JOURNAL OF THEORETICAL POLITICS

Article

Scope and precedent: judicial rule-making under uncertainty

Journal of Theoretical Politics
1–32

DOI: 10.1177/0951629814568397
jtp.sagepub.com
SAGE

Tom S Clark
Department of Political Science, Emory University, Atlanta, GA, USA

Abstract

I develop a formal model of Supreme Court opinion-writing in an environment of uncertainty. In particular, the model captures how the Supreme Court will optimally design the specificity of its legal rules. The model focuses on the tradeoff between more precise rules which are controlling in a smaller subset of cases against less precise rules, which have wider applicability but yield less certain outcomes. When the basic model is considered in a dynamic world in which the Court is able to hear multiple cases, it yields insights about how the factual representativeness of a case and the clarity of existing precedent jointly affect its optimal opinion-writing and willingness to hear new cases. These last implications provide theoretical foundations for theoretical and empirical questions about rule-making, case selection, and the construction of doctrine.

Keywords

Judicial rule making; judicial hierarchy; opinion writing

1. Introduction

The study of judicial decision-making has been undergoing a sea change. In moving away from simple "attitudinal" approaches, scholars have begun to take seriously the role of legal rules and doctrine (see Lax, 2011, for a review). The key component that distinguishes this contemporary development from debates in recent decades is the shift from perceiving law as a constraint on judges' pursuit of their policy goals (e.g. Gillman and Clayton, 1999; Bailey and Maltzman, 2011; Segal

Corresponding author:
Tom S Clark, Department of Political Science, Emory University, 327 Tarbutton Hall, 1555 Dickey Drive, Atlanta, GA 30322, USA.
Email: tom.clark@emory.edu

and Spaeth, 2002) to law as an instrument that judges use to achieve their policy goals (e.g. Friedman, 2006; Sunstein, 1999). Central to this development has been the formal modeling of legal rules, which serve to partition cases, according to the facts of the case, into discrete outcomes. Intended to capture the nature of adjudication in most common law jurisdictions, this theoretical apparatus, known as the *case space model*, considers the two functions courts serve: to apply (or create) legal rules and use them to dispose cases into discrete dispositions (Kornhauser, 1992b). A recent trend in this literature has been to consider how high courts communicate legal rules to lower courts, charged with applying those rules in future cases (e.g. Fox and Vanberg, 2011; Jacobi and Tiller, 2007; Lax, 2012; Staton and Vanberg, 2008; Sunstein, 1999).

An often overlooked, but nevertheless critical, aspect of the task facing high courts is the difficulty of clearly and completely communicating with lower courts (see, for example, Jordan, 2011). Scholars have grown increasingly interested in the Court's decision to craft a more or less clear rule (e.g. Owens and Wedeking, 2011; Staton and Vanberg, 2008), and this paper pushes that line of inquiry in a new direction. When judges write opinions that contain instructions for how lower courts must decide cases, they face an inherent, unavoidable tension. Opinions that are closely and thoroughly embedded in the factual specifics of a case are relatively more able to articulate a clear, unambiguous statement of precisely the way to resolve that case and similar future cases than are opinions that seek to announce principles that are broader and more far-reaching than the specific factual circumstances of the instant case. However, precisely because a less ambiguous rule requires greater embedding in factual specifics, such precedents are more easily distinguished from future cases (future layers and courts can more effectively argue the precedent is not dispositive and "controlling") than are precedents that are perhaps more ambiguous in their meaning but are more broadly applicable because they are not tightly connected to particular factual scenarios. In other words, there exists a continuum where opinions at one end are more *narrowly tailored*, which clearly and precisely communicate their authoritative interpretation of the law. At the other end are more *principle-based* opinions, which are less precise. Any opinion a court might write entails a particular balancing of the benefits of narrowly tailored and principle-based opinions. Simply put, the tradeoff is one between precise instructions that have limited future applicability and less precise instructions with wider future applicability.

The decision about how best to communicate a legal rule through a series of related opinions, doctrine, goes to the heart of the function high courts serve. While there is variation in the types of jobs supervisory courts perform, high courts in many jurisdictions are primarily concerned with overseeing the law applied by lower courts, with a particular emphasis on ensuring the law is consistently applied by lower courts. For example, this is the stated main function served by the US Supreme Court. In essence, then, high courts serve as managers of massive bureaucracies, which interpret and implement the high court's instructions, as communicated through their opinions. This paper studies a challenge that is especially important in the judiciary but is also a more general issue in rule- and opinion-

writing. In particular, once one writes down a rule, there is inherent uncertainty in how it will be applied in the future. Importantly, that uncertainty is itself a function of how precisely written a rule is and how closely tied it is to the specific factual context in which it was originally developed. That tension is, more generally, a fundamental feature of policy-making in the form of case-by-case individual dispute resolution.

I develop a formal model of Supreme Court rule-making that captures the tension created by this relationship. I model a court's decision about how best to write an opinion, knowing that future applications of the opinion will be variable. Future courts and actors subject to the court's rule must interpret how to apply the rule in future cases. The analysis proceeds in three stages. First, I analyze the optimal opinion to write in a single-shot decision. Second, I consider how the existence of past precedent and optimal opinion-writing drive the court's decision about which cases are best to hear in order to best communicate doctrine to lower courts and external actors. Finally, I consider how the prospect of future cases affects opinion-writing in the first instance. Together, these analyses provide novel insights into the nature of doctrine creation and management by a supervisory court and the optimal strategy to pursue in order to effectively manage those who interpret and apply judicial doctrine. As I describe below, those insights have implications for a number of theoretical and empirical questions in the literature, including how non-ideological features of law affect the incentives judges face, how non-ideological features of law influence case selection, and how judges can use a sequence of cases to construct a line of doctrine. Because such questions are central to what Lax (2011) calls "the new judicial politics of legal doctrine," their implications go to the core of a growing line of inquiry in the study of law and courts.

2. Judicial rules and policy communication

One of the most notable features of judicial decision-making in the American system is the extensive written records that are produced by courts. The Supreme Court, which sits at the apex of a large, complex bureaucracy of courts is charged with maintaining legal clarity and consistency throughout the country and relies heavily on that body of written records. Independent judicial systems in each state work along with the federal judiciary to apply law in the hundreds of thousands of cases (those that do not settle to begin with) each year. In doing so, courts look to the authoritative interpretations of the law that are contained in the written opinions handed down by superior courts. As a consequence, perhaps the single most difficult problem facing higher courts, especially the Supreme Court, is communicating its authoritative interpretation of the law for applications in a wide variety of cases, within the confines of a single legal dispute.

Scholars have long been interested in the process by which a superior court, such as the US Supreme Court, can communicate its interpretation of the law through a series of opinions written in individual cases, and recent research has sought to formalize the tradeoffs and mechanisms at play in that process. For example, Bueno de Mesquita and Stephenson (2002) develop a formal model in which lower court

applications of the law are represented by random draws from an accumulated body of law, which is represented itself by a distribution; the Supreme Court's opinions determine the body of law from which lower courts draw, i.e. the shape of that distribution. Bueno de Mesquita and Stephenson's model, like the one developed below, is decision-theoretic; it focuses on the Supreme Court's decision about what kind of body of precedent to establish. In that model, the Supreme Court must decide whether it is worth continuing to modify doctrine and potentially "muddying the waters" for lower courts or instead to break from precedent and create a new distribution of doctrine for the lower courts.

A series of related papers develop formal models of the clarity and specificity of the doctrine a superior court will create. Central to those models is the Supreme Court's effort to have future outcomes map to its preferred outcome. In the model of Staton and Vanberg (2008), the high court wants to avoid publicly observable non-compliance by lower courts or external enforcers and so may want to employ vague holdings in order to help conceal defections by opposed enforcers. Thus, the "value of vagueness" for the courts turns out to be that by intentionally remaining vague in their holdings, they can avoid the negative repercussions that come with explicit political defiance of judicial will.

Closer to the approach taken here is the consideration, primarily in the legal literature, of the distinction between "rules" and "standards" (e.g. Sullivan, 1992). Lax (2012) and Jacobi and Tiller (2007) are recent formalizations of the central tension in this debate. In Lax's model, the Supreme Court must decide whether to write (i) a "bright line" rule which cannot take into account subjective dimensions of a case's factual circumstances or (ii) a standard which does take into account the subjective dimension. The tension for the Supreme Court is that while it wants the outcomes to reflect its own preferences as closely as possible (meaning it wants cases decided as it would decide the cases itself), once the subjective dimensions of the case are taken into account (i.e. a standard is adopted rather than a bright line rule), even faithful lower courts may be imperfect enforcers of a balancing test that employs a subjective element. The problem may be worsened by the presence of non-faithful lower courts who will take advantage of the subjective dimension to intentionally evade the Supreme Court.

Jacobi and Tiller (2007) explicitly consider this latter case, developing a decision-theoretic model in which the Supreme Court must choose between a determinate and indeterminate doctrine. The indeterminate doctrine allows a potentially ideologically divergent lower court to adopt some policy that diverges from the Supreme Court's most preferred policy, albeit within the confines determined by the Supreme Court's indeterminate doctrine. A determinate policy, by contrast, compels all lower courts to use the Supreme Court's preferred rule, even though that rule may sometimes result in outcomes with which the Supreme Court disagrees.

One common feature of the literature on legal rules is that it is often assumed that the Supreme Court has the *ability* to clearly and effectively communicate its preferences to the lower courts: the question is only how to deal with potentially conflicted enforcers of the Court's will. However, there is reason to believe that assumption may mask important dynamics in the rule-making process. First,

human capacity to communicate policy is inherently limited, which may pose a strategic dilemma to a principal looking to communicate a complete policy mapping to a subordinate (Jordan, 2011). Second, and perhaps more important, a court's ability to authoritatively control the law applied to future cases is correlated with those future cases' factual similarity to the precedent case. As future cases become increasingly dissimilar, lower courts/enforcers have a more legitimate claim that the precedent in question is not binding on their decision-making. As Maltz (2000, 1402) notes,

> "[B]ecause of the nature of the judicial process, the Court does not formally adjudicate more than one case at a time. It is in the course of such adjudication that the Court creates a legal rule that is tailored to the specific fact situation before it. This legal rule in turn binds lower courts faced with analogous fact situations in the future."

Thus, it is unclear how future outcomes will relate to the rule announced in a precedent, especially as those future cases diverge in their similarity.[1]

Indeed, the primary task for lawyers and lower court judges in arguing or deciding cases is to situate the new case relative to those that have come before. If it can be reasonably argued that a new case is different in a material way from a past precedent, then that past precedent will be less "controlling" or dispositive of the new case than a precedent that is less materially different. Professional advocates, i.e. lawyers, have as their charge the task of distinguishing precedents whose legal rules will lead to adverse outcomes for their clients. Outcome-oriented judges seeking to reach their preferred outcomes similarly must engage the body of precedent in order to evade past rulings whose implications for the current case they oppose. Consequently, a central problem for a high court seeking to establish an authoritative statement of the law is that in future cases, lawyers and judges below will attempt to evade the high court's past rulings by claiming the new case is different from the past one and that the past precedent can be legitimately evaded.

It is important to note here that the model sets aside a critical component of the process of doctrinal creation by the courts: judicial learning. The process by which litigants bring cases to the courts (e.g. Baird, 2007), allow issues to percolate (e.g. Clark and Kastellec, 2013), and signal to the courts about the legal significance of a case (e.g. Caldeira and Wright, 1988) is an important one for how the Court identifies which cases are the best vehicles for them to etch their views into the law. Here, I focus on how the Supreme Court uses one particular criterion, the factual relationship of a case to others, to strategically select among cases and set aside other features of that choice, including the political salience and legal ripeness of a given case.

The model developed below addresses the process of doctrinal creation in the Supreme Court by modeling a previously unexamined feature of doctrine creation. One can think of doctrine as a mapping from factual situations to outcomes, as depicted in the left panel of Figure 1. When deciding a case, the Court announces a doctrine that maps that specific case (given by a point on the x-axis) to a particular outcome (given by a point on the y-axis). For example, one may think of the y-axis as representing the level of deference to be paid to an administrative agency. In a

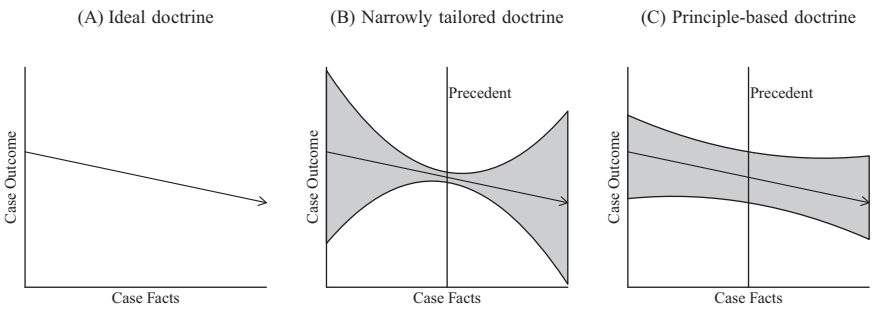

Figure 1. Visual representation of doctrine as mapping from fact space to outcomes. The left panel shows ideal doctrine, which clearly and authoritatively maps factual situations to outcomes. The middle figure shows a specific doctrine which clearly and authoritatively maps very similar cases to outcomes but whose binding authority decreases quickly in the dimension of factual dissimilarity. The right panel shows a broad doctrine which is less predictive of even factually very similar cases but whose predictiveness is less sensitive to factual dissimilarity.

given case, the Court ideally would announce a doctrine that specifies an authoritative, exact mapping from all factual situations to the Court's ideal outcome. In other words, for every possible factual scenario, the Court's doctrine would specify the exact level of deference to be paid to the agency.

A central question in the study of legal rules, both in the political science literature and in the legal academy, concerns the choice by a judge to avoid reaching a broad decision and to instead limit a decision (e.g. Epstein et al., 1996; McGuire and Palmer, 1995; Sunstein, 1999). In terms of Figure 1, this literature examines why a court may prefer to avoid specifying the full range of the line that represents doctrine. The model developed here cabins off those concerns to focus instead on a related, but previously unexamined, feature of rule choice by a court.

Most formal models of rule-making by courts assume that the Court is able, in a single case, to specify a complete mapping from all factual scenarios to a particular outcome, even if the Court may choose to make the mapping intentionally vague or future courts may modify or distinguish the rule (e.g. Gennaioli and Shleifer, 2007; Lax, 2007, 2012; Staton and Vanberg, 2008). As noted above, the Court's announced legal rule is most binding in the most factually similar future cases, and less so in less similar cases. Moreover, the extent to which the Court's holding in any given case extends to other cases is a function of the opinion the Court writes. Research on judicial opinion-writing has explicitly noted this feature of doctrine, distinguishing broad or wide opinions from narrow ones (e.g. Fox and Vanberg, 2011; Sunstein, 1999). The extent to which a given precedent is predictive of future applications is a function of the specificity of the rule: more narrowly tailored and specific rules will be very predictive of very similar cases and not very predictive of cases that are even slightly factually dissimilar. Broader, more vague rules, by contrast, will be less predictive of very similar cases (than will be a narrowly tailored, specific rule), but their predictive power will not be as sensitive to factual discrepancies. Thus, there exists a tradeoff between writing a rule that is precise and not

subject to interpretation but is, by necessity, constrained in its applicability to future cases, and writing a rule that is oriented around a broad principle and more subject to interpretation but is therefore more broadly applicable to future cases. It is in this feature of the model developed here that we find the particular contrast between this model and the "rules v. standards" debate, whereas that literature focuses on the relationship between subjective and objective elements of a legal rule, this model focuses on the limitations of establishing and communicating even a perfectly objective legal rule.

To develop an intuition for this tradeoff, consider two examples from the doctrine relating to class action lawsuits, both decided in 2011. In *Erica P. John Fund v. Halliburton* (2011), the Supreme Court reviewed a securities fraud case in which a group of investors was suing Halliburton for artificially inflating its stock price, leading the investors to lose money. The Court was asked to decide whether the group had to prove that Halliburton's actions caused their losses in order to be certified as a class for purposes of a class action lawsuit. In doing so, the Court narrowly tailored its decision to one factual scenario concerning class action certification, whether loss causation must be demonstrated, a group of investors who have not demonstrated loss causation but wish to be certified as a class. The rule provides clear and direct instructions in the specific case of a group of investors seeking to sue in a class action lawsuit without first demonstrating loss causation. However, what the case means for even marginally different cases, such as a case where investors want to sue without first proving other elements of the suit, remains very unclear. The key point is that the rule is one whose precise application is closely connected to a narrowly specified set of circumstances.

By contrast, consider *Wal-Mart Stores v. Dukes* (2011), in which the Supreme Court held that all female employees of Wal-Mart could not be certified as a class for purposes of suing Wal-Mart for employment discrimination. The crux of the Court's holding was that the women did not have enough "common interests" to sue as a single group in a class-action lawsuit. This finding, by way of contrast to the *Erica* case, has broad implications, reaching potentially any group of litigants seeking to sue as a class for any purpose. However, what "common interests" means is something that remains vague and allows considerable leeway for interpretation, even in nearly identical cases in the future. This is especially true in relation to the "loss causation" question in *Erica*, which has a specific, clear meaning in the law. A more narrowly tailored rule in *Wal-Mart* might have specified the nature of the class members' relationship (employees at specific levels of employment in a retail organization) and the nature of their interests that are asserted to be held in common (issues related to maternity leave, pay scales, and other benefits). Such a rule would have made clear precisely how to handle similar claims in the future, by defining what makes the claimant's interests common. Crafting the more narrowly-tailored rule, though, necessitates limiting the applicability of the legal principle by connecting it to the particular facts of the *Wal-Mart* case. Thus, we can see in this example the central tradeoff I examine. The clarity and specificity of a legal rule is connected to the extent to which is tailored to the factual scenario in the case being decided.

Of course, the differences between narrowly tailored and principle-based rules is one of degree. However, these two examples illuminate the crux of the distinction I draw. Returning to our graphical representation of an opinion's mapping from facts to outcomes, Figure 1(B) shows an example of how binding (i.e. predictive) a very narrowly tailored rule is across a range of potential future cases. Figure 1(C) shows how binding (i.e. predictive) a more principle-based rule is across a range of potential future cases. Taken in the context of Lax (2012), we can think of the breadth of a rule as somewhat analogous to "standard-ness." A rule that is very broad leads to variable outcomes for all cases, but that variability is not increasing very much in the dimension of factual dissimilarity. A rule that is less broad is more like the "bright line" rule of Lax, but its applicability, and therefore its authoritativeness and predictive utility, decreases very quickly as cases become increasingly factually dissimilar.

In the following section, I formalize these feature of doctrine in a model of Supreme Court policy-making. The model yields insights about the specificity of the rule the Court will create in the context of any given case. Central to the Court's decision calculus is the case's representativeness of the larger class of cases to which it anticipates the rule being applied in the future. Building from that insight, I extend the model to identify how existing precedent will affect the cases the Court will select for discretionary review. It is well established that the Court strategically selects cases in order to most effectively communicate its authoritative interpretation of the law and minimize variation in how the law is applied across the various lower courts' jurisdictions (e.g. Perry, 1991). Moreover, given that the Court is not randomly selecting cases for review but is instead selecting cases with particular factual contexts, it is important to have a theoretical understanding of how that decision is made in order to interpret the relationships between the facts of a case and the Court's underlying policy motivations (e.g. Carrubba and Clark, 2012; Kastellec and Lax, 2008).

3. The model

To isolate and analyze the tension described above, I develop a formal model of judicial opinion writing in which future interpretations of the Court's opinion are a function of the specificity with which the decision is communicated. In particular, I assume that while the Court can announce a rule in a given case, even one that has general applicability, the application of that rule becomes more variable as the facts of the cases in which the rule is applied become increasingly dissimilar. That is, because differences in the legally relevant facts (i.e. the facts that are mapped into case outcomes) allow future courts, litigants, or advocates to more effectively distinguish future cases, legal rules from factually dissimilar cases will have less compelling force as binding authority.

The critical feature of this model is therefore that the Court is uncertain about how its opinion will be interpreted or applied in future cases, and the extent of the uncertainty is itself a function of how factually similar future cases are to the case in which the rule was announced. Note that I am explicitly ruling out uncertainty

on the part of the Court about its preferences over doctrine. I assume throughout that the Court knows what doctrine is wants but that its uncertainty is in how its instructions will be implemented. To formally capture this tension, I import the model of variance-in-variance captured by the Brownian motion technology Callander (2011) uses to model policy-making. In particular, I model the consequences of applications of Supreme Court rules as a variable outcome: one in which there is variance in the variability of the outcomes. The model captures the tension described above which commonly faces judges. An opinion may be more narrowly tailored and therefore result in highly predictable outcomes in very similar future cases but highly *un*predictable outcomes in dissimilar future cases. Alternatively, a rule may be more principle-based and therefore result in relatively more variable outcomes in similar future applications but be less sensitive to factual dissimilarity in its predictability.

3.1. Elements of the model

The Court's informational environment is as follows. Any *case* is a set of facts that are represented by a point in $\mathbb{R}$. A generic (and as yet unobserved) future case is denoted by p and the case at hand (i.e. currently observed by and before the Court) is denoted by q. Given that I consider the Court's incentives regarding how to dispose of this case in terms of how the disposition will affect future cases, the relevant consideration is how the Court believes the facts of *future* cases will be realized. I assume that the Court correctly believes that the facts of any future case, denoted by $p \in [a, b] = X \subset \mathbb{R}$, will be drawn from a continuous, twice differentiable density function, $f(\cdot)$, with support on the closed interval X. I further assume f is concave.

Upon observing q, the Court decides the case. Court decisions include both rules for deciding the disposition of future cases and precision in the rule articulated. I leave the rule component of the Court's decision unmodeled.[2] Thus, in deciding the case, the Court chooses a single parameter, ν, representing the precision of the Court's opinion. The precision is a measure of how clearly the Court's decision is communicated and, as a consequence, how much variability there will be in future interpretations by lower courts, bureaucrats, etc., and by implication how closely connected the disposition is to the particular facts of the case. Very precise decisions (low ν) offer little question about the particular disposition in case q, but they are inherently bound up in the facts of the case, i.e. the specific value of q. Less precise decisions (high ν) provide less clear specific guidance about what dispositions in case q are consistent with the decision but are also broader in nature and therefore more easily applied to other, future cases.

Any future case, p, is mapped into an interpretation, $z \in \mathbb{R}$, and distributed according to a cumulative distribution function $D(q, \nu)$ with variance given by

$$\mathrm{Var}_{D(q,\nu)}[z|p] = \nu + \frac{(p-q)^2}{\nu} \tag{1}$$

Recall Figure 1; here the black sloping lines represent the ideal rule, and the gray areas show the variability in outcomes associated with the precedent. In the middle panel we see an example of a high-precision, i.e. low ν, opinion. Cases very similar to the precedent establishing the rule are subject to little variability, whereas dissimilar cases are easily distinguished from the precedent and therefore subject to greater variability. By contract, in the right-hand panel we see an example of a lower-precision, i.e. high ν, opinion. In this case, future cases very similar to the precedent establishing the rule are subject to relatively higher variance, but the magnitude of variance is relatively constant across all future cases.

I assume the Court receives (dis)utility equal to the variability of interpretations when its decision is applied in future cases. Interpretations of the Court's decision, either by lower courts or by outside actors, are draws from the distribution with variance given by Equation (1). Consider a new case p being heard by a future lower court. The greater the variance at case p, while the outcomes may on average be consistent with the rule picked, the variability of the interpretations will be greater. More variance allows, in other words, more interpretations of the rule, more error or confusion in application, or more successful efforts by lawyers to invoke a different standard or rule. Specifically, I assume the Court's utility function is given by

$$U(\nu; q) = -\int_X \left(\nu + \frac{(p-q)^2}{\nu}\right) f(p)\, dp \tag{2}$$

It bears describing in greater detail three features of the underlying theory of rule application by lower courts and external actors. The first concerns the nature of the outcome space. In general, legal rules are thought to be mappings of facts into dichotomous outcomes. For example, a legal rule in tort law may specify a range of actions that are mapped into the "liable" outcome and a range of facts that are mapped into the "not liable" outcome (Kornhauser, 1992b). Here, though, I have modeled an opinion as a mapping into a continuous "interpretation space." There are two natural ways to think about this representation of legal opinions. First, one might think that the continuous interpretations, z, are instructions about how much weight to give a particular consideration when seeking to reach a dichotomous outcome. For example, when an alleged tort is committed by a corporation, a doctrine may instruct lower courts to put more weight on the duty of care than in cases where the alleged tort was committed by an individual. The doctrine, then, is a mapping from factual scenarios to instructions about how to evaluate a claim. Second, though related, because I have assumed the fact dimension here is unidimensional, whereas cases are surely highly dimensional phenomena, one might interpret the doctrine as a mapping from one dimension to the set of hyperplanes that divide the additional dimensions. For example, in any given case, there may be one factual issue at hand, and the doctrine articulated in that case may specify how that factual dimension relates to other factual dimensions (for a related, but distinct, conceptualization, see Lax, 2012). In either interpretation, the doctrine consists of an ideal statement (the linear function), complicated by the precision with

which that statement is communicated (the variance in the random component of the doctrine).

The second feature of lower court interpretations that bears being made explicit is that the lower courts are assumed to be non-strategic, passive appliers of the legal doctrine, who do not learn about doctrine. The courts are generally characterized by a division of labor, whereby lower courts specialize in fact-finding and rule application, while higher courts specialize in rule articulation (e.g. Kornhauser, 1995). In this model, lower courts take their instruction from higher courts and can only learn when the higher court intervenes and decides a new case. In this sense, the lower courts are bound by the principles of vertical precedent: that a set of instructions from a higher court is unclear (i.e. variable) can only be corrected by further instruction from the higher court.

A third important feature of this model is that it sets aside the litigants as actors in this process. As noted above, litigants are important because they decide which cases to bring, provide the means by which courts can experiment with and learn about alternative legal rules, and reveal to the courts information about which types of cases are resulting in controversial application of the law. Related, whereas scholars have debated whether the courts simply adopt rules suggested by litigants (e.g. McGuire and Palmer, 1995) or instead restrain themselves to the narrow questions presented by the litigants (e.g. Epstein et al., 1996), here I do not directly model the choice about whether to expand a case beyond the questions presented and instead focus on the trade-off a Court makes when crafting broad rules that cover expansive types of cases as opposed to narrow rules that are more factually bound. One implication of these latter two assumptions is that they allow me to set aside considerations of which mean to pick for the function (i.e. the target rule) and focus analytic attention on the variance component (i.e. the extent to which the rule is narrowly tailored).

3.2. Analysis

I analyze the model in three steps. First, I consider the model as a single-shot decision. The Court must decide what opinion to write in a case before it. Second, I consider the world where there exists some previously set precedent, and the Court must decide whether to hear a new case it is requested to hear. This extension allows us to examine the effect of lower court opinion application on the Court's decision about which cases to hear. Third, I consider the world where the Court, in deciding the first case, anticipates what kinds of cases it is likely to encounter in the future. This third variant of the model allows us to examine how the prospect of being able to revisit a line of doctrine affects the Court's opinion-writing in the first instance. Taken together, these three analyses provide insight into the incentives created by hierarchical application of judicial rules. All proofs of formal results are gathered in the appendix.

3.2.1 Baseline model. The primary tension in the model described above concerns the Court's decision about how principle-based versus how narrowly tailored to make a particular decision, given its expectations about the set of future cases (the

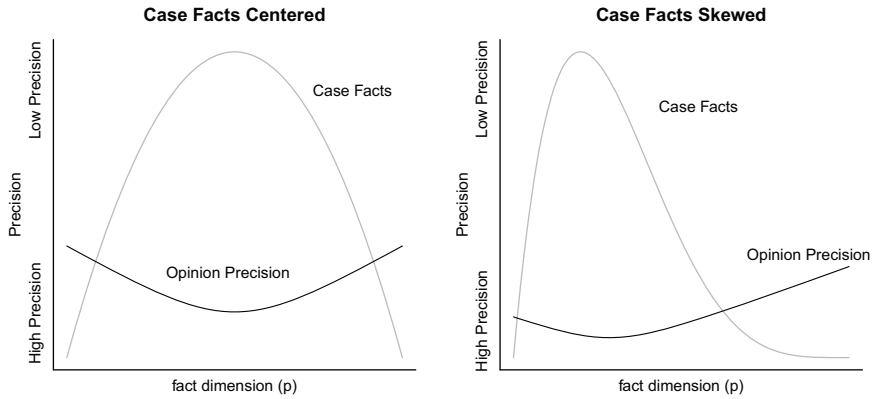

Figure 2. Optimal precision in a legal rule as a function of case facts in am instant case. The gray line shows the distribution of case facts; the black line shows the optimal level of variance as a function of which case facts are drawn for the Court.

distribution of p) and how the decision will be interpreted in those cases. The first result, then, concerns the optimal precision to include in any decision. The optimal level of precision that maximizes Equation (2), as one would expect, is a function of the facts of the particular case.

Proposition 1. *The optimal precision the Court should choose is*

$$\nu^* = \sqrt{\frac{\int_X (p-q)^2 f(p)\,dp}{\int_X f(p)\,dp}}$$

Proposition 1 shows that when the case the Court is deciding is representative of the overall distribution of case facts, then the optimal opinion has relatively high precision (i.e. low ν).[3] However, as a case becomes increasingly less representative (i.e. further from the mean of the distribution of case facts), the optimal opinion is less precise. Figure 2 depicts the optimal level of precision as a function of the facts the case of the case is deciding (q) under two different distributions of case facts.

It is also helpful to note that ν^* is simply the root mean squared error of the opinion, when applied to future cases and is the standard deviation of the distribution of case facts, $f(p)$. This value is smallest when the instant case, q, is at the mean of the distribution of case facts, $\mathbb{E}[p]$. In this sense, the results here parallel standard results in statistical decision theory, the implication of which is that precision should greatest when the instant case is the most representative of the underlying distribution of case facts.

Result 1. *The optimal opinion for case q becomes increasingly precise (i.e. lower ν^*) as the facts of the case q become more representative of the underlying distribution of cases, $f(p)$.*

The intuition behind this result is straight-forward. When deciding a case, if the case is representative of many other cases, a very precise opinion means that similar cases will be decided very much in line with the Court's announced rule, while less-similar cases will be subject to variability: variability increasing in their dissimilarity. However, because the instant case is representative, the tradeoff of having the relatively few dissimilar cases subject to greater variability is offset by the benefit of having many similar cases decided in line with the Court's opinion. By contrast, if the instant case is not representative of the larger distribution of cases, then writing an opinion with less precision but variability that is less sensitive to case dissimilarity is desirable. The reason is that it is preferable to the Court that all cases be decided with a comparable degree of variability rather than allow variability to increase as the cases become increasingly dissimilar. Because the instant case is unrepresentative, as case dissimilarity increases one approaches more representative cases. This is the range of case facts where the Court *does not* want increasing variability. Thus, the Court prefers to set a level of variability in its initial opinion that does not exacerbate the problem in the bulk of cases.

3.2.2 The decision to hear a case. We now move from an analysis of the Court's decision in a single-shot opinion-writing model to its decision to hear a case. As noted above, the Court's discretion over its docket provides it with great flexibility in terms of choosing which cases to hear and how to focus its energies; it also provides the Court with a strategic dilemma about how what kind of opinion to write in those cases it has agreed to decide. The model provides some traction on the strategic dilemma concerning which cases to hear.

I invoke an additional assumption: I assume that when lower courts apply a line of doctrine, they rely on the precedent from the most factually similar case. This assumption rests on the method of legal reasoning, which holds that the most controlling precedent is the one from the most similar past case (e.g. Levi, 1949). Of course, the assumption is particularly stark and does not capture the richness of legal reasoning by analogy. It is important to acknowledge that lower courts often rely on multiple precedents and balance different rules in an effort to decide each case. However, many, if not all, of the results below could be generalized to a situation in which courts reason by taking a weighted average of all possible relevant precedents. I impose the assumption I adopt for maximum tractability. In particular, given precedents from two cases, q and q', when deciding a new case p, the lower courts will apply whichever precedent, q or q' is closer to p. In addition, I also assume there is some positive cost, $k > 0$, to hearing a new case; this cost represents an opportunity or resource cost associated with deciding a new case. Models of judicial hierarchy and oversight often invoke such an assumption to represent the idea that it is not effortless to decide a new case (e.g. Cameron et al., 2000; Clark and Carrubba, 2012). Without any such cost, the Supreme Court's decision about hearing cases would be trivial: it would simply decide to hear all cases brought to it, which is clearly not an empirically supported proposition.

Thus, assuming $q' > q$, we can represent the Court's utility from deciding a new case q' as

$$-\int_{a}^{\frac{q+q'}{2}} \nu + \frac{(p-q)^2}{\nu} f(p)\,dp - \int_{\frac{q+q'}{2}}^{b} \nu'^{*} + \frac{(p-q')^2}{\nu'^{*}} f(p)\,dp - k \tag{3}$$

where ν is the precision of the previous opinion from case q, and ν'^{*} is the optimal level of precision the Court will set in case q'. The objective function when $q' < q$ is symmetric.[4] Note that this condition does not require, and I do not assume, optimal precision was set in the precedent.

The key tension set up here is whether the Court can gain enough from reducing variation in interpretations of its doctrine in cases dissimilar to the precedent case, q, to justify its allocation of resources to the new case, q'. In other words, if Equation (2) evaluated at the precedent precision is less than Equation (3) evaluated at the precedent precision and the new case optimal precision, then the Court will want to take the new case, q'. This decision will depend upon the precision of the past precedent, the dissimilarity of the new case, and how much more (or less) representative the new case is than the precedent case. Algebraically, the Court will prefer to take the new case if

$$\int_{\frac{q+q'}{2}}^{b} \left(\nu - \nu'^{*} + \frac{(p-q)^2}{\nu} - \frac{(p-q')^2}{\nu'^{*}}\right) f(p)\,dp > k \tag{4}$$

Substantively, this condition demonstrates that the Court's improvement in utility, which is defined as the difference in outcomes in the range where the new precedent will apply, must be large enough relative to the cost of reviewing the new case.

As noted, both the precedent's precision and the new case's similarity to the precedent will affect this calculation. Note that changing the precision of the precedent presents competing incentives for the Court. On the one hand, decreasing the precision of the precedent (increasing ν) means that cases similar to the precedent case will be subject to *greater* variability. At the same time, cases that are dissimilar from the precedent will be subject to *less* variability. These incentives, though, are potentially further complicated by an additional feature of the Court's problem: simply selecting cases that are further away from the precedent case does not necessarily resolve the tension, because the further away the new case is from the precedent, the *smaller the range of future cases* to which the new case will apply, thereby potentially undermining the utility of taking the new case to minimize variability in outcomes. To see how these two incentives interact with each other, I consider first the effect of changing the precision on the incentive to hear a new case and then the effect of changing the similarity of the new case to the precedent.

The effect of precedent precision on the decision to hear the new case. The effect of changing the precision of the precedent, as the preceding discussion suggests, depends on both how similar the two cases are and the level of precision itself. Specifically, as precision in the precedent decreases (as ν increases), so too does the incentive to hear a new case. However, there reaches a point where further decreases in precision decrease the incentive to hear the new case. To understand why this result holds, note that decreases in precedent precision imply that the

variability of outcomes in sufficiently dissimilar cases actually *decreases*. This is because under a more principle-based (i.e. high-variance) rule, very dissimilar cases, while still subject to variability, are subject to less variability than they would be under a narrowly tailored precedent, which would have little predictive power for those dissimilar cases. Thus, there reaches a point at which the adverse consequences of any further decreases in precision are offset by the benefits of decreasing the variability of interpretations in very dissimilar future cases.

Proposition 2. *Decreasing the precision of precedent has a non-monotonic effect on the Supreme Court's incentive to hear a new case. When precision in the precedent is sufficiently high (low ν), decreasing precision increases the incentive to hear any case. When precision is sufficiently low (high ν), decreasing precision decreases the incentive to hear all cases.*

The logic behind this result is subtle but intuitive. As the precision of the precedent decreases, so too does the variability of the interpretations in other cases. This is true for similar as well as dissimilar cases. At first, the decrease in precision disproportionately affects similar cases. That effect creates an incentive to hear a new case. However, because the marginal effect of decreasing precision is decreasing and ultimately potentially negative (by increasing the precision in very dissimilar cases), there exists a tipping point at which the more principle-based (i.e. low-precision) precedent decreases the incentive to hear a new case because the adverse consequences of having little predictability in new cases under a narrowly tailored (i.e. high-precision) precedent are mitigated by the lower variability induced in dissimilar cases by the principle-based precedent. In short, under a narrowly tailored rule, anything can happen in dissimilar cases; decreasing precision makes that problem worse. However, sufficiently large decreases in precision mitigate the problem by introducing a limited degree of predictability from the more principle-based precedent.

Critically, the point at which further decreases in precision stop increasing the incentive to hear a new case and start decreasing the incentive depends on where the precedent falls in the larger distribution of cases. To see this, note that the effect of decreasing precision of the precedent is positive only when

$$\nu > \sqrt{\frac{\int_{\frac{q+q'}{2}}^{b} (p-q)^2 f(p)\,dp}{1 - F\left(\frac{q+q'}{2}\right)}}$$

(see the proof of Proposition 2) and is negative when the converse condition holds. That critical value of ν is decreasing in q *as long as q is sufficiently typical of the overall distribution of new cases, relative to q'*. In other words, when the precedent case is a typical case, then the critical point at which additional decreases in precision start to diminish the incentive to hear a new case is lower than when the precedent case is in a less typical case. As a consequence, for any given level of precision in the precedent, it is less likely that the Court will be sufficiently incentivized to hear a new case. Thus, Proposition 2 demonstrates that precedent precision presents a potentially complicated set of incentives for the Court.

Corollary 1. *When the precedent case is sufficiently typical of the overall distribution of cases relative to the new case it is considering hearing, for any given level of precision the Supreme Court will be less likely to want to hear the new case.*

Corollary 1 shows that, quite intuitively, the Court is "better off" when precedents have been set in typical cases and "worse off" when precedents have been set in unusual cases. However, the rationale for that finding is not what one might expect. It is not *simply* that the precedent in the typical case means that most cases are already governed by a good precedent: recall that we have not assumed that the precedent was written with optimal precision. Instead, when the precedent is in a typical case, choosing any other case is not as likely to affect lots of future cases as when the precedent is in an atypical case; consequently, the point at which there are no longer benefits from decreasing precision arises very quickly. This holds because the Court's decision is whether to write an opinion in the new case: when the precedent is more typical than the new case, the new case necessarily has less of an effect on overall outcomes than does the precedent case.

The effect of case similarity on the decision to hear the new case. We now turn to an analysis of how the relationship between the factual similarity between the precedent case, q, and the new case under consideration, q', affects the Court's decision calculus. Note first that, as we have seen several times, the effect of varying different parameters on inequality (4) is ambiguous. As one increases the factual dissimilarity of the new cases (i.e. as one increases q'), the effect on the optimal level of precision, ν^* depends on the underlying distribution of cases. If q is more representative than q', then the optimal precision strictly increases as q' moves further away from q. However, if q is less representative than q', then the optimal precision initially decreases and then ultimately increases as q' moves further away from q. This follows from Proposition 1.

Proposition 3. *The effect of increasing factual dissimilarity between the new case q' and the precedent case q on the optimal precision for case q' depends on the underlying distribution of cases. As q' moves away from q, the optimal level of precision decreases as q' becomes more factually representative of the underlying distribution of cases and increases as q' becomes less factually representative of the underlying distribution of cases.*

Note also that as the factual dissimilarity between the precedent and new case increases, the region to which the new precedent applies decreases. The consequence is a complex effect of factual dissimilarity on the left-hand side of inequality (4). As the new case becomes increasingly dissimilar, the potential benefit from hearing the case actually decreases, as it will have less *ceteris paribus* impact because it applies to fewer cases. However, the impact the new case will have on those cases to which it does pertain in application (i.e. those case that are governed by q' as opposed to q) depends itself on how representative q' is relative to how representative q is of the underlying distribution of cases facts. The reason is that,

by Proposition 3, how different the precision of the new case will be from the existing precedent depends on that relationship.

The consequence of this complex interaction is that increasing the factual dissimilarity between the new case and the precedent case has a weakly non-monotonic effect on the incentive to hear the new case, q'. Initially, increasing the factual dissimilarity between the new case and the precedent can either increase or decrease the incentive to hear the new case; however, there reaches a point at which further increases in the factual dissimilarity have a strictly negative effect on the incentive to hear the new case. Whether this effect is non-monotonic, i.e. whether the initial increase in factual dissimilarity does in fact increase in the incentive to hear the new case, depends on whether the new case presents an opportunity to "improve" outcomes in enough cases.

Proposition 4. *Increasing factual dissimilarity between two cases has a weakly non-monotonic effect on the Supreme Court's incentive to hear a new case. Initial increases in the factual dissimilarity between two cases can either increase or decrease the likelihood the Court will prefer to hear a new case, while further increases in factual dissimilarity strictly decrease the likelihood the Court will prefer to hear the new case.*

The effect of Proposition 4 is illustrated visually in Figure 3. In this figure, the cost of review, the factual location and variance of the precedent case, the distribution of case facts, and the cost of review are all fixed at arbitrary values. The x-axis measures the set of possible case facts; the dashed vertical line shows the factual location of the precedent case, q; and the black line shows the distribution of case facts. The gray-shaded region shows the range of cases $q' > q$ that the Court would be willing to hear. As the figure demonstrates, the Court is not willing to hear the cases most similar to the precedent case: the logic again is that the variance in the precedent does not increase that much for initial factual deviations. However, provided sufficient factual discrepancy, the Court is willing to hear the new case because with enough factual dissimilarity, there is sufficient variance in the precedent, relative to the amount of cases to which the new precedent could apply, to justify hearing the case. By the same token, though, taking very dissimilar cases is not justified, because while the precedent is not much more variable than it was in its initial case, q, taking the new case yields relatively very little return, as it would apply to so few cases.

This finding yields an implication for the factual relationship among cases the Court will hear. In particular, it demonstrates that rather than jumping around through disparate cases, the Court's optimal strategy for creating a body of doctrine among cases is to strategically take cases that are neither too similar nor too dissimilar in their factual circumstances. Thus, the law will be constructed in a way that is somewhat endogenously path dependent. The Court's decision to hear a case today is determined in part by the type of decisions it made in past cases (Proposition 2) and the particular problem (facts) of the past cases it has heard (Proposition 4). While these relationships are certainly intuitive, Propositions 2 and 4 provide specific predictions about how they will drive those patterns.

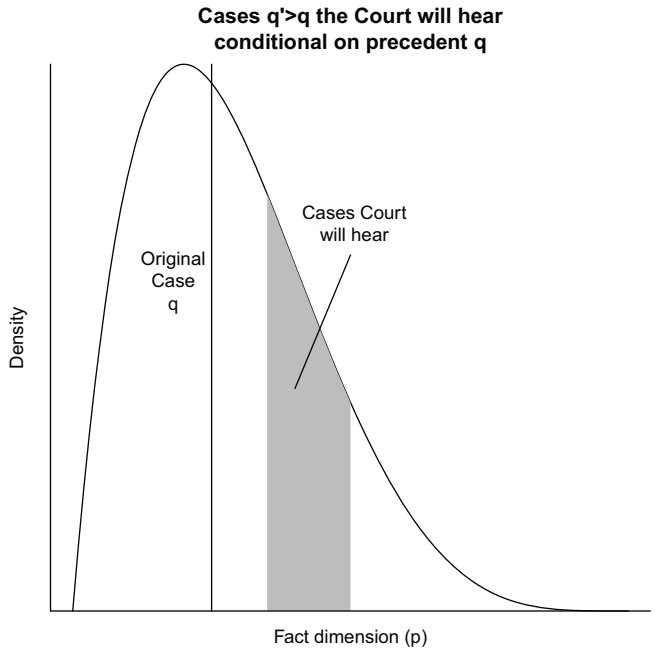

Figure 3. Cases that the Court is willing to hear, conditional upon the location and variance of the precedent case and cost of review such that the effect of case similarity is non-monotonic (Proposition 4). The *x*-axis measures the facts of the cases; the black line shows the distribution of cases; the dashed vertical line shows the factual location of the precedent case, *q*, and the gray region shows the set of cases that the Court would be willing to spend resource cost *k* to hear and announce a new, optimal, precedent.

Result 2. *The Supreme Court will prefer to decide new cases that are factual complements to their precedents: the best new cases are close enough to past cases to benefit from the precedent but are sufficiently different to allow the Court to further expound the law through the new case*.

From this result, we can see how the model provides theoretical underpinnings for how the Court can optimally create a line of doctrine through a series of cases. The optimal strategy is not to randomly select cases, nor is it to consistently decide similar cases. Instead, the Court's optimal strategy is to strategically select cases that present complimentary factual scenarios in order to "cover" as much of the "fact space" as it can with well-tailored legal rules.

3.2.3. Opinion writing in a multi-stage game. Finally, we turn to how the prospect of hearing future cases affects the opinion the Court will write in a current case. The previous two analyses isolated, respectively, the incentives the Court faces when it has the opportunity to structure the specificity of doctrine and how existing doctrine shapes the decision about what future cases to hear. In this final analysis, we

put those pieces together to examine how expectations about future cases will affect the structure of doctrine in the first instance. In essence, we return to the initial analysis, which considers the optimal level of precision for the Court to include in its opinion, while incorporating the focus of the second analysis: the potential for the Court to revisit the legal issue by hearing another case.

Consider the Supreme Court's choice in a simple two-period example in which nature presents the Supreme Court with a case q in the first period and then another case q' in the second period. In the event it does not want to hear case q' in the next period, its expectation about the utility of choosing precision ν in case q is determined by its expectation about the application of the opinion with precision ν to all future cases. However, in the event that the Court does want to hear case q', then the Court's utility from choosing precision ν is the combination of the utility from ν, applied to all of the cases for which case q will be the controlling precedent, plus the utility from choosing its optimal precision in case q', applied to all of the cases for which case q' will be the controlling precedent. I assume that the Court discounts the value of applications after the second period by a factor $\delta \in (0, 1)$.

Of course, the matter is complicated by the fact that, as we saw above, the Court's preference over the decision to hear case q' is endogenous to its selected precision in case q. Let $\pi_{q'<q} = \int_{\underline{\underline{q}}^*}^{\bar{\bar{q}}^*} f(p)\,dp$ be the probability that the Court wants to hear a new case $q' < q$, conditional on the value of precision the Court chooses in the instant case, $\nu(q)$. Similarly, let $\pi_{q'>q} = \int_{\underline{q}^*}^{\bar{q}^*} f(p)\,dp$ be the probability that the Court wants to hear a new case $q' > q$, conditional on the value of precision the Court chooses in the instant case, $\nu(q)$. The values $\underline{\underline{q}}^*$ and $\bar{\bar{q}}^*$ are those that solve the condition in Proposition 4 at indifference for $q' < q$; the values $\underline{q}^*$ and $\bar{q}^*$ are those that solve the condition in Proposition 4 at indifference for $q' > q$. For example, the integral $\int_{\underline{q}^*}^{\bar{q}^*} f(p)\,dp$ corresponds to the gray region in Figure 3. The integral $\int_{\underline{\underline{q}}^*}^{\bar{\bar{q}}^*} f(p)\,dp$ corresponds to the complementary gray region that could exist in Figure 3 to the left of q. Finally, for notational simplicity, let $\Pi = \int_{\underline{\underline{q}}^*}^{\bar{\bar{q}}^*} f(p)\,dp + \int_{\underline{q}^*}^{\bar{q}^*} f(p)\,dp$. The Court's expected utility, conditional on the level of precision it selects in the instant case, and optimal behavior in the second period (both the decision to hear the case the level of precision to include in any second period opinion), is given by

$$-\int_X EU(\nu) f(p)\,dp - \delta\left[\pi_{q'<q}\left(\int_a^{\frac{q+q'}{2}} EU(\nu'^*) f(p)\,dp + \int_{\frac{q+q'}{2}}^b EU(\nu) f(p)\,dp\right)\right.$$
$$\left. + \pi_{q'>q}\left(\int_a^{\frac{q+q'}{2}} EU(\nu) f(p)\,dp + \int_{\frac{q+q'}{2}}^b EU(\nu'^*) f(p)\,dp\right) + (1-\Pi)\int_X EU(\nu) f(p)\,dp\right] \tag{5}$$

where ν is the precision the Court sets in the instant case, q, ν'^{*} is the optimal precision the Court would set in case q' in the next period, and $EU(\nu) = \nu + \frac{(p-q)^2}{\nu}$. The analytic question is: what value of ν maximizes Equation (5)?

Obtaining a closed-form solution to this maximization problem is intractable. However, we can analytically demonstrate there exists a finite value of ν that maximizes the Court's expected utility and the way that value changes with respect to other model parameters. First, just as in the single-shot model, the optimal level of precision is a function of the case q's factual representativeness. The Court prefers less precision (i.e. higher ν) for more unusual cases and greater precision (lower ν) for more typical cases. The intuition behind this result is straightforward.

Proposition 5. *In a multi-period opinion-writing model, the optimal precision has a similar relationship to the case facts as in the single-shot model. The optimal precision decreases as a case becomes increasing factually unusual.*

Second, contrasted with the single-shot game, the Court prefers to write a more precise opinion (lower ν) for any given case. The intuition behind this result is similarly straightforward. The adverse consequence of deviating from optimal precision to greater precision is that dissimilar cases will have a very variable rule applied to them. The positive incentive to deviate to a increase precision, though, is that is results in less variable rules being applied to similar cases in the future. Optimal behavior is defined as the balance of those two incentives such that the marginal cost of further decreases in variance is not justified by the marginal benefit of further increases in precision. However, the probabilistic opportunity to return to the doctrine in the future disrupts that balance in a predictable way. Specifically, the potential ability to return to the doctrine at least marginally decreases that adverse consequence of a deviation to greater precision without affecting the benefit. As a consequence, the Court will weakly prefer greater precision in a model where it can return to the doctrine in the future.

Corollary 2. *Given the chance to return to a line of doctrine in the future, the optimal precision to include in an opinion is weakly higher (i.e. lower ν) than in a model where the Court cannot return to a line of doctrine.*

A final formal result concerns the discount factor applied to future cases (δ). This parameter, in conjunction with the cost of reviewing a case, k, captures the Court's impatience or interest in an area of law, as well as the nature of litigation in the area of law, such as the extent to which new questions brought to the Court often and the consequences of the rule's application in the courts below. The larger δ, the more value the Court places on future periods of play; the smaller δ, the more short-sighted the Court. The optimal precision to include in an opinion is always increasing in δ.

Proposition 6. *The higher the value the Court places on future periods, the more narrowly tailored the opinion the Court will write. Alternatively, the less the Court cares about future rounds of play, the more principle-based the opinion it will write.*

That the Court will write more narrowly-tailored opinions when it cares more about rounds may seem on its face counter-intuitive. Indeed, one might suspect that when the Court places a great deal of weight on future applications of its rule (i.e. when δ is large), the Court would write more principle-based opinions to guard against highly variable future applications of the rule. However, the intuition behind the result here is that when the Court cares a great deal about future rounds, it will be more wiling to take new cases in the next period and therefore will prefer to write a more narrowly tailored opinion in the first instance because it expects to be more likely to hear a new case in the future and thereby complement the precedent and offset the variability in distant cases induced by the narrowly tailored precedent. Notice the parallel between this result and Corollary 2. That finding made the comparison between a situation where $\delta = 0$, i.e. future rounds of play do not affect the Court's utility, and a situation where $\delta > 0$. Proposition 6 generalizes this implication in the context of the two-period model.

Taken all together, these findings give rise to an important result. In an area of law where the Court is forward-looking, and willing to invest its time and resources in a series of cases, the Court will be more likely to craft narrowly tailored opinions, taking advantage of its future opportunities to construct precedent through a series of opinions. By contrast, if the Court is more myopic and discounts future opportunities, then the Court should be more likely to write broad, principle-based opinions. The implication is that in areas of the law where there is likely to be a stream of "good" cases, where the Court anticipates good opportunities to hear new cases, then doctrine will be composed of a series of fact-bound rules; when it is not likely that there will be such a stream of cases, doctrine will be composed of few opinions, with principle-based rules seeking to cover wide ranges of cases.

Result 3. *The extent to which an opinion is more principle-based or more narrowly tailored is a function of the Court's patience. As the Court increasingly values future opportunities to revise precedent, the Court will write more narrowly tailored opinions. When the Court does not anticipate opportunities to return to a doctrine or where the Court does not value future opportunities, the Court will write more principle-based opinions.*

4. Discussion and implications

The preceding analysis yields several implications for the study of judicial politics and discretionary review in a rule-making political hierarchy. Specifically, the model has implications for three features of Supreme Court rule-making that go to the heart of contemporary studies of doctrinal politics. First, the analysis highlights how managerial incentives over the future application of precedents will affect the scope and breadth of the Court's precedents. Second, the analysis provides theoretical foundations for case-selection by the Supreme Court as it seeks to articulate doctrine through a series of cases with different factual circumstances. Third, the model yields original insights about the way in which the Court builds doctrine dynamically, with a particular insight about the effect of the Court's forward-lookingness on the structure of judicial opinions.

4.1. Scope and precedent

Motivated in part by calls from the legal academy to "take law seriously" (Friedman, 2006; Kornhauser, 1992b), the literature has witnessed increasingly sophisticated models of law and doctrine over the past decade (Gennaioli and Shleifer, 2007; Landa and Lax, 2009; Lax, 2007; Staton and Vanberg, 2008). Perhaps the debate most closely connected to the legal academy concerns the specificity with which a legal rule is written in a judicial opinion. This debate is often framed in terms of "rules" versus "standards" (e.g. Lax, 2012). At its core, it is centrally about the extent to which the principal-agent problem associated with having one Supreme Court overseeing many lower courts affects the types of instructions the Supreme Court prefers to give to lower courts in order to best achieve its policy goals. While not principally about rules versus standard, the formal analysis above adds an additional level of nuance to that line of inquiry and raises new questions new avenues for examination. In particular, the model leverages the fact that the Supreme Court's instructions to lower courts in the context of any given case are only partial. They are limited by the Supreme Court's ability to authoritatively determine the outcome in future cases (e.g. Levi, 1949; Maltz, 1988), the inherent variability in outcomes that follows from lower court implementation of its opinions (e.g. Bueno de Mesquita and Stephenson, 2002), and the Court's practical limitations on its ability to clearly communicate (e.g. Jordan, 2011).

The model developed in this paper brings those features of the judicial environment to the fore in the study of rule-making. When crafting legal rules, the Supreme Court faces a problem replete with uncertainty. It must choose between minimizing uncertainty in outcomes in a limited set of cases and accepting great uncertainty in the wider set of cases, by crafting a more narrowly tailored opinion, or inducing a more evenly distributed level of uncertainty across most cases, by crafting a more principle-based rule. The analysis above reveals that when the Supreme Court expects there will be lots of factually different cases in the future (for example, when the case it is deciding is factually unusual or when there is no such things as a "typical" case), then it will prefer to write more principle-based rules. By contrast, when the case is very typical, the Court will be best served by writing a narrowly tailored opinion, accepting there may be divergent outcomes in the unusual future case. These results may be intuitive, but what is not is the logic that emerges from the model.

Because principle-based rules have a relatively high "fixed" level of variance but exhibit less variance in variance, when deciding a factually unrepresentative case, the Court is incentivized to write a principle-based rule because of the adverse consequences that would obtain for the bulk of cases. However, the Court prefers a more narrowly tailored rule when it is deciding a factually representative case, because the higher variance is tolerable in distant cases, as there are not many such cases (relatively speaking). The crucial feature here is that the Court would always prefer to write a narrowly tailored rule were it not for that rule's lack of predictability in dissimilar future cases. That lack of predictability creates the incentive to tolerate imprecise outcomes in those cases that would otherwise be precisely governed (i.e. similar cases).

4.2. Rules and standards

As noted above, the tension I study is connected to, though distinct from, the study of rules and standards. Especially in the law, but in any rule-making setting, there is often the capacity for policy-makers to craft statements that are relatively easier or harder for those applying the policy to exercise discretion. For example, one might craft a more rule-like policy that identifies clear, objective criteria that give unequivocal answers in all (or many) instances. Alternatively, one might craft a standard-like policy that incorporates more subjective judgments, such as the desire to balance multiple criteria. The key tension this model contemplates is between how *broad* a policy is intended to apply and how *precisely* the policy is written. Rules can be broad statements that apply to all cases: the speed limit is a rule that applies almost without exception to all driving situations. Standards can be fact-bound, such as the standards for determining negligence in the event of personal injury in a home. The analysis here, then, highlights a second dimension along which policies can vary. Taking account of both the *scope* of a rule and the rule-standard dimension can create potentially complex strategic dynamics for the judicial process.

4.3. Case selection by the Supreme Court

A second set of implications that follows from the formal analysis above concerns the Supreme Court's certiorari decisions. Since the early 20th century, the Supreme Court has had virtually complete discretion over which cases it will hear. This has enabled the Court to selectively focus its limited resources on those cases that the Court deems most worthy. That institutional power has spurred a great deal of scholarly attention (e.g. Caldeira and Wright, 1988; Caldeira et al., 1999; Cameron et al., 2000; Clark and Kastellec, 2013; Perry, 1991). One of the lessons of that research, which should not come as too surprising, is that the Court does not randomly select cases but instead chooses cases to strategically achieve its goals. The research, though, has not yielded must insight into the strategic dynamics of selection as they pertain to the Court's desire to cover as many types of cases with its rules; rather, the research has focused on inducing compliance with its preferences by raising the risks of being "caught cheating." That we do not know much about how the factual scenarios in the cases the Court does hear relate to the overall distribution of types of cases raises serious problems for our ability to draw inferences about the meaning of the Court's doctrine and the effects that particular factual patterns have on the application of legal rules (Kastellec and Lax, 2008).

The model developed here provides theoretical foundations for the process by which the Supreme Court decides which types of factual patterns it should take in order to optimally communicate its preferred doctrine to lower courts and external enforcers. While the literature increasingly contemplates how the Supreme Court can use is discretion over which cases to hear in order to maximize strategically its influence over the content of legal rules (e.g. Clark and Carrubba, 2012; Lax, 2012; Perry, 1991), less work has focused on the importance of hearing cases in sequence to build a body of doctrine. Those studies that have emphasized the sequential nature of doctrinal development have been more concerned with the consequences

of the sequential nature of case law than with the properties of optimal choice by the courts (e.g. Bueno de Mesquita and Stephenson, 2002; Gennaioli and Shleifer, 2007; Kornhauser, 1992a). The theoretical model developed in this paper represents a step toward the construction of a more comprehensive understanding of how the Court uses case selection to shape doctrine.

4.4. Opinion-writing under the shadow of the future

A third, though related, implication of the above analysis concerns how the Court's relative future-orientation affects the nature of its opinions. Specifically, as Result 3 demonstrates, when the Court is more future-oriented, it will write more narrowly tailored opinions than when it is more myopic. The important implication of this finding is that a Court may be willing to tolerate short-term adverse outcomes (because of high variance in dissimilar cases) in order to construct a body of precedent through a series of narrowly tailored opinions. This dynamic has parallels in the a number of features of the US Supreme Court's judicial process, such as its willingness to tolerate conflict among lower courts in order to allow an issue to "percolate" (Clark and Kastellec, 2013). Perhaps more important, it suggests that when the Court is thinking about constructing doctrine through a number of cases, i.e. is patient, it will do so though a set of very narrowly tailored opinions. However, when the Court discounts any future opportunities or has no interest in returning to an issue, the Court will write broad, principle-based opinions. This incentive may explain the dynamics we observe in lines of doctrine being constructed. For example, throughout the 1970s and 1980s, the US Supreme Court issued a steady stream of cases concerning abortion rights, most of which were very narrowly tailored and factually constrained (see, for example, the discussion of the *Roe* precedent above). However, in 1992, the Court moved to a more principle-based doctrine in *Planned Parenthood v. Casey*. This move might have signaled the Court's intention to no longer invest time and energy in abortion doctrine; in other words, its discounting of any future opportunities to return to the doctrine. Indeed, far fewer cases have been added to abortion doctrine during the past 20 years than were during the preceding 20 years.

This last implication is of particular import because, in conjunction with the preceding results, it demonstrates that the model provides both microfoundations for patterns of doctrine-building that we often observe. It also gives rise to insights about the way in which the Court builds doctrine, while balancing concerns of clarity in the law with effective governance. A Court constrained in its ability to clearly and completely articulate a doctrine, must decide how much it values the opportunity to return to a line of doctrine in the future and must also think about the types of cases that are likely to be presented in the future.

5. Conclusion

Scholarship on the judicial process is increasingly turning to questions of how the courts construct legal rules, how doctrine evolves through a series of cases, and

how judges use their opinions to communicate with lower courts and external actors (e.g. Baker and Mezetti, 2012; Gennaioli and Shleifer, 2007; Landa and Lax, 2009; Lax, 2007). An important feature of judicial policy-making that is generally not incorporated into these theoretical models is that judges are limited in their ability to communicate their preferred doctrine. While some models contemplate a judge's choice to intentionally introduce vagueness into her ruling (e.g. Fox and Vanberg, 2011; Lax, 2012; Staton and Vanberg, 2008), these models assumes that a judge could, in principle, make a fully articulated announcement of her ideal rule. However, judicial opinions are limited by the ability of lawyers, lower courts, or enforcing actors to distinguish the precedent on factual grounds, as well as by the sheer limitations of human language. As such, judges cannot perfectly communicate their ideal doctrine in a single opinion.

The model developed here isolates a tension that arises in this setting. In contrast to previous models that have contemplated judges' choice about the binding nature of rules (e.g. Jacobi and Tiller, 2007; Schwartz, 1992), this model highlights the mechanism by which such tradeoffs exist and uncovers implications that cannot arise in those previous models. Judges face a trade-off between writing a narrowly tailored opinion that contains clear, precise guidance in factually similar future cases or writing instead a more principle-based opinion that is less precise but more broadly applicable. Narrowly tailored opinions are by their nature closely connected to the facts of a case and therefore easily distinguished by future courts. As a consequence, the variability in outcomes under a narrowly tailored doctrine is very sensitive to the factual similarity between the precedent case and the case where the precedent is being applied. Principle-based opinions, by contrast, are less easily distinguished along factual grounds, because they are less tied to the facts of the precedent case. However, in order to untie a precedent from the facts of the particular case, a court must allow for greater variability in outcomes in even similar cases, because to be principle-based, the opinion must make use of more general principles and instructions.

Isolating this tension allows us to investigate a number of features of the judicial process and identify the incentives created by these features of doctrine creation. In particular, the model identifies three findings that provide new insight into the construction of judicial doctrine. First, the model predicts that the more typical a case (from a factual perspective), the more narrowly tailored an opinion. As a consequence, when cases are "lumpy", we should expect to find narrowly tailored opinions in the case of typical cases. However, when the distribution of cases is flat, and there are no typical cases, we should find doctrine that is more principle-based. Second, previous research does not give us much theoretical traction on which types of cases the Court will take (from a factual perspective) in pursuit of constructing doctrine. Scholars have invested considerable energy into understanding how lines of doctrine build and evolve through judicial decision-making (e.g. Clark and Lauderdale, 2012; Wahlbeck, 1997), but that research's theoretical orientation has been primarily concerned with the political factors that affect judicial choice, rather than the "judicial" factors, such as the role that the factual relationships among cases play in doctrinal development (but see Gennaioli and Shleifer, 2007; Kornhauser, 1992b,a). Without such theoretical grounding, our ability to make

inferences from the empirical patterns in the Court's decisions is limited (Kastellec and Lax, 2008). This model provides some such grounding. Finally, the model highlights the relationship between the Court's future orientation and the doctrine it will construct. Simply because of the trade-off between narrowly tailored and principle-based opinions modeled here, the Court has an incentive to construct narrow-tailored opinions in areas of doctrine where it intends to be active, while it has an incentive to issue principle-based opinions when it has a more myopic orientation. These theoretical results can provide foundations for future empirical scholarship examining the structure of legal doctrine.

Acknowledgements

I thank Deborah Beim, Steve Callander, Cliff Carrubba, John Kastellec, Jeff Lax, Jon Nash, John Patty, and Jeff Staton for helpful comments and discussions. An earlier version of this paper was presented at the annual meeting of the Midwest Political Science Association, Chicago, IL, April 12-15, 2012.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Notes

1. Kornhauser (1995, 1617–1618), for example, discusses how the extent to which cases are "clustered together" can affect the effort a judge will invest into crafting precedent. When there are many similar cases, i.e. when cases are clustered together, then Kornhauser says a precedent has greater scope, because the precedent will control a wider range of cases.
2. One can consider this as a special case of a more general model in which the Court picks both a rule and a variance. In this model, we simply assume the rule is always optimally set to its own ideal rule.
3. Note that the denominator in the expression for v^* is $\int_X f(p)\,dp$. When considering the entire distribution of cases and a single rule, this quantity is 1. However, below I extend the analysis to instances where the rule is applied to a subset of future cases.
4. An alternative to the assumption that the courts rely only on the closest precedent would be the assume all cases are decided with a weighed average of the relevant precedents: perhaps, weighting by the similarity between cases. In that case, Equation (3) would be an integral over the entire set X of a weighed combination of both precedents' variances and the density function.

References

Ashworth S and Bueno de Mesquita E (2006) Monotone comparative statics for models of politics. *American Journal of Political Science* 50(1): 214–231.

Bailey MA and Maltzman F (2011) *The Constrained Court: Law, Politics, and the Decisions Justices Make*. Princeton, NJ: Princeton University Press.

Baird VA (2007) *Answering the Call of the Court: How Justices and Litigants Set the Supreme Court Agenda*. Charlottesville, VA: University of Virginia Press.

Baker S and Mezetti C (2012) A theory of rational jurisprudence. *Journal of Political Economy* 120(3): 513–551.
Bueno de Mesquita E and Stephenson M (2002) Informative precedent and intrajudicial communication. *American Political Science Review* 96(4): 1–12.
Caldeira GA and Wright JR (1988) Organized interests and agenda setting in the U.S. supreme court. *The American Political Science Review* 82(4): 1109–1127.
Caldeira GA, Wright JR and Zorn CJW (1999) Sophisticated voting and gate-keeping in the supreme court. *Journal of Law Economics and Organization* 15: 549–572.
Callander S (2011) Searching for good policies. *American Political Science Review* 105(4): 643–662.
Cameron CM, Segal JA and Songer DR (2000) Strategic auditing in a political hierarchy: An informational model of the Supreme Court's certiorari decisions. *American Political Science Review* 94: 101–116.
Carrubba CJ and Clark TS (2012) Rule creation in a political hierarchy. *American Political Science Review* 106(3): 622–643.
Clark TS and Carrubba CJ (2012) A theory of opinion writing in the judicial hierarchy. *Journal of Politics* 74(2): 584–603.
Clark TS and Kastellec JP (2013) The supreme court and percolation in the lower courts: An optimal stopping model. *Journal of Politics* 75(1): 150–168.
Clark TS and Lauderdale BE (2012) The genealogy of law. *Political Analysis* 20(3): 329–350.
Epstein L, Segal JA and Johnson T (1996) The claim of issue creation on the U.S. supreme court. *American Political Science Review* 90(4): 845–852.
Fox J and Vanberg G (2011) Law as a discovery procedure: In informational rationale for broad judicial decisions? Working paper. Available at: http://www.princeton.edu/csdp/events/Fox01192011/Fox01192012.pdf
Friedman B (2006) Taking law seriously. *Perspectives on Politics* 4(2): 261–276.
Gennaioli N and Shleifer A (2007) The evolution of common law. *Journal of Political Economy* 115(1): 43–68.
Gillman H and Clayton C (eds.) (1999) *Supreme Court Decision-Making: New Institutionalist Approaches*. Chicago, IL: University of Chicago Press.
Jacobi T and Tiller EH (2007) Legal doctrine and political control. *Journal of Law, Economics, and Organization* 23(2): 326–345.
Jordan S (2011) Policymaking by rules and cases. Working paper, University of Rochester.
Kastellec JP and Lax JR (2008) Case selection and the study of judicial politics. *Journal of Empirical Legal Studies* 53(3): 407–446.
Kornhauser LA (1992a) Modeling collegial courts *i*: Path dependence. *International Review of Law and Economics* 12: 169–185.
Kornhauser LA (1992b) Modeling collegial courts *ii*: Legal doctrine. *Journal of Law Economics and Organization* 8: 441–470.
Kornhauser LA (1995) Adjudication by a resource-constrained team: Hierarchy and precedent in a judicial system. *Southern California Law Review* 68: 1605–1629.
Landa D and Lax JR (2009) Legal doctrine in collegial courts. *Journal of Politics*, forthcoming.
Lax JR (2007) Constructing legal rules on appellate courts. *American Political Science Review*. 101(3): 591–604.
Lax JR (2011) The new judicial politics of legal doctrine. *Annual Review of Political Science* 14: 131–157.
Lax JR (2012) Political constriants on legal doctrine: How hierarchy shapes the law. *Journal of Politics* 74(3): 765–781.

Levi E (1949) *An Introduction to Legal Reasoning*. Chicago, IL: University of Chicago Press.
Maltz E (1988) The nature of precedent. *North Carolina Law Review* 66(2): 367–394.
Maltz E (2000) The function of supreme court opinions. *Houston Law Review* 37: 1395–1420.
McGuire KT and Palmer B (1995) Issue fluidity on the U.S. supreme court. *American Political Science Review* 89(3): 691–702.
Milgrom P and Shannon C (1994) Monotone comparative statics. *Econometrica* 62(1): 157–180.
Owens RJ and Wedeking JP (2011) Justices and legal clarity: Analyzing the complexity of U.S. supreme court opinions. *Law & Society Review* 45(4): 1027–1061.
Perry HJr (1991) *Deciding to Decide: Agenda Setting in the United States Supreme Court*. Cambridge, MA: Harvard University Press.
Schwartz EP (1992) Policy, precedent, and power: A positive theory of supreme courtdecision-making. *Journal of Law, Economics, and Organization* 8(2): 219–252.
Segal JA and Spaeth HJ (2002) *The Supreme Court and the Attitudinal Model Revisited*. New York: Cambridge University Press.
Staton JK and Vanberg G (2008) The value of vagueness: Delegation, defiance, and judicial opinions. *American Journal of Political Science* 52(3): 504–19.
Sullivan KM (1992) The justices of rules and standards. *Harvard Law Review* 106(1): 22–123.
Sunstein CR (1999) *One Case at a Time: Judicial Minimalism on the Supreme Court*. Cambridge, MA: Harvard University Press.
Wahlbeck PJ (1997) The life of the law: Judicial politics and legal change. *Journal of Politics* 59(3): 778–802.

Appendix

Proof. Proposition 1 The first derivative of Equation (2) with respect to ν is

$$-\int_a^b \left(1-\frac{(p-q)^2}{\nu^2}\right) f(p)\,dp$$

Setting the derivative equal to zero, and solving for ν, we have

$$\begin{aligned}
0 &= -\int_a^b \left(1-\frac{(p-q)^2}{\nu^2}\right) f(p)\,dp \\
0 &= -\left[\int_a^b f(p)\,dp - \int_a^b \frac{(p-q)^2}{\nu^2} f(p)\,dp\right] \\
\int_a^b f(p)\,dp &= \int_a^b \frac{(p-q)^2}{\nu^2} f(p)\,dp \\
\nu^2 &= \frac{\int_a^b (p-q)^2 f(p)\,dp}{\int_a^b f(p)\,dp} \\
\nu^* &= \sqrt{\frac{\int_a^b (p-q)^2 f(p)\,dp}{\int_a^b f(p)\,dp}}
\end{aligned}$$

To verify that ν^* is a maximum, note that the second derivative is given by

$$-\frac{2}{\nu^3}\int_a^b (p-q)^2 f(p)\,dp<0 \qquad \square$$

Proof. Result 1 The derivative of ν^* with respect to q is given by

$$\frac{\partial \nu^*}{\partial q} = \frac{2\int_a^b (p-q) f(p)\,dp}{\sqrt{\int_a^b (p-q)^2 f(p)\,dp}}$$

which demonstrates ν^* is minimized at $\mathbb{E}[p]$. $\square$

Proof. Proposition 2 Assume $q' > q$. The Court prefers to hear a case with facts q' iff

$$\int_{\frac{q+q'}{2}}^b \left(\nu + \frac{(p-q)^2}{\nu}\right) f(p)\,dp - \int_{\frac{q+q'}{2}}^b \left(\nu'^* + \frac{(p-q')^2}{\nu'^*}\right) f(p)\,dp > k$$

The derivative of the left-hand side of the preceding condition, with respect to ν is

$$\int_{\frac{q+q'}{2}}^b \left(1 - \frac{(p-q)^2}{\nu^2}\right) f(p)\,dp$$

That derivative is positive whenever

$$\nu > \sqrt{\frac{\int_{\frac{q+q'}{2}}^b (p-q)^2 f(p) dp}{1 - F\left(\frac{q+q'}{2}\right)}}$$

Thus, for ν large enough relative to the case q's representativeness of the underlying distribution of case facts, the derivative is positive; for ν small enough, the derivative is negative. The case where $q' < q$ is symmetric. $\square$

Proof. Corollary 1 Assume $q' > q$. The Court's marginal benefit from taking a new case is given by prefers to hear a case with facts q' iff

$$\int_{\frac{q+q'}{2}}^b \left(\nu - \nu' + \frac{(p-q)^2}{\nu} - \frac{(p-q')^2}{\nu'^*}\right) f(p)\,dp$$

In order to take a new case, this marginal benefit must be greater than $k > 0$. Thus, as the density $\int_{\frac{q+q'}{2}}^b f(p)\,dp$ decreases (i.e. as the precedent q is more typical of the distribution of case facts), the Court will be less likely to take the new case. $\square$

Proof. Proposition 3 The proposition is a direct implication of Proposition 1 and the assumption that the lower courts apply the precedent from case q when $p < \frac{q+q'}{2}$ and the precedent from case q' when from case q when $p > \frac{q+q'}{2}$. $\square$

Proof. Proposition 4 Assume $q' > q$. Recall that the Court will only hear a new case q' when inequality (4) is satisfied. Inequality (4) is given by

$$\int_{\frac{q+q'}{2}}^{b} \left(\nu - \nu'^{*} + \frac{(p-q)^2}{\nu} - \frac{(p-q')^2}{\nu'^{*}} \right) f(p)\, dp > k$$

Consider $q' = q + \varepsilon$. For sufficiently small ε, then inequality (4) is satisfied only if $\int_{\frac{q+q'}{2}}^{b} f(p)\, dp$ is large enough. In particular,

$$\lim_{\varepsilon \to 0} \int_{\frac{q+q'}{2}}^{b} \left(\nu - \nu'^{*} + \frac{(p-q)^2}{\nu} - \frac{(p-q')^2}{\nu'^{*}} \right) f(p)\, dp = \int_{q}^{b} \left(\nu - \nu'^{*} + \frac{(q)^2}{\nu} - \frac{(p-q')^2}{\nu'^{*}} \right) f(p)\, dp$$

which implies that there must be enough cases to the right of q to justify hearing the new case. A complementary way of describing this condition is that the improvement in precision, i.e. the extent to which the optimal precision, ν'^{*}, improves the Court's expected utility, must be sufficiently large to justify the new case. Because variance in outcomes under the precedent variance, ν, is increasing in factual dissimilarity, i.e. ε, the Court is most likely to benefit enough to justify the cost (k) for ε sufficiently large. Thus, the incentive to hear a new case $q' = q + \varepsilon$ must be weakly increasing in ε. Now, consider sufficiently large ε. Note that

$$\lim_{\varepsilon \to \infty} \int_{\frac{q+q'}{2}}^{b} \left(\nu - \nu'^{*} + \frac{(q)^2}{\nu} - \frac{(p-q')^2}{\nu'^{*}} \right) f(p)\, dp = 0$$

which implies that for any $k > 0$ and sufficiently large q', increasing q' decreases the incentive to hear a new case. Taken together, these calculations imply increasing factual dissimilarity between two cases has a non-monotonic effect. Initial increases in factual dissimilarity increase the incentive to hear a new case q', whereas sufficient dissimilarity decreases the incentive to hear a new case. □

Proof. Proposition 5 Reorganized, Equation (5) can be written as

$$\begin{aligned} &- (1 - \delta(1-\pi)) \int_{0}^{1} EU(\nu) f(p)\, dp - 2\delta \int_{\underline{\underline{q}}^{*}}^{\bar{\bar{q}}^{*}} f(p)\, dp \\ &- \delta \left(\int_{\underline{\underline{q}}^{*}}^{\bar{\bar{q}}^{*}} f(p)\, dp - \int_{\underline{q}^{*}}^{\bar{q}^{*}} f(p)\, dp \right) \cdot \left(\int_{0}^{\frac{q+q'}{2}} EU(\nu'^{*}) f(p)\, dp + \int_{\frac{q+q'}{2}}^{1} EU(\nu) f(p)\, dp \right) \end{aligned} \tag{6}$$

Let

$$W = -(1-\delta(1-\pi))\int_X EU(\nu)f(p)\,dp$$

$$X = -2\delta\int_{\underline{\underline{q}}^*}^{\bar{\bar{q}}^*} f(p)\,dp$$

$$Y = -\delta\left(\int_{\underline{\underline{q}}^*}^{\bar{\bar{q}}^*} f(p)\,dp - \int_{\underline{q}^*}^{\bar{q}^*} f(p)\,dp\right)$$

and

$$Z = \int_a^{\frac{q+q'}{2}} EU(\nu'^*)f(p)\,dp + \int_{\frac{q+q'}{2}}^b EU(\nu)f(p)\,dp$$

By the proof of Proposition 1,

$$\frac{\partial^2 W}{\partial^2 \nu} < 0$$

Similarly, by the proof of Proposition 4,

$$\frac{\partial^2 X}{\partial^2 \nu} < 0$$

Note that Y captures the difference between the size of the range of cases that the Court would prefer to hear to the right of q and the range of cases the Court would prefer to hear to the left of q (multiplied by $-\delta$. We know from the proof of Proposition 4 that these two quantities must be strictly increasing in ν. Therefore,

$$\frac{\partial^2 Y}{\partial^2 \nu} = 0$$

Thus, the entire second derivative of Equation (6) is negative. Therefore, Equation (6) is concave and therefore has a finite maximum on the closed interval $X \equiv [a, b]$. □

Proof. Proposition 6 Note first that Equation (5) satisfies the single crossing with respect to the variance, ν, and the discount factor, δ. To see this, note that for two discount factors, $\delta > \delta'$, the effect of changing the variance, ν, is smaller with δ than with δ', because the larger δ, the more the utility from a probabilistic future case compensates for the change in utility associated with increasing variance. Formally,

$$\frac{\partial f(\nu,\delta)}{\partial \nu} - \frac{\partial f(\nu,\delta')}{\partial \nu} \leq 0 \Rightarrow \frac{\partial^2 f(\nu,\delta)}{\partial \nu \partial \delta} \leq 0$$

thereby satisfying single-crossing (see Ashworth and Bueno de Mesquita, 2006, 218–219). Given that the Court's objective function satisfies single crossing and that $\frac{\partial^2 f(\nu,\delta)}{\partial \nu \partial \delta} \leq 0$, $\nu(q)^*$ is weakly decreasing in δ (Ashworth and Bueno de Mesquita, 2006; Milgrom and Shannon, 1994). □