The Information Dynamics of Vertical *Stare Decisis*

Thomas G. Hansford

Associate Professor of Political Science

UC Merced

thansford@ucmerced.edu

James F. Spriggs II

Sidney W. Souers Professor of Government

Washington University in St. Louis

jspriggs@wustl.edu

Anthony A. Stenger

Ph.D. Candidate in Political Science

Washington University in St. Louis

astenger@wustl.edu

The Information Dynamics of Vertical *Stare Decisis*

We propose a dynamic model of precedent in a judicial hierarchy which incorporates a "bottom-up" informational component. When a high court establishes precedents it has uncertainty regarding how they will play out when applied to future legal disputes. Lower court implementation of these precedents can inform the high court about the contemporary policy implications--i.e., the ideological location--of the precedents. If lower court usage of a precedent is informative, the high court will consider the revealed location of the precedent when contemplating reducing the precedent's authority and applicability to future cases. Using data on U.S. Supreme Court precedents and U.S. Courts of Appeals citations to these precedents, we estimate a model of the Court's negative treatment of precedent. We find lower court usage of precedent can provide new, useful information on the policy content of a precedent, helping the Court shape law in a way consistent with its preferences.

Delegation is at the heart of politics. Entities endowed with policy-making authority (referred to as principals in the commonly-used principal-agent framework) typically entrust some of this authority to subordinates (the agents in this framework). Though most work on principal-agent questions examines delegation within and between the legislative and executive branches of government (e.g., Carpenter 1996; Shipan 2004), there is also a substantial literature applying the principal-agent framework to judiciaries. Many of these studies assess the existence of vertical *stare decisis*, meaning they investigate the degree to which lower courts (i.e., judicial agents) follow the precedents established by higher courts (i.e., judicial principals) (e.g., Baum 1980; Luse et al. 2009; Songer, Segal, and Cameron 1994). Other studies examine whether lower courts take into account the current preferences of higher courts (e.g., Haire, Songer, and Lindquist 2003; Klein 2002; Randazzo 2008; Westerland et al. 2010). While they differ in terms of whether judicial agents look to existing policy set by a judicial principal or anticipate the future decisions of this principal, both research strands take a "top down" approach and are concerned with the degree to which judicial principals control judicial agents.[1]

We propose there is much to be gained by considering a "bottom-up" perspective and examining how the behavior of judicial agents can provide important information to a judicial principal. Specifically, we consider the following question: Can lower court usage of high court precedent inform the high court about the actual policy content or implications of a precedent and thus influence the high court's subsequent doctrinal choices? Instead of solely focusing on the downward movement of information in the judicial hierarchy, we explore the possibility that information about precedent also moves upwards.

We begin this paper with a brief discussion of existing approaches to understanding judicial hierarchies. We then propose a theory of precedent which incorporates an important

"bottom-up" component to the flow of information. We contend that when a judicial principal (high court) establishes precedents (i.e., legal policies) it has a degree of uncertainty regarding how these precedents will actually play out when applied to current and future legal disputes. The implementation of these precedents by judicial agents (lower courts) can inform the principal about the "true" nature of the precedent. Specifically, the ideological nature of the lower court decisions citing a precedent can provide information to the high court about the implications of the precedent, as it is applied to contemporary disputes. The high court can then use this information to correct its body of precedent. The less desirable the location of a precedent is to the high court, the more likely it is that this court will choose to reduce its authority and applicability to future legal disputes (i.e., will treat the precedent negatively).

We test this prediction with data on more than 6,000 precedents established by the U.S. Supreme Court over nearly six decades and all U.S. Courts of Appeals citations to these precedents. We estimate a pair of models of the Supreme Court's negative treatment of precedent and find that the Court responds to the lower courts' use of the precedent in the manner we predict. Lower court use of Supreme Court precedents can thus provide meaningful information to the Court about the actual policy consequences of these precedents, better allowing the Court to shape the state of the law by altering precedent. This result has implications for our understanding of precedent, judicial hierarchies, and, most generally, principal-agent relationships in politics.

Existing Models of Judicial Hierarchy

Theories of judicial hierarchy typically employ a principal-agent framework in which a high court is the principal and lower courts are the agents. The high court has the ultimate policy-making authority, but much of this authority is effectively delegated to lower courts as

they are the first, and often only, courts to apply the legal policy to new disputes. The high court monitors these decisions and reverses them when it desires. Even with the assistance of litigants, attorneys, and organized interests, this monitoring is likely to be imperfect due to the large information asymmetry between a lower court deciding a specific case and the high court for which this case is but one of hundreds or thousands of cases that could be reviewed. This information asymmetry, combined with lower court preferences that will not always align with those of the high court, means lower courts may not be perfect agents of the high court and can decide cases in a way that conflicts with the policy or preferences of their judicial principal.

With this general principal-agent model underlying it, existing research on judicial hierarchies focuses on lower court compliance or high court monitoring and correction of errant decisions. While sharing a top-down approach to understanding judicial hierarchies, extant studies can be divided into two general types of model, which we label *hierarchy without stare decisis* and *vertical stare decisis with decision correction*. Figure 1 presents a simplification of the judicial decision process in these two models.

*** Figure 1 Here ***

In the *hierarchy without stare decisis* model (decision sequence A in Figure 1), it is assumed the judicial principal wants judicial agents to decide cases in a manner consistent with the principal's current preferences over legal/policy outcomes. The principal monitors the agents and reverses decisions that deviate from its preferences. If these judicial agents dislike being reversed then they avoid reversals by falling in line with the high court's preferences. There is no true vertical *stare decisis* in this model, as lower courts are guided by the current preferences, not the precedent, of the high court. There is the potential for agency loss, which in this context means lower courts deciding cases based on preferences that diverge from those of the high

3

court. Numerous studies either propose or test versions of this model (e.g., Haire, Songer, and Lindquist 2003; Klein 2002; Owens and Black N.d.; Randazzo 2008; Westerland et al. 2010).

The *vertical stare decisis with decision correction* model (sequence B) also proposes a judicial principal-agent relationship, but does so while incorporating the norm of *stare decisis*. The high court moves first by establishing a precedent. Lower courts then confront new cases and choose how to apply the precedent. The high court can subsequently review and correct cases in which the precedent was improperly applied. Lower courts may generally follow high court precedent either due to the fear of being reversed by the high court or for role-theoretic reasons, though agency loss can occur when these courts fail to properly apply the precedent. Studies of lower court compliance often specify, or implicitly assume, this model (Baum 1980; Songer, Segal, and Cameron 1994).

A New View of Information in a Judicial Hierarchy

While these two approaches to understanding judicial hierarchies differ as to whether it is the preferences or precedent of the high court that might constrain lower courts, they share a few key features. Both types of model emphasize the top-down nature of these hierarchies, assume the high court wants to correct errant lower court decisions, and focus on the extent to which information about the high court's preferences or precedents flows down to the lower courts and influences their decisions. We propose a different informational theory of judicial hierarchy which we term *vertical stare decisis with precedent correction*. As represented in decision sequence C, our model of judicial hierarchy emphasizes the importance of precedent and the uncertainty surrounding the actual consequences of precedent. The decision sequence starts with the high court setting a precedent which is then used by lower courts to resolve some disputes or at least to justify these resolutions. This lower court usage of high court precedent then provides

4

information to the high court about the true nature or implications of the precedent, which then affects whether the high court modifies the precedent.

Unlike the *hierarchy without stare decisis* model, the model we propose incorporates precedent in the form of vertical *stare decisis*. Lower courts use high court precedent to reach or justify their decisions, but they can exhibit informationally useful ideological bias while doing so. Unlike both of the existing models of judicial hierarchy, our model emphasizes a high court's interest in correcting its precedents, instead of simply correcting lower court decisions. We will now turn to explicating *vertical stare decisis with precedent correction* in more detail.

Underlying our model is the relatively uncontroversial assumption that judges are motivated by their desire to shape the law in a manner consistent with their preferences regarding legal policy (e.g., Epstein and Knight 1998). High court judges want to establish precedents that ultimately lead to the legal outcomes they desire, and they will have an expectation for the subsequent consequences or applications of a given legal rule. However, they will not have perfect information about the downstream effects of a legal rule. This may be particularly true as time passes and a precedent is applied to novel, unanticipated legal questions. Put differently, the court setting the precedent cannot be sure about the true policy content of the precedent as it is not clear exactly what type of outcomes will result from or be justified by all possible applications of the precedent. Will the precedent typically lead to the outcomes the high court envisions or will it be used to justify undesirable outcomes? The judges on the high court cannot be fully confident in the answer to this question when engaging in legal rulemaking.

The consequences of a precedent set by the high court play out in the lower courts. Evidence from the U.S. judicial system indicates lower court judges typically comply with high court precedent (see Songer, Segal, and Cameron 1994). But, compliance with high court

precedent is not perfect, the policy preferences of lower court judges influence their decisions

(Goldman 1975), and there is some room for these judges to be selective when applying

precedent or using it to justify their preferred outcomes.[2] Lower court judges will vary in how

they use high court precedent and this variation is in part a result of the ideological nature of the

lower court judges and the nature of the precedent as it may apply to lower court cases. To be

clear, we are not arguing that lower court usage of precedent is solely a function of the

preferences of the lower court judges. All we assume is that there is an ideological component to

how these judges use precedent. All else equal, judges will prefer to use or cite precedent that

allows them to reach the conclusions they prefer.

Furthermore, recent work emphasizes that citations to precedent provide meaningful

information about the meaning, scope, and authority of the cited cases. As Cross et al. (2010,

493) note, "An opinion's citations are the operationalization of the practice of stare decisis."

Research, for example, shows that citations can be used to measure the legal importance of

Supreme Court opinions over time (Fowler et al. 2007), the ideological location of Court

precedents (Clark and Lauderdale 2010), ideological bias in lower court judges (Choi and Gulati

2008), and the depreciation of precedent (Black and Spriggs 2013).

Following our assumptions that a high court will not have perfect information regarding

the future policy implications of a precedent, that there is an ideological component to the use of

precedent, and work demonstrating the informative nature of citations to precedent, we contend

the manner in which lower courts use a precedent can help reveal the policy content of the

precedent. If a precedent, for example, is often used by conservative judges to reach

conservative decisions then this pattern reveals the conservative nature of the precedent. Lower

court applications of a precedent can thus provide information to the high court about the actual

policy content of the precedent as it is applied to contemporary disputes. A degree of ideological bias in how lower courts use a high court precedent (a form of agency loss) is useful to the high court, in the sense that this bias better informs the high court about the nature of the precedent.

To be clear, it is neither necessary nor realistic to assume the high court will closely monitor every citation to precedent that occurs in the lower courts. Instead, we simply assume the high court will be aware of the general patterns of precedent usage. In the U.S. context, the Supreme Court's awareness of precedent usage will be facilitated by litigants, lawyers, *amici*, and law clerks, all of whom communicate to the Court the key components of relevant lower court decisions, including the precedent used to justify these decisions.

Furthermore, we should note that within any judicial hierarchy there will typically be a great deal of variation in the rate at which precedents are cited in the lower courts. At one extreme, there are precedents that are almost never cited in the lower courts while at the other there are precedents that are continually cited. It is very likely (as described below) that the extent to which the type of judge or type of decision citing a precedent provides information about the precedent will vary with the volume of these citations. In the data section, we will present and then test two alternative conceptions of the relationship between citation rates and the informational content of these citations.

High Court Response to Informative Citations

What will the high court do with information about the ideological location or policy content of a precedent? A high court wants its body of precedent to reflect its preferences, broadly defined (Epstein and Knight 1998; Hansford and Spriggs 2006). When there exists a precedent that the high court finds disagreeable, it can weaken the legal status of this precedent by treating it negatively in a legal opinion. The U.S. Supreme Court, for instance, regularly opts

to undercut the authority of its existing precedents through negative treatments of these precedents in its majority opinions (Hansford and Spriggs 2006). When the Court distinguishes, limits, or overrules a precedent, for example, it reduces the degree to which this precedent will apply to, influence, or justify future decisions by lower court judges. These negative treatments of Supreme Court precedent are driven, in part, by the extent to which the precedent is compatible with the preferences of the median Justice (Hansford and Spriggs 2006).[3]

If lower court usage of a precedent reveals the contemporary policy implications of the precedent diverge from the preferences of the high court, then the high court will consider negatively treating the precedent. For example, if lower court use of a precedent reveals the precedent has a liberal location (as measured by the ideology of the citing judges) then this should cause a conservative high court to be more likely to negatively treat the precedent. This negative treatment of the precedent then reduces the precedent's applicability to future lower court cases and the degree to which the precedent can be used to justify lower court decisions.

In sum, we offer a twist on the typical principal-agent model of the relationship between a high court and the lower courts - lower court use of a precedent can provide new and useful information on the nature of the precedent, and this information assists the high court in its effort to shape the law in a manner consistent with its preferences. Specifically, lower court usage of a high court precedent can provide information about the policy content of the precedent as applied to contemporary legal disputes and this information affects whether the high court then diminishes the legal authority of the precedent.

Our argument also represents a modification, and we contend an improvement, on prior work seeking to explain a high court's decision to revisit or treat one of its precedents. Hansford and Spriggs (2006), for example, use the ideological distance between the coalition of Supreme

Court Justices setting a precedent and subsequent Courts as a predictor of the likelihood the Court will negatively treat a precedent. Their approach assumes the policy content and subsequent implications of a precedent are perfectly known by the enacting coalition. Our model, in contrast, allows for uncertainty regarding the true policy content or downstream implications of a precedent. The Court's understanding of the location of the precedent can therefore change over time as a result of information flowing up from the lower courts.

Data and Model

Broadly speaking, we are interested in testing the effect of lower court usage of high court precedent on whether the high court subsequently "corrects" the precedent by limiting its scope in some manner. While our theory is intended to be general to judicial hierarchies, empirical tests are necessarily specific and the high court we focus on is the U.S. Supreme Court. Our dataset includes all orally argued Supreme Court opinions (i.e., precedents) from the 1946 Court Term through the 2003 Term (N = 6,680). The unit of analysis is the precedent-term dyad, meaning we have an observation for each precedent for each annual term of its existence, starting in the term after it was handed down by the Court and ending with the 2004 Term. Note we thus assume the Court can treat any precedent in any term, which we believe to be reasonable given the Court's control of its docket and the tendency of litigants and *amici* to provide a wide variety of potentially applicable precedents for each case.

To avoid the possibility of lower courts responding to how the Supreme Court has treated a given precedent (see Hansford and Spriggs 2006), we only include each precedent until the term of its first treatment by the Court.[4] Thus, if the precedent-setting case is decided in the 1960 Term and then first treated by the Supreme Court in the 1970 Term, we have 10 observations for the precedent – one for each term from 1961 to 1970. If we continued to

include this precedent in our data past the 1970 Term, it is possible that lower court usage of the precedent from that point on is affected by the Supreme Court's treatment of the precedent. In other words, it becomes less reasonable to assume that the lower court usage of the precedent is exogenous to Court treatments of the precedent after the first Court treatment. This research design yields a total of 116,629 precedent-term dyads.[5]

Our dependent variable is binary - whether the Supreme Court, in one of its majority opinions, negatively treats the precedent in the given Court term. We obtain these data from Fowler et al. (2007) and Black and Spriggs (2013), who gathered them from *Shepard's Citations*. *Shepard's* codes a majority opinion as negatively treating a precedent if there is language in the opinion that has a negative effect on the legal authority of the precedent. These codings are highly reliable (see Hansford and Spriggs 2006).

Measures

Our key independent variable is the ideological distance between the Supreme Court and the precedent, as revealed by lower court use of the precedent. To measure *Distance from Citing Decisions* we need to define both the ideological location of the Supreme Court in term t and the location of the lower court decisions citing precedent p. For the former, we simply use the Judicial Common Space (JCS) score for the median Justice on the Court in term t (Epstein et al. 2007). We measure the latter as the three-year moving average of the median JCS score for the U.S. Appeals Court judges on panels publishing opinions that cite Supreme Court precedent p.

This latter measure needs a good deal of unpacking. We begin by using Fowler et al.'s (2007) and Black and Spriggs' (2013) data to identify the 276,981 U.S. Appeals Court decisions that, according to *Shepard's Citations*, cite one of the Supreme Court precedents in our data (for a total of 978,683 citations to Supreme Court precedent).[6] We then identify the judges involved

in each of these cases and match them with their JCS scores, which measure their ideological locations in the same space as the Justices (see Epstein et al. 2007; Giles, Hettinger, and Peppers 2001).[7] Guided by the assumption that appeals court decisions will to some degree reflect the ideological leanings of the involved judges, we use the median member of the appeals court panel as the measure of the ideological nature of the panel's decision.[8]

For each Supreme Court precedent in our data, we are then able to calculate the mean of these panel medians for all the appeals court cases that cited the precedent in the three years preceding the Court term under analysis.[9] The final step is to aggregate these means into three-year moving averages (for the three years preceding the Court term under analysis). To generate *Distance from Citing Decisions*, we take the absolute value of the difference between the JCS score for the median Justice on the Court in term *t* and the three-year moving average of our measure of the ideological location of the judges on panels that cited the precedent.

We also include a host of control variables in our model. While the central theoretical argument we make in this paper is that a high court can learn about the contemporary policy content of a precedent by observing ideological bias in the types of lower court decisions citing the precedent, the high court likely also has a prior belief about the location of the precedent. We therefore include *Distance from Precedent Setting Coalition* in our model and measure it as the absolute value of the difference between the median Justice in the majority coalition that set the precedent and the median Justice on the Court in term *t*.[10] Again, we use JCS scores to measure the ideological location of the Justices.

Our model also includes characteristics of the precedent that might affect whether a precedent is treated negatively by the Supreme Court (see Hansford and Spriggs 2006).[11] *Constitutional Precedent* equals one if the precedent interpreted the U.S. Constitution (from

Spaeth 2007). We measure *Precedent Breadth* as the number of legal issues and legal provisions involved in a case (see Spaeth's "Issue" and "Law" variables, respectively). *Precedent Vote Margin* equals the number of justices in the majority minus the number of dissenters in the precedent-setting case (from Spaeth 2007). *Per Curiam* equals one if the precedent-setting opinion was unsigned and zero otherwise, as taken from Spaeth (2007). We capture the salience of the precedent at the time it was decided by the number of *amicus curiae* briefs filed on the merits. Since the volume of briefs increases over time, we create a term-standardized measure (*Precedent Amici*) consisting of the number of standard deviations above or below the mean amicus brief filings in all cases for a term. To control for changes in the Court's docket over the time period analyzed, we include *Agenda Relevance* and measure it as the number of cases in Term *t* that deal with the same general issue that was central to the precedent.[12]

To account for potential duration dependence, we include the age of the precedent in the model. Experimentation reveals that the best specification is to include the natural log of the age of the precedent in term *t*: *ln(Precedent Age)*. With the issue of non-independent residuals in mind, we estimate robust standard errors that allow for clustering on the precedents. We also include fixed effects for the decade in which the precedent was decided by the Court. These fixed effects address temporal heterogeneity and the possibility of informative censoring.

Model Specification

Our central argument is that lower court usage of high court precedent can inform the high court about the contemporary ideological location of the precedent and thus influence the high court's decision to correct its body of precedent through negatively treating the precedent in question. *Distance from Citing Decisions* is thus our key independent variable. The number of appeals court citations underlying this measure, however, varies dramatically from precedent-to-

precedent and from year-to-year for a given precedent. Supreme Court precedents can generate anywhere from zero to literally thousands of lower court citations during any three-year time window, and this wide variation in the usage of precedent likely has important implications for how we incorporate *Distance from Precedent Setting Coalition* into our model.

How might variation in the extent to which appeals courts use a precedent affect the informational content flowing from the type of judge using the precedent, and thus our model specification? The extent to which ideological bias in lower court usage of a precedent informs the Supreme Court about the contemporary policy content of the precedent will likely depend upon how frequently the precedent is being cited in the lower courts. Very few lower court citations to a high court precedent allows for less confidence about the location of the precedent. As the "sample size" increases, the uncertainty surrounding the citation-indicated location of the precedent may diminish. To put this in different terms, any apparent ideological bias in the type of judge citing a precedent could simply be random noise if there are only one or two instances of citation to the precedent. This noise will decrease and the underlying ideological signal could become clearer, at least initially, as the number of citations increases, meaning that lower court usage of precedent could become increasingly informative to the high court.

With any sort of sampling, an increase in precision occurs rapidly at first as sample size increases above one. There is then a diminishing increase in precision as sample size increases further. We expect, however, that there should also be a decrease in the potential for citations to be ideological biased as the number of citations becomes particularly large. If, for example, all lower court decisions cite a precedent in a given year then there can be no ideological bias in these citations as the average citing decision would the same as the average decision. A less extreme scenario is one in which all *relevant* lower court decisions in a given year (a decision

subset of unknown size) cite the precedent. Again, this should lead to the non-informative

circumstance of the typical citing judge approaching the location of the typical judge. To put

this differently, our assumption that judges are citing cases based in part on ideological

motivations is less plausible if a case is cited at a very high volume as compared to other cases.

To test whether the potential for informative ideological bias in the type of judge citing a

precedent might decrease with larger numbers of citing decisions, we estimate an ancillary model

in which the dependent variable is the distance between the three-year moving average of the

ideological location of the median judges for the appeals court decisions that cite the precedent

and the overall median appeals court judge at that time. The natural log of the number of appeals

court citations to the precedent during the three-year window (*Citations*) is the independent

variable in this model.[13]   The estimated model is:

|location of citing judges - location of overall median| = -.046 × ln(*Citations*) + .203 + *e*.

The negative coefficient estimate and *t*-statistic ($t = -107$) for ln(*Citations*) strongly support the

contention that increases in lower court citations correspond with a smaller potential for these

citations to reveal a pattern of ideological bias (i.e., deviation from the median appeals court

judge) in how the precedents are used.

Putting together these two countervailing influences on the informational content of

lower court citations of precedent implies that a moderate number of lower court citations to a

precedent should cause the nature of the lower court usage of precedent to be most informative to

the high court. To be clear, by "moderate" we mean somewhere in the middle of the range of

citation rates (i.e., non-extreme values of this range), which is not necessarily the same thing as

"average." There need to be enough cites during a period of time to have a meaningful signal of

the location of the precedent. But too many citations will simply lead the location of the average

citing decision to collapse upon the location of the average decision in the lower courts, which is not informative to the Supreme Court.[14]

We therefore specify our first model so that the effect of *Distance from Citing Decisions* is allowed to vary according to the number of citations that underlie the measure. This positive effect may disappear at very low and high rates of citation in the lower courts, due to the decrease in the informational value of the citations. The most straightforward way to test this hypothesis is by interacting *Distance from Citing Decisions* with a quadratic specification of the frequency with which the precedent is cited by the lower courts (*Citations*). The first model we estimate (Model 1.1) is:

Pr(*Negative Treatment of Precedent p in Term t*) $= f\{\beta_1$(*Distance from Citing Decisions*) $+$

β_2(*Distance from Citing Decisions* $\times$ *Citations*) $+ \beta_3$(*Distance from Citing Decisions* $\times$

Citations^2) $+ x\beta_c + e\}$,

where $x\beta_c$ is a set of control variables (including *Citations* and *Citations^2*) and their associated coefficients. Our expectations are that β_2 is positive, β_3 is negative, and the full, conditional effect of *Distance from Citing Decisions* is positive for at least non-extreme values of the range of *Citations*. We do not have an expectation for β_1, since on its own this coefficient only reveals the effect of *Distance from Citing Decisions* when there are, in fact, no citations to the precedent.

The key conditioning variable is *Citations*, which is measured as the number of appeals court decisions citing precedent *p* over the three years leading up to the term under analysis. As indicated above, *Citations* and *Citations^2* are then multiplied with *Distance from Citing Decisions*. *Citations* and *Citations^2* are also included separately as constituent terms (i.e., "main effects") in the model, though we have no theoretical expectations for their coefficients.

It is possible, however, that *Citations* might exert a simple, linear conditioning effect on *Distance from Citing Decisions*, meaning that increasing the number of citations underlying the signal of the precedent's location will always make this signal more informative. Moreover, the interpretation of interaction terms involving quadratic functions - such as included in Model 1.1 specification - is not necessarily intuitive and thus to the extent that that this simpler specification appropriately captures the conditional effect of *Distance from Citing Decisions* a model without the *Distance from Citing Decisions* $\times$ *Citations^2* term could be viewed as superior. Thus our second, alternative model specification (Model 1.2) is:

Pr(*Negative Treatment of Precedent p in Term t*) = $f\{\beta_1$(*Distance from Citing Decisions*) +

β_2(*Distance from Citing Decisions* $\times$ *Citations*) + $x\beta_c + e\}$,

where $x\beta_c$ is the same set of control variables and their associated coefficients.

Our expectation with this specification is that β_2 is positive and that the conditional effect of *Distance from Citing Decisions*, i.e., $\beta_1 + \beta_2$(*Citations*), is positive for much of the range of values for *Citations*. We again have no expectation for β_1 on its own since it captures the effect of *Distance from Citing Decisions* when there are zero lower court citations to the precedent. Note that this specification imposes a linear relationship on the conditioning effect of *Citations*.

Results

The results of our two logit models of the negative treatment of Supreme Court precedent are presented in Table 1. Both the statistical significance of the estimate for *Distance from Citing Decisions* $\times$ *Citations^2* and a Wald test suggest that Model 1.1 is the better fitting model of the two.[15] Thus, the following discussion largely focuses on Model 1.1, though we conclude with a brief discussion of Model 1.2.

*** Table 1 Here ***

The coefficient estimates for *Distance from Citing Decisions* × *Citations* and *Distance from Citing Decisions* × *Citations²* in Model 1.1 are fully consistent with our expectations, as there is a positive, statistically significant estimate for the former variable and a negative, statistically significant estimate for the latter.[16] Importantly, the statistical significance test for the estimate for *Distance from Citing Decisions* × *Citations²* represents a direct test of whether the conditioning effect of *Citations* is curvilinear or linear. The results of this test support the former as it allows us to reject the null hypothesis of *Citations* exerting a linear conditioning effect. The estimate for *Distance from Citing Decisions* is negative and insignificant, but we have no expectation for this parameter since it captures the effect of *Distance from Citing Decisions* when there have not been any citing decisions (i.e., *Citations* equals zero).

To best illustrate the conditional effect of *Distance from Citing Decisions* in this model (Model 1.1), Figure 2 plots the conditional coefficient for this variable over a wide range of values for the conditioning variable – *Citations*.[17] The solid line contains the point estimates for the coefficient for (i.e., effect of) *Distance from Citing Decisions*, conditional on the average number of appeals court citations to the precedent over the previous three years.[18] The shaded region depicts the 95% confidence intervals around these conditional point estimates.[19]

*** Figure 2 Here ***

As revealed by this figure, the effect of *Distance from Citing Decisions* is positive for almost all of this range of *Citations*. Moreover, the 95% confidence interval does not include zero when the average number of citations to the precedent ranges between 13 and 93, meaning that the coefficient for *Distance from Citing Decisions* is statistically significant for this range of *Citations*.[20] *Distance from Citing Decisions* has its greatest effect when there are 79 *Citations*.[21] Lower or higher values of *Citations* decrease the effect of *Distance from Citing Decisions*.[22]

17

This result supports our central hypothesis that lower court usage of Court precedent (as measured by the ideological nature of the lower court judges citing the precedent) can provide information to the Court about the location of a precedent and thus influence the Court's decision to weaken the precedent. When there are enough citations to a precedent to ensure the mean ideological position of the decisions citing the precedent is meaningful, but not so many citations that the mean location of citing decisions begins to approach the center of the distribution of all appeals court decisions, then the ideological position of the typical citing decision is informative. If this information reveals the ideological location of the precedent is at odds with the current Court then the Court will be more likely to weaken the precedent.

To be clear, our data suggest that for most precedent-term dyads the Court is not reacting to the appeals courts' use of its precedents. Most precedents receive relatively few citations in any given three-year period--the median value of *Citations* is two and the interquartile range is zero and eight--and thus the Court often does not receive a reliable or meaningful ideological signal. Nonetheless, we submit that *Distance from Citing Decisions* manifests a substantively significant relationship with the Court's negative treatment of its precedent. First, we do not expect the Court to react to lower court decisions unless there are "enough" of them to produce a reliable ideological signal. Our theory is agnostic as to how many citations are "enough" to generate an informative signal, and our data indicate that at least 13 citations in a given three-year period are required. Second, *Distance from Citing Decisions* matters for a nontrivial number of observations. It is in the direction we anticipate and statistically significant for approximately 16% of the precedent-term dyads. In addition, 45% of the precedents in our data experience a positive and significant effect during the time span analyzed.

Third, prior research indicates one of the principal reasons the Court negatively treats

precedent is the Justices' ideological distance from a precedent (Hansford and Spriggs 2006). In

our analysis, the positive coefficient on *Distance from Precedent Setting Coalition* confirms this

relationship by showing the greater the ideological disparity between the Supreme Court in Term

t and the Justices who established the precedent, the more likely the Court is to weaken the

precedent. Importantly, our analysis also demonstrates that when *Citations* ranges between 28

and 111 (which include 5.9% of the observations) *Distance from Citing Decisions* has a larger

substantive effect than *Distance from Precedent Setting Coalition*.

To further illustrate the effect size of *Distance from Citing Decisions* and compare it with

that of *Distance from Precedent Setting Coalition*, Figure 3 plots four sets of predicted

probabilities of negative treatment (predicted by Model 1.1). One set (the solid line) contains the

probabilities of negative treatment predicted when *Distance from Precedent Setting Coalition*

ranges from its minimum to slightly more than its maximum value.[23] The other three sets

contain the probabilities of negative treatment predicted when *Distance from Citing Decisions*

ranges from its minimum to slightly less than maximum value. For these three curves, *Citations*

is set at 8 (mean), 31 (one standard deviation above), and 54 (two standard deviations above).[24]

*** Figure 3 Here ***

This figure reveals that *Distance from Citing Decisions* has a larger effect size than

Distance from Precedent Setting Coalition when *Citations* is set at two standard deviations

above its mean. The substantive implication of this result is that when lower courts cite

precedent at a particularly informative rate the type of judge citing the precedent has a bigger

effect on whether the Court treats that precedent negatively than the type of Justice who

established the precedent. Prior work on the treatment of precedent at the Supreme Court (e.g.,

19

Hansford and Spriggs 2006) focuses only on the distance between the Court in Term *t* and the precedent-setting coalition. Our results indicate that under certain conditions the lower court usage of the precedent can have at least as big an effect on negative treatments by the Court.[25]

Turning now to the alternative specification of Model 1.2, the estimate for *Distance from Citing Decisions × Citations* is positive and statistically significant. Note, however, that the magnitude of this coefficient estimate is smaller than the analogous estimate in Model 1.1, which is not surprising given that Model 1.2 places a linear constraint on the non-linear conditioning effect of *Citations*. In Model 1.2, the full, conditional effect of *Distance from Citing Decisions* is $-.166 + .021 \times Citations$. The effect for of *Distance from Citing Decisions* is thus positive when there are more than eight appeals court citations and is statistically significant when there are more than 35 citations. This result generally supports our claim that lower court usage of a precedent can provide information to the Court about the effective location of the precedent, as long as there are enough citations for the ideological type of citing judge informative to the Court. As Model 1.1 reveals, however, this effect goes away at high levels of citation.

Conclusion

Studies of the delegation of policy-making authority typically focus on either determining the extent to which the agents comply with the directives of the principal or the effect of institutional design on this compliance. When applying the principal-agent framework to judicial hierarchies, scholars forward models in which the high court is concerned with correcting errant lower court decisions, which might in turn make lower courts more likely to comply with either the high court's preferences or precedent. In contrast, we suggest a new view of judicial hierarchy in which the high court observes how lower courts use a precedent and relies on information from this lower court behavior to gain a better sense of the precedent's

policy content as it is applied to current disputes. If this information reveals the precedent is at odds with the high court's preferences then the high court corrects its body of precedent by weakening the offending precedent. Our analysis of the U.S. Supreme Court's response to appeals court citations to its precedent provides support for our view of the information flow and policy making in a judicial hierarchy.

This result yields three central implications, the most general of which is that agency loss can provide a useful informational byproduct to the principal. Principal-agent models of delegation tend to focus on the obvious costs of agents not fully implementing the policies or preferences of the principal. To the limited extent that scholars note that agency loss can correspond with beneficial outcomes, it is in the sense that there may be a trade-off between agency loss and the technical expertise of the agents (e.g., Bawn 1995). Our results suggest there can also be an informational silver lining to the dark cloud of agency loss. Preference-based selection effects in the type of agent that chooses to implement a policy constitute a form of agency loss, but this type of loss can inform the principal about the current nature of the policy. The principal can then use this information to decide whether to alter the policy in question.

Second, studies of Supreme Court decision making tend to assume that the Justices have perfect information about the actual consequences of their decisions. To the extent they do not have perfect *ex ante* information about the consequences of policies they could adopt, scholars point to experienced attorneys (McGuire 1995), oral argument (Johnson, Wahlbeck, and Spriggs 2006), and amicus curiae briefs (Collins 2008) as sources of information. This paper provides evidence for another important source of information regarding the nature of existing precedents - the lower federal courts. Our findings thus suggest lower courts may play an even more significant role in the judicial system than generally has been acknowledged. Certainly, scholars

long have recognized that, particularly in the modern federal system, most lower court decisions are left undisturbed, implying these courts have the last word on the vast majority of legal disputes in this jurisdiction. Our study, though, suggests those courts can also have an unintentional informational influence on the Supreme Court and the evolution of its legal doctrine (see also Corley, Collins, and Calvin 2011).

Finally, prior work on the Supreme Court's treatment of its existing precedents shows these treatments affect how lower courts then use the precedents (Benesh and Reddick 2002; Hansford and Spriggs 2006). Our study reveals the causal arrow does not just point in this one direction. The treatment of a precedent by the Court can be influenced by how the lower courts have used the precedent. The U.S. federal judicial hierarchy thus manifests an interesting informational dynamic, with Supreme Court precedents informing lower courts how to resolve legal questions and lower court use of these precedents revealing additional information about their actual policy content or legal consequence as they are applied to new factual circumstances.

Endnotes

[2] For example, Choi and Gulati (2008) find partisan bias in appeals court citation patterns.

[3] Hansford and Spriggs (2006) find the vitality of a precedent conditions the effect of the ideological distance between a precedent and the Court on the probability of negative treatment. We assume lower court usage of Court precedent does not alter the vitality of that precedent. It simply provides information about the ideological location of the precedent.

[4] The average length of time it takes the Court to treat a precedent for the first time is 8.8 terms. There are 1,737 precedents with a negative treatment before any other Court treatment of the precedent (i.e., these are the negative treatments included for Models 1.1 and 1.2)

[5] As discussed in the On-Line Appendix (pp. 2-3), when assessing the robustness of the results obtained with these data we also re-estimate the model while including all precedent-term dyads, regardless of subsequent treatments. The inferences remain the same.

[6] This includes all the appeals court decisions from the numbered and DC circuits. We include appeals court citations to Supreme Court precedent that are coded as positive treatments, neutral

treatments, or cites, but exclude citations coded as negative treatments by *Shepard's*. We assume lower court judges are citing and positively treating precedents that are ideologically close to them, which means the ideology of the citing judges provides a reasonable proxy for the revealed location of a precedent. However, if lower court judges negatively treat cases that are ideologically distant from them, the ideology of the citing judges does not indicate the location of a precedent since these treatments can result from a precedent being either far to the right or far to the left of the median judge. Furthermore, appeals courts infrequently treat Supreme Court precedent in a negative manner. In our data, the ratio of non-negative citations to negative treatment is 22 to 1. The scarcity of these treatments thus constitutes a second difficulty in using them to construct a meaningful measure of the location of a precedent.

[7] The On-Line Appendix has additional details on data collection and the use of the JCS scores.

[8] *Shepard's* makes a distinction between citations to precedent and the explicit following of precedent. Like Clark and Lauderdale (2010), we err on the side of considering all citations informative and treat cites and "follows" as being equivalent. In part, this decision is based on *Shepard's* fairly restrictive coding rules that almost certainly undercount "follows." Nonetheless, decisions that explicitly follow a precedent could be more informative than decisions that simply cite the precedent. We conduct a robustness check in which we weight more heavily all instances in which an appeals court panel "follows" a precedent (per *Shepard's Citations*) when calculating the revealed ideological location of the precedent (i.e., the mean panel median for appeals court decisions citing/following a Court precedent). The results of this model estimation are very similar to those presented in Table 1 (On-Line Appendix, pp. 4-5).

[9] Given that that Supreme Court terms do not start until October and the great preponderance of the majority opinions are not written or published until the next calendar year, we use, for

example, lower court usage of precedent in the 1998, 1999, and 2000 calendar years to construct *Distance from Citing Decisions* for the 2000 Term. If a precedent has not been cited by an appeals court in the three years prior to term t then the location of the precedent is set at zero, which is the center of the JCS scale.

[10] Clark and Lauderdale (2010) and Carrubba et al. (2012) find that the median Justice in the precedent-setting majority has the most influence over the content of the precedent.

[11] We do not include the legal vitality of a precedent - as measured by the Supreme Court's subsequent treatment of a precedent (see Hansford and Spriggs 2006) - because we censor our data after a precedent's first treatment by the Court.

[12] We use Spaeth's (2007) "Value" variable to define general legal issue areas.

[13] Again, we do not include negative treatments by the appeals courts as citations. The mean value of *Citations* for a precedent-year dyad is 8.0 and the standard deviation is 23.

[14] We have no *ex ante* theoretical expectation regarding the location of the maximum, but we are able to estimate this location for the U.S. Supreme Court in the empirical section of the paper.

[15] The Wald statistic is 6.4 (one degree of freedom), which allows for the rejection of the null hypothesis ($p = .01$) that *Distance from Citing Decisions* $\times$ *Citations^2* fails to improve model fit.

[16] The variance of the ideological location of citing decisions is an alternative measure of the informational value of the citation-based measure of the current policy content of a precedent. The On-Line Appendix contains a version of Model 1.1 in which we also include *Distance from Citing Decisions* $\times \sigma_{Citing\ Decisions}$. The estimate for this interaction term is in the expected direction (negative) and is statistically significant. Its inclusion in the model has little effect on the substantive inferences regarding the other variables of interest.

[17] The conditional effect/coefficient is: $-.514 + .091(Citations) - .0007(Citations^2)$.

[18] We vary *Citations* from its minimum value (0) to 100, which is the relevant range for the vast majority (99.3%) of the observations in our data.

[19] See Brambor, Clark, and Golder's (2006) guide to conditional coefficients and standard errors.

[20] 15.5% of the precedent-dyads fall in this range of *Citations*. Thirteen is slightly less than one-half of a standard deviation above the mean of *Citations*, and 93 is nearly four standard deviations above the mean.

[21] There are 1,585 precedent-term dyads experiencing more than 79 *Citations*.

[22] Are precedents with very high citation rates procedural in nature and thus less inherently ideological? An examination of the 10 Supreme Court precedents in our data with the highest values of *Citations* reveals that they were established by divided Courts, though several of them could be considered primarily procedural (e.g., dealing with jury selection, right to counsel, civil procedure). Three of these precedents were decided by a five-to-four vote split while another four were decided six-to-three or five-to-three. Yet another was decided by plurality opinion. Procedural or not, these cases engendered ideological disagreement amongst the justices.

[23] *Citations* is set to 54 for this set of predictions to make them comparable to the most pronounced curve for *Distance from Citing Decisions*.

[24] All the predicted probabilities are for a five-year old, constitutional precedent established by a minimum winning coalition. All other independent variables are held at their means.

[25] The results for the other control variables reveal that constitutional precedents, broader precedents, and precedents decided by smaller vote margins are more likely to receive negative treatment. *Per curiam* precedents are less likely to be weakened by the Court. Cases with many *amicus curiae* and that dealt with legal issues active on the Court's docket are more likely to be treated negatively. As a precedent ages, it becomes less likely to receive this type of treatment.

References

Baum, Lawrence. 1980. "Response of Federal District Judges to Court of Appeals Policies: An Exploration." *Western Political Quarterly* 33(2):217-224.

Bawn, Kathleen. 1995. "Political Control versus Expertise: Congressional Choices About Administrative Procedures." *American Political Science Review* 89(March): 62–73.

Benesh, Sara C., and Malia Reddick. 2002. "Overruled: an Event History Analysis of Lower Court Reaction to Supreme Court Alteration of Precedent." *Journal of Politics* 64(2):534-550.

Black, Ryan C., and James F. Spriggs II. 2013. "The Citation and Depreciation of U.S. Supreme Court Precedent." *Journal of Empirical Legal Studies*, Forthcoming.

Brambor, Thomas, William Roberts Clark, and Matt Golder. 2006. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 14(1):63-82.

Carpenter, Daniel. 1996. "Adaptive Signal Processing, Hierarchy, and Budgetary Control in Federal Regulation." *American Political Science Review* 90(2):283-302.

Carrubba, Clifford J, Barry Friedman, Andrew D. Martin, and Georg Vanberg. 2012. "Who Controls the Content of Supreme Court Opinions." *American Journal of Political Science* 56(2):400-412.

Choi, Stephen, and Mitu Gulati. 2008. "Bias in Judicial Citations: A Window Into the Behavior of Judges." *Journal of Legal Studies* 37:87-129.

Clark, Tom S., and Benjamin Lauderdale. 2010. "Locating Supreme Court Opinions in Doctrine Space." *American Journal of Political Science* 54(4):871-890.

Collins, Paul M., Jr. 2008. *Friends of the Supreme Court: Interest Groups and Judicial Decision Making*. New York, NY: Oxford University Press.

Corley, Pamela C., Paul M. Collins, Jr., and Bryan Calvin. 2011. "Lower Court Influence on

U.S. Supreme Court Opinion Content." *Journal of Politics* 73(1):31-44.

Cross, Frank B., James F. Spriggs, II, Timothy R. Johnson, and Paul J. Wahlbeck. 2010.

"Citations in the U.S. Supreme Court: An Empirical Study of Their Use and

Significance." *University of Illinois Law Review* 2010(2):489-575.

Epstein, Lee, and Jack Knight. 1998. *The Choices Justices Make*. Washington, DC: CQ Press.

Epstein, Lee, Andrew D. Martin, Jeffrey A. Segal, and Chad Westerland. 2007. "The Judicial

Common Space." *Journal of Law, Economics, & Organization* 23:303-325.

Fowler, James H., James F. Spriggs II, Timothy Johnson, and Paul J. Wahlbeck. 2007.

"Network Analysis and the Law: Measuring the Legal Importance of Precedents at the

U.S. Supreme Court." *Political Analysis* 5(3):324-46.

Giles, Micheal W., Virginia Hettinger, and Todd Peppers. 2001. "Picking Federal Judges: A

Note on Policy and Partisan Selection Agendas." *Political Research Quarterly*

54(September):623-641.

Goldman, Sheldon. 1975. "Voting Behavior on the United States Courts of Appeals, Revisited."

American Political Science Review 69:491-506.

Haire, Susan B., Donald R. Songer, and Stefanie Lindquist. 2003. "Appellate Court Supervision in

the Federal Judiciary: A Hierarchical Perspective." *Law & Society Review* 37(1):143-168.

Hansford, Thomas G., and James F. Spriggs, II. 2006. *The Politics of Precedent on the U.S.

Supreme Court*. Princeton, NJ: Princeton University Press.

Johnson, Timothy R., Paul J. Wahlbeck, and James F. Spriggs, II. 2006. "The Influence of Oral

Arguments on the U.S. Supreme Court." *American Political Science Review* 100:99-113.

Klein, David E. 2002. *Making Law in the United States Courts of Appeals*. New York: Cambridge University Press.

Luse, Jennifer K., Geoffrey McGovern, Wendy L. Martinek, and Sara C. Benesh. 2009. "Such Inferior Courts... Compliance by Circuits with Jurisprudential Regimes." *American Politics Research* 37(1):75-106.

McGuire, Kevin T. 1995. "Repeat Players in the Supreme Court: The Role of Experienced Lawyers in Litigation Success." *Journal of Politics* 57:187-196.

Owens, Ryan C., and Ryan J. Black. N.d. "Consider the Source (and the Message): Supreme Court Justices and Strategic Audits of Lower Court Decisions." *Political Research Quarterly*, Forthcoming.

Randazzo, Kirk. A. 2008. "Strategic Anticipation and the Hierarchy of Justice in U.S. District Courts." *American Politics Research* 36(5):669-93.

Shipan, Charles R. 2004. "Regulatory Regimes, Agency Actions, and the Conditional Nature of Congressional Influence." *American Political Science Review* 98(August):467–480.

Songer, Donald R., Jeffrey A. Segal, and Charles M. Cameron. 1994. "The Hierarchy of Justice: Testing a Principal-Agent Model of Supreme Court-Circuit Court Interactions." *American Journal of Political Science* 38(August):673-696.

Spaeth, Harold J. The Original United States Supreme Court Judicial Database, 1953-2006 Terms (Michigan State University 2007).

Westerland, Chad, Jeffrey A. Segal, Lee Epstein, Charles M. Cameron, and Scott Comparato. 2010. "Strategic Defiance and Compliance in the U.S. Courts of Appeals." *American Journal of Political Science* 54(4):891-905.

Biographical Statement

Thomas G. Hansford is an Associate Professor of Political Science at UC Merced, Merced, CA 95343.

James F. Spriggs, II is the Sidney W. Souers Professor of Government at Washington University in St. Louis, St. Louis, MO 63130.

Anthony A. Stenger is a Ph.D. Candidate in Political Science at Washington University in St. Louis, St. Louis, MO 63130.
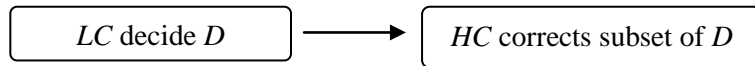
Table 1. Logit model of the negative treatment of precedent at the Supreme Court

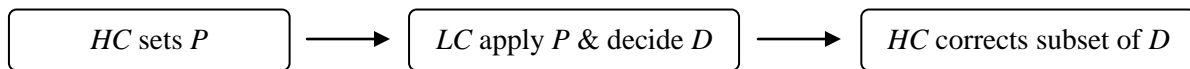| Independent Variable	Model 1.1	Model 1.2
Distance from Citing Decisions	-.514	-.166
	(.269)	(.225)
Distance from Citing Decisions × Citations	.091*	.021*
	(.027)	(.010)
Distance from Citing Decisions × Citations^2	-.0007*	---
	(.0003)	
Citations	.023^	.030^
	(.005)	(.004)
Citations^2	-.0001	-.0001^
	(.0000)	(.0000)
Distance from Precedent Setting Coalition	1.44^	1.45^
	(.168)	(.169)
Constitutional Precedent	.615^	.618^
	(.053)	(.053)
Precedent Breadth	.129^	.130^
	(.028)	(.028)
Precedent Vote Margin	-.020^	-.019^
	(.009)	(.009)
Per Curiam Precedent	-.755^	-.762^
	(.138)	(.138)
Precedent Amici	.188^	.187^
	(.019)	(.019)
Agenda Relevance	.013^	.014^
	(.002)	(.003)
ln(Precedent Age)	-.651^	-.651^
	(.023)	(.023)
N	116,629	116,629
Wald Test Statistic (19/18 d.f.)	2,076*	2,049*

Note: Entries are logit coefficient estimates (and robust standard errors). Fixed effects for the decade the precedent was decided are also included in these models. *$p \leq .05$ (one-tailed, for directional hypotheses). ^$p \leq .05$ (two-tailed, for control variables).

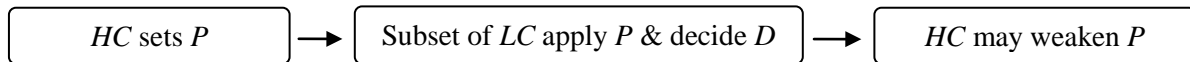**Figure 1. Decision sequences in models of judicial hierarchy**

A. Hierarchy without Stare Decisis

| *LC* decide *D* | ⟶ | *HC* corrects subset of *D* |

B. Vertical Stare Decisis with Decision Correction

| *HC* sets *P* | ⟶ | *LC* apply *P* & decide *D* | ⟶ | *HC* corrects subset of *D* |

C. Vertical Stare Decisis with Precedent Correction

| *HC* sets *P* | ⟶ | Subset of *LC* apply *P* & decide *D* | ⟶ | *HC* may weaken *P* |

Note: *HC* is a high court, *P* is a precedent, *LC* is the set of lower courts, and *D* is the set of lower court decisions.

Figure 2. The coefficient for *Distance from Citing Decisions*, conditioned by *Citations*
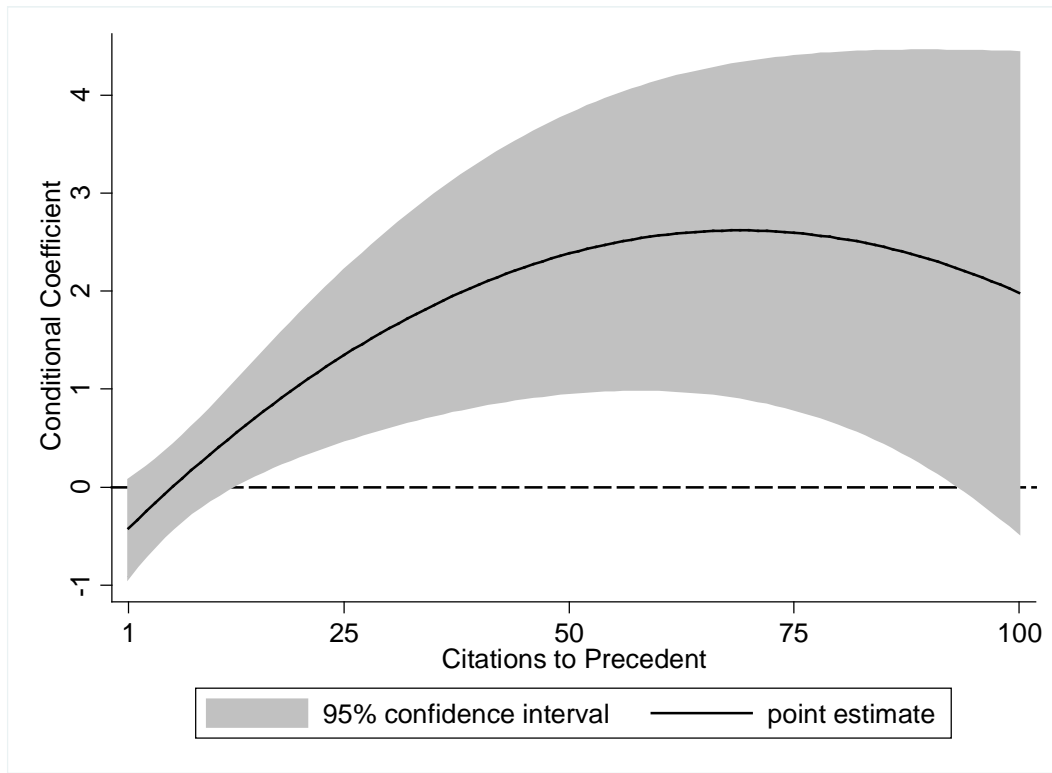
Note: Model 1.1's coefficient estimate for *Distance from Citing Decisions* (which measures the revealed ideological location of a precedent) is plotted on the *y*-axis and *Citations* (which is a three-year moving average of the number of Courts of Appeals cases citing a precedent) is plotted on the *x*-axis.

Figure 3. Predicted probabilities of negative treatment
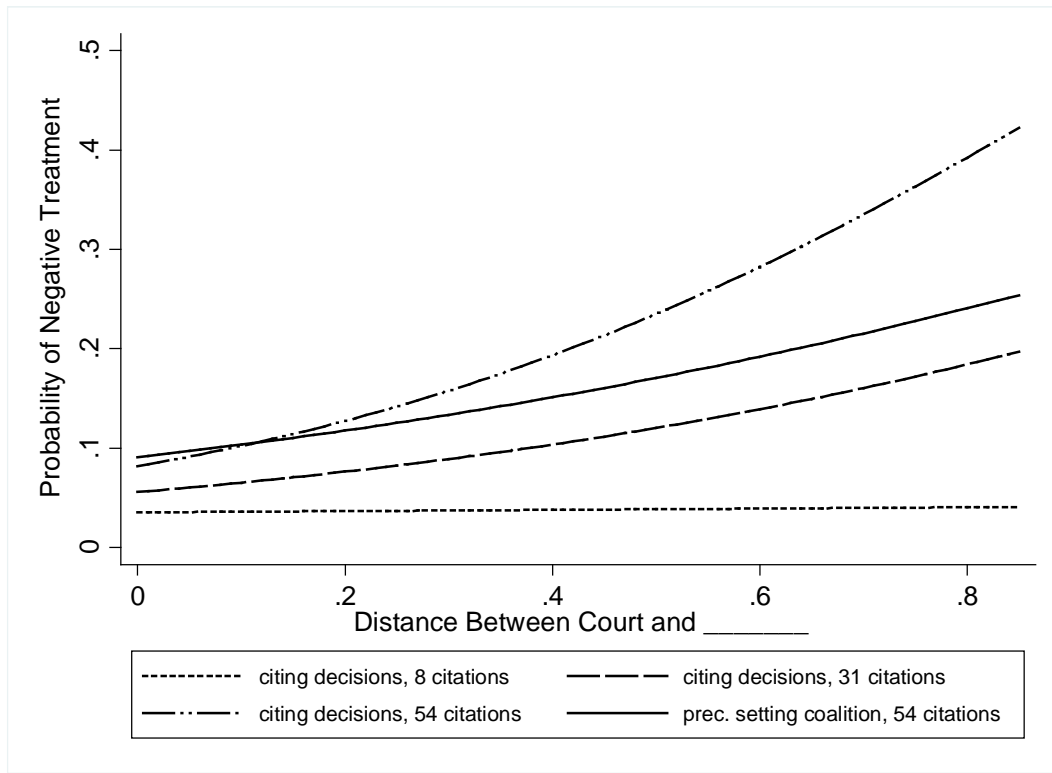
Note: Predicted probabilities of negative treatment are generated by Model 1.1. Probabilities are for a five-year old constitutional precedent established by a minimum winning coalition. For the three dashed lines, *Distance from Citing Decisions* is plotted on the *x*-axis and varies from its minimum to near-maximum value found in the data. For the solid line, *Distance from Precedent Setting Coalition* is plotted on the *x*-axis and varies from its minimum to a value slightly beyond the maximum found in the data. *Citations*, the conditioning variable, varies from its mean (8) to one (31) and two (54) standard deviations above its mean, as indicated in the legend. All other independent variables are held at their means (or modes for binary variables).