Doctrinal Choice in the Judicial Hierarchy

Jeffrey R. Lax
Department of Political Science
Columbia University
JRL2124@columbia.edu

April 8, 2008

Prepared for the Law and Economics Seminar, University of Texas Law School, April 2008.

Abstract

How does the structure of legal doctrine as chosen by a higher court affect the application of and compliance with that doctrine by lower courts? How does the Supreme Court’s need to rely on the lower courts affect the choice among doctrinal structures? Specifically, how does it drive the choice between determinate doctrines (rules) and more flexible or indeterminate doctrines (standards)? I present a case-space model of both rules and standards, wherein I define a standard as incorporating a factual dimension that is only imperfectly observable or specifiable. I show that the choice between rules and standards will depend on the alignment of preferences between levels of the judicial hierarchy, the transparency of decision-making, the precision with which doctrine can be specified, the sensitivity of the desired doctrine to changing case facts, and issue salience and complexity. PRELIMINARY DRAFT.
1 Introduction

A judge makes policy by resolving legal disputes, that is, by deciding cases. These cases present themselves as bundles of facts, discovered and revealed through legal processes such as trials. The primary task for appellate courts, however, is not the discovery of these case facts ("truth"), but rather the "correct" application of the legal rules that structure the relevance and impact of these case facts. At the Supreme Court, and even in the intermediate appellate courts, policy-making focuses far more on the disposition of numerous future cases than on the disposition of any single case. When appellate courts address judicial policy more generally, they typically do so in opinions that establish (new or modified) legal rules for deciding cases. That is, the top tier of the judicial hierarchy concentrates far more on rule creation and articulation, while application is the primary task of the bottom tier.

Of course, even the most ideological appellate judge must make policy by telling lower court judges what facts to consider and what those facts mean for case outcomes, just as a judge bound by legal principle would. And even a dictatorial judge could not list every possible case along with its desired dispositions. Rather, she must provide some framework for guiding lower court judges, a legal rule that sorts out winning cases from losing cases. In short, judge-created case-sorting rules are the heart of judicial policy-making.

A judge seeking to implement his or her preferences over case outcomes faces choices both over both the content and structure of legal rules. But are there any constraints on these choices?

The answer is by no means obvious. Indeed, the central claims of Segal and Spaeth's Attitudinal Model (the dominant model or foil in most of the judicial politics field) are that federal judges in the U.S. are largely unfettered policy-makers and that Supreme Court justices in particular are almost completely free to vote as they
wish in pursuit of their policy goals. But can legal policy-making really be this easy?

Segal and Spaeth are right that the justices’ choices are largely free from concerns for achieving higher office, free from electoral and political accountability, and, sitting as court of last resort, free from fear of reversal in a higher court (2002, 92). In particular, they see the final vote on the merits in each case as the end of the game, so that no sophisticated choices are necessary. But, as they themselves note (2002, 357), it is the opinions that accompany these final votes that “constitutes the core of the Court’s policymaking process.” This suggests that the final vote is not really the end of the game. It may end for a particular plaintiff and defendant, but it is only the start of the next stage of the larger game, wherein the Court’s policy is applied by lower courts and wherein other actors, political and public, react to the Court’s policy.

Even if the Supreme Court does have significant power, discretion, and freedom, the justices still need to worry about compliance from lower courts and still need to worry about how to best communicate with lower courts to get what they want. How do these complications affect legal policy-making? Moreover, while casting a vote may be a trivial act, crafting an effective legal opinion is not. Opinions do many things. One of them, which I am largely setting aside here, is the justification of the case disposition with a legally principled argument. Another thing opinions do, the focus of this paper, is to shape the law, which is to say that they enable and structure the application of legal rules by other actors. Crafting a good opinion is no trivial task. Rather, it takes the wielding of considerable time and expertise, and this too represents a constraint on judicial power and choice.

In this paper, I ask whether there are any incentives for the Supreme Court justices to announce a rule other than their sincerely preferred rule. In other words, is there any reason for justices to announce a rule to govern case outcomes that is different from the rule they would choose if they had dictatorial control over all case outcomes
themselves? Does the separation of tasks—rule creation and rule application—affect rule creation itself? This breaks down into further questions. Does the structure and content of legal doctrine affect application of that doctrine? And does the need for lower court application and compliance then constrain doctrinal choice in the higher court? What sort of trade-off exists between alternate doctrinal structures and what factors affect doctrinal choice, given this trade-off? Specifically, I ask, how does the need for rule-application and compliance drive the choice between determinate doctrines (bright-line rules) and flexible, indeterminate doctrines (standards)?

To answer these questions, I present a model of legal doctrine which is designed to capture the incentives, constraints, and complications attending doctrinal choice. I present a model of rules and standards, which, while simple, captures some of the trade-offs between these instruments of hierarchical control. I define a standard as incorporating a factual dimension that is qualitatively different, in that it lacks full transparency or specificity, from dimensions that are capable of greater precision, specificity, and transparency. I consider the role that such factors play in doctrinal choice, and the roles played by the ideological heterogeneity across the judicial hierarchy, judicial expertise, issue complexity, issue salience, and the sensitivity of doctrine to varying case facts.

In the next section, I take a stance on the role of law in judicial policy-making, one that drives the choices I make in modeling doctrinal choice. I then place this paper in the context of existing work on doctrinal choice and informally discuss issues of doctrinal choice centered on the choice between bright-line rules and standards. The concepts which appear here are then modeled formally in Section 4, which analyzes optimal doctrinal choice, presents formal results, and draws out their substantive implications. Section 5 concludes. Proofs and supplemental formal results are gathered.

\footnote{In related work, I consider how the collegial (multi-member) nature of the Court affects incentives as well. Here, I set aside collegiality for the most part.}
in the Appendix.

2 Law and Judicial Politics

Below, when I present a formal model of doctrinal choice, I will abstract away from the full complexity of law. But I begin by arguing for a way of thinking about the role of law in theories of judicial politics, and I argue that this perspective raises questions for which the tools of political science might be useful.

2.1 Cloak or Constraint?

The traditional debate between the legal and political views of judicial behavior recognizes only two roles for law, as a constraint on or as a cloak for policy-making. In the former camp are the traditional legalists (or at least the caricatures thereof), who see the choices of the justices as bound by existing precedents under stare decisis and by (sometimes mechanical) principles of legal interpretation. Others in the constraint school argue that law is merely a normative constraint, but a constraint nonetheless.

On the other hand, the best known political science view of the Supreme Court is Segal and Spaeth's Attitudinal Model (2002), which sees law—the application of legal doctrines and principles, the citing of precedent, and the like—as a mere cloak for policy-making. It is a distraction the justices use to disguise their ideological behavior. Law is used not to make policy, but to hide politics, so the justices can get away with policy-making.

Rules, cases, and case facts are quite significant concepts—yet political models of judicial policy-making, formal and otherwise, often pay little attention to them. This is, perhaps, understandable. The founding debate of judicial politics—whether judges make law or find law—yielded an uneasy relationship with legal doctrine. In
rejecting law as an exogenous constraint on choice, any role for legal rules or doctrine became suspect. In stressing ideological goals, the structure and substance of judicial preferences were set aside. In emphasizing the freedom of Supreme Court justices to pursue policy goals, the instruments by which they might do so were neglected. Perhaps for these reasons, even some legal scholars who agree that judging can be political argue that much political science trivializes law and the legal enterprise (e.g., Friedman 2006).

When law and ideology interact in political science, it is as competitors. “Attitudinal” variables are pitted against “legal” variables (and sometimes against “strategic” variables\(^2\)) to form “integrated” models of judicial behavior, with the balance sorted out empirically, but rarely theoretically. The question remains as to how to explicitly define a mutually compatible theoretical relationship between legal concepts and policy goals.

2.2 Cloak AND Constraint

I will argue that there is a way to do so—law as instrument. While legal discourse may indeed serve as a cloak for policy-making, this cloak does more than just hide the ideological nature of judicial decisions; it is how judges give body to their preferences. Law is both the substance and means of policy-making in the judiciary. It is the tool for making policy, and, like any tool, its limitations can constrain its use. In short, law is instrumental.

One reason that political science has not recognized law in this way is that our models, formal and otherwise, tend to treat legal policy-making as almost a trivial act. In the Attitudinal Model, for example, the final vote is considered to be the pivotal act.\(^3\)

\(^2\)Another mainstream political science view is the “Strategic Model” (e.g., Epstein and Knight 1998; Maltzman, Spriggs and Wahlbeck 2000), in which law is usually only recognized as important in being the output of the Court’s bargaining process.

\(^3\)To be sure, Segal and Spaeth emphatically acknowledge the importance of the Court’s opinion.
In the Strategic Model, the justices bargain over the Court’s opinion and that opinion becomes law and hence policy. Formal models, meanwhile, tend to reduce opinion writing to writing down a policy point, perhaps the “ideal point” of the median justice.

These abstractions serve a purpose, but they also ignore the difficulty of achieving policy goals. The justices sit atop the judicial hierarchy, hearing only a tiny sample of the cases that are brought to court, and a tiny fraction of the cases that could potentially be brought. The decision in a given case may be directly binding on litigants in the immediate case, but the Supreme Court’s power rests not so much on the relatively few cases it hears, but on the multitudes of actual and potential cases it never has to hear. Having the desired effect on actual political outcomes requires inducing the lower court judges who handle the vast majority of cases to decide these cases “correctly.” It requires that latent litigants and other political actors fall into line as well.

To effect their preferred policies, the justices have to (at the very least) communicate them to others. Thinking of judges as political creatures does not obviate the need to think about cases and rules. Even if the lower courts were perfect agents of the Supreme Court, it is no trivial matter to convey the exact set of outcomes that one might desire in each possible case. One may represent a justice’s preferences as an ideal point, but expressing a desired policy to others is not so straightforward. Formal models thus can cause us to overlook a crucial aspect of policy-making—a justice may act as though she has an ideal point of, say, .28, but she cannot simply write an opinion telling the lower courts to set policy at “.28”, nor can she list her preferred outcome in every possible case.

Rather, to make policy, a justice must use “law”—the language of judicial policy-making—to articulate a policy statement that will serve as an abstraction for the out-

Nonetheless, the target of the model is the final vote on the merits in each case.
comes she desires. She must make use of the standard mechanisms of legal discourse as inculcated in law school, not to figure out the right answer but to enact her "right" answer.

These rules must translate vague constitutional and statutory language or general ideological attitudes into real-world applications. As Fallon (2001) puts it, at least in the context of constitutional cases, the judicial task is that of implementation. A judge seeking to implement her preferred set of case outcomes must choose from a varied judicial toolbox, containing "rules," "formulas," and "tests," which themselves break down into "bright-line rules," "standards," "balancing tests," multiple-"prong" tests, and the like.

She must make use of, and is constrained by, the entire legal web in which judicial behavior is embedded. She can cite precedents (positively or negatively) as examples, reason through analogies to other cases or areas of the law, use or modify existing rules, and rely on or reject legal arguments and principles. She can define equivalence classes of cases to be treated similarly or dissimilarly (see Kornhauser 1992b) and assign degrees of importance to potential facts. She can make use of previous opinions (including dissents), the briefs in a case, and even law review articles that unravel issues of doctrine and the application thereof. Her task will be easier to the extent of her expertise in the relevant area of the law and to the extent of the development of the relevant body of law—these will make it easier to articulate and achieve her preferences. New or less familiar areas of adjudication will leave a justice more at sea in this regard.

The accuracy of policy-making through judicial opinions and decisions and even the

---

4To modify Bueno de Mesquita and Stephenson (2002), it is not that judges care about law instead of or in addition to policy; judges care about law because they care about policy.

5Some scholarship has argued that precedents serve purposes of information, guidance, and communication (e.g., Bueno de Mesquita and Stephenson 2002; Kornhauser 1989; Macey 1989; Rehnquist 1986; Richards and Kritzer 2002; Shapiro 1972).
compliance of lower courts with these opinions and decisions are, therefore, at least partially endogenous to the justices' choices. The justices actively craft their opinions so as to achieve their desired policy outcomes. They are inhibited in their policy-making by uncertainty, ambiguity, and complexity. Crafting an opinion to induce a desired set of policy outcomes is not so simple. Legal discourse is inherently ambiguous. Using the law well is both difficult and costly, requiring significant expertise, and this itself will constrain judicial policy-making (and bargaining). Justices (like other political actors) are imperfect policy-makers.

This perspective informs the modeling choices below.

While I have argued that political science has paid insufficient attention to the accoutrements of legal policy-making, there is one thread of research in particular that has taken such factors into account, and it is to this line of work I turn next.

3 Theories of Doctrinal Choice

The analysis of doctrinal choice is emerging as a vibrant new frontier in the bridging between legal theory and political science. Previous work on doctrinal choice has looked at the usage of precedent for communicative accuracy (Bueno de Mesquita and Stephenson 2002); inducements for lower courts to comply by granting them discretion (McNollgast 1995); resorting to vagueness to obscure noncompliance or to delegate (Staton and Vanberg 2007); and, more generally, see Tiller and Cross (2006). Other work includes that on incremental doctrine formation (Cameron 1993; Gennaioli and Shleifer 2007), Lax (2007) which studies doctrine formation on collegial courts and the effects of judicial institutions thereon; and Landa and Lax (2007), which, along with Lax (2007), explores the coherence of legal doctrine on collegial courts and the impact of collegiality on doctrinal structure. Lax and Cameron (2007)
explore bargaining over doctrine on a collegial court, incorporating the costs and uncertainty of rule-crafting. Last and most, for the purposes herein, Jacobi and Tiller (2007) model the choice between determinate and indeterminate doctrines, given how indeterminate the latter doctrine is, lower court conflict, and bias for certain types of litigants.\(^6\)

Like Jacobi and Tiller, the model I present makes explicit the grounds for distinction between rules and standards. The model below, however, focuses on somewhat different issues—most specifically, the transparency of lower-court decision-making, the precision of doctrine, and the role of non-compliance. Perhaps the main points of departure here are that (1) the nature of the doctrine’s component factual dimensions becomes the key driving force behind doctrinal choice and that (2) the analysis below makes various features endogenous, such as the degree of lower court discretion and the set of rules available to the higher court from which to choose. I assume lower courts are only able to achieve a degree of discretion given a lack of full transparency as to whether they have complied with higher court doctrine or given the limits of doctrinal precision. I allow the higher court to actively work to affect how the doctrine is applied through carefully crafting the opinion that announces it.

To do this, I make use of a “case-space” model, a variant of the more common policy-space model, one tailored substantively to the institutional features of judicial policy-making. This model provides a structure for thinking about cases, dispositions, and legal rules, without rejecting a role for judicial preferences.\(^7\)

\(^6\)To Do: Note connections to literature on bureaucracy, agency theory, and optimal contracts.

\(^7\)This model has its origins in Kornhauser (1992b,a), which specified the idea of a doctrine as a function that establishes equivalence classes of cases to be decided similarly. Cameron (1993) suggested a geometric version of this approach, with a spatial representation of cases and rules, treating courts as unitary actors. Grofman (1993) briefly considered unstructured rules from a collegial perspective. Cameron, Segal and Songer (2000) used a unidimensional, unitary-actor case-space to study auditing in the judicial hierarchy. Lax (2003) analyzed a collegial, but still unidimensional, case-space model to explore the impact of the Supreme Court’s institutional rules on compliance in the lower courts. Kastellec (2007) extended this model to explore whistle-blowing in the lower courts. Lax and Cameron (2007) study bargaining on a collegial court in a unidimensional case-space. Lax (2007) studied a
3.1 Types of Doctrine

One challenge to studying doctrinal choice is that scholars vary widely (wildly?) in their definitions of rule types, with much conflict in particular over the differences—or lack thereof—between rules and standards. For some sense of the contours of this debate, see Ehrlich and Posner (1974, 258-60), Sullivan and Amar (1992, 57-69), and Kaplow (1992, 559-62). The most common definitions seem to be that a rule is a determinate form of legal doctrine, in which the line of permissible conduct is specified in advance (ex ante), whereas a standard establishes a more flexible doctrine, in which the adjudicator of the case at hand need take additional facts or factors into account in applying the doctrine (adjudication is at least partially ex post). The former is seen as more clear-cut, the latter as more flexible.

It is possible that, as scholars have sought nuance, the terms have lost some of their usefulness. For example, both rules and standards can be clear or unclear, simple or complex, transparent or not, determinate or not, etc. For the purposes of this paper, I will adopt a somewhat crude distinction, no doubt omitting much that might be of interest in other contexts. And, to be sure, my usage differs from some of the most common definitions in some ways. (No part of my argument rests on the names “rules” and “standards,” and so it not a problem if the reader rejects my analogies to those terms, so long as he or she will forgive my borrowing of some of their connotations.)

In my terminology, a doctrine or rule is any logical classification of cases into winners and losers. (Kornhauser (1992b) draws distinctions between these terms which I set aside here.) A bright-line rule does so clearly and cleanly. I will define a bright-line rule (or, with some terminological abuse, a rule, for short, when the intent is clear)
as a doctrine which is based on a straightforward factual dimension or dimensions. It can be communicated so that it can be applied precisely by faithful agents and monitored for compliance by non-faithful agents (potentially non-compliant lower courts), because it is defined (only) with respect to factual dimensions that are specifiable and transparent in their application. In these senses, a bright-line rule is determinate.

A bright-line rule can be simple or even more complicated, by balancing different factual dimensions to reach a disposition: a balancing test incorporates competing factual dimensions which must be weighed against each other. A balancing test need not be indeterminate, if the balancing is clear and straightforward—if it only includes factual dimensions that are clean and clear in the sense above. Rather, whether a balancing test is determinate will depend on what the relevant factual dimensions are—and whether these dimensions themselves are problematic. That is, a balancing test still can be fully rule-like, if all dimensions can be transparently and precisely specified along with the relationship between them. What makes a balancing test a standard is that it incorporates some dimension over which the doctrine cannot be so cleanly specified or applied. Thus, what separates rules from standards is specificity and transparency, which in turn depend on the characteristics of the factual inquiries that the doctrine invokes.

A simple and familiar example will assist this exposition. The archetypal (bright-line) rule is a speed limit: “You may drive no faster than 55 miles per hour.” But a speed limit could also be defined as follows: “You must drive at a reasonable and prudent speed” (as Montana once did). What counts as reasonable will depend on situational context, and this cannot be stated with the same precision as a numerical speed limit nor as transparently applied as a numerical speed limit (that is, it would not be immediately obvious whether Judge A’s application would be the same as what Judge B would mean by “reasonable”). This makes the latter speed limit a standard,
not a rule.

To determine whether a lower court applied this reasonableness standard as desired by the higher court, the higher court would have to look closely at numerous factual circumstances, including, say, weather conditions. To check whether a rule was “correctly” applied, the higher court would only have to know the objective speed of the car. A standard thus contains a degree of subjectivity (due to the limits of specificity), raises the possibility of lower-court error (again due to imprecision in defining the limit), and even of noncompliance (due to the lack of transparency). If weather could be objectively and transparently defined, then the “reasonable speed” doctrine would be a rule, not a standard (in my terminology).

Note that lower court discretion that accrues from the setting of a standard arises, not from higher court choice itself (that is, I am not studying why lower court discretion might be good for the higher court), but as an inevitable side effect from the usage of a standard instead of a bright-line rule (more precisely, from using a doctrine that can be applied only as a standard). Below, I separate transparency from precision.

A substantive example of the rule-standard divide can be seen in the Supreme Court’s rulings on the death penalty as applied to minors or the mentally retarded. *Thompson v. Oklahoma* (1988) barred execution of offenders under the age of 16. *Roper v. Simmons* (2005) extended this to those under the age of 18, overruling *Stanford v. Kentucky* (1989). *Penry v. Lynaugh*, 492 U.S. 302 (1989), similarly permitted the execution of the mentally retarded, but this was overruled in *Atkins v. Virginia* (2002). Across these cases, the debates centered on whether “evolving standards of decency” were relevant; what these standards were *if* they were relevant; as well as the maturity and sense of responsibility of the defendants. It is the latter issue to which I wish to draw attention.

For minors, the Court adopted bright-line rules (first 16 years as the cut-off, then
18), whereas the Court could have told lower court judges to inquire as to the maturity of each defendant according to some set of guidelines or perhaps a multi-pronged test, of which age would play some part. Any such inquiry, however, would inevitably be imprecise and somewhat opaque—maturity is subjective and inherently lacking in specificity, whereas age is not. The mentally retarded present an even greater problem in terms of precision and transparency, as mental age and ability cannot be determined as simply and straightforwardly as calendar age, so that any doctrine relying on the degree of mental retardation seems inevitably rather far along on the standard end of the rule-standard continuum. On the other hand, where a particular IQ test result is used as the cut-off, again this would become a bright-line rule.

Or, consider differences in the application of the Fourth Amendment’s prohibition of unreasonable search-and-seizures. In Maryland v. Wilson (1997), the Supreme Court explicitly chose a bright-line rule (police can require passengers to exit a vehicle during a legal, routine traffic stop) instead of a balancing test that would dig into the reasons for the traffic stop and the situation faced by the police officer. Such a balancing test would likely be a standard, because (1) the Court would have to craft language, likely abstract language, to capture precisely the circumstances in they wanted to permit such exit commands, and likely winding up somewhat over- and under-inclusive in so doing even given eagerly faithful lower courts, and (2) since compliant application of the new test would not be obvious at first glance, the Court would have to look closely at cases decided under the new test to assess compliance and lower courts could use this to their advantage in evading full compliance.

Meanwhile, as Fallon (2001, 83) points out, in other areas of Fourth Amendment doctrine, the Court uses a “totality of the circumstances” test (e.g., Illinois v. Gates (1983), for whether an informant’s tip establishes probable cause, and Ohio v. Robinette, (1996), for assessing voluntariness of a search after a traffic stop if the police
have not informed the detainee of the right to refuse). This test is obviously a standard.

Or, consider the Roe v. Wade trimester structure of permissible abortion restrictions. It is not a one-dimensional rule, but rather one with three tiers, with some judgment calls in the second and even third trimesters. Still, this seems more rule-like than the "undue burden" standard first formulated in Justice O'Connor's dissent in Akron v. Akron Center for Reproductive Services.

Note that it is likely that most sincerely preferred doctrines—that is, the doctrines the higher court would really like to implement—take the form of balancing tests or flexible rules. Few judges truly believe in absolute rules for all circumstances. Few areas of the law truly reduce to a single objective dimension like "speed." Even Justice Scalia will join or author decisions with flexible rules or balancing tests, despite his well-known preference for bright-line rules (Fallon 2001, 104). This suggests that, regardless of the doctrines we observe the justices handing down, and regardless to what they may say about the desirability for bright-line rules, the trade-offs I study herein may be pervasive. But what are these trade-offs? I discuss this intuitively, before proceeding to the formal analysis. In short, where a flexible rule is desired, a bright-line rule is both over- and under-inclusive. A standard, on the other hand, is more prone to erroneous or non-compliant application.

3.2 The Rule-Standard Trade-off

Let us return to the speeding example. Suppose cases have two dimensions: how fast the driver was going and what the weather conditions were at the time. Figure 1 shows three cases: the driver in case $x^3$ was caught going faster than the driver in $x^2$ who was going faster than in $x^1$; but the weather was the worst in $x^1$ and the best in $x^2$. Each case has to be decided as safe driving (Y, a winner) or speeding (N, a loser).

If the higher court ($H$) truly prefers a constant speed limit, then she can simply
announce that directly. \( H \) could announce a bright-line rule, such as 55 m.p.h. (rule A in the figure), under which only driver 3 is speeding and drivers 1 and 2 are off the hook. But suppose that she wants weather conditions taken into account, such that what speed is permitted depends on the weather dimension. She would now prefer a lower speed limit the harsher weather conditions were (rule B in the figure). This sliding scale has more flexibility than a constant speed limit of 55, in that it varies with weather conditions. Under rule B, only driver 1 is speeding.

Note that, were \( H \) to prefer rule B but still announce rule A as a bright-line rule, only case 2 would be correctly decided (i.e., decided consistently with \( H \)'s sincerely preferred doctrine). Only case 2 would get the same result under both rules. Case 1 would be improperly decided as a reasonable speed, and case 3 would be improperly decided as speeding, at least given \( H \)'s referred definition of a reasonable speed (rule B). In short, a bright-line rule inevitably incurs losses in terms of incorrect case dispositions, if the preferred doctrine is actually a flexible rule or balancing test.

Of course, if \( H \) heard all cases herself, she could simply apply her preferred balancing test. Suppose instead that \( H \) must delegate case decisions to lower courts. If speed and weather conditions could be observed straightforwardly, then the balance between the dimensions could be laid out clearly and cleanly. There would be no wiggle room for lower courts to avoid compliance. Lower courts would apply \( H \)'s test as easily as would \( H \) herself and this test would serve as as a bright-line rule.

However, while speed is a hard and fast measurement, weather conditions are not. The weather dimension differs in observability and specificity. The crucial point is that, if \( H \) can observe the case's position on the speed dimension perfectly, but on the weather dimension only imperfectly, \( H \) will not be able to tell directly whether her preferred doctrine has been applied in every case. Many speeds are reasonable under some weather conditions but excessive under others. Lower courts that disagree with
H's doctrine could evade the H's preferred doctrine on the margins, and H would have to review a case herself to be sure. In short, the informational gap will make noncompliance possible. It also means that a doctrine incorporating weather would be a standard, not a bright-line rule.

Or, suppose that the second dimension can be perfectly observed, but not perfectly quantified or specified. Then, even friendly lower courts will decide some cases in opposition to H's preferred doctrine, not due to willful noncompliance but due to the inherent ambiguity of the second dimension. Again, a doctrine that incorporated a reasonableness requirement with respect to weather would be a standard, not a bright-line rule. Of course, both of these problems, imprecision and the lack of transparency, can be at play in the same case or issue area (indeed, this seems likely).

To be clear, it is the nature of the factual dimension that creates the difference between rules and standards, in this framework at least. A rule constant across the weather dimension can be applied directly and perfectly. A rule that varies with the weather dimension—given the nature of this associated inquiry—cannot be. In short, a "reasonable" speed limit will be a standard, not a rule. H does not purposely adopt a standard so as to allow the lower courts discretion; rather the nature of the issue area permits "discretion" such that H's preferred doctrine takes the form of a standard. Either a lack of transparency or the inherent imprecision of a legal test in this area will make the doctrine a standard.

This means that H must make a choice between the problems that arise in announcing a standard, such as a reasonable speed test, and the problems that arise in sticking with an over- and under-inclusive bright-line rule. Bright-line rules lead to some "incorrect" dispositions due to inflexibility where flexibility is desired. Standards do so because of noncompliance and error.

9While I would not rule that out as an incentive in other contexts, it is not my focus here.
The next step is to formalize this trade-off and derive the conditions under which a rule or standard is preferred.

4 Optimal Doctrinal Choice

I assume throughout that the Court is concerned with, inter alia, minimizing the size of the set of cases that are decided in opposition to its preferences. In particular, I assume that the Court suffers quadratic loss with respect to the area of the case space that is incorrectly decided: if the area is \( A \), then the Court gets a payoff of \(-A^2\) (i.e., there is increasing marginal loss).

4.1 Balancing Tests

If the Court’s desired doctrine were already a simple bright-line rule, then it could simply announce it directly. Assume then that the Court’s preferred rule is a balancing test across two dimensions.\(^{10}\) If both dimensions were purely objective, then the balancing test between them would be in effect a bright-line rule, albeit one more complicated than a unidimensional bright-line rule. Or, if the Court only cared about a single objective dimension, it could just announce a bright-line rule. The bite in doctrinal choice comes from the inclusion of a subjective dimension (the descriptor ambiguous would serve equally well). So, I assume that the Court’s preferred doctrine does incorporate a subjective factual dimension, which makes it difficult to identify case positions or rule limits. I draw out results for a balancing test in a two-dimensional case space and then show that results for a conjunctive two-prong test are similar.

In a two-dimensional case space, with cases given as a pair \( \{x, y\} \), a balancing test takes the form \( \hat{y} = b - ax \), where a case gets a Yes if and only if \( y \leq \hat{y}(x) \). The

\(^{10}\)If there is only one choice, the Court does not really have a choice between a rule and standard, given the framework here.
parameters $a$ and $b$ determine both the tradeoff between the two dimensions as well as whether some values along a given axis are strictly determinative (so that no value on the other axis can rescue the case). The slope captures the relative weights between the two dimensions.\textsuperscript{11}

Let the first dimension (the $x$ axis) be the subjective dimension, with the rule defined as the cut-off on the second dimension for a disposition of Yes. If $a = 0$, then the Court already prefers a bright-line rule (given that only the straightforward second dimension affects case outcomes). The larger $a$ is, the greater the sensitivity of desired case dispositions to factual dimension 1 (the weather dimension, say).

4.2 Choosing a “Rule”

Since the first dimension is subjective, and a bright-line rule cannot include a subjective dimension, a bright-line rule for this case space can only include the second dimension. It must be a horizontal line that divides the case-space with all cases below getting a Yes. It can be represented by a fixed cut-off $\bar{y}$ that does not vary with $x$. Under this rule, the lower courts decide cases only according to dimension 2, ignoring dimension 1. (See Figure 2.) Using a bright-line rule instead of the preferred balancing test means that there can be both over- and under-inclusion, since the bright-line rule does not take into account $x$ when disposing of cases, but rather only $y$, where $y$ is the straightforward dimension—the rule is “Yes if and only if $y \leq \bar{y}$”.

Any case above $\bar{y}$ but below $\hat{y}$ should get a Yes but does not and any case below $\bar{y}$ but above $\hat{y}$ should get a No but does not. (See Figure 3.) The loss to the Court from using a bright-line rule is based on the area of the region of the case space that is wrongly allocated.

\textsuperscript{11}I start with a straightforward configuration in which $a$, the slope, directly captures the tradeoff between the two dimensions. In this configuration, for any value of $x$ there is some interior value of $y$ that yields a Yes disposition.
Among the possible bright-line rules the Court might choose, the optimal bright-line rule is a function of both \(a\) and \(b\):

**Lemma 4.1.** The optimal bright-line rule is \(y^* = b - \frac{a}{2}\).

The losses are the shaded regions shown in Figure 3. The combined area of these regions is \(\frac{a}{4}\). Let the Court value this issue area with a salience weight \(s\). The higher \(s\) is, the more the Court suffers when cases are disposed of incorrectly. Then, the payoff from the optimal bright-line rule is \(-s \left(\frac{a}{4}\right)^2\). Where \(a = 0\), there is no loss. The higher the value of \(a\), the steeper the slope, the higher the desired tradeoff between dimensions, and the greater the cost of ignoring those tradeoffs and reducing the doctrine to a bright-line rule.

How does simplifying doctrine in this way compare to sticking with the preferred balancing test?

### 4.3 Choosing a “Standard”

If the Court includes the subjective first dimension in its announced doctrine, then it is invoking a standard (at least, given my usage). There are two potential sources of trouble. The common thread is that there will be a region of cases near the cut-line that might be decided improperly by lower courts and in which lower court decisions can vary. The first possibility is that lower courts will be faithful but imperfect agents of the higher court and thus will make mistakes in case dispositions (mathematically, this will turn out to be a special case of the second situation). This might occur because of the inherent difficulties in defining the cut-line (what I define below as imperfect precision). The second source of trouble is purposeful non-compliance, because the inherent ambiguity of the first dimension allows lower courts to strategically evade the Court’s preferences (due to imperfect transparency, again defined below). In short, the Court faces a principal-agent problem.
Let the width of the troublesome region extend \( v \) in either direction from the proposed standard, as shown in Figure 4, so that the higher \( v \) is, the larger the more "standard-like" the standard is. The area of this region is \( a(2v - v^2) \). How this compares to choosing a bright-line rule will depend on (a) whether \( v \) is exogenously fixed or endogenous in that the justices can actively engage in affecting it through careful drafting of their opinions and (b) why case dispositions in this region might be problematic.

Within this region, let the probability of the wrong decision be \( p \), so that the expected losses due to sticking with this balancing test are \( ps(a(2v - v^2))^2 \). When concerned with the agency problem, \( p \) reflects the conflict between higher and lower courts in terms of preferred case dispositions. When concerned with precision, \( p \) will be exogenously fixed to represent the chance of error.

### 4.4 Exogenous Transparency

Assume that when a lower court decides a case, the higher court can easily observe its position on dimension 2, but can only imperfectly observe its position on dimension 1, the subjective dimension. Call the higher court’s ability to assess the case on dimension 1 “transparency.” The parameter \( v \) then captures the discretionary region in which lower courts can take advantage of this lack of complete transparency to evade the higher court and apply their own “discretion.” 12 This discretion is of course not a willing choice of the higher court, but a result of the difficulty of monitoring compliance given the subjective dimension.

---

12 Consider a model of higher court/lower court interaction similar to that of Cameron, Segal and Songer (2000), in which the higher court cannot tell without actually taking a case whether the lower court has complied, and so there exists just such a region in which a hostile lower court will not comply and it is not worth it for higher courts to audit their decisions. The parameter \( v \) can be understood as emerging from such a model. Since that model is one-dimensional, the Court does not have the choice of dodging the subjective dimension entirely. Indeed, doctrine is treated as exogenous in the paper.
Let $p$ capture the likelihood of a hostile lower court and thus the percentage of case decisions in the discretionary region that will be in opposition to the higher court's preferred outcome. When $p = 0$, $v$ is irrelevant. If $v = 0$, then both dimensions are perfectly observable, and the balancing test is already a bright-line rule. Otherwise, this region has positive width, and within it, lower courts can decide as they desire. Outside this region, compliance is perfect. The losses when using this balancing test are $ps(a(2v - v^2))^2$.

We can now compare the utility from the optimal bright-line rule to the use of the balancing test:

**Proposition 4.2.** Given an exogenous level of transparency, the optimal bright-line rule is preferable to the balancing test (standard) if and only if $p > \frac{1}{16v^2(2-v)^2}$ (equivalently, if $v > 1 - \frac{1}{2} \sqrt{4 - \frac{1}{p}}$).

The choice between accepting the non-compliance associated with the balancing test and the automatic over- and under-inclusiveness of the bright-line rule depends on both the level of transparency and the extent of ideological conflict between higher and lower courts; it does not depend on either the sensitivity to the subjective dimension nor on issue salience. The greater the conflict between the two (higher $p$), the greater the desirability of the bright-line rule. The lower the degree of transparency (higher $v$), the greater the desirability of the bright-line rule. (Where conflict is low enough ($p \leq \frac{1}{16}$) or transparency high enough ($v \leq 1 - \frac{\sqrt{3}}{2}$), the balancing test is preferred no matter the value of the other parameters.)

The trade-off between conflict and transparency in driving doctrinal choice is shown in Figure 5. These results make sense intuitively: where ideological conflict is a problem, and so compliance is a problem, bright-line rules are preferred. The sensitivity of the balancing test to the subjective dimension and salience are both irrelevant because they affect equally the losses under the optimal bright-line rule and under the balanc-
ing test, so they drop out of the doctrinal choice calculus—at least when transparency is exogenous.

When affecting transparency is beyond the control of the Court, the substantive implications are that less transparent areas of the law will get bright-line rules; during times of ideological conflict between upper and lower courts, bright-line rules are preferable; for issues where there is hierarchical conflict, bright-line rules are preferred; standards are more likely to be chosen when the judicial hierarchy is more homogeneous; and the greater lower court opposition is, the greater transparency must be before a standard-like balancing test can be safely chosen.

Some of these incentives change if the Court can affect the transparency of compliance by investing costly effort into crafting the balancing test.

4.5 Endogenous Transparency

I argued earlier that the justices can actively craft their opinions to affect lower court application. Assume that the Court can invest in reducing discretion and increasing the transparency of lower court compliance (that is, in reducing $v$). Instead of dropping the subjective dimension altogether, the Court can work harder on specifying the doctrinal requirements so as to force the lower court to discuss the various aspects of the case facts that will help the higher court assess where precisely the case falls on both dimension and whether the lower court has been compliant, thus increasing the incentives for the lower court to preemptively comply.

Formally, I assume that the investment of costly effort reduces $v$. I then ask, what is the optimal level of $v$? And, given this optimal balancing test, should the Court instead choose the optimal bright-line rule? It will be seen that the answers to these questions now depend on the cost of increasing transparency, the sensitivity to the subjective dimension, issue salience, and hierarchical conflict. Let the cost of writing
a balancing test be \( c(1 - v)^{4} \), so that the payoff due to choosing the balancing test is
\[-ps(a(2v - v^{2}))^{2} - c(1 - v)^{4} \]. Then, we obtain the following results:\(^{13}\)

**Proposition 4.3.** The optimal endogenous transparency of the balancing test occurs at
\[ v^{*} = 1 - \frac{\text{pre}}{\sqrt{ps(c + \text{pre})}}. \]

**Corollary 4.4.** Optimal transparency is decreasing in cost and increasing in sensitivity, conflict, and salience.

Figure 6 shows the effects of the various parameters on optimal transparency. As cost increases, optimal transparency naturally drops. As conflict or sensitivity or salience increases, the curves shift upwards, so that for any fixed cost \( c \), optimal transparency is higher. The intuition is that greater lower court hostility puts pressure on the higher court to raise transparency to monitor them, while higher sensitivity to the subjective dimension means that a wider range of cases are subject to potential non-compliance, again pushing towards a greater investment in transparency.

These results define the optimal balancing test, which can then be compared to the optimal bright-line rule, yielding the following results:

**Proposition 4.5.** Given optimal transparency, the optimal bright-line rule is preferable to the balancing test (standard) if and only if \( c > \frac{\text{pre}^{2}}{16p - 1} \).

**Corollary 4.6.** The incentive to choose the optimal-transparency bright-line rule over the optimal balancing test increases with higher costs and conflict and it decreases with higher sensitivity and salience.

As cost increases, optimal transparency decreases, until eventually a balancing test is no longer itself optimal. In Figure 6, higher conflict, higher sensitivity, and higher

\(^{13}\)This functional form suffices to make the optimal level of \( v \) an interior solution. If, say, a quadratic cost term were used, then the optimal choice would sometimes be a corner solution, which would not change the substantive conclusion, but would unnecessarily complicate the presentation of results.
salience shifted the curves upwards—but they have different effects on the transition to a bright-line rule. Higher conflict lowers the transition point at which the cost of a balancing test becomes prohibitive, while higher sensitivity or salience raises this threshold, enlarging the range of costs for which a balancing test will be chosen.

In short, a bright-line rule is preferable where costs are high, conflict is high, salience low, and sensitivity low. Figure 7 reveals the more complicated trade-offs between these parameters in optimal doctrinal choice. The three panels represent three perspectives on this choice. Each panel considers a pair of parameters and shows the regions in which a bright-line rule should be chosen as opposed to choosing the balancing test, along with dividing lines between these regions. A dividing line is shown for higher or lower values of the remaining parameters.

While these pictures speak for themselves to some degree, some implications are more subtle than others. It is clear that sensitivity and salience push in the same direction, and that they each push in the opposite direction from conflict. Less apparent at first glance, there is another type of trade-off between the two parameters. In each panel, the dividing line is the line of indifference between the two doctrinal forms. Note in the final panel that sensitivity and salience are substitutes for each other; the higher one is, the lower the other can be while still maintaining indifference between the two doctrinal forms. More than that, each has a diminishing effect, in that the higher salience is, the less increasing salience matters for optimal doctrinal choice and the more sensitivity matters. The higher sensitivity is, the less increasing sensitivity matters and the more salience matters in doctrinal choice. To put these points another way, salience is more likely to affect the choice of doctrinal form when sensitivity is high than when sensitivity is low. Sensitivity is more likely to flip doctrinal choice when salience is high than when salience is low. Turning to the top panels, we find that higher salience and higher sensitivity both dull the effects of conflict in driving
doctrinal choice. And when conflict is low, the effects of salience and sensitivity are of lower consequence. Also note that for any level of conflict, there exists some level of salience such that the balancing test is preferable (this might seem true for sensitivity too, except that sensitivity is bounded by a maximum value).

4.6 Discretion and Effort

These formal results can be interrogated further to answer two additional questions. To what extent will there exist the potential for non-compliance given optimal doctrinal choice? Or, to put this another way, what degree of discretion will remain after such choice? Secondly, what level of effort will be expended given optimal rule choice?

The answers are sensitive to the parameters highlighted above and can be derived from Lemma 4.3 and Proposition 4.5.

Define residual discretion $d^*$ as the leeway a lower court will have given optimal rule choice. When a bright-line rule is chosen, it will take the value zero. When the balancing test is chosen, discretion will be inversely related to $v^*$, taking the value $1 - v^*$, (that is, it maps to transparency, which is larger when the resulting region of potential non-compliance is larger). Effort is the cost paid by the Court (zero when the bright-line rule). The results are easiest to see graphically, as shown in Figure 8. The level of discretion which the lower courts will maintain after optimal doctrinal choice will depend on the sensitivity of the balancing test, the costs of reducing discretion, and the degree of conflict between the courts.

The most striking features of these results are that discretion and cost are non-monotonic and discontinuous functions of the key parameters. For example, as the cost of crafting doctrine increases (or as judicial skill decreases), discretion (and the possibility of non-compliance) increases as the Court’s balancing test loses efficacy, but then the optimal choice transitions to a bright-line rule, under which some cases
are automatically decided incorrectly but lower court discretion is nil. Hierarchical conflict reveals a similarly discontinuous effect when the Court breaks in favor of the bright-line rule, but at least here the effect of conflict is monotonic. The sensitivity to the subjective dimension reveals a pattern similar to that for cost, but in the other direction. As sensitivity increases, there is at first no effect on discretion as the Court will still make use of a bright-line rule, until the point at which the balancing test is optimal, but one where the residual discretion is high. Residual discretion then decreases as sensitivity continues to increase.

What do these formal comparative statics mean substantively? In short, doctrinal choice creates strange bedfellows. Low residual discretion is associated with both the lowest-skilled and highest-skilled justices (those who face, respectively, the highest or lowest costs of generating transparency). The former choose bright-line rules and thus get full compliance (albeit with a non-ideal doctrine); the latter will craft high-quality balancing tests that still manage to rein in lower courts to a significant degree. The same is true for justices with high or low (as opposed to middling) concern for the substantive dimension—both types yield relatively low levels of residual discretion as compared to their more moderate brethren (the former invoke good balancing tests; the latter prefer bright-line rules). Shifted the focus from judicial ability to the nature of the issue area, the simplest and most complex areas (lowest and highest costs) will be associated with low residual discretion.

4.7 Ideology, Moderation, and Doctrinal Choice

What role does ideology play in doctrinal choice? Some argue that more extreme justices such as Justice Scalia tend to prefer bright-line rules whereas standards or balancing tests are associated with moderate justices such as Kennedy or O'Connor, those likely to be situated as swing justices in close cases [cites]. Why would this be
so? A coincidence? Or can it be explained by incentives for doctrinal choice? I argue that it can be.

In this section, I use the framework developed above to answer these questions and to clarify points that are sometimes conflated or omitted in discussions of ideology and doctrine. There are actually two discussions worth having, of perhaps equal importance. The first considers whether liberals and conservatives (the extremes) are similar as compared to moderates. The other considers whether doctrinal choice corresponds to ideology directly, in that liberals and conservatives will disagree over doctrinal form. Is it a matter of “ends-against-the-middle”? Or does ideology directly correlate to preferred structure? I consider both of these possibilities.

First, one possibility is that such justices are so extreme that they truly prefer a rule of “always” or a rule of “never.” If that is so, then they already prefer a bright-line rule and so face no doctrinal trade-off. More moderate justices will prefer balancing tests, but the extreme justices would prefer bright-line rules. But that makes the debate rather trivial, not the least of which because this effect could only be a feature of the jurisprudence of the most extreme justices. The more interesting questions of doctrinal choice arise for those justices short of such extremes, even barely so, where there is still “room” in the case space to place a balancing test across the two dimensions.

So, to continue, assume that each justice does indeed prefer a balancing test. Let them agree on sensitivity (the slope), and be otherwise equivalent in all respects except for differing in how high or low the line is drawn (the intercept). The justices with the higher intercepts want more Yes outcomes; those with the lower intercepts/lines want more No outcomes. But they each prefer the same tradeoff between the two dimensions. In such a configuration, what role does the differing intercept play in preferences over doctrinal form? None, as by Proposition 4.5. The intercept parameter ($b$) drops out of the doctrinal calculus, which is to say that it does not affect
whether a bright-line rule or a balancing test is preferred. All else equal, either they all prefer bright-line rules (albeit set at a different “height”), or they all prefer parallel balancing tests (again, set at different “heights”). Ideology is orthogonal to the choice of doctrinal structure.

Suppose that the justices instead vary in terms of sensitivity to the subjective dimension. In that case, we should not find ends-against-the-middle, but rather that, all else equal, those with higher sensitivity prefer balancing tests (as shown in Figure 7 and by Corollary 4.6). It seems reasonable to suppose that liberalism might indeed positively correlate with sensitivity to such a dimension (for example, whether exonerating factors matter for sentencing), but so might conservatism in other areas of the law (how much should good faith exonerate an improper search and seizure?). Either way, we would expect a split along party lines on doctrinal structure.

Recall Figure 8, which shows that high and low levels of sensitivity induce lower levels of residual discretion than do moderate levels. When ideology correlates with sensitivity, this suggests that liberals and conservatives would both leave less discretion in the hands of the lower court than moderates: one wing because it would choose a bright-line rule, the other because it would construct a high-quality balancing test.

The point remains, however, that it is not extremism in terms of the absolute number of “yes” or “no” dispositions that directly affects optimal doctrinal form, but rather sensitivity to the subjective dimension. Note also that the effects of sensitivity depend on the justices having influence over transparency—if they do not, then we return to the solution for exogenous transparency and Proposition 4.2, wherein sensitivity to the subjective dimension does not figure into doctrinal choice even if sensitivity does vary by ideology.

The formal results highlight two further concepts that might shed light on the ideology/doctrine debate. That is, two other parameters shown to affect the justices’
incentives might vary with ideology: salience and conflict.

Salience might vary with ideology because the distribution of cases in the case space might vary. That is, there might be many more cases falling in the middle of the case space than at the extremes. If that is the case, then the troublesome region might capture a greater number of potential lower court cases when it is in the moderate region of the case space than when it lies near one extreme or the other. The number of cases for which transparency might create a problem might be much lower for Scalia's preferred balancing test, which would lie much higher in the case space than for Kennedy's preferred balancing test. In the formal model, this is easily captured by the salience weight for this issue—Scalia would have to worry less about this region than Kennedy and so $s$ would be lower for Scalia than for Kennedy. If transparency is fully exogenous, this is irrelevant, as salience is irrelevant for choice. If transparency is endogenous, then higher salience suggests a balancing test so that moderate justices would indeed be more likely to prefer the balancing test over the bright-line rule.

Next, when justices are more extreme, they are likely to be positioned differently with respect to lower courts than are moderate justices. Much depends on the distribution of lower court judges. Suppose that, as might seem likely, the lower courts are roughly distributed around the center of the Supreme Court, and that the distribution of lower courts follows a bell curve, with many more lower courts concentrated near the center of the higher court with fewer in the wings. If that is the case, then a more extreme justice will find few allies in the courts below, and a more moderate justice will find it far more likely that a random lower court will resemble her own preferences for case dispositions. This suggests that, all else equal, moderate justices face a lower likelihood of a hostile lower court (lower $p$) than a more extreme justice.

\footnote{If the lower courts are polarized with respect to the Supreme Court, then the analysis that follows might only hold for one wing of the Supreme Court and not the other. If the distribution of lower courts is stranger still, the implications for doctrinal choice are even more ambiguous.}
In turn, this means that, if the concentration of lower courts is sufficiently high, then moderate justices should prefer balancing tests and more extreme justices should prefer bright-line rules (see Corollary 4.6 or Figure 7) [add formal proof].\textsuperscript{15} The intuition is that O'Connor is more willing to allow lower courts to make use of this residual discretion since she has a greater expectation they will decide as she would. Scalia will find fewer allies below and so will do better to accept the losses under a bright-line rule than to turn to a balancing test.

Note that the salience argument and the compliance argument are mutually reinforcing, in that both suggest moderates will be more likely to prefer balancing tests than will more extreme justices. The patterns of residual discretion, however, would be less clear-cut (in Figure 8, discretion is monotonic with respect to conflict, but not so with respect to salience).

These arguments can be extended to consider how polarization might affect doctrinal outputs. A Court full of moderate justices would tend to produce a greater number of balancing tests, all else equal. But, in a polarized Court, with relatively extreme justices on both sides, each side would prefer bright-line rules. Indeed, even if only one wing of the Court is extreme, if it has the majority, it would be more likely to produce bright-line rules.

One final complication remains, one which to the best of my knowledge has not been explored in the literature on rules versus standards: the varying perspectives of opinion authors versus other members of the majority. On a collegial court such as the Supreme Court, justices face issues that a judge acting alone does not. When we say that Justice Scalia prefers bright-line rules and Justice O'Connor prefers balancing tests, do we mean that those choices represent what each would choose were he or

\textsuperscript{15}If they are uniformly distributed, then it is true that Scalia will face a greater number of relatively liberal courts (liberal relative to him) than would Kennedy, but this would be balanced by the fact that Kennedy would need to worry about conservative opposition as well, so that extremism would not affect the conflict parameter $p$. 

31
she dictator? Or that Justice Scalia wants Justice O'Connor's preferred balancing test to be converted into the nearest bright-line rule? Such a bright-line rule would obviously differ from Scalia's preferred bright-line rule. That is, is he disagreeing with her incorporation of the subjective dimension or the placement of the rule (upwards or downwards) in the case space? Does he simply want a more extreme test or one that rejects one of the constituent dimensions of the rule? [examples]

Moreover, suppose that the production of legal quality (transparency) is in the hands of the opinion author. The optimal choice of doctrinal form will depend on who exactly is crafting the doctrine. It is largely the opinion author's ability alone that matters—so that Justice Scalia might prefer a balancing test if he himself were doing the writing, but might prefer the doctrine to take the form of a bright-line rule if Justice Kennedy is to be the justice responsible for crafting it. Or, since it is the opinion author that must bear the costs of authorship, one justice might want another to invest in a balancing test when he himself would simply go with the cheaper bright-line rule.\textsuperscript{16} In any case, the results above clarify how these issues might play out.

Next, I explore issues of doctrinal complexity.

\section*{4.8 Doctrinal Complexity}

Legal complexity has an effect on doctrinal choice, but the incentives for doctrinal choice also have an effect on legal complexity. The simpler relationship has already been mentioned, but I highlight it again here. The more complicated the area of the law, the greater the cost of increasing transparency (if transparency is endogenous) or the lower transparency will be outright (if transparency is exogenous). Either way, this relationship pushes towards the use of bright-line rules. On the other hand, where

\textsuperscript{16}I will simply note the further complication that for any of these reasons justices might differ over the rule form to use, which could suggest bargaining over such issues.
transparency is exogenously high or "cheap" to produce because the legal issues are more clear-cut, balancing tests can be safely used.

But there is another set of potential implications. The results above have been discussed in terms of a choice between a bright-line rule that simplifies a balancing test from two dimensions back to one dimension. Extending the logic to multiple dimensions, the same incentives should drive the choice to add a marginal dimension to any area of the law—at least where that dimension is subjective in nature (where it is fully objective and transparent, it would seem costless to include it in one's announced doctrine). Suppose that the current doctrine is itself a clear-cut balancing test and the Court is considering incorporating a third subjective dimension. Then, the results above should speak to when to add further complexity to the doctrine governing this issue area.

4.9 Precision

How do the results above differ when transparency and compliance are not the focus of the Court's doctrinal choice? In the analysis of transparency, it was assumed that incorrect decisions made under a balancing test were due to lower courts actively and intentionally evading the dictates of the higher court to whatever extent possible. Suppose instead that lower courts are faithful agents of the higher court. Do the dilemmas of doctrinal choice above disappear? Do balancing tests become "safe" choices, given the good intentions of lower courts? Is active non-compliance the only concern of a higher court in a judicial hierarchy? No—the difficulties in the application of legal rule to real-world cases extend beyond such concerns.

Another source of doctrinal misapplication remains—simple error due to the challenges of converting the dictates of broadly stated doctrines into the dispositions of simple cases. That is, as discussed earlier, the challenges in constructing and commu-
nicating a doctrine clearly and precisely will limit the ability of even a faithful agent to decide cases accurately. This is particularly true in the context of a more subjective dimension, as studied here. Lower courts, while seeking to obey the higher court, will not be sure exactly where the line is drawn. I call the degree to which lower courts can apply the higher court doctrine correctly precision.

Given faithful agents, we can simply assume the chance of an incorrect decision, rather than varying with lower court ideology, is fixed at some constant exogenous to the other parameters. Without loss of generality, I simply let \( p = \frac{1}{2} \).\(^{17}\) We now can extend the transparency results to analyze exogenous precision and endogenous precision in turn.

**Exogenous Precision** Again let the parameter capturing the width of the troublesome region be \( v \). Within \( v \) of the cut-line, the lower courts cannot tell what the correct disposition is. The harder it is to specify a limit on the troublesome dimension, given its complexity or inherent ambiguity, the higher \( v \) will be. For now, \( v \) is assumed to be exogenous, beyond the control of the higher court. Parameters \( a \) and \( s \) capture sensitivity and salience as before. The loss from sticking with the balancing test is then \( \frac{a}{2}(2v-v^2)^2 \), so that we can derive the following result:

**Corollary 4.7.** Given exogenous precision in specifying the balancing test, the optimal bright-line rule is preferable to the balancing test (standard) if and only if \( v > 1 - \frac{\sqrt{4v^2 - 2v}}{2} \) (approximately .2).

This result means that when the difficulty in announcing the desired balancing test is solely a matter of the exogenous precision with which it can be specified, (1) the Court should announce a bright-line rule instead of this standard if the degree of inherent imprecision is sufficiently high and (2) this choice does not depend on how sen-

---

\(^{17}\)One could think to let \( p \) vary to capture higher or lower chances of error, but in effect \( v \) already does this by denoting the size of the region of cases that could be erroneously decided.
sitive the preferred balancing test is to the subjective dimension. The intuition is that if the degree of inherent imprecision is sufficiently high, then the guaranteed losses due to the bright-line rule are preferred to the mistakes made under the standard. As in the exogenous transparency case, the sensitivity to dimension 1 affects both types of loss equally when imprecision is the cause of problematic doctrinal application.

What happens when \( v \) is endogenous instead of exogenous?

**Endogenous Precision.** Suppose that the more legally skilled the higher court is and the greater the legal work invested in crafting the balancing test, the greater the precision with which it can be applied (so that \( v \) will be lower). Assuming the same cost structure and weight \( c \), now for increasing precision rather than increasing transparency, the payoff when choosing a balancing test is now \( -\frac{\alpha}{2}(a(2v - v^2))^2 - c(1 - v)^4 \), so that we reach the following results:

**Corollary 4.8.** Optimal endogenous precision of the balancing test occurs at \( v^* = 1 - \frac{\alpha^8}{\sqrt{2c + a^2}} \).

**Corollary 4.9.** Optimal precision is decreasing in cost and increasing in sensitivity and salience.

**Corollary 4.10.** The optimal bright-line rule is preferable to the optimal balancing test (standard) if and only if \( c > \frac{\alpha^8}{14} \).

**Corollary 4.11.** The incentive to choose the optimal bright-line rule over the optimal balancing test increases with higher costs, and decreases with higher sensitivity and salience.

The greater the sensitivity to the subjective dimension and the lower the cost of specifying the cut-line, the more work will be invested in lowering imprecision and the lower the optimal level of imprecision, \( v^* \). The higher the cost of precision, the
higher the optimal level of imprecision, but, once the cost is high enough, then it no longer is optimal to stick with the balancing test.

Having active control over the quality of the opinion and thus the precision of the balancing test again means that sensitivity to the subjective dimension does have an impact on doctrinal choice (like the endogenous transparency case, but unlike exogenous transparency or precision). When precision is endogenous, the tradeoff between choosing the balancing test/standard and the bright-line rule depends both on the cost of specifying the test and on the sensitivity to the subjective dimension. Where precision is exogenous, the balancing test’s sensitivity to the first dimension is irrelevant; where it is endogenous, higher sensitivity pushes towards choosing the bright-line rule outright.

Overall, this suggests that the justices will choose balancing tests (even when they will be somewhat standard-like) for legal issues in which they have greater expertise or skill and where they have a greater bank of precedent to draw upon, whereas bright-line rules will make the better tactical choice in more complex or newer areas of the law, or in areas that resist quantification (think Justice Stewart’s “I know it when I see it” doctrine).

Figures 9 and 10 shows these results. Figure 9 shows the optimal precision level falling as the cost of precision rises as well as the cost level beyond which a bright-line rule becomes optimal. As cost increases, the level of precision that will be achieved will naturally decline, until it drops, in effect, to zero when it becomes optimal to switch to a bright-line rule (that is, the judge would give up on making the balancing test more precise and simply jettison the troublesome dimension entirely).

Figure 10 shows how the preferred choice between the optimal balancing test and the optimal bright-line rule varies with sensitivity to the troublesome dimension and to the cost of precision. While judges who face higher costs (perhaps due to lower
expertise in the issue area in question or greater complexity of that area of the law) will prefer bright-line rules, a greater concern for the first dimension will push towards a balancing test. When sensitivity is low, it is only those judges who face low costs or those issue areas where precision is easier to achieve that will yield balancing tests.

4.10 Conjunctive Rules

What happens when the preferred rule takes the form of a conjunctive test requiring various prongs to be met? Again, the bite comes in those situations where one of the component dimensions is standard-like in that there are concerns of transparency or imprecision. [To be added. The results are highly similar to those for balancing tests.]

5 Conclusion

Justices (or other appellate judges) concerned about the impact of their opinions on future case outcomes are indeed constrained in their choices. In short, the separation of rule creation and rule application creates striking incentives for rule creation.

I have offered a model that is simple, but which captures some of the fundamental tensions driving doctrinal choice, specifically that between bright-line rules and standards. The resolution of these tensions depends in a nuanced way on the degree of lower court conflict, the salience of the issue at hand, the complexity of that issue, the expertise and abilities of the rule crafter, the sensitivity of the preferred doctrine to varying case facts, and the nature of the doctrine's constituent factual dimensions.
6 Appendix: Proofs and Supplementary Formal Results

This requires \( b \leq 1 \) and \( 0 \leq a \leq b \), so that we have the configuration shown in Figure 3, in which the rule does not intercept the \( x \)-axis.

**Proof of Lemma 4.1.** If \( \bar{y} \) is set above the upper limits or below the lower limits of \( \bar{y} \), then the full region of area \( \frac{a}{2} \) is lost yielding a payoff of \( u = \frac{-s a^2}{4} \). Otherwise, the area of these regions is \( \frac{1}{2} (b - \bar{y}) \left( \frac{(b - \bar{y})}{a} \right) + \frac{1}{2} \left( 1 - \frac{b - \bar{y}}{a} \right) \left( \bar{y} - (b - a) \right) \). The payoff is then \( u = -\frac{1}{4a^2} \left( a^2 - 2a(b - \bar{y}) + 2(b - \bar{y})^2 \right) \) and \( \frac{\partial u}{\partial \bar{y}} = -\frac{1}{2a^2} \left( 2a - 4(b - \bar{y}) \right) \left( a^2 - 2a(b - \bar{y}) + 2(b - \bar{y})^2 \right) \) which is maximized at \( \bar{y}^* = b - \frac{a}{2} \) yielding a payoff of \( \frac{-s a^2}{16} \) (so that an interior value of \( \bar{y} \) is superior).

**Proof of Proposition 4.2.** The payoff using the bright-line rule \( \left( \frac{-s a^2}{16} \right) \) is greater than that using the balancing test \( \left( -s p a^2 (2v - v^3)/2 \right) \) which occurs when \( p > \frac{1}{16a^2(2-v)^3} \), which is equivalent to \( v > 1 - \frac{1}{2} \sqrt{4 - \frac{1}{\sqrt{b}}} \). \( \square \)

**Proof of Proposition 4.3.** The utility from the balancing test is \( -ps(a(2v - v^3))^2 - c(1 - v)^4 \), so that \( \frac{\partial u}{\partial v} = -4(-1 + v) \left( c(-1 + v)^2 + a^2ps (-2 + v) v \right) \) which is zero at \( v^* = 1 - \frac{psa}{\sqrt{ps(c + psa^2)}} \) (this is unique for interior \( v^* \)). The second derivative is \( \frac{\partial^2 u}{\partial v^2} = -4 \left( 3c(-1 + v)^2 + a^2ps (2 + 3(-2 + v) v) \right) \), which at \( v^* \) is \( -8psa \), which is negative making \( v^* \) a maximum. \( \square \)

**Proof of Corollary 4.4.** The comparative statics follow from the partial derivatives and can be signed as follows: \( \frac{\partial v^*}{\partial a} = -\frac{a}{c} \sqrt{\frac{ps}{c + psa^2}} < 0 \); \( \frac{\partial v^*}{\partial p} = -\frac{ac}{2} \sqrt{\frac{s}{ps(c + psa^2)^3}} < 0 \); \( \frac{\partial v^*}{\partial s} = -\frac{ac}{2} \sqrt{\frac{p}{s(c + psa^2)^3}} < 0 \); and \( \frac{\partial v^*}{\partial c} = \frac{ae^2p^2}{2(sp(c + psa^2))^2} > 0 \). \( \square \)

**Proof of Proposition 4.5.** The payoff using the bright-line rule \( \left( \frac{-s a^2}{16} \right) \) is greater than that using the bright-line rule \( \left( \frac{-c p a^2}{c + psa^2} \right) \) when \( c + psa^2 < 16cp \). If \( p < \frac{1}{16} \), then this cannot occur for positive values of the parameters, but otherwise it holds when \( c > \frac{p a^2}{16p - 1} \). \( \square \)
Proof of Corollary 4.6. If \( c \) is higher, the condition is directly easier to satisfy. Otherwise, the partial derivative of the right-hand side of the condition on cost \( \left( c > \frac{psa^2}{16p-1} \right) \) generates the comparative statics (assuming \( p > \frac{1}{16} \)): \( \frac{\partial}{\partial a} = \frac{2spa}{16p-1} > 0 \), \( \frac{\partial}{\partial p} = \frac{-sa^2}{(16p-1)^2} < 0 \), and \( \frac{\partial}{\partial s} = \frac{pa^2}{16p-1} > 0 \). The boundary on \( c \) for a bright-line rule is thus higher for higher \( a \), lower \( p \), and higher \( s \), each of which make the balancing test more attractive. \( \square \)

Proof of Corollaries 4.7, 4.8, 4.9, 4.10, and 4.11. These follow from Proposition 4.2 through Corollary 4.6, given fixed \( p = \frac{1}{2} \). \( \square \)
References


Cases Cited

_Roe v. Wade_, 410 U.S. 113 (1973)

7 Figures
Figure 1: Cases

Dimension 2
(speed limit)

55 m.p.h.

$x^1$

$x^2$

$x^3$
Figure 2: Bright-Line Rules vs. Balancing Tests

Dimension 2
(speed limit)

Rule A

Rule B

Dimension 1
(bad weather)

55 m.p.h.

$x^1$

$x^2$

$x^3$

Rule A is a constant bright-line rule (a fixed speed limit). Rule B is a flexible rule (the speed limit depends on the severity of weather conditions).
Figure 3: Losses using a Bright-Line Rule

The true preferred rule is \( y = b - \alpha x \). The shaded regions show the cases that will be decided incorrectly (due to under-inclusion and over-inclusion respectively) if the (optimal) bright-line rule is used instead of the true preferred rule of the higher court. (The case space runs from \((0,0)\) to \((1,1)\).)
The true preferred rule is \( y = b - ax \). The shaded regions can represent either the cases in which lower courts can potentially hide non-compliance (\( \nu \) indicates the potential amount of manipulability) or those cases that are subject to error due to imprecision of doctrinal specification.
Figure 5: Exogenous Transparency and Doctrinal Choice

The dotted line shows the division between choosing a bright-line rule versus the balancing test, with higher $\nu$ denoting lower transparency and $\rho$ denoting the hostility of lower courts.
The curves show, the optimal level of transparency the Court will invest in, for various levels of conflict or sensitivity (to the subjective dimension), given the cost of increasing transparency (lowering $v$). As conflict increases or as sensitivity increases, the curve shifts upward, lowering $v$ and increasing optimal transparency. The dotted portion of the curves show the region where the cost is so high that the Court should instead choose a bright-line rule instead of the balancing test (with this optimal degree of transparency). Unlike the effect on optimal transparency, the two parameters have opposite effect. High conflict shifts this cut-off leftward (extending the region in which a bright-line rule is chosen), but high sensitivity shifts this cut-off rightward (shrinking the region in which a bright-line rule is chosen).
Three perspectives on the optimal choice of doctrine is shown given the tradeoff between cost, salience, sensitivity, and conflict.
Figure 8: Equilibrium Residual Discretion and Effort
Figure 9: Optimal Precision of the Balancing Test

The curve shows the optimal level of precision of the balancing test given the cost $c$ of increasing precision (reducing $\nu$). The curve is solid where the balancing test would be chosen and dotted where the bright-line rule would be superior to the balancing test with optimal precision. This occurs once the cost is sufficiently high, relative to the sensitivity to the subjective factual dimension.
Figure 10: Doctrinal Choice given Optimal Precision

The optimal choice of doctrine is shown given values of sensitivity ($s$) and cost ($c$), with the bright-line rule optimal below the line (where the cost of precision is high and sensitivity to the subjective dimension low) and the balancing test optimal above (where the cost is low and sensitivity high).