

Political Constraints on Legal Doctrine: How Hierarchy Shapes the Law

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

July 21, 2011

Abstract

When higher court judges attempt to assert control over lower-court decision-making, do such hierarchical politics shape legal doctrine? Using a “case-space” model of choice between determinate doctrines (rules) and more flexible doctrines (standards), I argue that the structure of doctrine affects the application of and compliance with doctrine by lower courts, and this in turn affects choice among doctrinal structures. Doctrinal choice, legal complexity, lower court discretion, and the allocation of judicial resources are shown to depend on hierarchical conflict, the transparency of decisions, sensitivity to case facts, judicial expertise, salience, and issue complexity. These incentives have counterintuitive effects on lower court discretion and on doctrinal specificity, and they create odd patterns of ideological and doctrinal alignment. Ignoring these incentives undercuts our understandings of lower court compliance, of judicial ideology, and of the effects of collegiality on law.

Forthcoming, *Journal of Politics* (note: this is not the final version)

The Supreme Court and the lower federal courts play distinct roles in the production of legal policy. The top tier of a judicial hierarchy concentrates far more on doctrine—rule creation and articulation—while the bottom tier concentrates on application of rules to specific cases. As (Shapiro 2006, 273) puts it, the Supreme Court itself does not “routinely apply the rules and standards it announces. Instead, the Court has cast itself in an ‘Olympian’ role—announcing rules and standards from on high.” The justices of the Supreme Court rely on the legal opinions they hand down as vehicles for their legal policy goals; the judges on the Courts of Appeals, in the middle level of the federal judicial hierarchy, rely on their opinions to govern dispute resolution by the district court judges the bottom tier. Lower court judges can disagree with their hierarchical superiors about case outcomes and the rules determining them, for reasons of ideology or legal philosophy. Even when the lower courts are largely aligned with the higher court or simply seek to obey their superiors, it may still be difficult for the higher court to convey exactly what it wants in the full range of possible cases that can arise. How do the content and structure of legal doctrine affect compliance and rule application in the lower courts? And, as higher court judges construct legal doctrine to best get the case outcomes they want, does reliance on hierarchical application of doctrine affect choice of doctrine in the first place? In short, do hierarchical politics shape law?

I study the incentives driving the choice between determinate doctrines (bright-line rules) and more flexible or indeterminate doctrines (standards), in the face of concerns over hierarchical doctrinal application. Both bright-line rules and standards tell lower courts which factual dimensions to take into account when deciding cases and how. I argue that the key difference is that a standard incorporates a factual dimension that is qualitatively different—lacking full transparency or specificity—from dimensions that are capable of greater precision, specificity, and transparency.

Consider the Supreme Court’s rulings on death sentences for minors, which have focused on, *inter alia*, the maturity and sense of responsibility of the defendants. The Court could have instructed lower courts to explore the maturity of each defendant according to some set of guidelines

or perhaps a multi-pronged test, of which age or an IQ test could play some part as balanced against other information about maturity. Such a balancing test would be difficult to specify in the abstract and hard to monitor, and lower court judges could take advantage of this ambiguity to decide particular cases against the wishes of the higher court. It is indeed these very features that define a standard in the analysis below. In *Thompson v. Oklahoma* (1988), the Court (by plurality) implicitly rejected a maturity test, choosing instead a clear determinate cut-off based on calendar age alone, barring execution of offenders under 16. This bright-line rule was extended to those under 18 in *Roper v. Simmons* (2005), overruling *Stanford v. Kentucky* (1989), which had allowed execution of those under 18 and over 16.

Perhaps the best known bright-line rule is that suspects must be informed of their rights, *Miranda v. Arizona* (1966). Subsequent cases established further bright-line rules requiring suspects to explicitly request counsel after getting *Miranda* warnings, in *Davis v. U.S.* (1994), and to explicitly invoke the right to silence, in *Berghuis v. Thompkins* (2010). Heytens (2008, 2094) gives further criminal procedure examples of bright-line rules (in particular, of rules replacing previous standards). However, the Court has often rejected a bright-line approach in favor of more nuanced doctrines. They have done so even in the same areas of the law and for seemingly similar questions such as the right to counsel and the maturity-age question. In *Fare v. Michael C.* (1979), the Court established a “totality of the circumstances” test for juvenile waiver of Fifth Amendment rights mandating “evaluation of the juvenile’s age, experience, education, background, and intelligence, and into whether he has the capacity to understand the warnings given him.” In *Bellotti v. Baird* (1979), the Court invoked a maturity standard and explicitly rejected a bright-line calendar age rule, for unmarried women under 18 seeking an abortion without parental consent.

The justices themselves often specifically conflict over doctrinal form. In *Miranda*, Justice Clark wrote separately opposing a bright-line rule in favor of a “totality of the circumstances” standard. Returning to the issue of death sentences for minors, in *Stanford*, Justice O’Connor

wrote a concurrence arguing a proportionality analysis should be conducted, but still joined parts of the majority opinion, agreeing with the conservative plurality that there was no bar against the execution of those 16 and 17 years of age. Justice Brennan, dissenting on behalf of the liberal minority, argued that “the Eighth Amendment requires that a person who lacks that full degree of responsibility for his or her actions associated with adulthood not be sentenced to death,” which sounds like a standard, but he argued for a bright-line rule of 18 years of age even though there “may be exceptional individuals who mature more quickly than their peers, and who might be considered fully responsible for their actions prior to the age of 18.” Justice Scalia, usually portrayed as a strong proponent of bright-line rules, dissented in *Roper*, arguing in favor of a maturity standard. (Note that Justice Kennedy was in the majority in both cases, and his majority opinion in *Roper* explicitly discusses the over- and under-inclusiveness of a bright-line age rule.)

The puzzle is why the Court chooses standards in some areas of the law or at some points in time but not others, and why some justices prefer rules while others prefer standards, often inconsistently across different areas of the law. I present a model of legal doctrine that is designed to capture key incentives, constraints, and complications attending doctrinal choice, given the separation of rule creation and rule application across levels of the judicial hierarchy. My goal is to show how we can rationalize preferences over doctrinal structures from preferences over case outcomes in case space. In the model, doctrine is endogenous, with the content, structure, and even “legal quality” of the doctrine in the hands of the justices themselves. I argue that they choose instruments of hierarchical control based on the strategic trade-off between them, showing how this trade-off is affected by ideological conflict across the levels of the judicial hierarchy, judicial expertise, issue complexity, issue salience, and the sensitivity of the desired doctrine to varying case facts.

The separation of rule creation and rule application creates striking and counterintuitive incentives for doctrinal choice and legal policy. These findings raise new questions and shed new light on old ones, and I draw out the implication of the model’s results to address a series of substantive

issues. First, there is the question of how much discretion lower courts will have given optimal doctrinal choice by the higher court. Next, why are some doctrines far more detail specific or complex than others? To put this differently, how will the higher court allocate resources to the development of legal doctrine? Third, I discuss how the model can be interpreted slightly differently, to capture concern over legal precision as opposed to willful non-compliance. Finally, there is the relationship of ideology to doctrinal choice. In particular, I analyze the conventional wisdom that moderate or centrist justices prefer standards or balancing tests and that extreme justices tend to instead prefer bright-line rules. I discuss how collegiality complicates rule choice, with some thoughts on the implications of polarization within the higher court.

Theories of Doctrinal Choice

Opinions, rules, cases, and case facts are quite significant concepts—yet traditional political models of judicial policy-making, formal and otherwise, often paid little attention to them. One goal of the paper at hand is to continue the development of a model, the “case-space” model, that brings these concepts into sharp relief. The case-space model is a variant of and supplement to the policy-space model common in political science. It has its origins in Kornhauser (1992*a,b*) with further development in Cameron (1993), Grofman (1993), Cameron, Segal and Songer (2000), Lax (2003, 2007), Lax and Cameron (2007), Kestelleg (2007), and Landa and Lax (2009). It is tailored to capture the substance and institutional features of *judicial* policy making, putting case facts and legal doctrine at the analytic center, without rejecting a role for judicial preferences. A case-space model recognizes that 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, and organized by legal doctrine, which in turn can be a function of ideological preference. This model allows us to think about constituent elements of judicial choice and the politics of legal doctrine in new ways, opening up questions of the structure of choice and preference that are

obscured in more traditional approaches. Indeed, the analysis of doctrinal choice is emerging as a vibrant frontier in the bridging between legal theory and political science.

See Lax (2011) for a discussion of this “new judicial politics of legal doctrine” (also see the influential essay Tiller and Cross 2006). Six works are particularly relevant. First, Jacobi and Tiller (2007) model the choice between determinate and indeterminate doctrine given exogenous degrees of indeterminacy, conflict, and bias towards litigants. I move beyond this important work by unpacking these concepts further and focusing on different questions of doctrinal choice. I argue that the nature of constituent factual dimensions and the active crafting of doctrinal efficacy play key roles in doctrinal choice. The degrees of transparency, indeterminacy, precision, and lower court discretion then emerge endogenously. The relationship of ideology to doctrinal choice here is not one of judicial bias towards types of litigants but rather a complex relationship based on the desire for control case outcomes, conflict with lower courts, and the distribution of cases.

Second, in McNollgast (1995), doctrine is modeled as an announced policy point and the range of points around it that the higher court will accept from the lower courts. The Supreme Court might induce lower courts to comply by granting permission to be non-compliant in a wider range of cases, thus isolating the remaining lower courts for review. The Supreme Court cannot tell the location of a lower court choice until it takes a case for review, but it can tell whether the lower court has been compliant (randomly reviewing any non-compliant lower courts). This model is not meant to explain which factual dimensions a doctrine will include nor the choice between rules and standards, but rather focuses more directly on conflict and discretion. As will be seen, it yields different conclusions about this relationship than the model below.

Third, Staton and Vanberg (2008) is a policy-space model in which justices are strategically vague in their opinions to build institutional strength and prestige, manage limited resources, and defer to those with informational advantages. This trades off against diminished policy control. Specificity can increase compliance by making noncompliance more visible to external monitors

such as the public, but can risk laying bare judicial weakness if other actors will still not comply. They identify incentives new to the analysis of hierarchy and delegation. Some findings complement those here. For example, in their model, high clarity maximizes judicial leverage over policy implementation and low clarity delegates authority to non-judicial policymakers; I find that the higher court invests in clarity to better control lower courts. It is intriguing that both models reveal non-monotonic relationships between clarity and political factors (though not the same ones).

Fourth, Friedman (2005, 295) writes, “Not much scholarship is devoted to the impact that lower courts have on the Supreme Court’s exercise of its power of judicial review. Positive theory, however, indicates that the lower courts exert substantial influence over the Supreme Court. To the extent this is true, existing theories of judicial review are necessarily incomplete because they fail to take account of the gravitational pull of the lower courts.” Normative scholars neglect “how the Supreme Court governs the judicial hierarchy,” “likely because of their assumption that lower courts simply follow precedents” (302). Intriguingly, he posits that “whether the Supreme Court can rely on ‘rules’ or ‘standards’ when it decides cases—much mooted as a normative matter—may turn as much on questions of lower court compliance as on jurisprudential preferences.”

This possibility is the central focus of Heytens (2008), a law review article which notes that legal scholars rarely consider the high court’s pragmatic “need to craft rules that can and will be faithfully implemented by the lower court judges who have the last word in the overwhelming majority of litigated cases” (2046). In particular, “the existing literature about the choice between rules and standards has generally neglected its implications for the relationship between the Supreme Court and lower courts” (2057). On the other hand, political scientists have spent much time trying to measure lower court compliance, but “have been unable to advance a convincing explanation for why their own studies almost invariably find high levels of compliance” (Heytens 2008, 2047). Heytens then argues that the Supreme Court has “at its disposal a number of doctrinal tools that can be used to shape and direct lower court behavior” (2047) and presents numerous

examples of the Court so doing (e.g., the development of a more constraining doctrine for punitive damages in light of state court non-compliance). He suggests that anticipatory doctrine-crafting could explain high compliance with stated doctrines. My paper is a formalization of some of the intuition in this piece, which allows for a deeper analysis of the incentives for rule choice, showing how they depend on various (strategic and legal) contextual features. For example, Heytens's comment that a higher court might do better to "bar trial courts from relying on criteria it deems... to difficult to verify on appeal" (2048) connects nicely to the formal model here where higher court might drop a more subjective dimension from a legal doctrine.

Finally, Segal and Spaeth (2002, 92) argue—in response to the founding debate in judicial politics on the nature of judicial choice and the role of ideology in such choice—that U.S. federal judges are largely unfettered policy-makers and that Supreme Court justices in particular are almost completely free to decide the cases they hear as they wish. This perspective treats the final vote in each case as the end of the game, which implies that no sophisticated choices are necessary, that the justices are free to cast sincere votes. The number and incidence of liberal or conservative votes have thus become the focal points of most empirical analysis of the Court, often reducing even *Roe v. Wade* to just seven 1's and two 0's. But the final vote is *not* really the end of the game, except perhaps for a particular plaintiff and defendant. It is the start of the next stage of the larger game in which others react to the Court's opinion, most notably the lower courts which apply it in future cases. Even if the Court does have significant power and discretion, the justices still need to worry about compliance from lower courts and how to best communicate what they want from them. This suggests that the justices, even if casting sincere votes in the cases they hear, might craft rules that are insincere or strategic. They might announce a rule other than their truly preferred rule, thus leading to future cases being decided differently than the justices might themselves want.

As Segal and Spaeth themselves emphatically note (2002, 357), it is the opinions that accompany these final votes that "constitutes the core of the Court's policymaking process." While

casting a vote may be a relatively trivial act, crafting an effective opinion is not. Opinions do many things. One which I am largely setting aside here is justification 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 doctrine 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. 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 outcomes that one might desire in each possible case. And of course, lower courts are not perfect agents—as Friedman (2005, 295,300-1) put it, much legal scholarship “take[s] for granted that lower courts... follow the mandate of higher courts” even though “study after study... makes [it] clear that ideology plays a role in lower court decisions.... Law may hold sway in the lower courts, but ideology plainly does as well.”

A key argument below is that the justices *actively* craft their opinions to achieve their desired policy outcomes (see Lax and Cameron 2007, Heytens 2008). They are inhibited in their policy-making by uncertainty, ambiguity, and complexity. The Court itself recognizes that there can be “endless variation in the facts and circumstances, so much variation that it is unlikely that the courts can reduce to a sentence or a paragraph a rule that will provide unarguable answers” (*Florida v. Royer* (1983)). 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). The efficacy of policy-making through opinions and decisions and the compliance of lower courts with them are, therefore, endogenous to the choices justices make.

Types of Doctrine

Usually, a rule is said to be a determinate form of legal doctrine, in which the line of permissible conduct is clear-cut and specified in advance (*ex ante*), whereas a standard establishes a more flex-

ible 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*). However, definitions of rule types vary widely, with much conflict in particular over the differences—or lack thereof—between rules and standards (see Sullivan and Amar (1992, 57-69), Kaplow (1992, 559-62), Fallon (2001), and Johnston (1995)). As scholars have sought nuance, these terms have lost some distinctiveness: both can be clear or unclear, simple or complex, transparent or opaque, determinate or indeterminate. To be clear, no part of my argument rests on the labels “rules” and “standards,” and I acknowledge that the definitions I use might omit that which is of interest in other contexts.

In my terminology, a doctrine or rule is any logical partitioning of cases into equivalence classes (yes vs. no, winner vs. loser, legitimate search vs. unreasonable search, etc.). A *bright-line rule* does so clearly and cleanly. I will define a bright-line rule (or 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 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—it depends on whether the relevant factual dimensions are themselves problematic. A balancing test will still be rather “bright-line” if all dimensions can be transparently and precisely specified along with the relationship between them. What makes a particular balancing test a *standard* is the incorporation of a dimension over which doctrine cannot be so cleanly specified or applied. That is, what separates rules from standards is specificity and transparency, which in turn depend on the characteristics of the factual inquiry demanded.

The archetypal bright-line rule is a speed limit: “Drive no faster than 55 miles per hour.”

But a speed limit could also be defined as follows: “Drive at a reasonable and prudent speed” (as Montana once did). What counts as reasonable cannot be defined with the same precision as a numerical speed limit nor as transparently applied as a numerical speed limit (Judge A’s “reasonable” speed may be “unreasonable” to Judge B and whether this is so might not be so easy to observe). This makes a “reasonable” 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 assorted factors, including weather conditions. To check whether a rule was “correctly” applied, the higher court only has to know the car’s objective speed. A standard thus contains a degree of subjectivity which raises the possibility of lower-court error (due to imprecision) and 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 (by my terminology). For example, if “bad weather” simply meant low temperatures, then one could have a balancing test that was still bright-line. Lower court discretion under a standard then arises, not from purposeful delegation (which might be useful for other reasons), but as an inevitable side effect of using a doctrine that operates as a standard.

To return to the death-penalty-for-minors example, consider an inquiry into maturity and responsibility beyond a simple calendar age rule. This would inevitably be imprecise and somewhat opaque—maturity is subjective and inherently lacking in specificity, whereas age is not. Or consider execution of the mentally retarded. In *Penry v. Lynaugh* (1989), the Court permitted the execution of the mentally retarded, overruled by *Atkins v. Virginia* (2002). 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, were a particular IQ test result used as the cut-off, this would again be a bright-line rule.

Or, consider Fourth Amendment search-and-seizure doctrine. 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 explore 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, for other Fourth Amendment issues, the Court uses a “totality of the circumstances” test, which is obviously a standard: e.g., whether an informant’s tip establishes probable cause, *Illinois v. Gates* (1983), replacing the bright-line two-prong Aguilar-Spinelli test; whether a post-traffic-stop search is voluntary without being informed of the right to refuse, *Ohio v. Robinette* (1996); and whether a person reasonably believes himself “seized,” *Schneckloth v. Bustamonte* (1973). While the *Bellotti* abortion case mentioned earlier invoked a bright-line rule, consider again the *Roe v. Wade* trimester system for abortion restrictions. It is not a one-dimensional rule, but rather has three tiers, bright-line overall, but with standard-like judgment calls in the second and third trimesters. Still, this seems more rule-like than the “undue burden” standard formulated in Justice O’Connor’s dissent in *Akron v. Akron* (1983) and adopted by the plurality in *Planned Parenthood v. Casey* (1992).

It is quite unlikely that a given bright-line rule exactly captures what the justices of the Court would themselves do if they heard every case themselves. 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 standards (e.g., *Montejo v. Louisiana*, 2009), 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 may be pervasive.

Returning to the speed limit example, suppose cases have two dimensions: speed and weather conditions at the time. Figure 1 shows three cases: the driver in case x^3 was 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 acceptable (Y, a winner) or speeding (N, a loser). If the higher court (H) truly prefers a constant speed limit, then she can simply announce a bright-line rule, such as 55 mph (the horizontal line in Panel 1a), under which 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, creating a sliding scale. She would now prefer a lower speed limit the harsher weather conditions were (the balancing test line in Figure 1, under which only driver 1 is speeding, even though she is going the slowest).

Were H to prefer the balancing test but still announce the 55 mph bright-line rule, only case 2 would be correctly decided (i.e., consistent with H 's preferences). 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. 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.

If H heard all cases herself, she could simply apply her preferred balancing test. Suppose instead that H must delegate cases to lower courts. If speed and weather 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 and the test would be a bright-line

rule. Lower courts would apply H 's test as easily as would H herself. However, while speed is a hard and fast measurement, weather conditions are not. The weather dimension differs in observability and specifiability. 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. The informational gap will make noncompliance possible. It also means that a doctrine incorporating weather would be a standard, not a bright-line rule. And, given such ambiguities, even friendly lower courts might 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. This means that H must make a choice between the problems that arise in announcing a standard and the problems that arise in using a bright-line rule. Bright-line rules lead to some "incorrect" dispositions due to over- and under-inclusive where flexibility is desired. Standards do so because of noncompliance and error. The next step is to formalize this trade-off.

Optimal Doctrinal Choice

I assume throughout that the High Court H (treated for now as a unitary actor) wants to minimize the area of the case space that is incorrectly decided according to its preferences, suffering increasing marginal loss with respect to this area: if the area is A , then the Court gets a payoff of $-A^2$ (quadratic loss).¹ The Appendix contains a reference list of the various parameters defined below, the "moving parts" of the model.

¹A more complicated assumption would be that a particular case is weighted by how close to the case lies to the dividing rule. This complication would make for an intractable model (and has only been considered in simpler models of rule choice), and so I set this aside here.

Balancing Tests. If the Court’s desired doctrine were actually a simple bright-line rule, it could just announce it. Assume then that the Court’s preferred rule is a balancing test across two dimensions. 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, or ambiguous, dimension. 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. Hereafter, when I say balancing test, I mean one with a subjective dimension, and not a fully objective bright-line balancing test.

In a two-dimensional case space, with cases given as a pair $\{x, y\}$, a simple 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 slope captures the relative weights between the two dimensions. 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 subjective dimension 1 (the weather dimension).

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*, a fixed cut-off \bar{y} that does not vary with x , ignoring dimension 1. (See Figure 1.) 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 Panel 1a.) For this result and subsequent results, I assume that $0 \leq a \leq b \leq 0$ to match

the particular configuration shown, in which for any value of x there is some case that yields a *Yes* disposition. Other configurations are similar. 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. The optimal bright-line rule is a function of both a and b , but not s :

Proposition 1. *The optimal bright-line rule is $\bar{y}^* = b - \frac{a}{2}$.*

The minimized losses are the shaded regions in Panel 1a. If $a = 0$, there is no loss. For higher a (steeper slope), the more over- and under-inclusiveness there is under the optimal bright-line rule.

Choosing a “Standard.” If the Court includes the subjective first dimension in its announced doctrine, then it invokes a balancing test which is a standard. There are two potential sources of trouble. 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 problem is that lower courts might 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 (a principal-agent problem), 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).

Let the width of the troublesome region extend v in either direction from the proposed standard, as shown in Panel 1b, so that the higher v is, the larger the more “standard-like” the standard is. Low v means higher transparency. As v goes to 0, this “standard” devolves to a bright-line rule. Choice between doctrinal forms will depend on whether v is exogenously fixed or endogenous in that the justices can actively engage in affecting it through careful drafting of their opinions. I discuss varying interpretations of the v parameter below.

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.” 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.²

Let there be two types of lower court, an allied lower court with the same preferences as the higher court and a hostile lower court that wants all cases in the troublesome region to be decided *No* (or, equivalently as far as equilibria, all such cases to be decided *Yes*; the model can be extended to other configurations of hierarchical disagreement as shown in the online appendix). Let the hostile court occur with probability p . 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. The gray region in Panel 1b underneath the balancing test is the set of cases in which the courts disagree and the lower court will disobey if possible; in the gray region above the balancing test, the courts agree so compliance will occur even though lower courts could successfully disobey. Outside this gray region, lower courts comply. Next, we compare the optimal bright-line rule to the balancing test:

²In the Cameron, Segal and Songer (2000) model of higher court/lower court interaction, 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. Here, I am, in effect, taking this game back one step, asking how the Supreme Court can set up a better compliance game for itself by crafting legal doctrine in anticipation of potential non-compliance.

Proposition 2. *Given an exogenous level of transparency, the optimal bright-line rule is preferable to the balancing test (standard) if and only if $p > \frac{1}{8v^2(2-v)^2}$.*

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 transparency and hierarchical conflict, but not on the sensitivity to the subjective dimension nor on salience. The greater the likelihood of conflict (and thus noncompliance), the greater the desirability of the bright-line rule; the lower transparency, the greater the desirability of the bright-line rule. 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. 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 balancing test, so they drop out of the doctrinal choice calculus, at least when transparency is exogenous.

Endogenous Transparency. Assume instead that the Court can invest in the quality of legal doctrine and increasing the transparency of lower court compliance. 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.³ Greater legal quality, costly to produce, might also increase the impetus to comply with higher court doctrine.

³Heytens 2008 (2062, 2089) thus explains the Court's establishment of such a "duty of explanation" in a sentencing case, *North Carolina v. Pearce* (1969), and for summary judgment based on qualified immunity, culminating in *Saucier v. Katz* (2001). More generally, Friedman (2005)

Formally, I assume that the investment of costly effort reduces v . The cost of transparency can depend on the ability, experience, and expertise of the judge/justice crafting the doctrine; how intensively fact-laded the issue area is; and its inherent complexity. That is, there are two interpretations of the cost of writing more effective opinions (or, for that matter, of the source of v where exogenous). One is that some judges (and perhaps their clerks) are different from others in their ability to craft high quality legal rules, perhaps varying across issue areas. The other is that it is mainly variation across issue areas that matters, with some areas of the law more amenable to precise delineation of rules. Since I set aside collegiality for now, the latter seems the better interpretation to keep in mind. The formal technology is similar, of course.

What is the optimal level of v ? And, given this, should the Court choose this optimal balancing test or the optimal bright-line rule? I assume a cost function with increasing marginal cost of transparency such as makes optimal v^* an interior solution. Let the cost of creating an endogenous-transparency balancing test be $c(1 - v)^4$, with c the cost weight for producing transparency.⁴

Proposition 3. *Optimal balancing test endogenous transparency is $v^* = 1 - \frac{psa}{\sqrt{ps(2c+psa^2)}}$. Optimal transparency decreases in cost and increases in sensitivity, conflict, and salience.*

Panel 2a shows the effects of the various parameters on optimal transparency. As cost increases, optimal transparency naturally drops. As conflict, sensitivity, or salience increases, the curve shifts upwards, so that for any fixed cost weight c , optimal transparency is higher. The intuition is that greater lower court hostility pressures the higher court to invest in transparency, while higher sensitivity to the subjective dimension means that a wider range of cases are subject to potential notes, “some empirical work, and some experience, suggest that [sometimes] the Court will have to employ very specific tests to ensure its mandates are followed” (304-5). Indeed, Heytens (2008, 2079) relates how Justice Kennedy established guideposts in *State Farm v. Campbell* in such great detail that Justice Ginsburg accused him of giving “marching orders” to the lower court.)

⁴Quadratic cost yields the same substantive findings, but requires an additional assumption for an interior solution and yields a less elegant mathematical explication. See Reviewer’s Appendix.

non-compliance, again pushing towards an investment in transparency. These results define the optimal balancing test, which can then be compared to the optimal bright-line rule:

Proposition 4. *The optimal bright-line rule is preferable to the optimal balancing test (standard) if and only if $c > \frac{psa^2}{16p-2}$. The incentive to choose the optimal bright-line rule 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 worthwhile. In Panel 2a, higher conflict, higher sensitivity, and higher salience all shift 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 point. A bright-line rule is preferable where costs are high, conflict is high, salience is low, and sensitivity is low.

Panel 2b reveals more complicated trade-offs between these parameters in optimal doctrinal choice. The two panels represent two 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. A dividing line is shown for higher or lower values of the remaining parameters. While judges who face higher costs (perhaps due to lower expertise in the issue area in question or greater legal complexity) 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 low complexity (given their own ability or the issue area's features) that will choose balancing tests. Overall, this suggests that the judges will choose balancing tests (even when they will be somewhat standard-like) for legal issues in which they have greater expertise or skill, where they have a greater bank of precedent to draw upon, or where the issue area is more amenable to delineation. Bright-line rules will make the better tactical choice in more complex or newer areas of the law, or given dimensions that resist easy quantification.

Discretion and Effort. 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? Next, what level of effort will be expended given optimal rule choice? The answers can be derived from Propositions 3 and 4. 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 take the value v^* , (that is, it maps to transparency, which is smaller when the resulting region of potential non-compliance is larger). Effort is the cost paid by the Court (zero for a bright-line rule). The results are shown in Figure 3. The level of residual discretion given optimal doctrinal choice depends on the sensitivity of the balancing test, the costs of reducing discretion, and the degree of conflict between the courts. The most striking feature of these results is that discretion and cost are non-monotonic, discontinuous functions. For example, as the cost of crafting doctrine increases, discretion given to the lower courts initially increases, because the Court economizes by writing lower quality rules—but eventually, as the costs continue to rise (say, in highly technical areas where supervision is difficult) the optimal choice transitions to a bright-line rule, under which some cases are automatically decided incorrectly and lower court discretion vanishes. Hierarchical conflict reveals a similarly discontinuous effect when the Court breaks in favor of the bright-line rule, but 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: the Court will still use a bright-line rule, until the point at which the balancing test is optimal, but one where the residual discretion is high. As sensitivity increases further, residual discretion increases then decreases.

Substantively, these formal comparative statics mean that 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. Finally, note that increasing the higher court’s resources (or choosing justices with greater legal ability increases higher court control but also can affect doctrinal choice, leading to great use of standards instead of bright-line rules. More resources will also lead to greater doctrinal complexity.

Higher conflict within the hierarchy should lead to less discretion for lower courts, as the higher court increasingly turns to bright-line rules or at least invests in more constraining standards (cf. McNollgast 1995, which argues that the Court relaxes its hold when faced with conflict). The relationship of conflict to doctrinal development is more complicated given the non-monotonic relationship discussed earlier. We might expect to see the most development not when the lower courts are so allied with the higher court that specificity is unnecessary, nor when the lower courts are so hostile that standards are too dangerous, but rather in the middle ground when there are lower courts to constrain and it is necessary and feasible to do so.

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 transparency is exogenously high or “cheap” to produce because the legal issues are more clear-cut, balancing tests can be safely used. Figure 3

shows how complexity is affected given changes in the model parameters. Not all these effects would be intuitive without the model to guide analysis. For example, for any given cost parameter, as conflict increases, doctrinal complexity, in the sense of detailed specification, will increase to better control the lower courts... until it is so costly to do so given rising conflict that the higher court instead drops down to a straightforward bright-line rule. There is another potential implication. 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.

Precision. Setting aside any principal-agent problems, suppose instead that lower courts are faithful but imperfect agents of the higher court. One can call the degree to which lower courts can apply the higher court doctrine correctly *precision*. The formal results and most discussion above can be adjusted to suit this interpretation of the v parameter.⁵ Thus, the model can speak to both the “agency” (or principal-agent) perspective on hierarchy, as above, and the “team” perspective (see Kornhauser 1999), by recognizing issues of error correction and reduction.

Ideology, Collegiality, and Polarization. Work on judicial choice can conflate what is best for a given justice with what is best for the Court (Vermuele 2005). So far, I have not considered how the

⁵One way to do this is to let the probability of correct choice within the v region be $p = \frac{1}{2}$, so that both dispositions are equally likely within the imprecise region. Then, if the balancing test has exogenous precision, the optimal bright-line rule is preferable to the balancing test if and only if $v > 1 - \frac{\sqrt{4-\sqrt{2}}}{2}$. For exogenous precision, the Court should announce a bright-line rule if it is sufficiently low. If precision is endogenous, then optimal endogenous precision of the balancing test occurs at $v^* = 1 - \frac{as}{\sqrt{s(2c+a^2s)}}$, and the optimal bright-line rule is preferable to the optimal balancing test (standard) if and only if $c > \frac{a^2s}{14}$. These follow from the proofs for transparency.

implications of the model apply differently to individual justices as opposed to the Court. While full consideration of the complicated relationship between hierarchy and collegiality is beyond the scope of this paper (see Kestelec 2011, connecting lower court collegiality and hierarchy), I informally discuss how the model above can improve our understanding of this relationship.

I begin by considering how judicial ideology relates to doctrinal choice. There are actually two potential relationships. The first is that doctrinal choice corresponds to ideology directly, so that liberals and conservatives will split over doctrinal form. The more complicated possibility would be that liberals and conservatives (the extremes) are similar as compared to moderates. For example, it is argued that more extreme justices such as Scalia tend to prefer bright-line rules whereas moderates such as O'Connor prefer standards or balancing tests (e.g., Shapiro 2006, 274). Can this be explained by the incentives shown here? When should we expect "ends-against-the-middle" doctrinal choice versus more straightforward ideological splits over doctrine?

One possibility is that such justices are so extreme that they already directly prefer a rule of "always" or of "never." More moderate justices might still prefer balancing tests, but this makes the debate rather trivial. The more interesting questions of doctrinal choice arise for those extreme justices short of such extremes, who might still prefer a balancing test. Assume then that that is the case. Let them agree on sensitivity (slope) but differ in how high or low the line is drawn (the intercept). The justices with higher intercepts want more *Yes* outcomes; those with lower intercepts/lines want more *No* outcomes. But they each prefer the same tradeoff between the two dimensions. In such a configuration, the role of the differing intercept plays a simple role in doctrinal choice: none, given Proposition 4. 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 then orthogonal to the choice of doctrinal structure.

Suppose that the justices instead vary in terms of sensitivity to the subjective dimension. Then, we should not find ends-against-the-middle, but rather that those with higher sensitivity prefer bal-

ancing tests (see Figure 2b and Proposition 4). 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 (the degree to which good faith exonerates an improper search and seizure). Either way, we would expect a split along party lines on doctrinal structure.

Recall Figure 3, 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 2, in which sensitivity to the subjective dimension does not affect doctrinal choice even if sensitivity does vary by ideology.

There are two further wrinkles to the ideology/doctrine debate, as 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. 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 set of cases for which transparency might create a problem would then be much smaller 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 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 relative.

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. 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 Figure 2b). The intuition is that O'Connor is less worried about lower court discretion since she has a greater expectation they will decide as she would anyway. Scalia will find fewer allies below and so will do better accepting the losses of a bright-line rule than using a balancing test. 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 discretion, however, would be less clear-cut: discretion is monotonic with respect to conflict but not 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.

However, one further complication is the varying perspectives of opinion authors versus other members of the majority (see, e.g., Heytens 2008, 2092 on the split over the granting of discretion in summary judgment-qualified immunity cases, in *Crawford-El v. Britton* (1998)). 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 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 in the case space? Does he simply want a more extreme test or one that rejects one of the constituent dimensions of the rule?

Note that a more extreme justice, say more conservative than the author, might actually like some non-compliance from conservative lower courts. Such a justice might prefer a foggier standard, then, depending on how the lower courts are configured vis-a-vis the higher court and the standard in question. For that matter, say there is a justice sufficiently extreme so as to be higher than the troublesome region of a bright-line rule as shown in Figure 1, Panel 1a. The most extreme justice would be one who prefers the entire speed case space to be treated as *Yes*. Suppose he has to choose between the bright-line rule as shown and the balancing test as shown, both for a more moderate opinion author. While the author would prefer the balancing test representing her true preferences to be enacted, the extreme justice will be indifferent between the bright-line rule and the balancing test: The shaded triangle on the left would now be handled correctly (*Yes*) according to the extreme justice's own preferred rule at the expense of the shaded triangle on the right, which is now incorrectly decided (*No*) but these two triangles are of equal area.⁶ (For the author,

⁶Equal impact only holds under the area-utility assumption; if the extreme justice cared more about the righthand triangle because those cases are further away from his preferred rule along the top of the square, then the bright-line rule would actually be better.

both these triangles now have wrong dispositions.) Similarly, to extend the polarization argument above, while each wing may prefer its own bright-line rule to a moderate justice's preferred standard, these will differ, with the liberal one being "high" and the conservative one "low" (or vice versa)—and each wing might prefer the moderate standard to the "wrong" bright-line rule.

Moreover, suppose that the production of legal quality (transparency) *is* in the hands of the opinion author. The optimal choice of doctrinal form can depend on who exactly is crafting the doctrine. It could be largely the author's ability that matters—so that Justice Scalia might prefer a balancing test if he himself were crafting it, but might prefer a bright-line rule if Justice Kennedy is to craft 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. There is also the burden of authorship to be taken into account. As shown in Lax and Cameron (2007), the most preferred opinion can vary depending on whether a justice personally has to pay the cost of crafting the rule or whether another justice will. Justice A can want Justice B to shrink the non-compliance region under a standard to zero by investing in opinion quality, but Justice A might not be willing to invest the required amount of effort himself or herself.⁷

Another question is what will happen in the end if justices do prefer different doctrines. One treatment of this is Lax (2007) showing that a median rule exists even when there is not a single median justice: there is an "implicit median rule" that is the baseline prediction, so that collegiality alone is shown to be insufficient for non-"median" outcomes. (Institutional features and incentives also affect collegial rule construction). For now, note that where the justices agree about the slope but not the intercept of the balancing test, the implicit median rule would also be a balancing test with that slope and the median intercept. Where they agree on the intercept and not the slope, the implicit median rule would have that intercept and the median slope. Where there is more compli-

⁷A justice might, of course, assist with such efforts with suggestions and comments—what is often seen as bargaining at the expense of the author could work to his/her benefit.

cated variation, the implicit median rule can take a very different structural form, and, as shown in Lax (2007) need not be a straightforward balancing test. Connecting that paper (wherein the issues of hierarchy-based incentives rule choice are set aside except for a result on enforceability) with this paper (which focuses on those very incentives), extending and integrating both lines of analysis, is a natural next step in this research. For now, this informal discussion sketches some intriguing aspects of the collegiality-hierarchy intersection, suggesting that this type of modeling, focusing on the judicial politics of legal doctrine, provides a promising way to move forward. It also demonstrates how much more complicated judicial choice can be than some accounts suggest.

Conclusion

The primary debate in judicial politics is whether judicial decisions depend on more than legal precedents, principles, and text—and whether and how external, collegial, or hierarchical politics shape law. I have argued that justices (or other appellate judges) concerned about the impact of their opinions on future case outcomes are indeed constrained in their choices, and I have shown that the separation of rule creation and rule application creates striking incentives for the former. In short, hierarchical politics do shape law.

Moreover, while Court observers often attribute judicial use of different types of rules to idiosyncratic preference, I show that there are clear incentives for rule choice, providing micro-foundations for the variation therein. The model captures some of the fundamental tensions driving doctrinal choice, specifically the choice 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, 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. These factors can induce insincere doctrinal choice and affect the care taken in crafting legal doctrine, that is, the investment in legal “quality.” Doctrinal choice, doctrinal complexity,

lower court discretion, and the allocation of judicial resources to an area of the law are all shown to be affected by hierarchical politics (thus connecting hierarchical politics to the rule of law).

The model suggests that our understanding of how ideology impacts the law may be incomplete or naive, because intra-Court collegial politics may be connected in striking ways to hierarchical politics. A justice's underlying liberalism or conservatism can lead to nonintuitive incentives for doctrinal choice. These incentives create odd patterns of ideological and doctrinal alignment, with strange bedfellows on both wings of the court opposing the middle. Some may find it disconcerting for the rule of law that the selection of winners and losers can depend on how liberal or conservative a given Supreme Court justice is. It is even more disconcerting that doctrine can depend on judicial strategy. While disagreement on the bench can sometimes be attributed to sincere beliefs in different legal philosophies rather than crude ideological differences, strategic incentives that undercut such sincere choices undermine even that defense.

The rules/standards model shares some insights with the traditional delegation literature, which tends to focus on congressional and bureaucratic institutions. For example, the connection between doctrinal choice and intra-hierarchy conflict resembles the "ally principle" in standard delegation models (one gives more discretion to those with similar preferences). These analogies might be explored, applying the rules/standards model to those contexts to the extent that legislatures and bureaucracies also enact preferences by constructing rules for dividing up the space of possible actions into permissible and impermissible.

Future work might extend the model to explore other forms of legal rules and other tradeoffs, as well as further exploring collegiality in the judicial hierarchy. One could explore how transparency and legal quality are achieved. Thinking about doctrine in the way suggested here, of course, also requires thinking about new ways to measure legal preferences and output empirically, going beyond unidimensional policy points or the percent of votes cast in the liberal direction, and even beyond recent advances in the ideological location of opinions.

Formal Appendix

The variables in the formal model are: a – sensitivity to the subjective factual dimension (the slope parameter, capturing the trade-off between the two dimensions); b – the intercept of the rule with the axis; s – salience of the issue area; v – lack of transparency (higher values mean a wider troublesome range, this can be exogenous or endogenous); p – likelihood of a hostile lower court; c – cost weight for producing transparency; a_L – sensitivity for the lower court when hostile; and b_L – intercept for the lower court when hostile. I assume that for any value of x , there is some value of y that yields a *Yes* disposition, which requires $b \leq 1$ and $0 \leq a \leq b$ (yielding the configuration shown in Figure 1. For constraining the lower court to be outside the troublesome region in the exogenous transparency solution, it suffices to assume that $b_L \leq b - av$ and $a_L \geq \frac{a-(b-b_L)}{1-v}$ (note the latter always holds if the two courts have the same sensitivity ($a_L = a$) but different slopes). For the endogenous transparency solution, this would invoke v^* , which is implicitly defined.

Proof of Prop. 1. If \bar{y} is set above the upper limits or below the lower limits of \hat{y} , then the full region of area $\frac{a}{2}$ is lost (payoff $u = \frac{-sa^2}{4}$). Otherwise, the area is $\frac{1}{2}(b-\bar{y})\left(\frac{b-\bar{y}}{a}\right) + \frac{1}{2}\left(1 - \frac{b-\bar{y}}{a}\right)(\bar{y} - (b-a))$. Utility is then $u = \frac{-((a^2-2a(b-\bar{y})+2(b-\bar{y})^2)^2 s)}{4a^2}$ with $\frac{\partial u}{\partial \bar{y}} = \frac{-((2a-4(b-\bar{y}))(a^2-2a(b-\bar{y})+2(b-\bar{y})^2)s)}{2a^2}$ with maximum at $\bar{y}^* = b - \frac{a}{2}$ (payoff $u^* = \frac{-sa^2}{16}$). The combined area of the loss region is $\frac{a}{4}$. \square

Proof of Prop. 2. The troublesome region area is $a(2v - v^2)$ (but only half is in contention). The bright-line rule payoff is greater than the balancing payoff $\left(-\frac{1}{2}spa^2(2v - v^2)^2\right)$ if $p > \frac{1}{8v^2(2-v)^2}$ (or, $v > 1 - \sqrt{1 - \sqrt{1/8p}}$). \square

Proof of Prop. 3. The balancing test payoff is $-\frac{1}{2}ps(a(2v - v^2))^2 - c(1 - v)^4$, so that $\frac{\partial u}{\partial v} = -4(-1 + v)(c(-1 + v)^2 + \frac{1}{2}a^2ps(-2 + v)v)$ which is zero at $v^* = 1 - \frac{psa}{\sqrt{ps(2c+psa^2)}}$ (unique for interior v^*). The second derivative is $\frac{\partial^2 u}{\partial v^2} = -4(3c(-1 + v)^2 + \frac{1}{2}a^2ps(2 + 3(-2 + v)v))$, which at v^* is $-4pa^2$, which is negative making v^* a maximum. Comparative statics follow from the partial derivatives and can be signed as follows: $\frac{\partial v^*}{\partial a} = -\left(2c\sqrt{\frac{ps}{(2c+psa^2)^3}}\right) < 0$; $\frac{\partial v^*}{\partial p} = -2ac\sqrt{\frac{s}{p(2c+psa^2)^3}} < 0$; $\frac{\partial v^*}{\partial s} = -ac\sqrt{\frac{p}{s(2c+psa^2)^3}} < 0$; and $\frac{\partial v^*}{\partial c} = \frac{as^2p^2}{(sp(2c+psa^2))^{\frac{3}{2}}} > 0$. \square

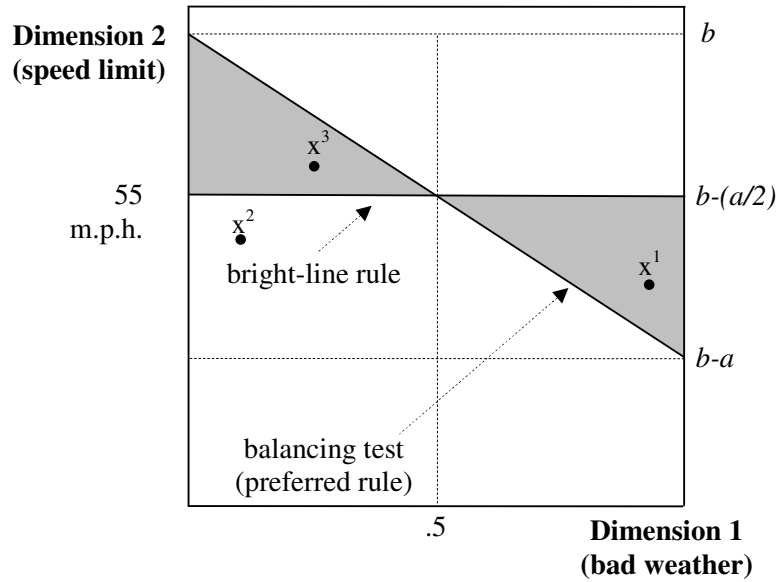
Proof of Prop. 4. The bright-line rule payoff $\left(\frac{-sa^2}{16}\right)$ is greater than that of the standard $\left(\frac{-cpsa^2}{2c+spsa^2}\right)$ when $c + \frac{psa^2}{2} < 8cp$. If $p < \frac{1}{8}$, then this cannot occur for positive values of the parameters, but otherwise it holds when $c > \frac{psa^2}{16p-2}$. 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 ($p > \frac{1}{8}$): $\frac{\partial}{\partial a} = \frac{spsa}{8p-1} > 0$, $\frac{\partial}{\partial p} = \frac{-sa^2}{(8p-1)^2} < 0$, and $\frac{\partial}{\partial s} = \frac{pa^2}{16p-2} > 0$. The boundary on c for a bright-line rule is thus higher for higher a or s , or lower p , each of which make the balancing test more attractive. □

References

- Cameron, Charles M. 1993. "New Avenues for Modeling Judicial Politics." Prepared for delivery at the Conference on the Political Economy of Public Law, Rochester, N.Y.
- Cameron, Charles M., Jeffrey A. Segal and Donald R. Songer. 2000. "Strategic Auditing in a Political Hierarchy: An Informational Model of the Supreme Court's Certiorari Decisions." *American Political Science Review* 94:101–16.
- Fallon, Richard H. 2001. *Implementing the Constitution*. Cambridge, MA: Harvard University Press.
- Friedman, Barry. 2005. "The Politics of Judicial Review." *Texas Law Review* 84:257–337.
- Grofman, Bernard. 1993. "Public Choice, Civic Republicanism, and American Politics: Perspectives of a 'Reasonable Choice' Modeler." *Texas Law Review* 71:1541–1587.
- Heytens, Toby J. 2008. "Doctrine Formulation and Distrust." *Notre Dame Law Review* 83:2045.
- Jacobi, Tonja and Emerson H. Tiller. 2007. "Legal Doctrine and Political Control." *Journal of Law, Economics, & Organization* 23(2):326–45.
- Johnston, Jason Scott. 1995. "Bargaining under Rules versus Standards." *Journal of Law, Economics, & Organization* 11(2):256–81.
- Kaplow, Louis. 1992. "Rules versus Standards: An Economic Analysis." *Duke Law Journal* 42(3):557–629.
- Kastellec, Jonathan P. 2007. "Panel Composition and Judicial Compliance on the United States Courts of Appeals." *Journal of Law, Economics & Organization* 23(2):421–41.
- Kastellec, Jonathan P. 2011. "Hierarchical and Collegial Politics on the U.S. Courts of Appeals." *Journal of Politics* Forthcoming.

- Kornhauser, Lewis. A. 1992a. "Modeling Collegial Courts I: Path Dependence." *International Review of Law and Economics* 12:169–85.
- Kornhauser, Lewis A. 1992b. "Modeling Collegial Courts II: Legal Doctrine." *Journal of Law Economics & Organization* 8:441–70.
- Kornhauser, Lewis A. 1999. "Appeal and Supreme Courts." *Encyclopedia of Law and Economics* 7200:45–62. Available at <http://encyclo.findlaw.com/7200book.pdf>.
- Landa, Dimitri and Jeffrey R. Lax. 2009. "Legal Doctrine in Collegial Courts." *Journal of Politics* 71(3).
- Lax, Jeffrey R. 2003. "Certiorari and Compliance in the Judicial Hierarchy: Discretion, Reputation, and the Rule of Four." *Journal of Theoretical Politics* 15(1):61–86.
- Lax, Jeffrey R. 2007. "Constructing Legal Rules on Appellate Courts." *American Political Science Review*. 101(3):591–604.
- Lax, Jeffrey R. 2011. "The New Judicial Politics of Legal Doctrine." *Annual Review of Political Science* 14:131–57.
- Lax, Jeffrey R. and Charles M. Cameron. 2007. "Bargaining and Opinion Assignment on the U.S. Supreme Court." *Journal of Law, Economics, and Organization* 23(2):276–302.
- McNollgast. 1995. "Politics and the Courts: A Positive Theory of Judicial Doctrine and the Rule of Law." *Southern California Law Review* 68:1631–89.
- Segal, Jeffrey A. and Harold J Spaeth. 2002. *The Supreme Court and the Attitudinal Model Revisited*. New York: Cambridge University Press.
- Shapiro, Carolyn. 2006. "The Limits of the Olympian Court: Common Law Judging versus Error Correction in the Supreme Court." *Washington and Lee Law Review* 63:271–337.
- Staton, Jeffrey K. and Georg Vanberg. 2008. "The Value of Vagueness: Delegation, Defiance, and Judicial Opinions." *American Journal of Political Science* 52(3):504–19.
- Sullivan, Kathleen M. and Akhil Reed Amar. 1992. "The Supreme Court, 1991 Term." *Harvard Law Review* 106(1):19–122.
- Tiller, Emerson H. and Frank B. Cross. 2006. "What is Legal Doctrine?" *Northwestern University Law Review* 100(1):517–34.
- Vermuele, Adrian. 2005. "The Judiciary is a They, Not an It: Interpretive Theory and the Fallacy of Division." *Journal of Contemporary Legal Issues* 14:549–84.

Panel 1a: Losses using a Bright-Line Rule



Panel 1b: Losses using a Balancing Test/Standard

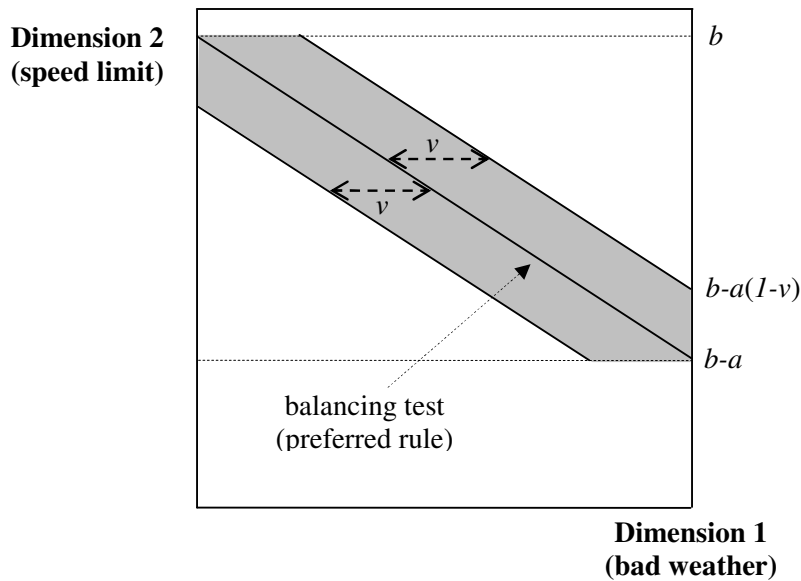
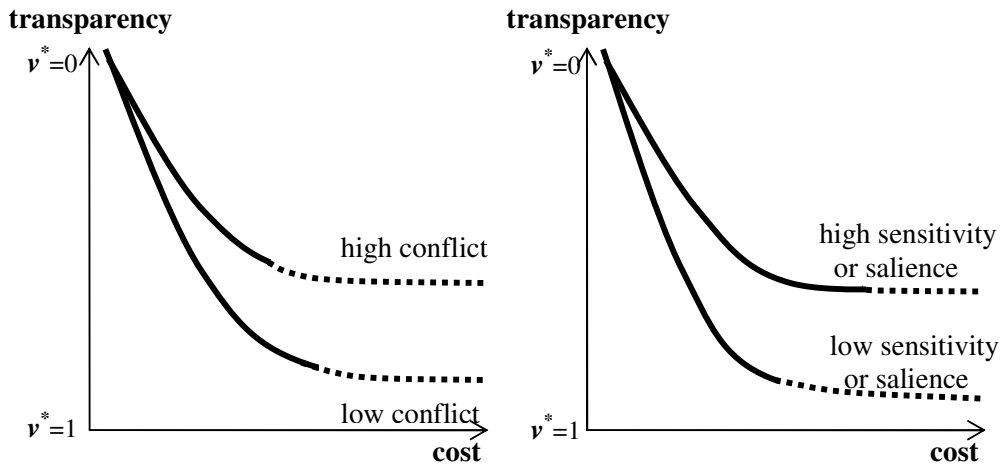


Figure 1: *Cases, Rules, and Tradeoffs*. Three cases are shown: x^1 , x^2 , and x^3 . The bright-line rule ignores weather (an absolute speed limit). Rule B is a balancing test which makes the limit depend on the severity of weather conditions. In the top panel, the shaded regions shows the cases decided incorrectly (due to under-inclusion and over-inclusion respectively) if the (optimal) bright-line rule is used instead of the true preferred rule. In the bottom panel, the shaded regions can represent either the cases in which lower courts can potentially hide non-compliance (v captures the potential amount of manipulability) or those cases that are subject to error due to imprecision of doctrinal specification.

Panel 2a: Optimal Transparency



Panel 2b: Doctrinal Choice given Optimal Transparency

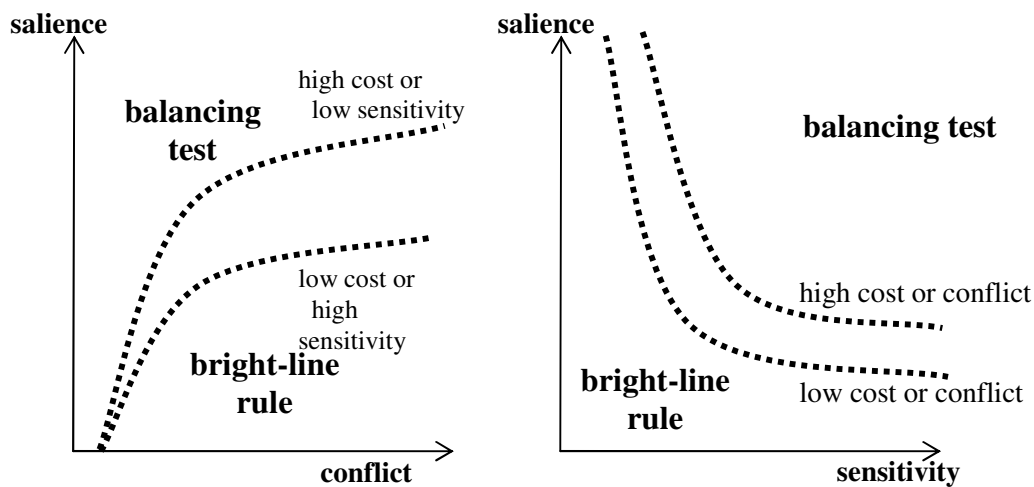


Figure 2: *Transparency and Doctrine*. The top panels show the optimal level of transparency the Court will invest in, for various levels of conflict or sensitivity (to the subjective dimension), given the cost of increasing transparency (lowering v). As conflict increases or as sensitivity increases, 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). The bottom panels show two perspectives on the optimal choice of doctrine is shown given the tradeoff between cost, salience, sensitivity, and conflict. In the left panel, the salience and sensitivity labels could be swapped to see yet another perspective.

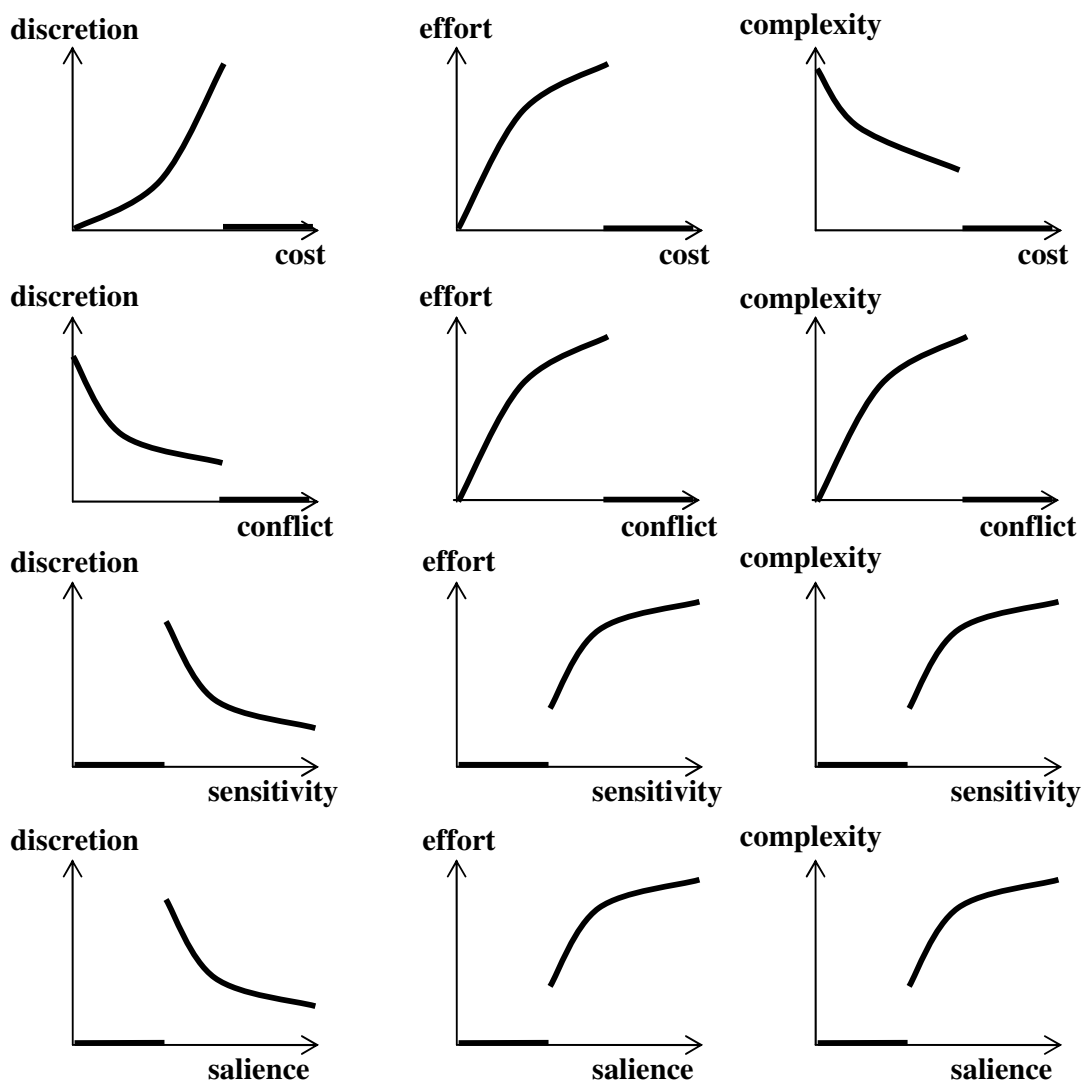


Figure 3: *Comparative Statics: Residual Discretion, Effort, and Complexity.* These curves show, for optimal doctrinal choice, the relationship between the parameters of the model (cost, conflict, sensitivity, and salience) and three quantities of interest—the amount of residual discretion lower courts will possess, the amount of effort the higher court will invest in crafting the doctrine, and the complexity of the doctrine announced (reduction of the troublesome region under a standard, zero for the bright-line rule).

Reviewers' Appendix

Quadratic Cost Term

Using a quadratic cost term instead of the quartic cost term requires an additional assumption to constrain the parameters to the interior solution in v ($c < psa^2$). The interior solutions for either cost function are more reasonable approximations to a general cost functions with appropriate assumptions about derivatives. In this section, I show that results are qualitatively similar using a quadratic cost term (the propositions are in parallel, Proposition 3' being the equivalent of Proposition 3 with the alternate cost function).

Proof of Proposition 3'. The balancing test payoff is $-\frac{1}{2}ps(a(2v - v^2))^2 - c(1 - v)^2$, so that $\frac{\partial u}{\partial v} = 2(1 - v)(c + psa^2(v - 2)v)$, which is zero at $v^* = 1 - \sqrt{1 - \frac{c}{psa^2}}$ (unique maximum for interior v^* , which requires $c < psa^2$). The partial derivatives with respect to a , p , and s are negative and with respect to c positive. \square

Proof of Proposition 4'. The bright-line rule payoff $\left(\frac{-sa^2}{16}\right)$ is greater than that of the optimal standard $\left(-c\left(1 - \frac{c}{2spa^2}\right)\right)$ when $16cpsa^2 - 8c^2 - ps^2a^4 > 0$, and so when $p > \frac{8c^2}{sa^2(16c - a^2)} \equiv Z$. Note for $p < \frac{1}{8}$, the rule cannot be chosen for any value of c (given conditions on c for choosing the rule). $\frac{\partial Z}{\partial s} = (sa^2 - 16c)^2(sa^2 - 8c)$ which is positive for $s > \frac{8c}{a^2}$. At the low boundary on s for interior v^* , $s = \frac{c}{pa^2}$ and then $p > Z$ so that the rule is chosen. For $\frac{c}{pa^2} < s \leq \frac{8c}{a^2}$, $\frac{\partial Z}{\partial s} < 0$, so that the boundary on p for choosing the rule weakly decreases, so that the rule is still chosen. Finally, for $s > \frac{8c}{a^2}$, $\frac{\partial Z}{\partial s} > 0$ so that there exists \bar{s} such that the standard is chosen if and only iff $s > \bar{s}$. The proofs for a and c are parallel to this one for s . The standard is chosen for sufficiently high s or a and for sufficiently low c or p , all else equal. \square

Other Hierarchical Configurations

The main text assumes that the higher court faces an allied or hostile lower court, in which the entire troublesome regions is in contention between the two levels: $b_L \leq b - av$ and $a_L \geq \frac{a - (b - b_L)}{1 - v}$ (note that the latter holds automatically if the two courts have the same sensitivity but different slopes,

with $a_L = a$). Obviously other configurations are possible, and I consider one here. Suppose that the higher court faces either an allied or *partially* hostile lower court, for which only part of the troublesome region is in contention between the two levels. Further assume the two courts (higher and hostile lower) share the same sensitivity $a_L = a$, but the intercept b_L is so high as to yield overlap with the troublesome region: $b - av < b_L < b$. Let $\Delta_b = b - b_L$, the difference in intercepts. Now the area in contention is $\Delta_b - \frac{(\Delta_b)^2}{2a}$. The payoff from using the balancing test $\left(-sp \left(\Delta_b - \frac{(\Delta_b)^2}{2a}\right)^2\right)$. The bright-line rule is then preferred if $p > \frac{a^4}{4(\Delta_b)^2(2a - \Delta_b)^2}$. Compared to the situation where the lower court is potentially hostile instead of potentially partially hostile, the likelihood of conflict still pushes towards the bright-line rule, but now the exogenous degree of transparency does not matter on the margins (it is already sufficiently large that it has no marginal impact). On the other hand, sensitivity would then matter. To put this another way, if exogenous transparency is sufficiently high, then sensitivity can matter (though in the case solved here it is the sensitivity for both higher and lower courts).

Other Configurations of the Transparency Region

Suppose that (exogenous) v is not constant across the level of the observable fact. I consider two possibilities in this vein here. Both yield the same comparative statics (in v and p) for the exogenous-transparency rule choice. First, let transparency be perfect at the highest level of the observable fact (so that the width of the troublesome region at that point is zero, and let the width of the troublesome region only reach full v at the lowest value of the observable fact. Equivalently, the width is zero for the lowest value of the subjective dimension ($x = 0$) and equals v at the highest value ($x = 1$)). The boundary line is $b - \frac{ax}{1-v}$. All other configuration details are as in the exogenous transparency case. Then, the area of the troublesome region is $\frac{av}{2}$ and the balancing test payoff is $-ps \left(\frac{av}{2}\right)^2$. The bright-line rule is preferred when $p > \frac{1}{4v^2}$.

Second, flip the region so that the width of the region is v at $x = 0$ and zero at $x = 1$, bounded by the the line $\left(b - \frac{av}{1+v}\right) - \frac{ax}{1+v}$. The troublesome region then has area $\frac{av}{2v+2}$ and the bright-line rule is preferred when $p > \left(\frac{1+v}{2v}\right)^2$. The right-hand side is decreasing in v .