law Review*). The team theory, briefly described in note 33, *infra*, has not been subjected to much systematic empirical testing and is not the focus of this Article. R

[13] *See, e.g.*, Bueno de Mesquita & Stephenson, *supra* note 4, at 757 (modeling trial judge as faithful agent of appellate court); Charles M. Cameron, Jeffrey A. Segal & Donald Songer, *Strategic Auditing in a Political Hierarchy: An Informational Model of the Supreme* R

enunciating policy but delegating the work of implementation to its agents, the lower federal courts. Law, in the form of court precedents, is a signal—the means by which the Supreme Court communicates its preferences to the lower courts, which are then expected to follow that doctrine. The principal-agent model focuses attention on the possibility that the interests of the principal and the agent will diverge and highlights the resulting need for mechanisms of supervision and control by the principal. Theorists commonly assume that the Supreme Court exercises such control through its power of reversal.[14] This explanation is insufficient, however, given the extremely small percentage of cases ultimately subject to Supreme Court review.[15] More importantly, a purely strategic account of law fails to take seriously internal perspectives—those of judges and lawyers who participate in the system—on what law is and how it influences court decisionmaking. Participants report that the law has independent normative force: Legal rules influence how cases come out, even though they may not determine the result in all cases.[16]

The failure to develop a more robust account of the law may have resulted from positive scholars' predominant emphasis on the Supreme Court, which operates in a unique institutional setting. The relationship of lower federal courts to law and precedent is quite distinct from that of the Supreme Court because of their different positions in the judicial hierarchy. Although the principle of stare decisis constrains the Court, it is not an "inexorable command,"[17] and the

Court's Certiorari Decisions, 94 AM. POL. SCI. REV. 101, 103–04 (2000) (using simplified model of upper and lower courts to explore principal-agent conflicts over policy); Tracey E. George & Albert H. Yoon, *The Federal Court System: A Principal-Agent Perspective*, 47 ST. LOUIS U. L.J. 819, 822 (2003) (asserting that Supreme Court and lower federal courts are in agency relationship); McNollgast, *supra* note 2, at 1632 (utilizing approach that "focuses on the competition and conflict that arise between higher and lower courts"); Songer et al., *supra* note 2, at 675 (modeling relationship between Supreme Court and lower courts in terms of principal-agent relationship). R R

14 *See infra* note 50 and accompanying text. R

15 *See infra* note 30 and accompanying text. R

16 *See* Jon O. Newman, *Between Legal Realism and Neutral Principles: The Legitimacy of Institutional Values*, 72 CAL. L. REV. 200, 208–14 (1984) (asserting that judges' views of courts, other sources of law, federalism, and adjudicatory process itself are at least as relevant to decisionmaking as are policy preferences); Patricia M. Wald, *A Response to Tiller and Cross*, 99 COLUM. L. REV. 235, 236–39 (1999) (commenting from her experience as court of appeals judge that law does have independent normative force beyond judge's personal policy preferences). In addition, Judge Harry Edwards has emphasized the importance of collegiality and deliberation about the law in shaping outcomes. *See* Harry T. Edwards, *Collegiality and Decision Making on the D.C. Circuit*, 84 VA. L. REV. 1335, 1335 (1998).

17 Planned Parenthood v. Casey, 505 U.S. 833, 854 (1992).

Court has on occasion overturned its prior precedents.[18] The courts subordinate to the Supreme Court, however, are subject to an absolute duty to follow its precedents.[19] Federal district courts are further required to follow the precedents of the court of appeals of the circuit in which they sit.[20] Thus, how law impacts a judge's decisionmaking process will depend very much upon the particular court on which she sits and where that court is located in the judicial hierarchy.

Despite the demand of hierarchical precedent, lower federal courts retain a substantial amount of discretion when deciding cases. By "discretion" I mean to indicate situations in which a judge is required to exercise judgment because the outcome of a case is not fully determined by existing legal materials. Even the judge who understands legal rules to be obligatory and faithfully attempts to follow precedent will find that she has the power to exercise choice in deciding a case. To some extent that discretion exists because it is unavoidable—legal language is at some point irredeemably indeterminate. But discretion may reflect certain value trade-offs as well: choosing flexibility over certainty by selecting a standard rather than a bright-line rule; or allocating certain powers to trial courts, rather than appellate courts, by establishing a deferential standard of review. Recognizing the presence and nature of that discretion is key to understanding how law shapes the decisionmaking environment of lower court judges and ultimately to developing better theoretical models of the judicial hierarchy.

The idea that judges have discretion is, of course, not new. Legal theorists have long argued that doctrine cannot fully determine out-

---

[18] *E.g.*, Garcia v. San Antonio Metro. Transit Auth., 469 U.S. 528, 531 (1985) (overruling Nat'l League of Cities v. Usery, 426 U.S. 833 (1976)); Gideon v. Wainwright, 372 U.S. 335, 339 (1963) (overruling Betts v. Brady, 316 U.S. 455 (1942)); *see also* LEE EPSTEIN, JEFFREY A. SEGAL, HAROLD J. SPAETH & THOMAS G. WALKER, THE SUPREME COURT COMPENDIUM: DATA, DECISIONS, AND DEVELOPMENTS 208 tbl.2-17 (4th ed. 2007) (listing Supreme Court cases overruled in subsequent opinions).

[19] *See* 18 JAMES WM. MOORE ET AL., MOORE'S FEDERAL PRACTICE § 134.02[2] (3d ed. 2006) (explaining that lower federal courts owe obedience to decisions of Supreme Court on questions of federal law); Evan H. Caminker, *Precedent and Prediction: The Forward-Looking Aspects of Inferior Court Decisionmaking*, 73 TEX. L. REV. 1, 12 (1994) [hereinafter Caminker, *Precedent and Prediction*] ("It is axiomatic that an inferior court must respect prior precedents created by its superior courts."). As the Supreme Court has said, "unless we wish anarchy to prevail within the federal judicial system, a precedent of this Court must be followed by the lower federal courts no matter how misguided the judges of those courts may think it to be." Hutto v. Davis, 454 U.S. 370, 375 (1982). For an exposition of the normative justifications for the duty to obey superior court precedent, see generally Evan H. Caminker, *Why Must Inferior Courts Obey Superior Court Precedents?*, 46 STAN. L. REV. 817 (1994).

[20] *See* 18 MOORE ET AL., *supra* note 19, § 134.02[2] (stating that district courts owe obedience to decisions of court of appeals in that circuit). R

comes and have debated what judges *should* do when called upon to decide a case in which prior authority is inconclusive. These discussions, however, are often carried on in rather abstract terms, theorizing about the actions of the single judge, acting alone and wholly apart from any institutional context. For example, Ronald Dworkin builds his theory of how judges decide hard cases through the example of "Hercules," "a lawyer of superhuman skill, learning, patience and acumen" who is "a judge in some representative American jurisdiction."[21] We do not know whether Hercules has colleagues, whether he votes as part of a multi-member court, or whether he has inferior courts he supervises or superior courts to which he must answer.[22] An institutional perspective suggests that each of these factors is important: We cannot begin to understand how Hercules, or any ordinary judge, makes decisions without paying attention to the context in which he decides.[23]

By contrast, the principal-agent model highlights the importance of the institutional context but neglects the discretionary spaces created by law. Scholars utilizing a principal-agent model have not examined closely the concept of discretion or how it fits with various theories of law.[24] To the extent that they write about discretion at all, they tend to use the term as a pejorative, suggesting that the exercise of discretion by lower courts is a form of shirking[25] or that the existence of any judicial discretion evidences a failure of law.[26] In doing so, they implicitly embrace formalist understandings, assuming that law can and should determine the outcome in every case. This approach ignores much recent jurisprudential thought and the wide-

[21] RONALD DWORKIN, TAKING RIGHTS SERIOUSLY 105 (1977). Sanford Levinson has argued that Dworkin's vision of "Hercules" tells us little about the appropriate role of the lower court judge in interpreting the Constitution. Sanford Levinson, *On Positivism and Potted Plants: "Inferior" Judges and the Task of Constitutional Interpretation*, 25 CONN. L. REV. 843, 844–45 (1993). He criticizes "a jurisprudential tradition that too often has looked only at a single abstract figure called 'the judge' rather than at the multiplicity of figures who, although all called 'judge,' nonetheless play significantly different roles and engage in significantly different practices when they identify and interpret 'the law.'" *Id.* at 852.

[22] We do know, however, that he is male.

[23] In a later work, Dworkin acknowledges that real judges must worry about things like docket pressures and how to convince colleagues to join an opinion. He continues to use the example of Hercules precisely in order "to abstract from these practical issues." RONALD DWORKIN, LAW'S EMPIRE 380–81 (1986).

[24] A recent exception is work by Jacobi and Tiller which models the decisions of upper courts when choosing between doctrines that allow differing levels of discretion to lower courts. Tonja Jacobi & Emerson H. Tiller, Legal Doctrine and Political Control (June 21, 2005) (unpublished manuscript, on file with the *New York University Law Review*).

[25] *See, e.g.*, Songer et al., *supra* note 2, at 674. R

[26] *See* SEGAL & SPAETH, *supra* note 1, at 33. R

spread understanding that discretion is inherent in the task of applying the law in particular cases.

Application of the principal-agent model to the judicial hierarchy entails some implicit normative assumptions as well. One of the central concerns of principal-agent theories is the "need for mechanisms of control."[27] Implicit in this concern is a normative claim that control over the judicial agents, the lower federal courts, appropriately rests in the principal, the Supreme Court. While it is widely accepted that judges have a duty to follow the *precedents* of a superior court, the claim that lower courts should also conform to the policy *preferences* of the Supreme Court, even when not expressed in binding decisional law, is quite distinct and open to serious question. Because the principal-agent model purports to be a positive description of upper and lower court interactions, these underlying normative assumptions have gone largely unexamined. However, paying attention to legal discretion—and the reasons for its presence—draws into question the implicit assumption that lower courts should strive to implement the Supreme Court's policy preferences.

This Article proceeds as follows: Part I begins by briefly introducing principal-agent theories of the judicial hierarchy and reviewing the empirical literature on lower court decisionmaking. It then examines—and rejects—the explanation common to principal-agent models that lower court judges follow superior court precedent because they fear reversal, and urges instead that models of decisionmaking include judges' legal preferences. Part II explores the concept of discretion, arguing that it should be understood as inherent in the operation of law and legal norms, rather than evidence of their failure. It then examines the different contexts in which lower court judges can exercise discretion consistent with legal rules and the different types of power they afford: power over case outcomes, law development, or both. Part III considers how paying attention to the nature and types of judicial discretion alters the principal-agent model commonly used to frame empirical studies of lower court behavior. Understanding discretion as an inherent part of the legal system elucidates features of the interaction between upper and lower courts that are overlooked in the principal-agent model. In addition, Part III highlights that model's implicit normative assumption that lower federal courts have a duty to follow the preferences—not merely the precedents—of the Supreme Court, and questions whether the resulting centralization of power in the hands of the Justices is normatively desirable.

[27] George & Yoon, *supra* note 13, at 821. **R**

I
LOWER COURT DECISIONMAKING

A. The Judicial Hierarchy

Understanding how lower court judges make decisions requires attention to the institutional context in which they operate. Within the federal system there are various types of lower courts, each situated differently within the judicial hierarchy.[28] Federal district court judges generally hear cases sitting alone. As trial judges, they hear evidence and manage the course of litigation in addition to deciding legal questions. If appealed by a party, final decisions (and those subject to interlocutory appeal) are reviewable by the court of appeals of the circuit in which they sit. Federal courts of appeals are collegial courts: Appellate judges generally hear cases in panels of three. Panel decisions may be overturned by the entire court of appeals sitting en banc or by the Supreme Court. En banc decisions are relatively rare,[29] however, and Supreme Court review is discretionary. Because the Supreme Court currently hears only about ninety cases per year, more than ninety-nine percent of federal court of appeals decisions are effectively the final decision in a case.[30] Thus, federal district courts and circuit courts operate in environments quite different from one another and from that of the Supreme Court, and these differences likely affect how decisions are made.

In addition, certain legal norms govern the work of judges at every level of the hierarchy. By "norms," I refer to rules—both formal and informal—which are understood by actors within the system as creating obligations to act or not act in a certain way. Thus broadly defined, legal norms include the substantive legal rules found in the precedential opinions, statutes, and regulatory provisions that

---

[28] In this Article, I consider only the federal district courts and circuit courts of appeals, leaving aside other types of federal judges, such as magistrate judges and Tax Court judges.

[29] *See, e.g.*, Tracey E. George, *Developing a Positive Theory of Decisionmaking on U.S. Courts of Appeals*, 58 OHIO ST. L.J. 1635, 1670 (1998) (finding that Fourth Circuit heard, on average, 7.8 cases en banc each year between 1962 and 1996); Tracey E. George, *The Dynamics and Determinants of the Decision to Grant En Banc Review*, 74 WASH. L. REV. 213, 250 n.171 (1999) (reporting that Second, Fourth, and Eighth Circuits decided 90, 284, and 151 en banc cases, respectively, between 1956 and 1996).

[30] For example, in 2005, the federal courts of appeals terminated 61,975 cases, *see* ADMIN. OFFICE OF THE U.S. COURTS, JUDICIAL BUSINESS OF THE UNITED STATES COURTS 2005, at 101 tbl.B, http://www.uscourts.gov/judbus2005/appendices/b0.pdf, while the Supreme Court decided only eighty-seven cases in the 2005 Term. Supreme Court of the United States, 2005 Term Opinions of the Court, http://www.supremecourtus.gov/opinions/05slipopinion.html (last visited Mar. 9, 2007). Even assuming that all cases heard by the Supreme Court come from the federal courts of appeals (though in fact the Court's docket every year includes appeals from state courts), the chance that a given court of appeals decision will be reviewed by the Supreme Court is approximately 0.14%.

judges are expected to apply to the facts of the cases before them. They also include a wide variety of both formal rules and informal practices that govern how judges should go about the task of deciding cases. These norms tell judges such things as which materials are relevant to consider, how they should treat the opinions of judicial colleagues above and below them in the hierarchy, when reasons are required to explain a decision, and what kinds of reasons are authoritative. Some of these norms—such as the rule of strict hierarchical precedent—are well established; others are highly contested—for example, the norm regarding the use of foreign legal materials.[31] Because at least some of the relevant legal norms depend upon a judge's place in the judicial hierarchy,[32] taking account of institutional context is crucial to understanding how judges decide.

Positive political theorists commonly use principal-agent models to describe the judicial hierarchy,[33] assuming that the Supreme Court is the principal and the lower federal courts are the agents. These models posit that judges have policy preferences and are motivated to decide cases in a manner consistent with those preferences. When the preferences of lower court judges diverge from those of the Supreme

[31] *See, e.g.*, Steven G. Calabresi & Stephanie Dotson Zimdahl, *The Supreme Court and Foreign Sources of Law: Two Hundred Years of Practice and the Juvenile Death Penalty Decision*, 47 WM. & MARY L. REV. 743, 743–56 (2005) (describing controversy over citation to foreign legal materials in Court opinions).

[32] In addition, some norms relating to the operation of the court—as opposed to the judge's decisionmaking process—may vary across courts at the same level of the judicial hierarchy. For example, practices regarding publication of opinions vary widely across circuits and from one district court to another. *See* Stephen J. Choi & G. Mitu Gulati, *Choosing the Next Supreme Court Justice: An Empirical Ranking of Judge Performance*, 78 S. CAL. L. REV. 23, 45 (2004) ("Statistically significant variation . . . exists across the circuits in terms of number of total published opinions."); Peter Siegelman & John H. Donohue III, *Studying the Iceberg from its Tip: A Comparison of Published and Unpublished Employment Discrimination Cases*, 24 LAW & SOC'Y REV. 1133, 1144 (1990) ("[P]ublication rates for employment discrimination cases vary dramatically across districts.").

[33] Lewis Kornhauser has suggested an alternative model—the team model—to describe the judicial hierarchy. Lewis A. Kornhauser, *Adjudication by a Resource-Constrained Team: Hierarchy and Precedent in a Judicial System*, 68 S. CAL. L. REV. 1605 (1995). In his model, judges are assumed to share a common goal of maximizing "the expected number of 'correct' answers." *Id*. at 1606. Regarding the judiciary as a resource-constrained team, the model suggests that maximizing the expected number of correct decisions will produce certain organizational features, such as a hierarchical court system and a principle of strict adherence to vertical precedent. *Id.* at 1628. Similarly, Steven Shavell, although not using the metaphor of a team, has focused on "error correction" as an explanation for why a system of adjudication would include an appeals process. Steven Shavell, *The Appeals Process as a Means of Error Correction*, 24 J. LEGAL STUD. 379 (1995). While useful for understanding why certain institutional structures might emerge, these theories are less helpful in explaining how, given the existing institutional structure of the court system, trial and intermediate appellate courts decide cases.

Court, value conflicts arise, raising the possibility of noncompliance by lower courts. Songer, Segal, and Cameron describe the model:

> The Supreme Court is the principal, whose subordinates, the courts of appeals, are the agents. If the circuit courts consisted of faithful agents, they would obediently follow the policy dictates set down by the Supreme Court. But utility maximizing appeals court judges also have their own policy preferences, which they may seek to follow to the extent possible.[34]

The existence of divergent policy preferences means that lower court judges will have an incentive not to follow the directives laid down by the Supreme Court—to "shirk" their duty as agents. Principal-agent theory is thus concerned with how hierarchical control is maintained: How and to what extent can the Supreme Court control the behavior of lower federal courts to ensure that its policy dictates are implemented?[35]

The principal-agent model of the judicial hierarchy draws on theories of economic organization.[36] In the context of a firm, the principal hires the agent to act on its behalf in order to reduce the costs of coordination and increase efficiency. However, the agent has her own interests, which may conflict with those of the principal. The central challenge for the principal is to design an incentive structure to ensure that the agent pursues the principal's objectives—a goal made more difficult by the fact that it typically lacks complete information about the agent's efforts and the context in which she acts.[37] Common mechanisms of control in hierarchical organizations include monitoring agent behavior and the use of sanctions and rewards to induce compliance. The principal-agent model has been extended to the political context as well, leading to theories of how Congress and the Executive maintain control over bureaucratic agencies to which responsibility for policy implementation has been delegated.[38] Regardless of the context, the principal-agent model focuses attention

---

[34] Songer et al., *supra* note 2, at 675. **R**

[35] *Id.* at 673.

[36] George & Yoon, *supra* note 13, at 820–22. **R**

[37] Terry M. Moe, *The New Economics of Organization*, 28 AM. J. POL. SCI. 739, 756–57 (1984).

[38] *See, e.g.*, Mathew D. McCubbins, Roger G. Noll & Barry R. Weingast, *Administrative Procedures as Instruments of Political Control*, 3 J.L. ECON. & ORG. 243, 243–44 (1987) (identifying problem of political control of agencies as principal-agent problem); Moe, *supra* note 37, at 758–69 (applying principal-agent theory to public bureaucracies and discussing some difficulties in doing so); Barry R. Weingast, *The Congressional-Bureaucratic System: A Principal Agent Perspective (with Applications to the SEC)*, 44 PUB. CHOICE 147, 151–58 (1984) (using principal-agent theory to describe congressional control over agency activities). **R**

on "issues of hierarchical control in the context of information asymmetry and conflict of interest."[39]

Applying the principal-agent model to the judiciary usefully highlights the possibility that judges may have conflicting goals. As discussed in the next section, considerable empirical evidence supports the notion that judges in fact are influenced to some degree by their policy preferences and that these preferences often diverge along ideological lines.[40] The principal-agent model focuses attention on the issue of how value conflicts will be resolved in the context of a hierarchical organization and the extent to which superior courts can effectively direct and control the actions of lower courts. It also asks how the institutional structure of the judiciary will affect the possibilities for and limitations on monitoring and sanctioning lower court behavior—for example, by focusing attention on the Supreme Court's lack of power over traditional economic incentives such as the salaries, job security, or promotion prospects of its subordinates.

Principal-agent models thus offer important insights into the decisionmaking context of lower court judges; however, the way in which positive political theorists have utilized these models to describe the judicial hierarchy currently suffers from certain limitations. The following sections explore these limitations, arguing that principal-agent models currently lack a satisfactory account of why lower courts comply with precedent and fail to take seriously the role of law as a normative force.

B. Empirical Evidence

Over the past several decades, a considerable body of empirical work has investigated the factors that explain the decisions of lower court judges. Many of the earlier studies were framed in terms of the debate between "legal" and "attitudinal" models of judicial decisionmaking, testing whether legal doctrine or political ideology accounted for observed case outcomes.[41] Some have examined whether lower federal courts comply with Supreme Court precedent. Despite prominent examples of defiance—most notably in the school desegregation

[39] Moe, *supra* note 37, at 757. **R**

[40] *See infra* Part I.B.

[41] *See generally* Charles A. Johnson, *Law, Politics, and Judicial Decision Making: Lower Federal Court Uses of Supreme Court Decisions*, 21 LAW & SOC'Y REV. 325 (1987) (testing competing legal and attitudinal models of lower court decisionmaking); Donald R. Songer & Susan Haire, *Integrating Alternative Approaches to the Study of Judicial Voting: Obscenity Cases in the U.S. Courts of Appeals*, 36 AM. J. POL. SCI. 963 (1992) (testing relative impact of various influences on court of appeals judges' votes in obscenity cases).

setting[42]—most systematic studies have found defiance to be rare and compliance the norm.[43] Other work has asked whether ideological preferences influence the decisions of lower court judges. Most commonly using the party of the appointing President as a rough proxy for political ideology, these studies have found that Democrat-appointed judges and Republican-appointed judges in fact have different patterns of voting, with the latter voting for conservative outcomes more often than their Democrat-appointed counterparts.[44] Taken as a whole, this body of work strongly suggests that *both* politics and law matter.

The empirical studies highlight the inadequacy of both a naïve legal model and a purely ideological model of judicial behavior. The naïve legal model cannot account for the pattern of case outcomes documented in these studies. Under strict formalist assumptions, case outcomes are entirely determined by legal doctrine, leaving no room for the influence of personal or political factors. Even relaxing the assumption that the law determines the result in all cases, the traditional legal model holds that where a judge confronts novel or unsettled questions, she should (and does) use analogic reasoning to reach the legally "correct" answer. Different judges, applying the same precedents, may disagree about which of several prior cases is most closely analogous to a pending case, but their disagreement is internal to legal forms of reasoning. Thus, a traditional legal model would not predict, and cannot explain, why divergent outcomes are observed to fall along ideological lines, with Democrat-appointed judges more

---

[42] *See* J.W. PELTASON, FIFTY-EIGHT LONELY MEN: SOUTHERN FEDERAL JUDGES AND SCHOOL DESEGREGATION 1–23, 93 (1961) (describing struggles against desegregation in Southern courts and resistance of some judges to implementing *Brown v. Board of Education*).

[43] *See generally* John Gruhl, *The Supreme Court's Impact on the Law of Libel: Compliance by Lower Federal Courts*, 33 W. POL. Q. 502 (1980) (finding lower courts overwhelmingly comply with Supreme Court decision in area of libel law); Johnson, *supra* note 41 (finding that lower federal courts generally follow Supreme Court decisions); Donald R. Songer, *The Impact of the Supreme Court on Trends in Economic Policy Making in the United States Courts of Appeals*, 49 J. POL. 830 (1987) (finding Supreme Court decisions in labor and antitrust areas had significant impact on decisional trends in courts of appeals); Donald R. Songer & Reginald S. Sheehan, *Supreme Court Impact on Compliance and Outcomes:* Miranda *and* New York Times *in the United States Courts of Appeals*, 43 W. POL. Q. 297 (1990) (finding nearly universal compliance in courts of appeals with two significant Supreme Court decisions). R

[44] *See, e.g.*, C.K. ROWLAND & ROBERT A. CARP, POLITICS AND JUDGMENT IN FEDERAL DISTRICT COURTS 24–57 (1996); Revesz, *supra* note 3, at 1738–46; Songer, *supra* note 43, at 836–37; Songer & Sheehan, *supra* note 43, at 310–11; Cass R. Sunstein, David Schkade & Lisa Michelle Ellman, *Ideological Voting on Federal Courts of Appeals: A Preliminary Investigation*, 90 VA. L. REV. 301, 303 (2004). R R

often deciding cases in a liberal direction and Republican-appointed judges doing the opposite.

On the other hand, ideological preferences also cannot fully explain lower court behavior. The theory that judges single-mindedly pursue their policy preferences is incompatible with the numerous studies concluding that judges shift the nature of their reasoning and the general trend of their decisions in response to Supreme Court precedent. Moreover, even studies documenting the influence of politics find that ideological factors explain only a fraction of case outcomes.[45] In other words, while many empirical studies have shown that political preferences influence lower court decisionmaking, they also suggest that the effect of ideology is modest.[46] Recent work that has included both legal and political variables has typically found that both types of variables appear to influence judicial decisionmaking.[47] Thus, the old "legal" versus "attitudinal" debate has not so much been won as transcended, with scholarly efforts now focused on understanding how legal and political factors interact.

C. Law as a Signal

Utilizing principal-agent models, positive political theory attempts to reconcile the dual influences of political and legal variables within a traditional rational choice framework. Starting with the assumption that judges seek to maximize their policy preferences, this approach suggests that law plays a role in the strategic interaction between the Supreme Court and lower federal courts as judges at each level struggle to advance their preferred policy outcomes within the constraints set by their institutional environment. More specifically, it

[45] *Cf.* Daniel R. Pinello, *Linking Party to Judicial Ideology in American Courts: A Meta-analysis*, 20 JUST. SYS. J. 219, 243 (1999) (conducting meta-analysis of many empirical studies and concluding that political party explains between thirty-one and forty-eight percent of ideological variance in case outcomes).

[46] Gregory C. Sisk & Michael Heise, *Judges and Ideology: Public and Academic Debates About Statistical Measures*, 99 Nw. U. L. REV. 743, 771 (2005).

[47] *See, e.g.*, Songer et al., *supra* note 2, at 687 (finding that court of appeals decisions are affected by both Supreme Court policy and ideology of court of appeals judges); Songer & Haire, *supra* note 41, at 973–75 (finding that variables based on attitudinal factors and Supreme Court precedent both contribute to explaining court of appeals judges' voting behavior); Nancy C. Staudt, *Modeling Standing*, 79 N.Y.U. L. REV. 612, 669 (2004) (finding that both formal legal rules and political preferences play important roles in judicial decisionmaking). In a recent study, Westerland and his coauthors examined how lower courts have treated Supreme Court precedent and found that circuit court judges appear to respond to both the Court's treatment of its own prior precedent *and* to the ideological preferences of the contemporaneous Supreme Court, as opposed to the Court that established the precedent. Westerland et al., *supra* note 12, at 12–13. This finding suggests that the traditional debate over whether legal or political preferences drive lower court decisionmaking is too simplistic.

treats law as a signal—the means by which the Supreme Court communicates its policy directives to lower court judges. For example, according to McNollgast, "judicial doctrine" signifies "a statement about the range of lower court decisions acceptable to the Court on an issue of law."[48] Thus understood, doctrine has no independent influence on lower court decisionmaking. To the extent that lower court judges comply with precedent that diverges from their own policy preferences, they do so because of the incentive structure created by the judicial hierarchy, not because the existence of the precedent itself motivates them in any way.

Crucial to any principal-agent account is an explanation of *how* the Supreme Court exercises control. It is commonly observed that few of the usual economic incentives are available to the Supreme Court to influence the behavior of lower courts.[49] Federal judges are appointed for life, and Congress, not the Supreme Court, determines their level of compensation. In the absence of traditional economic carrots and sticks, principal-agent models have relied heavily on the reversal power of the Supreme Court to explain lower court compliance.[50]

In its simplest form, the model posits that the risk of reversal works as an effective sanction, inducing lower courts to comply with higher court precedent even when their own political preferences would lead them to a different result. While the model has a certain facial plausibility, a closer examination suggests some practical and theoretical difficulties. On a practical level, the effectiveness of the sanction depends in part on the *actual* risk of reversal. For court of appeals judges, the chance of reversal by the Supreme Court is quite small. Even though the Supreme Court reverses a majority of the cases it accepts for review,[51] it only reviews a tiny fraction of the output of the federal courts of appeals—far less than one percent of

[48] McNollgast, *supra* note 2, at 1641. R

[49] *See, e.g.*, Cameron et al., *supra* note 13, at 102 (observing that superior courts "cannot promote, demote, or fire; they cannot cut salaries, give bonuses, or offer stock options"). R

[50] *See, e.g.*, *id.* at 102 (asserting that "[f]requent reversals bring the derision of colleagues and a decline in professional status"); George & Yoon, *supra* note 13, at 822 (arguing that Supreme Court's "obvious mechanism of control over lower court judges is reversal of their decisions"); Songer et al., *supra* note 2, at 675 (describing reversal as mechanism for monitoring lower courts). R R

[51] *See* EPSTEIN ET AL., *supra* note 18, at 244 tbl.3-6. R

the appellate cases decided in a given year.[52] As a result, the threat of reversal is quite low and its effectiveness as a sanction questionable.[53]

For district court judges, the analysis is further complicated by their position at the bottom of the judicial hierarchy. While they face a significant chance that their decisions will be reversed by a court of appeals, the likelihood of reversal, or even review, by the Supreme Court is extremely low.[54] Thus, their incentives for acting as faithful agents of the Supreme Court seem quite attenuated. More likely, if fear of reversal in fact motivates judges, the district court judge should be primarily concerned with avoiding reversal by the appellate court directly above it. Determining the preferences of the circuit court, however, is complicated by the fact that appeals are heard by panels of three judges whose identity is unknown at the time that the trial court makes its decision. It is further complicated by the possibility that the circuit court's decisions may be influenced by the panel judges' own desire to avoid reversal—either by the entire court of appeals sitting en banc or by the Supreme Court. Given these uncertainties, it is unclear to what extent fear of reversal motivates district courts to comply with Supreme Court precedent.[55]

Aware of the extremely low rate of Supreme Court review, scholars have elaborated on the basic principal-agent model to explain how the Court can nevertheless maintain control over lower federal courts. For example, Songer, Segal, and Cameron have suggested that

---

[52] *See supra* note 30 (citing sources for proposition that 0.14% of court of appeals decisions are reviewed by Supreme Court); *cf.* EPSTEIN ET AL., *supra* note 18, at 74 tbl.2-6 (demonstrating that since 1970 Court has granted review for fraction of petitions). R R

[53] Seventh Circuit Court of Appeals Judge Richard Posner writes that "so few court of appeals decisions are reviewed by the Supreme Court that the threat of reversal cannot operate as a significant constraint on circuit judges' decisions." Richard A. Posner, *Judicial Behavior and Performance: An Economic Approach*, 32 FLA. ST. U. L. REV. 1259, 1273 (2005) [hereinafter Posner, *Judicial Behavior*].

[54] In 2005, over 271,000 civil and over 66,000 criminal cases were terminated in the United States District Courts. ADMIN. OFFICE OF THE U.S. COURTS, *supra* note 30, at 14. Even though fewer than a third of federal civil cases ended by adjudication in 2000, Gillian K. Hadfield, *Where Have All the Trials Gone? Settlements, Nontrial Adjudications, and Statistical Artifacts in the Changing Disposition of Federal Civil Cases*, 1 J. EMPIRICAL LEGAL STUD. 705, 732 tbl.8 (2004), the proportion of district court decisions ultimately reviewed by the Supreme Court is still infinitesimally small. R

[55] Nancy Staudt has theorized that because a district court judge cannot know in advance which three-judge panel of the circuit court will review her decision on appeal, the judge "will take the safest route and adhere to the Supreme Court precedent that governs the circuit court." Staudt, *supra* note 47, at 639. Other scholars have made the opposite prediction. *See, e.g.*, SARA C. BENESH, THE U.S. COURT OF APPEALS AND THE LAW OF CONFESSIONS: PERSPECTIVES ON THE HIERARCHY OF JUSTICE 17–18 (2002) ("One might expect that trial courts will be more successful in ignoring or defying precedent than will appellate courts, as the trial court judge's professional scrutiny is not as great as, for example, that of the Court of Appeals judge."). R

litigant policing plays a crucial role in insuring effective monitoring.[56] They assert that litigants who have lost in the lower courts are more likely to seek Supreme Court review of noncomplying decisions, thereby sounding a "fire alarm" that alerts the Court to cases of "flagrant doctrinal shirking."[57] In their view, this risk is sufficient to induce substantial compliance with precedent by circuit judges.[58] However, even assuming that a litigant's petition for certiorari provides an accurate signal, the Supreme Court must still sort through approximately eight thousand petitions each year and only accepts about ninety of those cases for review.[59] Thus, even when a fire alarm is sounded, the risk of reversal remains quite low. Knowing this, circuit judges who are constrained only by the risk of reversal might respond by *increasing* their level of noncompliance, thereby diluting the effectiveness of the fire alarm and further reducing the reversal risk in any particular case.[60]

---

[56] Songer et al., *supra* note 2, at 693. R

[57] *Id*.

[58] In their study of circuit court decisions in search and seizure cases, Songer, Segal, and Cameron found that the impact of Supreme Court policy appeared to be greater, and the impact of the circuit panel's ideology correspondingly less, in cases that were not appealed than in those in which review was sought. From this, they concluded that litigant monitoring offers an explanation for the apparent "'paradox' of (relatively) effective control and rare reversals." *Id*. The finding that appealed cases appear to diverge more significantly from Supreme Court preferences than nonappealed cases supports the theory that *litigants* act rationally in deciding whether or not to seek review. However, it does not answer the question whether the risk of Supreme Court reversal is sufficient to influence judges' decisionmaking behavior. In addition, their study sampled only officially published cases, creating a risk of selection bias. A number of studies indicate that published and unpublished opinions differ in systematic ways. *See, e.g.*, ROWLAND & CARP, *supra* note 44, at 117–35 (finding that magnitude and consistency of political appointment effects R differ between unpublished and published decisions); Peter Siegelman & John J. Donohue III, *Studying the Iceberg from Its Tip: A Comparison of Published and Unpublished Employment Discrimination Cases*, 24 LAW & SOC'Y REV. 1133, 1156 (1990) (concluding that employment discrimination cases with published opinions differ significantly from those without published opinions).

[59] EPSTEIN ET AL., *supra* note 18, at 74 tbl.2-6. Even if *in forma pauperis* cases are R excluded, the Supreme Court grants review only in about three to four percent of the paid cases in which a petition for certiorari is filed. *Id.*

[60] Frank Cross explains how the rational lower court judge with a policy preference at point "L" would decide cases if the Supreme Court had a differing preference at point "S":

> Suppose the lower court can decide five cases, of which the Supreme Court can review at most one. In this circumstance, the lower court's strategy is clear. It would decide all five at point L and suffer one reversal, leaving that lower court with 80% of its maximum possible L-S utility. Deciding all five cases at point S would lose all of the lower court's L-S utility. If the lower court decided any one case compliantly, at S, the Supreme Court would simply reverse another decision, leaving three decisions at L and giving the lower court 60% of its maximum L-S utility. The best strategy is to ignore the risk of reversal and engage in total non-compliance.

McNollgast advance another theory to explain how low rates of reversal can nevertheless produce high levels of compliance with Supreme Court precedent. They start with the assumption that judges are strategic actors, each seeking to maximize his or her own policy preferences.[61] They further observe that the Supreme Court not only establishes precedent at its preferred decision point but also indicates the extent of deviation from that point that it will tolerate, creating what they call a "doctrinal interval."[62] So long as a lower court decision falls within that interval, the Supreme Court will not reverse. The Court tolerates deviations within a doctrinal interval because it has limited resources for reviewing lower court decisions that are not congruent with its own preferences. According to McNollgast, the Court can induce greater compliance by creating a doctrinal interval because some lower courts will choose outcomes within the doctrinal interval, at a point relatively close to their most preferred outcome, rather than deciding solely according to their own preferences and risking reversal all the way to the Supreme Court's ideal point.[63] Thus, by using a doctrinal interval, the Supreme Court can insure substantial compliance with its precedents, even when faced with a number of lower courts with differing preferences.

As Frank Cross has argued, McNollgast's theory of control unravels as its simplifying assumptions are relaxed.[64] In the basic model McNollgast use to construct their theory, there is one type of case to be decided, the Supreme Court faces three lower courts with divergent preferences, and resource constraints allow it to review only one decision.[65] If we assume that two different types of cases exist and the Supreme Court only has the resources to review one decision of either type, then the risk of reversal for each court on each issue would be halved. Manipulating the doctrinal interval might still induce compliance, but only if the Court significantly expanded the width of the interval.[66] Similarly, if the Court faced four noncomp-

Frank Cross, *Appellate Court Adherence to Precedent*, 2 J. EMPIRICAL LEGAL STUD. 369, 373 (2005). This analysis assumes that lower court judges do not experience reversal as a sanction beyond the loss of their preferred policy outcome. Below I consider the possibility that reversal also causes reputational harms or other forms of disutility that increase the sanction. *See infra* note 69 and accompanying text. R

[61] McNollgast, *supra* note 2, at 1636. R

[62] *Id.* at 1639.

[63] *Id.* at 1645–46.

[64] Cross, *supra* note 60, at 377–78. R

[65] *See* McNollgast, *supra* note 2, at 1644–45. R

[66] *See* Cross, *supra* note 60, at 377. McNollgast model the Court's preferences in one-dimensional space. As an example, they suppose a Supreme Court with a policy preference at 0.7 and three lower courts with preferences of 0.05, 0.2, and 0.4. They argue that by adopting a doctrinal interval such that decisions down to 0.44 are acceptable, the Supreme R

liant courts instead of three, a significant expansion of the doctrinal interval would be necessary to induce the same level of compliance.[67] As the model is made more realistic by including more issues and more lower courts, the doctrinal interval has to expand further in order to have the predicted effect. As Cross concludes, "Reversal probabilities as low as those currently existing would drive the interval to the point where it clearly could not be considered control at all."[68]

In the models described above, reversal by a higher court operates as a sanction because the lower court is frustrated in its pursuit of its preferred policy outcomes. The impact of reversal—apart from its policy consequences—is implicitly assumed to be zero. Under these assumptions, it is difficult to explain why lower courts with divergent preferences comply with Supreme Court precedent. An alternative explanation posits that reversal entails greater disutility for a lower court judge than merely the loss of a preferred outcome in a particular case. Thus, it has become commonplace to assert that judges avoid reversal because they believe it will harm their professional reputation or reduce their chances for promotion to higher office, and that this fear of reversal induces compliance with precedent.[69] This explanation, however, faces the same difficulty: The chance of reversal by the

Court can induce all three lower courts to comply with its doctrine. McNollgast, *supra* note 2, at 1645–46. Cross's analysis suggests that if two types of cases exist, halving the risk of review, then the Court will have to expand its doctrinal interval all the way down to 0.32 to get these three lower courts to comply with its precedent. Cross, *supra* note 60, at 377. R R

67 Once again, the decreased risk of review of any given lower court decision means that an expansion of the doctrinal interval will be necessary—in this case to 0.325—in order to induce compliance.

68 Cross, *supra* note 60, at 378. Cross further argues that McNollgast's model unrealistically assumes that the preferences of lower court judges and the outcomes of their decisions are transparent to the Supreme Court. In fact, the signals the Court receives are noisy and the unavailability of accurate information prevents the Court from "precisely . . . titrat[ing] its doctrinal interval." *Id*. at 380. R

69 *See, e.g.*, Cameron et al., *supra* note 13, at 102 (asserting that frequent reversal brings derision of colleagues and decline in professional status); Caminker, *Precedent and Prediction*, *supra* note 19, at 77–78 (suggesting that psychological and professional costs of reversal may influence lower court judges' decisions); Paul L. Colby, *Two Views on the Legitimacy of Nonacquiescence in Judicial Opinions*, 61 TUL. L. REV. 1041, 1051 (1987) (stating that lower court judges seek "to avoid the professional obloquy that attaches to a reversal"); Christopher R. Drahozal, *Judicial Incentives and the Appeals Process*, 51 SMU L. REV. 469, 477 (1998) (asserting that judges decide cases in order to avoid being reversed on appeal); George & Yoon, *supra* note 13, at 822 (claiming that lower court judges know their promotion to higher court depends in part on "number of times they have been reversed"). R R

Not everyone agrees with this assumption. Most notably Judge Richard Posner has written: "Judges don't like to be reversed (I speak from experience), but aversion to reversal does not figure largely in the judicial utility function." Richard A. Posner, *What Do Judges and Justices Maximize? (The Same Thing Everybody Else Does)*, 3 SUP. CT. ECON. REV. 1, 14 (1993) [hereinafter Posner, *What Do Judges Maximize?*].

Supreme Court is vanishingly small. Given the very low probability of reversal by the Supreme Court, the degree of disutility suffered as a result of a single reversal—independent of the loss of a preferred policy outcome—would have to be large indeed to change behavior. More to the point, it is not self-evident that lower court judges actually suffer reputational harm when they are reversed. Not all lower court judges will care about promotion. The prospects for promotion—at least to higher judicial office—are quite small for most federal judges.[70] Nor is it obvious that high reversal rates will in fact damage a judge's chances of promotion.[71] In some situations, the opposite might hold true. For example, frequent reversals of the decisions of a conservative district court judge by a more liberal court of appeals might enhance rather than diminish her chances for promotion if the President and Senate are more conservative than her appellate superiors.[72]

The handful of relevant empirical studies generally do not support the theory that fear of reversal motivates lower court compliance with precedent.[73] Higgins and Rubin tested whether judges who are

[70] *See* Daniel Klerman, *Nonpromotion and Judicial Independence*, 72 S. CAL. L. REV. 455, 461 (1999) ("[T]he chance that any particular appellate judge would be promoted [to the Supreme Court] during [the 1990s] was less than three percent. Similarly . . . the probability that a district court judge serving during the 1990s would be promoted to the court of appeals was only six percent."); Posner, *Judicial Behavior*, *supra* note 53, at 1265–66 ("[W]hile at present almost all the Supreme Court Justices were federal circuit judges previously, there are so few Justices, and they serve for such a long time, that the percentage of federal court of appeals judges who becomes Supreme Court Justices is miniscule."). R

[71] A desire for promotion may influence the behavior of lower court judges without necessarily causing them to avoid reversal. A recent empirical study of decisions from 1988 suggested that federal district court judges with greater promotion potential were more likely to have issued a written opinion on the constitutionality of the Federal Sentencing Guidelines. Andrew P. Morriss, Michael Heise & Gregory C. Sisk, *Signaling and Precedent in Federal District Court Opinions*, 13 SUP. CT. ECON. REV. 63 (2005). The authors theorized that "written opinions serve as a signal to appointing authorities." *Id*. at 92.

[72] As Judge Posner writes, "[T]he impact of a particular decision on the prospects for promotion is normally very slight," and the impact may be unpredictable because "the decision may offend as many influential people as it pleases." Posner, *What Do Judges Maximize?*, *supra* note 69, at 5. R

[73] An exception is Richard Revesz's study of voting on the D.C. Circuit in environmental cases, finding that for at least one subset of his data, ideological voting appeared more pronounced in cases involving procedural challenges than in those involving statutory challenges. Revesz, *supra* note 3, at 1747–50, 1760. Because questions of statutory interpretation are reviewed by the Supreme Court more frequently than cases involving procedural challenges, *id*. at 1729–30, Revesz concludes that this finding supports a "hierarchical constraint" hypothesis, i.e., that judges act more ideologically when their decisions are less likely to be reviewed, *id.* at 1729, 1750. He also acknowledges, however, that another possible explanation is that the legal standards for procedural challenges are malleable. *Id*. at 1731. Thus, it is difficult to disentangle whether the more pronounced ideo- R R

more likely to be eligible for promotion based on their age and tenure on the bench are more sensitive to the risk of reversal.[74] They found no correlation between age or tenure and reversal rates and little responsiveness in the probability of promotion to changes in a judge's reversal rate.[75] More recent studies have failed to find evidence that fear of reversal affects circuit court decisionmaking. Klein and Hume analyzed a sample of search and seizure cases, asking whether circuit courts were more likely to decide consistently with Supreme Court preferences in those cases most likely to be subject to review—such as cases raising issues of importance or presenting circuit conflicts.[76] They found little evidence that circuit court decisions were more likely to be congruent with Supreme Court preferences in cases facing an increased risk of review and reversal.[77] Songer, Ginn, and Sarver examined tort cases heard in federal court based on diversity of citizenship of the parties.[78] Under the *Erie* doctrine, federal courts are to apply state law in these cases. Federal circuit courts face virtually no risk of reversal for this category of cases, because state courts cannot directly review federal court decisions and the Supreme Court is extremely unlikely to hear them due to the lack of a federal question.[79] Songer, Ginn, and Sarver found that despite the lack of a realistic threat of reversal, circuit court decisions in these cases generally follow relevant state tort law rather than reflecting strongly ideological voting.[80] Thus, while it is plausible that judges dislike having their decisions reversed, it is questionable whether fear of reversal is a

logical effects he observes are the result of greater legal discretion or the lesser probability of review.

74 Richard S. Higgins & Paul H. Rubin, *Judicial Discretion*, 9 J. LEGAL STUD. 129, 129–30 (1980).

75 *Id.* at 135–37.

76 David E. Klein & Robert J. Hume, *Fear of Reversal as an Explanation of Lower Court Compliance*, 37 LAW & SOC'Y REV. 579, 584–86 (2003).

77 *Id.* at 594–97. *But see* Jonathan P. Kastellec & Jeffrey R. Lax, Can We Ignore Case Selection When We Study Judicial Politics? 7–8 (Dec. 15, 2006) (unpublished manuscript, on file with the *New York University Law Review*) (criticizing Klein and Hume study on grounds that selection effects severely bias their measure of compliance).

78 Donald R. Songer, Martha Humphries Ginn & Tammy A. Sarver, *Do Judges Follow the Law When There Is No Fear of Reversal?*, 24 JUST. SYS. J. 137, 138 (2003).

79 *Id.* at 141. Out of the 697 tort diversity cases analyzed by Songer and his coauthors, not one was reviewed by the Supreme Court on the question of whether its interpretation of state law was correct. *Id*. at 142.

80 *Id*. at 155. The authors did find that judges' votes were more closely correlated with their political preferences in nonunanimous cases—those which the authors speculate are not governed by clear law. *Id*. Their finding that political preferences appear to play a role in "hard" cases is not surprising given that numerous studies have documented such effects in court of appeals decisions involving federal questions that are much more likely to be reviewed. More notable is their finding that among votes in unanimous cases (1598 of the 1742 judicial votes they analyzed), measures of judicial ideology had no statistically

factor significant enough to explain lower court compliance with superior court precedent.[81]

D. Law as a Normative Force

If the fear of reversal is insufficient to explain judicial behavior, then the principal-agent model presents a puzzle. In the absence of any effective sanction, why would lower court judges—assumed to be motivated by their policy preferences—choose to follow legal authority rather than pursuing their own preferred outcomes? The simplest explanation for lower court compliance is that judges have *legal* preferences independent of their political preferences. More precisely, even if judges care about whether the outcome in a given case advances their preferred policy, they likely care about whether it conforms to legal norms as well. Judges may have a variety of legal preferences regarding matters such as the appropriate mode of interpreting statutes, or the relevance of foreign legal materials, and these preferences may vary from judge to judge. But their decisions are also guided by a set of widely shared norms—some of which are formulated as legal rules—regarding their role in the judicial hierarchy. One fundamental and widely accepted norm requires that lower federal court judges follow precedent established by a court directly in line above them in the judicial hierarchy. Adherence to this norm offers a straightforward explanation of why lower courts comply with superior court precedent, even that with which they disagree.[82]

Attitudinalists have resisted the idea that legal preferences influence judicial behavior, but their objections have been largely directed

significant effect on the direction of judicial votes, while state law variables were significant. *Id*. at 153 tbl.4.

[81] Cross also tested whether fear of reversal or legal factors explain circuit court decisionmaking. Cross, *supra* note 60, at 398–404. His empirical study, however, aggregated all types of circuit cases and used rather broad measures for the variables of interest. More specifically, he used the ideology of the median Supreme Court Justice using Segal/Cover scores as a measure of the risk of reversal and the ideology of the median Supreme Court Justice for the ten years prior to a decision as a measure of "legal model" requirements. Although he concludes that the risk of reversal appears to play no role in circuit court decisions, *id.* at 404, his study is at best a very indirect test of the theory. R

[82] Of course, there have been instances of explicit refusals by lower courts to follow established Supreme Court precedent. For example, the Fourth Circuit in *United States v. Dickerson*, 166 F.3d 667, 667 (4th Cir. 1999), held that the admissibility of confessions in federal court is governed by statute, not the Supreme Court's decision in *Miranda v. Arizona*, 384 U.S. 436 (1966), and the Fifth Circuit in *Hopwood v. Texas*, 78 F.3d 932, 944 (5th Cir. 1996), rejected the rationale of Justice Powell's plurality opinion in *Regents of the University of California v. Bakke*, 438 U.S. 265 (1978), and concluded that the goal of achieving a diverse student body can never be a compelling reason justifying the use of race in university admissions under the Fourteenth Amendment. These cases are notable precisely because the phenomenon they represent is unusual.

at formalist accounts of how the law operates. For example, Segal and Spaeth reject a legal model of judging as implausible, arguing that such claims "are obviously belied by the fact that different courts and different judges do not decide the same question or issue the same way" and that appellate decisions often include dissenting votes.[83] Because different courts and judges in fact reach divergent outcomes and opinions regarding the same legal issue, they assert that the notion that judges decide cases according to legal authority is pure "mythology."[84] However, their expectation that a single correct legal outcome should exist in every case rests on a naïve formalism not widely shared by judges, lawyers, or legal scholars.

An account of legal preferences does not require that the outcome in every case is determined in advance by existing legal materials. Rather, a judge's preferences regarding *legal* outcomes might be understood as an "attitude" in much the same way that Segal and Spaeth model political preferences. They state that "an attitude is nothing more than a set of interrelated beliefs about at least one object and the situation in which it is encountered."[85] They assume that the relevant "attitudes" are policy goals, which explain judges' behavior in response to a set of cases that present similar "objects" or "situations," such as criminal procedure or First Amendment issues. Judges might also be assumed to have "attitudes" toward sets of legal materials—such as statutes or judicial opinions by superior courts—that influence their behavior in cases that present those sorts of materials. Even the most conservative judge will not necessarily decide *every* case involving issues of criminal procedure in favor of the government's position; so too, legal preferences should not be assumed to predetermine an outcome in every case. Like political preferences, legal preferences will produce a behavioral tendency, a correlation, rather than an exact correspondence between attitude and behavior.

Although strategic models have added nuance to the simple behavioralist accounts of the attitudinalists, positive political theorists also resist the idea that the law has independent force.[86] As Cross observes, positive political theorists "seem extraordinarily determined to prove that lower court compliance results purely from political

[83] SEGAL & SPAETH, *supra* note 1, at 17–18. R

[84] *Id.*

[85] *Id.* at 69 (quoting DAVID W. ROHDE & HAROLD J. SPAETH, SUPREME COURT DECISION MAKING 76 (1976) (emphasis omitted)).

[86] *See, e.g.*, Bueno de Mesquita & Stephenson, *supra* note 4, at 755 (explaining that judges respect precedent not in addition to caring about policy, but "*because* they care about policy"). R

preference and not from any judicial preference for legal variables."[87] Similarly, Friedman asserts that as a result of this "almost pathological skepticism that law matters," they "simply fail to take law and legal institutions seriously."[88] Thus, positive political theorists attempt to explain lower court adherence to precedent purely in terms of political strategy, as if political preferences were a priori more rational than legal ones.

Although no definitive explanation exists for the formation of legal preferences, several plausible theories have been advanced in the literature. Some scholars argue that the socialization process involved in professional training or the role perceptions of judges shape their legal preferences.[89] Others contend that judges have self-interested reasons for following precedent, such as ensuring respect for their own decisions or for the judiciary more generally.[90] Judge Posner has suggested that judges gain inherent utility from following precedent, analogizing doctrine to the rules of a game to which they must adhere to make the game meaningful.[91] In any case, the assumption that judges have legal preferences is at least as plausible as the theory that they have policy preferences. Judges who decide cases in ways that advance their policy preferences do not benefit directly from those decisions, except to the extent that they derive utility from generating outcomes consistent with their attitudes, just as they might

[87] Cross, *supra* note 60, at 384. R

[88] Friedman, *supra* note 6, at 262. R

[89] *See, e.g.*, James L. Gibson, *Judges' Role Orientations, Attitudes, and Decisions: An Interactive Model*, 72 AM. POL. SCI. REV. 911, 922 (1978) (suggesting that role orientation of judges blocks or reduces influence of ideology on behavior); James L. Gibson, *Personality and Elite Political Behavior: The Influence of Self Esteem on Judicial Decision Making*, 43 J. POL. 104, 123–24 (1981) (finding that internal role expectations of judges are associated with restrained approach to decisionmaking); J. Woodford Howard, Jr., *Role Perceptions and Behavior in Three U.S. Courts of Appeals*, 39 J. POL. 916, 935 (1977) (finding that conceptions of judicial role influence, but do not control, judicial behavior).

[90] *See, e.g.*, William M. Landes & Richard A. Posner, *Legal Precedent: A Theoretical and Empirical Analysis*, 19 J.L. & ECON. 249, 273 (1976) (suggesting that judges follow precedent because failing to do so would undermine practice of following precedent and therefore reduce precedential significance of their own decisions); Eric Rasmusen, *Judicial Legitimacy as a Repeated Game*, 10 J.L. ECON. & ORG. 63, 81 (1994) (arguing that judges may obey existing law in order that new law they create will be obeyed by future judges). This explanation has been subject to a number of criticisms. *See, e.g.*, Bueno de Mesquita & Stephenson, *supra* note 4, at 756–57 (criticizing self-interested explanations of why judges follow precedent as resting on dubious assumptions and failing to explain when and why judges will not follow precedent); Cross, *supra* note 60, at 386 (criticizing Rasmusen's team theory on grounds that self-interested judges will defect in particular cases in absence of mechanisms for monitoring and enforcement). R R

[91] *See* Posner, *What Do Judges Maximize?*, *supra* note 69, at 28–30 (explaining how self-selection and screening assure that lawyers who become federal judges adhere to judicial "game"). R

gain utility from decisions that comport with legal norms. And, as with legal preferences, no authoritative account exists of how judges' policy preferences are formed. Some studies have attempted to explain judicial attitudes using background characteristics, with mixed success.[92] Moreover, the measures most often used for judicial ideology in empirical studies often indicate little more than a generalized liberal or conservative orientation, and cannot adjust for changes in attitude over time.[93]

Importantly, the assumption that judges have legal preferences does not exclude the influence of other motivations. A more plausible starting point—especially in light of existing empirical work—posits that judges care both about adherence to the law *and* about policy outcomes. The more interesting and difficult questions ask how these preferences interact to produce outcomes.[94] Cross proposes a model of judicial decisionmaking based on the assumption that "circuit court judges have two separate utilities when deciding cases: (1) an ideological policy utility and (2) a legal model utility."[95] By making various assumptions about the shape of a judge's political and legal utility curves and the relative intensity of these preferences, the model explains why a circuit court judge would comply with Supreme Court precedent despite her differing policy preferences.[96] Although Cross's model offers a plausible framework for understanding how legal and

[92] *See* Tracey E. George, *Court Fixing*, 43 ARIZ. L. REV. 9, 15–31 (2001) (reviewing empirical studies testing effect of personal attributes and social background on judicial behavior and describing mixed findings); Lee Epstein, Jack Knight & Andrew D. Martin, *The Norm of Prior Judicial Experience and Its Consequences for Career Diversity on the U.S. Supreme Court*, 91 CAL. L. REV. 903 app. (2003) (summarizing studies exploring links between prior occupation and judicial decisionmaking and reporting mixed results).

[93] An important exception is Martin-Quinn scores, which measure the policy preferences of Supreme Court Justices dynamically over time. *See generally* Andrew D. Martin & Kevin M. Quinn, *Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953–1999*, 10 POL. ANALYSIS 134 (2002) (modeling preferences of Supreme Court Justices from 1953 through 1999).

[94] *See* MALCOLM M. FEELEY & EDWARD L. RUBIN, JUDICIAL POLICY MAKING AND THE MODERN STATE: HOW THE COURTS REFORMED AMERICA'S PRISONS 213 (1998) (arguing that both ideology and law influence court decisions and that "it would be more productive to explore their interaction than to choose between them").

[95] Cross, *supra* note 60, at 389. **R**

[96] For example, if both utility functions are linear and the circuit court judge's political preferences are stronger than her legal preferences for a given issue, then she will not comply with precedent when her preferences diverge from the Supreme Court's. *Id.* at 390, 391 fig.1. If, however, the judge's legal utility curve is S-shaped, she may comply with Supreme Court precedent with which she disagrees. *Id*. at 394, 395 fig.2. Cross argues that such an S-shaped legal utility curve is plausible because relatively small departures from Supreme Court precedent will not be viewed as noncompliant; however, "at some point the departure becomes so great that it is . . . clear that the lower court is being disobedient," *id*. at 393–94, and further departures will not entail much greater loss of legal utility.

ideological preferences interact, it produces no clear predictions that could be tested empirically. We can neither observe the shape of the judges' utility curves nor the intensity of their preferences, and without some measure of these variables, the model does not generate any clear predictions of judicial behavior. Still, if as Cross suggests, judges' legal preferences help explain why they comply with superior court precedent,[97] then theoretical models must pay more attention to the nature of law and how it shapes the decisionmaking environment.

II
Lower Court Discretion

Reviewing the empirical literature on lower court decisionmaking, Judge Richard Posner expresses no surprise at its findings that judicial votes often correlate with political factors. Rather, he writes, it "shows nothing more than that there is a large open area in American law, that is, an area in which conventional legal materials do not dictate the outcome and the judge is forced to make a policy judgment, inevitably influenced by political or ideological preferences."[98] As Posner's comment suggests, the key to untangling the interaction between legal and ideological influences lies in recognizing and understanding this "open area" in American law. Those areas in which "conventional legal materials do not dictate the outcome" afford lower court judges—even when they adhere to legal norms—a measure of discretion in deciding cases. In this Part, I explore the nature of judicial discretion, where and why it exists, and how it shapes the decisionmaking environment of the lower court judge.

A. Law and Discretion

One of the difficulties of speaking about discretion is that the term is used in a wide variety of contexts in ways that indicate quite different things. If there is any common core of meaning, it is that "discretion" has something to do with choice.[99] Where someone acts under compulsion, she cannot be said to exercise discretion. But discretion also implies something more than mere choice. It suggests that a decision should be made not randomly or arbitrarily, but by exercising judgment in light of some applicable set of standards, guidelines, or values. Those standards or norms may rule out certain

[97] *See id.* at 389 (suggesting that judges actually gain utility from adhering to superior court precedents).

[98] Posner, *Judicial Behavior*, *supra* note 53, at 1272. R

[99] George C. Christie, *An Essay on Discretion*, 1986 Duke L.J. 747, 747; Maurice Rosenberg, *Judicial Discretion of the Trial Court, Viewed From Above*, 22 Syracuse L. Rev. 635, 636 (1971).

options while still permitting the decisionmaker to exercise some choice. Thus, Hart and Sacks imply the existence of some standard of reference when they define discretion as "the power to choose between two or more courses of action each of which is thought of as permissible."[100] Similarly, Dworkin asserts that "[t]he concept of discretion is at home in only one sort of context; when someone is in general charged with making decisions subject to standards set by a particular authority."[101] Thus, discretion can be thought of as the exercise of choice or judgment subject to certain constraints.[102]

Because the relevant constraints will frame the discretionary choice, the institutional setting in which it is exercised is critical. Discretion inheres in a range of legal settings, but its meaning and significance vary considerably from one context to another. A great deal has been written about administrative agency discretion and prosecutorial discretion, but the discretion exercised in those contexts is quite different than the discretion afforded federal district court and appellate judges when deciding cases. Insights from the administrative and criminal contexts cannot necessarily be extrapolated to the judicial context, and attempting to do so risks confusing rather than clarifying understanding.[103]

What, then, does discretion mean in the context of judicial decisionmaking? When judges decide cases, the body of applicable legal rules—statutes, regulations, and prior precedents—constitute the relevant constraints on their decisionmaking. In some cases, those rules can be applied rather mechanically to reach a determinate result. In other cases, however, application of the relevant authority to the case at hand requires the exercise of judgment, perhaps because the precise issue raised in the case is not addressed by the rule, or because the rule itself calls for the exercise of judgment, as in the case of a mul-

---

100 HENRY M. HART, JR. & ALBERT M. SACKS, THE LEGAL PROCESS: BASIC PROBLEMS IN THE MAKING AND APPLICATION OF LAW 144 (William N. Eskridge, Jr. & Philip P. Frickey eds., 1994).

101 DWORKIN, *supra* note 21, at 31. **R**

102 As Dworkin colorfully puts it: "Discretion, like the hole in a doughnut, does not exist except as an area left open by a surrounding belt of restriction." *Id*.

103 Edward Rubin makes a variant of this argument, asserting that the concept of discretion developed by Dworkin and others in the context of studying the judiciary creates confusion when applied to the administrative state. Edward L. Rubin, *Discretion and Its Discontents*, 72 CHI.-KENT L. REV. 1299, 1300–04 (1997). Although his concern is primarily with administrative agencies, he suggests that study of the judiciary would also be advanced if "discretion" were replaced by concepts of "supervision" and "policymaking." *Id.* at 1303. He may be right that questions regarding the administrative state are better cast in those terms; however, the considerable differences between the judiciary and administrative agencies do not necessarily mean that concepts developed in the bureaucratic context are usefully extended to the judicial context without modification.

tifactor balancing test. To say that a court has discretion, then, "is to say that it is not bound to decide the question one way rather than another. . . . [T]here is no wrong answer to the questions posed—at least, there is no *officially* wrong answer."[104] Or, as another scholar put it, judicial discretion is the "ability to decide a case either way consistently with the preexisting rule structure . . . ."[105]

As this account of discretion makes clear, the exercise of judgment under these conditions is *permissible* under authoritatively recognized rules. The judge who exercises discretion is doing so pursuant to and consistent with the various legal norms that govern the work of judging. More to the point, the fact that judges sometimes have discretion in deciding cases is not inconsistent with the existence of legal rules that are binding or obligatory on them. Rules are necessarily framed in general terms, and even judges who believe themselves to be bound by them, and who intend to follow them, will find that these rules do not reach every situation. As H.L.A. Hart has written, rules—whether expressed through precedent or legislation—"will, at some point where their application is in question, prove indeterminate; they will have what has been termed an *open texture*."[106] Because language is "irreducibly open textured,"[107] even judges with a strong legal preference for following superior court precedent will encounter cases in which "the guidance of authoritative legal rules runs out"[108] and therefore find themselves with discretion to decide. And when discretion in this sense exists, it "is quintessentially associated with variability of result."[109]

Contrast this understanding of the relationship between law and judicial discretion with Segal and Spaeth's characterization of the legal model: "[J]udges exercise little or no discretion; . . . they do not speak; rather, the Constitution and the laws speak through them. Accordingly, judicial decisions merely apply the law objectively, dispassionately, and impartially."[110] Their conception of the law entirely misses its inevitably open texture and the possibility that rules can *both* constrain decisionmaking *and* permit some variability of result. Because Segal and Spaeth define law as incompatible with the exer-

[104] Rosenberg, *supra* note 99, at 637. **R**

[105] Charles M. Yablon, *Justifying the Judge's Hunch: An Essay on Discretion*, 41 HASTINGS L.J. 231, 238 (1990).

[106] H.L.A. HART, THE CONCEPT OF LAW 124 (1961).

[107] *Id.* at 125.

[108] Yablon, *supra* note 105, at 239. **R**

[109] Christie, *supra* note 99, at 748. **R**

[110] SEGAL & SPAETH, *supra* note 1, at 33. **R**

cise of discretion, the fact that similar cases sometimes come out differently inexorably leads them to conclude that law is pure mythology.

Other empirical scholars have similarly failed to take seriously law's open texture. A number of studies have explored the influence of Supreme Court precedent on lower court decisions by asking whether lower courts decided similar cases in a liberal or conservative direction congruent with the political valence of the relevant Supreme Court decision.[111] If the decisional trend mirrored changes in Supreme Court decisions—for example, if more conservative court of appeals decisions were observed after the Supreme Court changed its doctrine in a conservative direction—they concluded that the Court's decisions have a significant impact.[112] Conversely, decisional trends in the lower courts that did not change, or changed only slightly, in response to shifts in Court doctrine are taken as evidence of the Court's limited impact.[113] This approach fails to acknowledge the possibility of selection effects, which may cause the type of cases heard by the courts to change over time. Particularly when the Supreme Court has shifted its policy direction, litigants are likely to assess their prospects of success differently, and therefore to settle or abandon cases with clear outcomes under existing law, and to pursue cases in which the application of the law is unclear. Thus, the choices that litigants make will alter the mix of cases over time, tending to concentrate them in those areas where the law "runs out" and legal norms permit lower court judges to exercise discretion.

By controlling for case facts, Songer, Segal, and Cameron take a changing case mix into account in their study of court of appeals' search and seizure decisions.[114] However, they do not consider that lower court judges might legitimately exercise their discretion when the facts of the case before them diverge significantly from those considered by the Supreme Court. Rather, upon finding that liberal and conservative panels differ significantly in their likelihood of upholding the validity of a search even when controlling for case facts, they conclude that lower court judges "were able to shirk, thereby partially advancing their own policy preferences, by interpretations of Supreme

111 *See* Donald R. Songer, *The Circuit Courts of Appeals*, *in* THE AMERICAN COURTS: A CRITICAL ASSESSMENT 35, 44–45 (John B. Gates & Charles A. Johnson eds., 1991) (reviewing numerous studies which examine change in decisional trends of lower courts in response to significant Supreme Court decisions).

112 *See id.* at 44 (citing studies in which statistically significant shift in decisional trends in courts of appeals followed major shift in Supreme Court precedent in certain areas).

113 *See id*. at 44–45 (finding that change in most courts of appeals' decisions after major Supreme Court precedent was "modest in magnitude," and concluding that judges "retain significant independence in a number of cases").

114 Songer et al., *supra* note 2, at 677–78. R

Court doctrine in ambiguous situations that were not directly noncompliant."[115] This characterization suggests that discretionary action is somehow illegitimate—that it evidences law's failure to control subordinate courts in the judicial hierarchy. Given the open texture of legal rules, however, lower courts necessarily exercise discretion when they apply them. Because variability of outcome will occur even when lower courts faithfully follow precedent, variability alone cannot be taken as evidence of "shirking."

Discretion is inevitable in judicial decisionmaking, not only because of the indeterminacy of language, but also because of the difficulty of anticipating future scenarios in which a rule of decision might be required.[116] As Jerome Frank wrote:

> [M]en have never been able to construct a comprehensive, eternalized set of rules anticipating all possible legal disputes and settling them in advance. . . . [N]o one can foresee all the future permutations and combinations of events; situations are bound to occur which were never contemplated when the original rules were made.[117]

Similarly, Hart argued that "we labour under two connected handicaps whenever we seek to regulate . . . . The first handicap is our relative ignorance of fact: the second is our relative indeterminacy of aim."[118] Without knowing in advance every situation that may call for the application of a rule, it is impossible to frame the rule in a manner that is never ambiguous in any future case. If a particular factual scenario is wholly unanticipated, a preexisting rule cannot clearly communicate its application in that situation. Of course, one could insist that every rule has a fixed, determinate meaning upon its announcement and thereby deny that judges have any discretion when applying it in particular cases. To do so, however, merely disguises the need for making a choice and attempts "to secure a measure of certainty or predictability at the cost of blindly prejudging what is to be done in a range of future cases, about whose composition we are ignorant."[119]

To argue, as Hart does, that legal rules both have a binding or obligatory quality *and* that they are open textured is to recognize that they must embrace a tension between competing aims. As Hart writes, all legal systems "compromise between two social needs: the need for certain rules . . . and the need to leave open, for later settlement by an informed, official choice, issues which can only be prop-

115 *Id*. at 693.
116 JEROME FRANK, LAW AND THE MODERN MIND 5–7 (1949).
117 *Id.* at 6.
118 HART, *supra* note 106, at 125. **R**
119 *Id.* at 126.

erly appreciated and settled when they arise in a concrete case."[120] On this view, judicial discretion exists not merely as an unavoidable by-product of the indeterminacy of legal rules, but also in response to a particular social need, albeit one in tension with the need for certainty. Judges may experience authoritatively enacted rules as obligatory—that is, they may believe that they are bound to follow them—but the inherent uncertainty of legal rules and the need for flexibility to respond to unanticipated situations means that rules cannot definitively determine what a judge should do in every case.

Consider how this plays out in the context of the judicial hierarchy. If the Supreme Court acts as the "principal," it must communicate its directions to its "agents," the lower federal courts. It does so by issuing written decisions that are then treated as binding precedent by trial and appellate court judges. These decisions, however, cannot provide unambiguous guidance across all potential future cases for the reasons discussed above. Even judges with strong legal preferences for following superior court precedent will sometimes find that they have discretion. Songer, Segal, and Cameron utilize a metaphor that nicely suggests the connection between Court decisions and lower court discretion. They write:

> The relationship [between the Supreme Court and the lower courts] is in some ways like that of persons walking their dogs. The dog on a leash is free to lead or follow the owner. The dog's position is not congruent with that of the owner, but the degree of incongruence is limited by the length of the leash selected by the owner. And when the owner changes direction and pulls on the leash, the dog follows (it is responsive to changes in the owner's position).[121]

By deciding cases, the Supreme Court directs the work of the lower courts, but the inevitable ambiguity of the Court's precedents permit them to exercise a measure of discretion while still complying with its basic directions.[122] Court precedent thus constrains lower

---

[120] *Id.* at 127.

[121] Songer et al., *supra* note 2, at 674–75. Songer, Segal, and Cameron have a different purpose than I do in using this metaphor. For them, it illustrates the difference between congruence—"the degree to which agents follow the wishes of principals"—and responsiveness—"the degree to which agents change their behavior as the desires of principals change." *Id*. at 674. In their view, when lower court decisions are responsive to changes in direction set by the Supreme Court, but not perfectly congruent, it shows that lower court judges have "considerable room for discretionary action or 'shirking.'" *Id*. As I argue in Part III.B, *infra*, the fact that lower court judges sometimes take discretionary action should not necessarily be viewed as "shirking," a term that is implicitly pejorative. **R**

[122] Friedman uses a similar metaphor to describe constraints on the Supreme Court. Arguing that if the Court strays too far from public opinion in its decisions, it will inevitably be brought back in line, he argues that "the Court operates on a leash—or perhaps a bungee cord provides a better analogy." Barry Friedman, *The Importance of Being Posi-*

court judges by ruling out many possible outcomes, while nevertheless allowing a zone in which they may legitimately exercise choice.

Elaborating on this metaphor, one might think of a Supreme Court opinion as the "leash" which defines a zone of discretion in which lower courts may legitimately exercise their judgment. Depending upon how an opinion is crafted, the "leash" may be longer or shorter, granting lower courts a greater or narrower zone of discretion in which to operate. Whenever the Court decides a case, it makes a variety of choices in framing its opinion that will affect how much discretionary space it affords lower courts.[123] It may choose to define the question before it narrowly or broadly; it may focus on the particular facts of the case before it or enunciate sweeping principles; it may dispose of a case on procedural grounds and avoid reaching a widely disputed substantive issue altogether.[124]

In some situations, the Court may not be able to avoid writing an opinion that permits a great deal of discretion in its application. As discussed above, even when the author of a rule has a clear purpose in regulating, communicating that intent in the form of a completely determinate rule is likely to be impossible. That difficulty is compounded when the author is a multi-member decisionmaking body like the Supreme Court, where the individual Justices have different, but overlapping, goals. The need to accommodate their differing preferences may require that an opinion announcing the decision of the Court contain ambiguities in order to garner the support of a majority of its members. If the Justices cannot agree on how to regulate across a broad class of related cases they may decide a pending case narrowly, using vague language to elide areas of disagreement.

Apart from the need for compromise, the nature of their preferences in particular cases may lead the Justices to grant lower courts a

tive: The Nature and Function of Judicial Review, 72 U. CIN. L. REV. 1257, 1279 (2004); *see also* Barry Friedman, *The Politics of Judicial Review*, 84 TEX. L. REV. 257, 327 (2005) ("[T]he diffuse support hypothesis suggests that the judiciary can stray a certain distance from public opinion but that ultimately it will be snapped back into line."). His claim that public opinion operates as an effective check on Supreme Court power remains a matter of controversy. Nevertheless, his bungee cord metaphor, like the dog-on-a-leash metaphor for lower courts, makes the point that the existence of discretion does not necessarily mean that decisions are wholly unconstrained.

123 As Justice Scalia has written, "[W]hen the Supreme Court . . . decides a case, not merely the *outcome* of that decision, but the *mode of analysis* that it applies will thereafter be followed by the lower courts . . . . [It] can either establish general rules or leave ample discretion for the future." Antonin Scalia, *The Rule of Law as a Law of Rules*, 56 U. CHI. L. REV. 1175, 1177 (1989).

124 *See* CASS R. SUNSTEIN, ONE CASE AT A TIME: JUDICIAL MINIMALISM ON THE SUPREME COURT 3–60 (1999) (describing and advocating for "judicial minimalism," in which judges decide cases on narrow grounds, leaving many issues undecided).

significant measure of discretion. They may not have strong preferences over some range of cases and can indicate their indifference by articulating a legal rule that gives lower courts substantial leeway in that area.[125] Alternatively, the Justices may not yet know what rule they would like to communicate regarding outcomes in related cases. They may lack sufficient information to form firm preferences regarding outcomes in cases not yet before them, or they may recognize that they do not yet know what the set of related cases will encompass.[126] In such situations, the Court may issue a narrow holding based only on the facts before it, leaving lower courts substantial room to exercise discretion in factually distinguishable cases, rather than a broader rule that substantially limits that discretion in future cases.[127]

Even when the Supreme Court is clear on the policy it wants to implement, it must choose what form its guidance to the lower courts will take. The classic debate over rules versus standards is in large part a disagreement about the amount of discretion appropriately given to lower courts.[128] As Kathleen Sullivan puts it:

> These mediating legal directives take different forms that vary in the relative discretion they afford the decisionmaker. . . . Rules aim to confine the decisionmaker . . . . A rule captures the background principle or policy in a form that from then on operates independently. . . . A legal directive is "standard"-like when it tends to collapse decisionmaking back into the direct application of the background principle or policy to a fact situation. Standards allow

[125] For example, Justice Scalia has expressed the view that the question of what searches are reasonable under the Fourth Amendment is one on which "we can tolerate a fair degree of diversity." Scalia, *supra* note 123, at 1186. R

[126] As Third Circuit Court of Appeals Judge Sloviter wrote: "When circumstances are either so variable or so new that it is not yet advisable to frame a binding rule of law, trial courts may be given discretion until the factors important to a decision and the weight to be accorded them emerge from the montage of fact patterns which arise." United States v. Criden, 648 F.2d 814, 818 (3d Cir. 1981).

[127] Note that this account differs from McNollgast's explanation as to why "doctrinal intervals" exist. *See supra* notes 61–63 and accompanying text (describing McNollgast's theory). They argue that doctrinal intervals are "part of the equilibrium interaction among the Supreme Court and the lower courts," with the Supreme Court setting the width of the interval in order to "induce the optimal pattern of compliance among the lower courts." McNollgast, *supra* note 2, at 1646. Thus, in their account, the Supreme Court *has* a fixed preference, but it nevertheless communicates a range of acceptable outcomes as a strategy to induce compliance. By contrast, the account I offer here admits the possibility that the Supreme Court Justices, individually or as a body, may be indifferent or uncertain of their preferences over a range of outcomes. R R

[128] It is also a disagreement about how much discretion a court should grant *itself* in the future. *See* Scalia, *supra* note 123, at 1179–80 ("[W]hen . . . I adopt a general rule . . . I not only constrain lower courts, I constrain myself as well. . . . Only by announcing rules do we hedge ourselves in."). R

> for the decrease of errors of under- and over-inclusiveness by giving the decisionmaker more discretion than do rules.[129]

Of course, "rules" and "standards" are not clearly demarcated categories. Legal directives exist on a continuum and may be more or less rule-like.[130] Where they should fall on that continuum, however, is essentially an argument about the values served or defeated by permitting discretion. By reducing discretion, rules reduce the risk of arbitrariness and promote consistency and predictability of outcomes.[131] On the other hand, standards promote fairness by permitting relevant similarities and differences to be taken into account. By granting the decisionmaker discretion, standards avoid the substantive unfairness and arbitrariness that sometimes result when rigid rules are applied in unforeseen circumstances. Unlike rules, standards are "flexible and permit decisionmakers to adapt them to changing circumstances over time."[132]

Individual Justices famously disagree about whether rules or standards are preferable. Justice Scalia explicitly laid out the case for preferring rules in a Holmes lecture given at Harvard University, and concluded by urging that "the *Rule* of Law, the law of *rules*, be extended as far as the nature of the question allows."[133] In deciding particular cases, he frequently advocates a rule-like approach over balancing or totality-of-the-circumstances tests.[134] By contrast, Justice O'Connor often resisted sweeping rules. For example, she wrote that "[e]xperience proves that the Establishment Clause, like the Free Speech Clause, cannot easily be reduced to a single test,"[135] and warned against "the flaws and dangers of a Grand Unified Theory that may turn out to be neither grand nor unified."[136] Even when applying rule-like directives, she denied that they could mechanically determine results, arguing, for example, that "strict scrutiny" was not

129 Kathleen M. Sullivan, *Foreword: The Justices of Rules and Standards*, 106 HARV. L. REV. 22, 57–59 (1992).

130 *Id.* at 61 (describing distinction between rules and standards as continuum, not divide).

131 *See id.* at 62–66 (summarizing arguments in favor of rules).

132 *Id*. at 66.

133 Scalia, *supra* note 123, at 1187. R

134 *See, e.g.*, Bendix Autolite Corp. v. Midwesco Enters., Inc., 486 U.S. 888, 897–98 (1988) (Scalia, J., concurring) (arguing that balancing governmental interests against needs of interstate commerce is responsibility of Congress, not judiciary).

135 Bd. of Educ. v. Grumet, 512 U.S. 687, 720 (1994).

136 Rosenberger v. Rector & Visitors of Univ. of Va., 515 U.S. 819, 852 (1995) (O'Connor, J., concurring).

"strict in theory, but fatal in fact," but required taking "relevant differences into account."[137]

Regardless of the reasons, the Supreme Court's opinions inevitably afford lower court judges some degree of discretion in deciding cases. And the existence of that discretion means that mere variability in outcome—even along political lines—is not necessarily inconsistent with lower court compliance with legal norms. This account of how law shapes lower court decisionmaking does not preclude the influence of other motivations. Factors such as judges' policy goals, their social background or life experiences, and their desire to clear their dockets are likely to play a role in the interstices of the law. The important point here is that the exercise of independent judgment by lower court judges may be fully consistent with legal norms requiring compliance with superior court precedent.

B. Discretion and Power

When existing legal authority permits lower court judges to exercise judgment, it affords them a type of power—the "power to choose."[138] But the power of the judge is also shaped by the institutional context in which she decides. Her ability to achieve her goals in judging depends not only on the substantive rules of law applicable to a given case, but also on where she sits in the judicial hierarchy. Principal-agent models have highlighted the importance of interactions between superior and inferior courts in shaping the decisions of each, and they have typically focused on the extent to which appellate courts are able to exercise control through their power of reversal. What these models often overlook, however, is that legal rules also *restrain* the use of that reversal power by reviewing courts.

Perhaps the most significant restraint is contained in 28 U.S.C. § 1291, which grants courts of appeals jurisdiction only over appeals from final decisions of the district courts.[139] Many types of trial court decisions are nonfinal, and, unless they fall within a handful of exceptions,[140] they cannot be immediately appealed even when they may have a significant impact on the course of the litigation.[141] Thus, a

[137] Grutter v. Bollinger, 539 U.S. 306, 326–27 (2003) (quoting Adarand Constructors, Inc. v. Peña, 515 U.S. 200, 228, 237 (1995)) (internal quotation marks omitted).

[138] HART & SACKS, *supra* note 100, at 144. R

[139] 28 U.S.C. § 1291 (2000). A final decision is "one which ends the litigation on the merits and leaves nothing for the court to do but execute the judgment." Catlin v. United States, 324 U.S. 229, 233 (1945).

[140] For example, 28 U.S.C. § 1292(a) (2000) permits appeals from nonfinal orders involving injunctions, receiverships, and admiralty cases.

[141] Although in theory interlocutory appeal is possible pursuant to § 1292(b), a district court must first state in writing that an otherwise unappealable order "involves a control-

defendant may be forced to litigate a suit even if the trial court erroneously ruled that it had personal jurisdiction over the defendant or denied a motion to dismiss that it should have granted. Some decisions, if not immediately appealed, may become effectively unreviewable.[142] For example, trial court rulings on issues such as the joinder of claims or parties, the discoverability of confidential information, or the bifurcation of issues at trial cannot be appealed until after a final judgment has been rendered. By delaying the appeal, the final judgment rule creates the possibility that any errors made may become moot if the case is settled or if the objecting party ultimately prevails on the merits anyway. Even when the objection is preserved, an appellate court may be reluctant to void a judgment reached after a case was fully tried on the basis of erroneous rulings early in the proceedings. This reluctance is consistent with the rule that appellate courts should not reverse for "errors or defects which do not affect the substantial rights of the parties."[143]

Legal rules prescribing standards of review also require reviewing courts to exercise self-restraint in the use of their reversal power in certain circumstances. Depending upon the type of decision appealed, the appellate courts take a more or less deferential stance toward trial court decisions. At one end of the spectrum, questions of law are reviewed de novo with the court of appeals taking a fresh look at the arguments on both sides and reaching its own judgment, without according the trial court's opinion any particular weight. At the other extreme, certain decisions—most often those relating to the management of the litigation at the trial level—are reviewed under a highly deferential "abuse of discretion" standard. Under this standard, the trial court's decisions are not wholly insulated from review; however, the appellate court should not reverse merely because it would have decided the issue differently. According to one classic formulation of the "abuse of discretion" standard, a trial court decision "cannot be set aside by a reviewing court unless it has a definite and firm conviction that the court below committed a clear error of judgment in the

ling question of law as to which there is substantial ground for difference of opinion and that an immediate appeal from the order may materially advance the ultimate termination of the litigation," and the court of appeals must agree to hear the appeal. In fact, few interlocutory orders are ever reviewed pursuant to § 1292(b). *See* Michael E. Solimine, *Revitalizing Interlocutory Appeals in the Federal Courts*, 58 GEO. WASH. L. REV. 1165, 1174 (1990) (reporting that in 1980s about 35% of § 1292(b) appeals were accepted, representing only about 0.3% of appeals).

142 *See* Stephen C. Yeazell, *The Misunderstood Consequences of Modern Civil Process*, 1994 WIS. L. REV. 631, 660–66 (explaining how expanded pretrial activity and rules of appeal have created significant set of lower court rulings that are effectively unreviewable).

143 28 U.S.C. § 2111 (2000).

conclusion it reached upon a weighing of the relevant factors."[144] Thus, deferential standards of review also create discretionary space by allocating decisionmaking authority between trial and appellate courts.

A number of legal scholars have distinguished the discretionary power afforded by the open texture of substantive legal rules from that afforded by the rules of review. For example, Maurice Rosenberg has differentiated primary from secondary discretion. Primary discretion exists when the judge has "a wide range of choice . . . free from the constraints which characteristically attach whenever legal rules enter the decision process."[145] As discussed in the previous section, this type of discretion exists because legal authority does not provide a determinate answer in every case. The existence of primary discretion thus turns on the applicable substantive legal rules. Secondary discretion, by contrast, "has to do with hierarchical relations among judges" and is a "review-restraining concept."[146] It exists when a particular court's decision is insulated to some degree from reversal, even if erroneous, because the reviewing court applies a deferential standard of review. As Rosenberg expressed it, secondary discretion grants the judge a limited "right to be wrong without incurring reversal."[147]

Judge Henry Friendly similarly distinguished two uses of the word discretion. One use, which corresponds with Rosenberg's "primary discretion," concerns the normative question faced by lawmakers choosing between rigid rules that promote certainty of results and more flexible standards that permit courts to exercise judgment.[148] The other use, like Rosenberg's "secondary discretion," deals with allocation of power within the judicial system, "namely how far an appellate court is bound to sustain rulings of the trial judge which it disapproves but does not consider to be outside the ballpark."[149] Thus, primary or normative discretion is a function of the inherent limitations of substantive legal rules and is present "at every level of the system."[150] Secondary or allocative discretion exists as "a matter of institutional arrangements and power relationships."[151] It too is a product of legal rules—specifically, the rules regulating when and how appellate courts should exercise their powers of

144 *In re* Josephson, 218 F.2d 174, 182 (1st Cir. 1954).

145 Rosenberg, *supra* note 99, at 637. R

146 *Id*.

147 *Id*. at 641.

148 Henry J. Friendly, *Indiscretion About Discretion*, 31 EMORY L.J. 747, 754–55 (1982).

149 *Id*. at 754.

150 Rosenberg, *supra* note 99, at 638. R

151 Yablon, *supra* note 105, at 250. R

review[152]—and can expand the power of trial courts by restraining that of appellate courts.

As Friendly's account suggests, these different meanings of discretion distinguish two distinct types of normative questions of concern to legal scholars. Jurisprudentially-inclined scholars focus on the extent to which judges have freedom in deciding cases (primary discretion), ask what doctrinal forms (rules or standards) are optimal, and how judges should decide cases when the doctrine gives them discretion. Process-minded scholars focus instead on which level of court is in a better position to decide what types of questions, and when appellate courts should defer to the judgments of trial courts (secondary discretion). *Both* types of discretion, however, are relevant to understanding lower court judging, because it is the combination of the two that defines the environment in which lower courts decide.

When deciding cases, judges do at least two things: They determine the outcome of the dispute before them, and they offer reasons for their decision that connect the facts of the case to applicable legal doctrine. Empirical scholars typically focus on outcomes when studying the behavior of judges; however, the second aspect of judging—reason giving—is a crucial part of what judges do. Just as when they decide outcomes, judges may exercise discretion when they explain the rationales for their decisions, for they are choosing which legal rules are relevant or articulating what those rules demand in a particular situation. In what follows, I explore a variety of situations in which lower court judges have discretionary power and consider how they might exercise it. In doing so, I examine both types of discretion discussed above and both aspects of judging. More specifically, I suggest that different kinds of discretion will afford different types of power, which in turn will influence how judges ultimately choose to exercise their discretion.

1. Novel Legal Issues

For both district and appellate courts, novel legal issues create primary discretion. Novel legal issues arise because changes in the

[152] Dworkin has also described several types of discretion. One form of what he called "weak" discretion roughly corresponds to Rosenberg's "primary discretion" and Friendly's "normative discretion" in that it refers to situations in which "the standards an official must apply cannot be applied mechanically but demand the use of judgment." DWORKIN, *supra* note 21, at 31. Dworkin has also referred to another form of "weak" discretion, which exists for a court of last resort whose decisions are nonreviewable as a practical matter, and to a form of "strong" discretion, which exists when an official is "simply not bound by standards set by . . . authority." *Id*. at 32. These latter meanings of the term "discretion" in Dworkin's work are not relevant to understanding decisionmaking by lower federal courts. R

law or technological or societal developments create unanticipated situations not addressed by any existing statute or precedent. For example, when Congress passed the Civil Rights Act of 1991, it failed to address the statute's retroactive effect, leaving lower courts to decide the question in the first instance.[153] Similarly, technological developments raised the question of whether accessing stored email messages met the definition of an "intercept" under the Electronic Communications Privacy Act (ECPA).[154] In each of these situations some relevant legal materials existed, such as the legislative history of the Civil Rights Act, or cases interpreting the term "intercept" in other contexts. However, lower courts confronting such questions for the first time had discretion in deciding them because the Supreme Court had not yet addressed them.

Even when the Supreme Court has definitively addressed a particular issue, some ambiguity will remain regarding its application. Thus, for example, the Supreme Court in *Miranda v. Arizona* held that incriminating statements produced as a result of custodial interrogation are inadmissible against a criminal defendant at trial unless the suspect knowingly and voluntarily waived his Fifth Amendment rights.[155] Although *Miranda* articulated a clear rule, even clear rules are ambiguous at their boundaries, and lower court judges therefore retained some discretion in determining what conditions constituted "custodial interrogation." For example, the lower court judges had primary discretion to decide whether questioning a suspect who is not "significantly restrained" is or is not a custodial interrogation[156]—subject to revision of the rule by a higher court.

When lower courts confront novel legal issues, they have the power not only to determine the outcome of the case before them, but also to declare law. Deciding the legal issue will obviously affect the outcome in the pending case, but the impact of the decision—the choice of a rule—will also influence the outcome of future cases and, perhaps more to the point, primary behavior outside the legal system. So, for example, if accessing stored email messages is held not to be an "intercept" under the ECPA, not only will fewer (or no) cases be

153 *See* Landgraf v. USI Film Prods., 511 U.S. 244, 286 (1994) (explaining that Civil Rights Act did not expressly address issue of its retroactive effect and concluding that provisions relating to damages should not apply to cases arising before its enactment).

154 18 U.S.C. §§ 2510–2522 (2000 & Supp. III 2004). *See generally* Katherine A. Oyama, *E-Mail Privacy After* United States v. Councilman*: Legislative Options for Amending ECPA*, 21 BERKELEY TECH. L.J. 499 (2006) (reviewing case law interpreting meaning of "intercept" under ECPA).

155 384 U.S. 436 (1966).

156 *See, e.g.*, United States v. Gibson, 392 F.2d 373, 376 (4th Cir. 1968) (affirming district court ruling that questioning by police on sidewalk was not custodial interrogation).

brought alleging violation of the ECPA on that basis, but companies are more likely to review stored email messages over which they have control. The lower court judge who finds herself with discretion to decide a novel legal issue may be more motivated by the opportunity to have a broader impact through the articulation of a legal rule than to affect the outcome in the particular case.

For district courts, the power to declare law is quite limited. They lack the authority to review the decisions of other courts and their opinions do not constitute binding authority on any other court. Although other judges may consider their reasoning when confronted with the same issue, their legal conclusions will be reviewed de novo if appealed and superceded once the appeals court decides the issue. Despite these limitations, the opportunity to decide novel issues of law in the first instance offers a district court the potential to influence the development of the law. Its opinion will not be the final word, but it has the advantage of speaking first on the issue. In doing so, it has the opportunity to shape, at least initially, how the issue will be framed and thereby to identify the range of relevant arguments and possible outcomes.

For courts of appeals, cases presenting novel or ambiguous issues of law afford considerable lawmaking authority. The decisions of an appellate panel create precedent binding on all district courts and subsequent appellate panels within the circuit unless reversed by the circuit court sitting en banc or by the Supreme Court. Because the Supreme Court's limited resources only allow it to review a fraction of appellate cases, a court of appeals' resolution of an issue will often be the final word in that circuit for a considerable period of time.

For the individual court of appeals judge, however, this potential lawmaking authority is constrained by the fact that she sits on a collegial court. In order to influence the direction of the law, she must convince at least one of the other judges on her panel (or in the case of an en banc hearing, a majority of her circuit colleagues) to join her reasoning. A number of empirical studies have documented the importance of panel effects, demonstrating that the identity of the colleagues with whom an appellate judge sits significantly influences her vote.[157]

[157] *See, e.g.*, Cross & Tiller, *supra* note 2, at 2172 (finding that presence of potential "whistleblower" on ideologically divided court of appeals panels improves chances court will comply with legal doctrine rather than pursuing partisan policy preferences); Revesz, *supra* note 3, at 1764 ("[W]hile individual ideology and panel composition both have important effects on a judge's vote, the ideology of one's colleagues is a better predictor of one's votes than one's own ideology."); Sunstein et al., *supra* note 44, at 315–25 (finding in **R** **R** **R**

2. *Applying Law to Facts*

Lower courts—both district and appellate—also have discretionary power when the relevant legal standard requires them to exercise judgment. The legal rule may be quite clear; however, the lower court judge has discretion to decide the case because the relevant standard is indeterminate until applied to a particular set of facts.[158] Balancing tests provide an obvious example. When a public employee claims her termination violated the First Amendment, the Supreme Court requires a balancing of her interests "in commenting upon matters of public concern" against the state's interest "in promoting the efficiency of the public services it performs through its employees."[159] Framing the question in this manner affords decisionmaking power to lower courts because individual cases might be resolved in different ways while nonetheless complying with the applicable legal test. So long as the relevant factors are fairly taken into account, the lower court judge can be said to have applied the law, whether or not she finds a First Amendment violation has occurred.

The discretion inherent in the task of applying legal doctrine to concrete facts primarily gives lower courts power over the outcome in the particular case before them. They are bound to follow the legal test established by the superior court but exercise control over case disposition in the way they apply the test. Applying law to fact, however, can also involve a modest form of lawmaking. Using the example of the First Amendment rights of public employees again, suppose a lower court holds that if the public employee's speech involves interrupting the work of a co-employee, it is not protected because the state's interest in efficiency will outweigh the employee's interest in speaking. Although the holding may not be framed in terms of a formally binding rule (the relevant doctrine remains the generalized balancing test), a subsequent court confronting a case involving the same fact (interrupting the work of a co-employee) may be more likely to reach the same outcome, particularly if the first court is one with revisory authority over the deciding court. In this way, even when "merely" applying established law to specific facts,

variety of substantive areas that judges' ideological voting patterns are influenced by party affiliation of other two judges on panel).

158 The Supreme Court describes such issues as mixed questions of law and fact: "questions in which the historical facts are admitted or established, the rule of law is undisputed, and the issue is whether the facts satisfy the statutory standard, or to put it another way, whether the rule of law as applied to the established facts is or is not violated." Pullman-Standard v. Swint, 456 U.S. 273, 289 n.19 (1982).

159 Pickering v. Bd. of Educ., 391 U.S. 563, 568 (1968).

lower courts may have an opportunity to influence the decisions of later courts.

The allocation of discretionary power between district and appellate courts over mixed questions of law and fact is somewhat unclear. No definitive rule establishes the appropriate standard of review for mixed questions across all types of cases. Rather, the rule varies by circuit and appears to turn to some extent on the particular question to be decided.[160] The more a decision turns on the resolution of factual issues, the more likely it is the appellate court will defer to the trial court decision; conversely, if the central issue is what legal implication follows from established facts, the appellate court is more likely to consider the issue de novo. Because of uncertainty over the standard, district and appellate courts may struggle over how mixed questions are characterized in order to expand their discretionary authority in a particular case.

3. Questions of Fact

Before law can be applied to facts, someone must decide what the facts are. If a case goes to trial, the fact finder is likely to be a jury, although in a significant number of cases it will be a judge.[161] When district court judges act as fact finders, they obviously wield considerable discretion. Fact-finding requires a judge to hear testimony, review documentary and physical evidence, and draw inferences. Although the law provides some guidance through both substantive and evidentiary rules, the process of fact-finding is largely an exercise of independent judgment. This process affords the trial judge significant discretionary power over the outcome of the case before it. This power, however, is a limited one, as the factual findings in one case will influence neither the outcome in other cases nor the development of the law more generally.

The district court's discretion to make factual findings is reinforced by the rules of review. Federal Rule of Civil Procedure 52(a) states that "[f]indings of fact . . . shall not be set aside unless clearly erroneous."[162] Thus, an appellate court cannot set aside the trial judge's finding of facts, even if it would have reached a different con-

160 *See* 19 MOORE ET AL., *supra* note 19, § 206.04[3][a], [b] (discussing varying standards of appellate review for mixed questions of law and fact). R

161 *See* Marc Galanter, *The Vanishing Trial: An Examination of Trials and Related Matters in Federal and State Courts*, 1 J. EMPIRICAL LEGAL STUD. 459, 462 tbl.1 (2004) (reporting data on federal civil trials in 2004 of which approximately one-third were bench trials).

162 FED. R. CIV. P. 52(a).

clusion on the same evidence.[163] The Supreme Court has explained the reasons for this allocation of fact-finding authority: Trial judges are not only in a better position to assess the credibility of live witnesses, but they also have superior expertise and greater experience determining facts.[164] Thus, the discretionary power afforded by courts' fact-finding function is primarily allocated to the trial courts.

4. Managerial Decisions

District court judges also exercise a great deal of discretion over nonfinal decisions made in the course of managing the litigation of individual cases. In civil cases, district courts routinely make decisions such as whether to permit the joinder of new claims; whether contested information is discoverable; or whether to bifurcate or consolidate the trial of issues or claims. Although these decisions are not completely unrestrained, the applicable legal rules typically frame the trial judge's decision in a way that affords a great deal of primary discretion. For example, the judge should permit amendments to the pleadings "when justice so requires,"[165] limit discovery when "the burden or expense . . . outweighs its likely benefit,"[166] and conduct separate trials "to prevent delay or prejudice."[167] These very open-ended standards afford district court judges significant power in shaping the course of litigation and, potentially, its ultimate outcome. The effect of this discretionary power is limited, however, to the immediate case. These decisions rarely result in publicly available opinions and may not even be accompanied by written reasons. By their nature, they turn on highly case-specific considerations and will likely have no influence on the conduct of future cases.

Although managerial decisions entail considerable power over the outcomes in particular cases, this power is concentrated primarily in the district courts. Such decisions typically do not satisfy the "final judgment" rule,[168] and therefore, they are rarely subject to review by an appellate court. Given the sheer volume of such decisions in the

---

163 "Where there are two permissible views of the evidence, the fact finder's choice between them cannot be clearly erroneous." Anderson v. City of Bessemer City, 470 U.S. 564, 574 (1985).

164 The Supreme Court has noted: "The trial judge's major role is the determination of fact, and with experience in fulfilling that role comes expertise. Duplication of the trial judge's efforts in the court of appeals would very likely contribute only negligibly to the accuracy of fact determination at a huge cost in diversion of judicial resources." *Id.* at 574–75.

165 FED. R. CIV. P. 15(a).

166 FED. R. CIV. P. 26(b)(2)(iii).

167 FED. R. CIV. P. 20(b).

168 28 U.S.C. § 1291 (2000).

district courts and the low probability of interlocutory appeal,[169] district court judges exercise largely unreviewed discretion in these areas. And in the rare cases when these types of decisions are immediately appealed, the reviewing court uses a highly deferential "abuse of discretion" standard.[170] Thus, appellate courts have limited ability to affect the outcomes of cases through review of these types of decisions.

The situations involving discretionary judgments are not always neatly demarcated. Litigants or judges may dispute whether a particular issue involves finding facts or applying law to facts. The important point, however, is that discretion is an inherent part of the lower courts' decisionmaking environment, both because of the nature of legal rules and because of the norms governing the judicial hierarchy. Moreover, the particular legal and institutional context is critical because it will shape what type of discretionary power is afforded a lower court judge, which may in turn influence how she chooses to exercise that power.

III
REVISITING THE PRINCIPAL-AGENT MODEL

A. Positive Implications

1. Testing the Assumptions

As discussed above, the principal-agent theories used to model the judicial hierarchy typically assume that legal doctrine is the means by which a superior court signals its preferences to lower courts.[171] On this account, lower court judges do not follow precedent because they view legal rules as binding, but because they fear sanction by their superiors in the form of a reversal. I have argued that a more plausible model starts with the assumption that law and legal norms create binding obligations on judges that shape their behavior independent of any fear of reversal. This binding quality of law is consistent with the variations in outcomes observed in many empirical studies, because the law inevitably affords lower court judges some measure of discretion in deciding cases.

Which of these competing assumptions is correct is ultimately an empirical issue, albeit a particularly intractable one to sort out. Asking judges why they decide as they do is not likely to resolve the

169 *See supra* note 141. R

170 *See* Yeazell, *supra* note 142, at 665 ("Many matters of pretrial process, ranging from the conduct of settlement negotiations to discovery to permissive joinder, are reviewed, when they are reviewed at all, under the 'abuse of discretion' standard."). R

171 *See supra* Part I.C.

question, given that they may have reason to conceal their true motivations.[172] And a problem of behavioral equivalence often plagues efforts to discern motivation by examining what judges actually do. Consider a situation in which the Supreme Court decides an issue on which the lower courts have been divided, such as whether the exclusionary rule applies if the police fail to "knock-and-announce" prior to executing a search warrant. The Supreme Court recently held that violation of the "knock-and-announce" rule does not necessarily require suppression of evidence found in the ensuing search,[173] and lower courts will likely confront similar situations in subsequent cases. If lower court judges feel bound to follow the Court's precedent because of legal norms, even those lower court judges who earlier demonstrated that they preferred the opposite rule will be observed to follow the Supreme Court's ruling. If, on the other hand, the Supreme Court's decision functions merely as a signal to lower court judges and they fear reversal, they will also be observed to follow the rule despite having expressed a contrary preference in the past. Thus, studying lower court behavior in response to a new or changed legal rule cannot necessarily distinguish the two explanations.

McNollgast's theory suggests a specific situation in which competing assumptions about the role of law might be tested. Recall that they model doctrine as "part of the equilibrium interaction among the Supreme Court and the lower courts, each acting to maximize its own preferences or ideology."[174] As described above in Part I, they posit that the Supreme Court will create a doctrinal interval around its preferred outcome in order to induce lower courts with preferences near the boundaries of that interval to comply. According to McNollgast, the Court manipulates the doctrinal interval to induce an optimal level of compliance, expanding the interval in order to maintain control when faced with a significant number of noncomplying lower courts.[175]

One implication of McNollgast's theory is that if many lower courts disagree with the Supreme Court's preferred policy, a Court decision narrowing a doctrinal interval could result in *decreased* compliance by lower courts. If the doctrinal change moves the outer edge of the acceptable interval too far from a lower court's ideal point, it may be able to maximize its preferred policy outcomes by *not* com-

172 SEGAL & SPAETH, *supra* note 1, at 33. R

173 Hudson v. Michigan, 126 S.Ct. 2159, 2165 (2006) ("Since the interests that [are violated by non-adherence to the rule] have nothing to do with the seizure of the evidence, the exclusionary rule is inapplicable.").

174 McNollgast, *supra* note 2, at 1646. R

175 *Id.* at 1646.

plying. In such a situation, the lower court would decide the case according to its own preferences and face some chance (significantly less than one) of reversal, rather than accept a certain outcome far from its ideal point. If, on the other hand, lower court judges are motivated by a legal norm requiring them to follow precedent, a Supreme Court decision that narrowed the range of acceptable outcomes would be predicted to induce *greater* compliance, even when there are a large number of lower courts with differing preferences. Such an opinion would reduce the area of discretion afforded lower court judges, leading them to conform more closely to the Supreme Court's preferred outcome.

For many purposes, it does not matter whether judges follow precedent primarily because they fear reversal or because they are adhering to legal norms. When the Supreme Court speaks, the lower courts will (largely) follow. However, on the margins, the extent to which lower court judges follow precedent depends upon their reasons for doing so. More importantly, the differing accounts of judicial motivation may suggest differing prescriptions for structuring judicial institutions. For example, if fear of reversal explains lower court compliance with precedent, then it may make sense, as some scholars have proposed, to require that circuit court panels be ideologically mixed.[176] The presence of a judicial colleague with a different ideological perspective might induce greater compliance with the law, because ideologically motivated judges would realize that a dissenting colleague could act as a "whistleblower," drawing attention to any deviation from precedent and increasing the risk of reversal by the Supreme Court.[177] On the other hand, if legal norms rather than fear of reversal motivate lower court compliance, then such a proposal may be counterproductive. By making salient an ideological role for the judge, the practice may have the effect of weakening legal norms and thereby reducing the likelihood of compliance over the long run.

2. Law Matters

If in fact law independently influences lower court judges, then understanding their decisionmaking environment requires taking account of the ways in which legal doctrine removes or creates discretionary space. One implication is that large "n" empirical studies of court decisionmaking that lack controls for legal factors will be of lim-

[176] *See, e.g.*, Emerson H. Tiller & Frank B. Cross, *A Modest Proposal for Improving American Justice*, 99 COLUM. L. REV. 215, 216 (1999) (proposing requirement that every three-member circuit court panel include both Republican- and Democrat-appointed judges).

[177] *Id.* at 228–29.

ited usefulness in advancing our understanding of judicial decisionmaking. They may show that ideological factors are at work, but cannot indicate the contexts in which those factors are most significant or how they interact with legal norms. Rather, paying attention to the legal context—specifically, the discretionary spaces created by doctrine—is crucial to untangling those interactions. In this section, I explore several ways in which this perspective could help frame empirical studies of lower court decisionmaking and some of the challenges that remain.

First, focusing on discretion suggests that legal norms will have varying force in different contexts. One obstacle to incorporating the assumption that legal norms are independently binding has been the difficulty of developing appropriate measures of judges' legal preferences. Empirical research has tended to focus on ideological explanations because of the relative ease of developing measures of judges' political preferences. One simple measure, used by both political scientists and legal scholars, looks at the party of the President who appointed a federal judge as an indicator of that judge's ideology.[178] A recently developed alternative measure utilizes information about the political preferences of both the President and the senators involved in the appointment process,[179] offering a more finely calibrated measure of ideology than the simple binary classification of Republican- or Democrat-appointed judge. Although these measures are not perfect, the challenge of developing a comparable scale of judges' legal preferences is daunting. Even if judges share a preference for complying with legal authority, they will differ regarding how best to "follow the law." In interpreting statutes, some may be strict constructionists, whereas others may be more willing to examine legislative history. Some may be more likely to see their decisions as bound by higher court precedent, while others will read those decisions narrowly, leaving them substantial discretion to decide the issue before them. Thus, legal preferences, like political preferences, are likely to vary across judges, and there is currently no way to observe directly the differing legal preferences of individual judges.

---

[178] Thus, federal judges appointed by Republican presidents are assumed to be conservative; those appointed by Democratic presidents are assumed to be liberal. Epstein and King have criticized this measure as overly simplistic, as it assumes that all Republican presidents are equally conservative (and all Democratic presidents equally liberal) and ignores the role of senatorial courtesy in the appointment process. Lee Epstein & Gary King, *The Rules of Inference*, 69 U. CHI. L. REV. 1, 88–89 (2002).

[179] *See* Micheal W. Giles, Virginia A. Hettinger & Todd Peppers, *Picking Federal Judges: A Note on Policy and Partisan Selection Agendas*, 54 POL. RES. Q. 623, 631, 638 (2001) (measuring judicial preference based on ideological scores of President and senators involved in appointment process).

Legal environments also differ in the amount and type of discretion they afford lower court judges. Empirical studies might take advantage of the fact that even though individual judges may vary in their interpretive philosophies, the aggregate behavior of judges is likely to differ systematically in different legal environments. More specifically, one might predict that lower court judges will behave differently in situations in which little or no discretion exists than in situations in which they have a great deal of discretion. When deciding cases in which a clear rule applies, one might expect to see very little variation in the outcomes. On the other hand, if legal authority allows a great deal of discretion—as when district courts are called on to interpret a new statute—one would predict significantly more variation. Because legal norms by definition have less force where discretion exists, other motivations will likely play a stronger role. To the extent that those motivations include policy preferences, the variation that is observed will correlate with the judges' ideology.[180]

This approach to observing legal preferences has its own difficulties. The prediction that a greater degree of variance in outcome will be observed if greater discretion exists only holds true if all else is equal. However, due to selection effects all else is not likely to be equal. Under different legal regimes, litigants will make different choices about which cases to bring and which cases to settle, and the resulting mix of cases heard by judges will affect what outcomes are observed.[181] One possible method of dealing with selection effects is to control for the presence or absence of relevant facts. A number of scholars have taken this approach recently, identifying the relevant facts from precedential court opinions and then testing whether a

[180] This expectation is consistent with the results of empirical studies that have controlled for both legal and ideological factors. *See, e.g.*, sources cited *supra* note 47. R

[181] Priest and Klein's well-known article on selection effects focused attention on the likelihood that litigated cases are unrepresentative of all filed cases because the parties' settlement decisions will not be random. *See generally* George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1 (1984). Since their article appeared, an immense literature, both theoretical and empirical, has explored the selection effects of parties' settlement behavior. *See, e.g.*, Lucian Arye Bebchuk, *Litigation and Settlement Under Imperfect Information*, 15 RAND J. ECON. 404, 407–09 (1984) (exploring settlement decisions under conditions of asymmetric information); Daniel Kessler, Thomas Meites & Geoffrey Miller, *Explaining Deviations from the Fifty-Percent Rule: A Multimodal Approach to the Selection of Cases for Litigation*, 25 J. LEGAL STUD. 233, 238–41 (1996) (reviewing numerous empirical studies testing Priest-Klein hypothesis); Steven Shavell, *Any Frequency of Plaintiff Victory at Trial Is Possible*, 25 J. LEGAL STUD. 493, 495–98 (1996) (same); Kathryn E. Spier, *The Dynamics of Pretrial Negotiation*, 59 REV. ECON. STUD. 93, 95–96 (1992) (developing dynamic model of settlement choices over time). This literature, however, has been largely passed over in empirical studies of judicial decisionmaking. *See* Kastellec & Lax, *supra* note 77, at 1 (observing that empirical studies of Supreme Court usually note problem of selection bias only in passing). R

case-facts model can explain lower court decisions.[182] Assuming that the relevant case facts can be identified,[183] this approach permits observation of variations in outcomes while controlling for a changing case mix.

Another obstacle to observing legal preferences across different legal contexts is the difficulty of determining *when* and *to what extent* judicial discretion exists. The amount of discretion a judge may legitimately exercise in a particular situation is often a matter of considerable dispute. Although the presence of discretion is difficult to measure, it may be possible to compare situations in which different degrees of discretion indisputably exist. For example, when the Court changes the governing legal regime from an open-ended standard to a rule, or when a previously ambiguous statutory provision is amended by Congress to clarify its intent, the scope of the lower court's discretion to decide that particular issue is undeniably narrowed. Importantly, such an approach requires close attention to the *content* of Supreme Court opinions, rather than measuring the Court's output solely in terms of the liberal or conservative direction of the outcome, as is typical of many empirical studies. For example, whether the Supreme Court chooses to apply a standard or a bright-line rule will influence how lower courts respond to a decision. The outcome—for example, "defendant wins"—might be the same under either alternative, but the effect on lower court behavior is likely to be quite different. Thus, a remaining challenge to observing behavior across different legal contexts is developing methods for systematically analyzing and coding opinion content.[184]

---

182 *See, e.g.*, Songer et al., *supra* note 2, at 677–90 (controlling for facts of case in testing whether courts of appeals are responsive to changes in Supreme Court doctrine and policy). R

183 Coding for relevant case facts is more difficult than is sometimes acknowledged. The presence of relevant facts cannot be coded from the court opinion to be explained by a case-facts model without risking circularity. Alternative approaches have coded case facts from the opinion of the court below or the parties' briefs. *See, e.g.*, Tracey E. George & Lee Epstein, *On the Nature of Supreme Court Decision Making*, 86 AM. POL. SCI. REV. 323, 328 (1992) (using briefs of parties as primary source of data for coding case facts); Jeffrey A. Segal, *Predicting Supreme Court Cases Probabilistically: The Search and Seizure Cases, 1962–1981*, 78 AM. POL. SCI. REV. 891, 894 (1984) (coding case facts as they were stated in lower court decision). Both approaches have limitations given the motivations a lower court or a party has to characterize the facts in a particular way. Particularly when the "fact" at issue involves a judgment—for example, whether probable cause existed for a search—relying on lower court opinions or briefs as evidence of the existence of that fact is problematic.

184 *See* Scalia, *supra* note 123, at 1177 ("[T]he modern reality . . . is that . . . not merely the *outcome* . . . but the *mode of analysis* . . . will thereafter be followed by the lower courts . . . ."); Frederick Schauer, *Opinions as Rules*, 53 U. CHI. L. REV. 682, 683 (1986) (pointing out that when lower courts apply Court opinions, "it is not what the Supreme R

A second way that paying attention to legal discretion helps frame empirical inquiries focuses on the different types of power afforded lower court judges in different situations. As seen above, a judge's discretionary authority may entail power over case outcomes or law development, depending upon whether she is a district or appellate court judge and whether the decision to be made is primarily one of law, fact, or procedure. Even assuming that judges are motivated in part by their policy preferences, *how* they pursue those preferences will likely depend upon the context. In some situations, a judge may pursue policy goals by focusing on which party wins in a specific case; in other instances, she may seek to advance her policy goals by influencing the development of the law, affecting not just that particular case, but all similar cases in the future.

The judge will pursue different strategies depending upon whether her goal centers on a case outcome or law development. For example, the district court judge concerned primarily about the outcome in a particular case is likely to frame the decision as a matter of factual, rather than legal, dispute, and may not publish her decision, in an effort to insulate the outcome from close scrutiny. On the other hand, if a district court judge wished to influence law development, she would frame the case as presenting a novel legal question, invest time and effort in developing and explaining the rationale for her preferred decision rule, and publish her opinion. When judges follow the latter strategy, they act as judicial entrepreneurs, advancing new approaches and seeking others' acceptance, rather than as agents who attempt to conceal evasion of their duty to obey the principal.

A third way that focusing on discretion might inform models of judicial decisionmaking is by revising our understanding of the impact of reversals. Principal-agent models tend to assume that all reversals are costly to the lower court judge in the same way.[185] If, however, a judge's discretionary authority differs from one context to another, then the impact of reversals might also differ depending upon the context. For example, when a district court judge decides a novel issue of law, she exercises discretionary power, often with the goal of influencing the development of the law. If the circuit court reverses her

Court held that matters, but what it *said*"); James F. Spriggs II & Thomas G. Hansford, *Measuring Legal Change: The Reliability and Validity of* Shepard's Citations, 53 POL. RES. Q. 327, 328 (2000) ("[T]o understand judicial decision-making fully we must move beyond votes and study what is arguably the judiciary's most important policy output—the precedents set by court opinions.").

185 Some scholars have noted the possibility that the manner of reversal may affect the degree of sanction imposed by a reversal. *See* WALTER F. MURPHY, ELEMENTS OF JUDICIAL STRATEGY 104–05 (1964) (suggesting that particularly critical reversal may be more effective sanction than ordinary reversal); Cross, *supra* note 60, at 388 (same). R

decision, she obviously suffers a loss because it did not adopt her preferred rule. However, if her exercise of discretion in that situation was consistent with legal norms—that is, if she chose among one or more *permissible* courses of action—then no stigma or reputational harm would attach to the reversal. Similarly, the court of appeals judge who is reversed by the Supreme Court after deciding an open question of law is unlikely to suffer loss beyond the rejection of her preferred rule. Because these types of reversals often "reflect ideological differences rather than error correction,"[186] they do not necessarily entail criticism of the lower court judge.

In these situations, the risk of reversal is less likely to motivate judges not only because reversal is not particularly costly, but also because the lawmaking opportunity cannot be realized without risking reversal. A district court cannot create binding precedent for any other court, and therefore its opinions can only have influence if they are read by other judges and perhaps adopted by a circuit court. Thus, a district court judge acting as a judicial entrepreneur would prefer the risk of reversal to having her decisions ignored altogether. Circuit courts have much greater lawmaking power, but again, this power cannot be exercised unless their decisions are visible both to the district courts that must follow them and to the other circuit courts that may be persuaded to do so. The goal of avoiding reversal by evading scrutiny is incompatible with acting as a judicial entrepreneur. As Judge Posner observes, "Reversal rate and creativity are likely to be positively correlated, since a judge who is creating precedents rather than just following them can be expected to be reversed more often than the unadventurous judge."[187]

On the other hand, other types of reversals may carry with them a heavier sanction. As discussed above, district court judges have a great deal of secondary discretion when finding facts or managing the litigation. Given the deferential standards of review in these situations, reversal will be relatively rare; however, when a reversal does occur—on the grounds that the court "abused its discretion"—it may have more bite, signaling as it does that the decision was completely out of bounds. An explicit rebuke for failure to conform to legal norms[188] would appear to be the most costly for the lower court judge.

---

[186] Posner, *Judicial Behavior*, *supra* note 53, at 1273. **R**

[187] *Id.* at 1277.

[188] For instance, see *Heath v. Varity Corp.*, 71 F.3d 256 (7th Cir. 1995):

> The district court's approach is not appropriate in a hierarchical judiciary. . . . To avoid heaping needless costs and delay on the litigants, a district court should apply existing precedents while explaining why it believes that innovation is in order. . . . Today's case shows why. The district court erred in

Although such instances are not common, a reversal that includes criticism of the lower court's actions is likely experienced as more costly than an appellate opinion that merely expresses disagreement with its reasoning.

B. Normative Implications

Given that judicial discretion is inevitable, the question of how judges *should* decide cases when existing law is not determinative is a matter of ongoing debate. Applying the principal-agent model to the judicial hierarchy offers an implicit answer to this normative question: Lower courts should decide in accordance with the preferences of the Supreme Court. This section explores how paying attention to the role of discretion in judicial decisionmaking undermines this implicit normative claim. Although the concept of discretion alone cannot support a fully worked-out theory of how judges should decide, it does suggest some considerations that push against the claim that the sole duty of lower court judges is to act as faithful agents of the Supreme Court.

The premise of the principal-agent relationship is that the agent owes a duty to act only in the interests of the principal. Thus, when used as a model for upper-lower court interactions, the model implicitly suggests that lower federal courts should decide cases according to the policy preferences of the Supreme Court.[189] This claim goes beyond the well-accepted legal norm that lower court judges must follow established precedent, suggesting instead that even when precedent is indeterminate, the judge ought to conform her decisions to the desires of the Supreme Court to the greatest extent possible. On this view, when lower court judges exercise their discretion in ways that appear to deviate from those preferences, they are "shirking" their duty.[190] Because the principal-agent model purports to be

> thinking [our prior decision] wrongly decided, this case has been sidetracked, and [the plaintiff] has been put to substantial expense simply to receive the benefit of settled law.

Id. at 257.

189 A recent study suggests that federal courts of appeals respond to the preferences of the contemporaneous Supreme Court, rather than the Court that first established a precedent, a finding consistent with the normative expectations of a principal-agent model. *See* Westerland et al., *supra* note 12, at 13. Caminker has also argued that lower court judges in fact sometimes act in conformity with his "proxy model." *See* Caminker, *Precedent and Prediction*, *supra* note 19, at 75–79. More empirical work remains to be done in order to sort out whether and to what extent lower courts conform to the preferences of the Supreme Court as opposed to responding to its doctrinal statements. In any case, the normative question of how lower court judges *should* decide cases when they have discretion is distinct from the positive question of how they in fact behave. R R

190 *See, e.g.*, Songer et al., *supra* note 2, at 693. R

merely a positive account of judicial behavior, its proponents never explain why lower federal courts owe such a duty of obedience to the Supreme Court.[191]

The principal-agent model was developed in other contexts, most notably to describe economic organizations and the bureaucratic state.[192] In describing the relationships between owners and managers or between Congress and administrative agencies, the model usefully highlights the challenges of aligning the actions of the agent with the interests of the principal and the need for effective mechanisms of supervision and control. In each of these settings, the basis for the agent's duty to act solely in the interest of the principal is clear. The relationship between managers and owners of a firm is established by agreement; the obligation to act in the interests of shareholders is a contractual one. In the case of administrative agencies, Congress creates them specifically in order to carry out its will. Because it owes its existence and its powers to the legislature, the agency's clear duty is to implement congressional policy.

Challenges of supervision and control are also present in the judicial hierarchy; however, the relationship between the lower federal courts and the Supreme Court is not parallel to the agency relationships found within a firm or between agencies and Congress. Lower court judges do not have any contractual relationship with the Supreme Court. Congress, not the Court, is empowered to create (or abolish) the lower federal courts,[193] and the President, with the advice and consent of the Senate, appoints judges to fill those courts.[194] Often, the law those judges are called upon to interpret is created by Congress, not the Supreme Court. When a lower court judge faces a novel issue of statutory interpretation, it would be plausible to view her as an agent of Congress, with a duty to follow its preferences, rather than those of the Supreme Court.

Even when a lower court judge is interpreting the Constitution or prior judicial precedent, it is not self-evident that the Supreme Court

---

191 The oath sworn by federal judges upon taking the bench says nothing about obeying the Supreme Court:

> I, _____ _____, do solemnly swear (or affirm) that I will administer justice without respect to persons, and do equal right to the poor and to the rich, and that I will faithfully and impartially discharge and perform all the duties incumbent upon me as _____ under the Constitution and laws of the United States. So help me God.

28 U.S.C. § 453 (2000).

192 *See supra* notes 36–39 and accompanying text. **R**

193 U.S. CONST. art. III, § 1 ("The judicial Power of the United States, shall be vested in one supreme Court, and in such inferior Courts as the Congress may from time to time ordain and establish.").

194 U.S. CONST. art. II, § 2.

is appropriately regarded as the principal. The judge might equally well be thought of as an agent of the President who appointed her, or of "the law."[195] And when lower federal courts hear diversity cases, it seems strange to conceptualize them as agents of the Supreme Court. If anything, in that context they are agents of the states, as they are required to apply the law of the state in which they sit,[196] and if the law is not clear, to try to determine what the highest court sitting in that state would decide.[197] The critical point here is that *who* or *what* is appropriately thought of as the principal of the lower federal courts is precisely the normative question that needs to be addressed openly. Without a clear answer to that question, the analogy between the judicial hierarchy and traditional agency relationships is incomplete, raising the question of why lower court judges should be obliged to follow the preferences of the Supreme Court when those preferences do not take the form of binding precedential opinions.

Although positive political theorists who utilize the principal-agent model have not explained why lower court judges have a duty to follow Supreme Court preferences, Evan Caminker has laid out a normative case for why they should do so. He argues for a "proxy model" of decisionmaking whereby "an inferior court discharges its duty . . . by applying the dispositional rule that the superior court enjoying revisory jurisdiction predictably would embrace."[198] In other words, the lower court judge should "act as a proxy for its superior court by 'attempt[ing] to replicate the result that would be reached if the [superior court] were faced with the same set of facts and allegations.'"[199]

Caminker argues that the proxy model can be defended as a normative matter because it serves institutional values of judicial economy, uniformity of interpretation, and decisional proficiency.[200] Economy is promoted because, by conforming their decisions to the predicted decision of a superior court, inferior courts will obviate the need for appeal.[201] Even if an appeal is taken, the lower court's judg-

195 Posner, *Judicial Behavior*, *supra* note 53, at 1272. R

196 Erie R.R. Co. v. Tompkins, 304 U.S. 64, 78 (1938) ("Except in matters governed by the Federal Constitution or by Acts of Congress, the law to be applied in any case is the law of the State.").

197 ERWIN CHEMERINSKY, FEDERAL JURISDICTION 326 (4th ed. 2003) (explaining that if no authoritative ruling exists, "[t]he law is now settled that a federal court [in diversity cases] must try to predict how the state's highest court is likely to decide the case . . .").

198 Caminker, *Precedent and Prediction*, *supra* note 19, at 16. R

199 *Id.* at 5 (quoting Earl M. Maltz, *The Concept of the Doctrine of the Court in Constitutional Law*, 16 GA. L. REV. 357, 399 (1982)).

200 *Id.* at 36–43.

201 *Id.* at 36.

ment will be affirmed such that no additional judicial proceedings will be necessary, as would be required if the appellate court had to reverse the decision below. In addition, lower courts following a proxy model will all embrace the same rule on a given issue—specifically, the Supreme Court's predicted rule—thereby ensuring national uniformity in the interpretation of federal law.[202] Finally, Caminker argues that the proxy model enhances judicial proficiency by recognizing the relative competencies of different courts.[203] He asserts that the Supreme Court is more likely to reach a "better" answer to legal questions because its work focuses on legal argument and reasoning rather than fact-finding, and its composition—nine members, rather than the usual three judge court of appeal panel or district court judge sitting alone—"ensures better decisionmaking."[204] By mimicking the Supreme Court's anticipated decision, the lower courts will incorporate the "'better' answer from the beginning, thus improv[ing] decisionmaking at all levels of the judiciary."[205]

Caminker's arguments are based on the paradigmatic case of a lower court judge confronting a novel legal question. However, as discussed above in Part II.B, lower court judges encounter a variety of situations in which they are called upon to exercise their discretion. Even when conforming to the legal obligation to obey superior court precedent, a lower court judge may find that she has a choice among permissible courses of action when applying an open-ended legal standard, finding facts, or managing litigation. In these situations, the values of uniformity and proficiency do not necessarily argue for following a proxy model of decisionmaking.

Whenever the Supreme Court adopts a flexible standard rather than a rigid rule, it is inevitably making a tradeoff between uniformity and other values. If uniformity of outcome were the sole object, a bright-line rule will always come closer to achieving it. By choosing a balancing test or a totality-of-the-circumstances test, the Court is opting to sacrifice some uniformity for the sake of avoiding the over- and under-inclusiveness of a determinate rule. Choosing a "discretion-conferring approach"[206] entails the judgment that fairness is better served in that situation by permitting judges to take a variety of relevant factors into account in each case, rather than by relying on a

202 According to Caminker, uniformity in turn promotes predictability, facilitates public law administration, ensures equal treatment, and secures popular respect for judicial authority. *Id.* at 38–40.

203 *Id.* at 41–42.

204 *Id.* at 42.

205 *Id.* at 41.

206 Scalia, *supra* note 123, at 1177. **R**

determinate rule whose application will prove arbitrary in some circumstances. To argue that lower court judges should exercise their discretion by following the preferences of the Supreme Court is to insist on a level of uniformity that was rejected as undesirable when the flexible standard was adopted in the first place.

In the same way, deferential standards of review grant lower court judges discretion (of the secondary type) because they are better situated to make certain types of decisions, even though such deference may entail a loss of uniformity in outcome. In arguing that the Supreme Court is more proficient at legal reasoning because of its functional role, Caminker also recognizes that district courts are more proficient at fact-finding.[207] Similarly, district courts are in a better position to make decisions regarding the management of litigation. Consistent with this view, institutional norms require appellate courts to give deference to these types of trial court decisions precisely in order to take advantage of their special competencies. It seems perverse, then, to argue, as principal-agent models implicitly suggest, that the district court judge ought to exercise that discretion by deciding the case as her appellate superiors would if they were hearing the case instead.

Even as applied to novel legal questions, Caminker's case for the proxy model is subject to two significant objections. The first objection is a practical one, having to do with the difficulty of prediction. The second questions his claim that the Supreme Court is more likely to reach a "better" answer to legal questions in the absence of any consensus regarding what makes one legal answer "better" than another.

The ability of lower court judges to predict *accurately* the decisions of the Supreme Court is crucial to Caminker's argument in favor of the proxy model. Inaccurate predictions will do little to advance the institutional values of economy, uniformity, and proficiency, yet there is reason to doubt judges' abilities to predict accurately. If, as Caminker suggests, lower court judges lack the Supreme Court's higher proficiency in discerning the "better" answer from primary legal materials, then they are also likely to have difficulty anticipating what the Supreme Court's "better" answer will be.

Nor does experience suggest that prediction will be easy. Lawyers, law professors, and other court watchers have long engaged in the parlor game of predicting the outcomes of cases pending before the Supreme Court, but in recent years, more public efforts at prediction have taken place. For example, during the 2002 Supreme Court

[207] Caminker, *Precedent and Prediction*, *supra* note 19, at 41. **R**

term, legal experts—both law professors and appellate practitioners—were recruited to predict the outcome of every orally argued case pending before the Supreme Court.[208] Their success rate in doing so was hardly impressive—less than sixty percent of the expert predictions regarding case outcomes were correct.[209] Although these efforts did not include judges, the lawyers and law professors attempting to predict the Court's decisions relied on much the same material that a lower court judge might: the Court's opinions in earlier cases, including "well-considered" dicta; the arguments raised by the parties; and information about the Justices' general ideological leanings.

Acknowledging the difficulty of predicting Supreme Court decisions, Caminker argues that the proxy model is nevertheless appropriate in cases in which "highly probative predictive data are available"[210]—for example, "when fragmented dispositional rules or well-considered dicta clearly foreshadow the Supreme Court's future direction,"[211] an admittedly small set of cases.[212] Thus, Caminker's defense of the proxy model, by its own terms, is quite narrow. Even if one fully accepts his arguments, they cannot justify the broader normative claim implicit in principal-agent models that lower court judges should always follow the preferences of the Supreme Court, not merely its written precedent.

The second difficulty with Caminker's defense of the proxy model lies in his claim that the Supreme Court is more likely to reach "better" answers to legal questions than lower courts. This argument implies the existence of some commonly agreed-upon metric by which to judge the quality or correctness of a legal rule, when in fact no such metric exists. Undoubtedly, differences between the institutional setting of the Supreme Court and the lower courts will affect the nature and quality of the opinions they produce; however, which decisions

---

208 *See* Theodore W. Ruger, Pauline T. Kim, Andrew D. Martin & Kevin M. Quinn, *The Supreme Court Forecasting Project: Legal and Political Science Approaches to Predicting Supreme Court Decisionmaking*, 104 COLUM. L. REV. 1150, 1167–69 (2004) (comparing legal experts' predictions of outcomes in Supreme Court cases with predictions generated by statistical model).

209 *Id.* at 1171. The experts' overall lack of success at prediction is more striking given that a statistical model that relied solely on a handful of easily observed case characteristics did better, correctly predicting the outcome in seventy-five percent of the Court's cases during that term. *Id.*

210 Caminker, *Precedent and Prediction*, *supra* note 19, at 73. R

211 *Id.* According to Caminker, "fragmented dispositional rules" exist when several Justices have endorsed a dispositional rule, but the rule is not a binding precedent because it is contained in different opinions, none of which was adopted by a majority of the Court. *Id.* at 17–18.

212 *Id.* at 73 ("This category is likely to be small, but I think not trivial in light of the frequency with which the Court issues split opinions and well-considered dicta.").

are "better" depends upon what values are applied. If, for example, the Supreme Court is comprised of strict constructionists, while judges that believe in a purposive mode of statutory interpretation predominate in the lower courts, then the judgment of which court makes "better" decisions will turn on the observer's own interpretive philosophy. Without some substantive account of what makes one legal answer "better" than another, the argument risks being reduced to a tautology—that is, the decisions of the Supreme Court are deemed better because they are the decisions of the Supreme Court.

If one assumes away the difficulty of prediction for a moment, the normative implications of the principal-agent model become clear. Imagine that lower court judges had accurate information about the Court's preferences across the universe of cases and conformed their decisions accordingly. In the extreme case, the Court's preferences would be completely transparent and lower courts would always comply, leaving nothing for the Supreme Court to do, except to have preferences. We might accept this as a normative ideal in the context of the firm; it would certainly be desirable if managers always acted in the best interests of shareholders rather than engaging in self-dealing. However, its attractiveness as a normative ideal for the judiciary is open to serious question.

What the extreme case highlights is the centralizing tendency of the principal-agent model. To assert that lower courts *should* decide according to the preferences of the Supreme Court, and not just their binding precedents, is to argue for a concentration of power in the hands of the Justices. From this normative perspective, the lower federal courts "function as geographically dispersed extensions of the Supreme Court . . . . [They] are merely intended to facilitate universal access to the Court's edicts."[213] This view may be justified if there are "right" answers, or at least "better" answers, to legal questions and the Supreme Court is more likely to arrive at them. However, judicial scholars who use the principal-agent model typically assume that judges—*especially* Supreme Court Justices—decide their cases in accordance with their preferred policy outcomes, rather than in pursuit of the "right" legal answer.[214] If they are right about the Justices'

[213] *Id.* at 16.

[214] *See, e.g.*, Bueno de Mesquita & Stephenson, *supra* note 4, at 755 (developing theory that assumes appellate judges are policy-oriented); Cameron et al., *supra* note 13, at 114 (modeling certiorari decisions as part of political struggle over doctrine); Jacobi & Tiller, *supra* note 24, at 2 ("[H]igher courts are policy seeking actors who . . . manipulate [doctrine] to maximize their own policy objectives."); McNollgast, *supra* note 2, at 1646 (assuming that Supreme Court acts to maximize its preferences or ideology).

motivations, then a vision of judicial power wholly concentrated in the Supreme Court is normatively troubling.

Of course, the doctrine requiring courts to comply with superior court precedent itself has a centralizing tendency because it locates final authority over law declaration in the Supreme Court. However, even accepting that certain institutional values, such as economy and uniformity, are served by obedience to hierarchical precedent, it does not necessarily follow that the discretionary authority afforded within a precedent-based system ought *also* be exercised in conformity with the Supreme Court's wishes. In other words, while some amount of centralization may be a good thing, complete centralization may not.[215] The presence of discretionary spaces in the lower court judges' decisionmaking authority reflects not only the inherent indeterminacy of language, but also the presence of competing institutional values—ones that push against such goals as uniformity and efficiency. Thus, the reasons that underlie the choice of a flexible standard over a rule, or a deferential standard of review, may also argue against requiring lower court judges to exercise their discretionary power by mimicking the preferences of the Supreme Court.

CONCLUSION

Principal-agent theories have become a common method of modeling relationships between courts in the judicial hierarchy. They improve on the traditional jurisprudential approach—focused on the abstract judge wrestling only with legal materials—by highlighting the importance of institutional context and the influence of policy goals. They also represent an advance from simple attitudinal models that relied solely on a psychological model of judging, without recognizing the possibility that interactions with other actors in the system might influence behavior. Thus, principal-agent models usefully focus attention on interactions between courts within a judicial hierarchy, emphasizing the potential value conflicts that can occur and the challenge of aligning judges' incentives with the needs of a legitimate and well-functioning judicial system.

---

[215] Scholars have debated whether or not it is beneficial to allow legal issues to "percolate" in the lower courts, thereby producing a divergence of approaches which may then inform the Supreme Court's ultimate resolution of an issue. *See, e.g.*, Caminker, *Precedent and Prediction*, *supra* note 19, at 54–61 (summarizing arguments in favor of percolation and arguing that its benefits for Supreme Court decisionmaking are exaggerated). However, apart from whether percolation offers any potential benefits to the Supreme Court's decisionmaking process, a larger question exists regarding the optimal balance between centralization and diffusion of judicial power. R

Those models fall short, however, in their failure to account for law and legal norms. Principal-agent models typically rely on fear of reversal to explain compliance with superior court precedent, yet both practical and theoretical considerations suggest that such an explanation is inadequate. As I have argued here, a more straightforward explanation recognizes that judges likely have preferences for complying with legal norms, as well as preferences regarding policy. Positive political theorists have been skeptical that legal norms in themselves motivate judges, in part because they find formalistic accounts of the law implausible. The law, however, need not be fully determinative in order to have a binding quality. Jurisprudential accounts recognize that law can be both binding on judges *and* permit them to exercise discretion in certain contexts. That discretion exists not only because legal rules will inevitably be indeterminate at some point, but also because social needs demand some measure of flexibility in the application of legal rules, and because institutional values argue for allocating different types of power between different levels of the judiciary.

Of course, all models necessarily simplify a complex reality. That simplification, however, creates the risk that some essential aspect of the process or phenomenon under study will be lost. In the case of principal-agent models, the failure to account for the nature of law and legal institutions obscures important aspects of the interaction between upper and lower courts. Principal-agent accounts emphasize control by the superior court and evasion by "shirking" lower courts. What they overlook is that lower courts sometimes have power over law development as well as case outcomes, and, depending upon the type of power afforded in a particular case, their goals in deciding may include publicity and persuasion rather than evasion.

Finally, principal-agent models elide important normative questions by assuming, without offering justification, that lower federal courts in fact have a duty to follow the preferences of the Supreme Court. Recognizing that legal norms inevitably create discretion for lower court judges highlights the fact that the duty to comply with superior court *precedent* is not the same as the duty to follow its *preferences*. How lower court judges should decide when they have discretion is a difficult and highly contested issue, one that goes to important questions about institutional design and the appropriate balance between centralizing judicial authority and sharing that power between levels of the hierarchy. Acknowledging the implicit assumptions that shape our models—and therefore our understanding of the world—permits us to recognize and address those questions explicitly.