

The Path of Law: Legal Uncertainty and Issues of First Impression in the U.S. Courts of Appeals

ANTHONY R. TABONI *University of Texas at Austin, United States*

When deciding new issues, judges face uncertainty about how cases map into their existing understanding of the law. This uncertainty can lead to conflicting decisions on the same legal question, generating inconsistent law. I develop a formal theory of judicial decision making, where courts learn about and rule on new legal issues. I find that courts learn most from their ideological allies; however, increasing the ideological distance between courts can either increase or decrease legal uniformity. Using an original dataset of cases of first impression in the U.S. Courts of Appeals, I find that increasing the ideological distance between two courts increases the probability of disagreement if the previous court's decision is in-line with their relative bias, and decreases disagreement when the decision runs counter to their relative bias. My findings highlight the ways that courts can use decisions from even ideologically distant peers to learn about new legal issues.

INTRODUCTION

Courts play a major role in formulating policy in the United States. While the Supreme Court has formal final authority, the vast majority of issues are resolved by lower courts. The U.S. Courts of Appeals, the federal intermediate appellate courts, occupy a central role in the development of law. As the final arbiter of over 99% of federal appeals, these courts are the ultimate decision-makers for many questions of law—deciding issues ranging from criminal procedure to artistic expression to environmental regulation.¹ The Courts of Appeals are divided into 12 geographic regions called circuits. Circuits function as independent entities and can reach decisions in direct conflict with one another. These conflicting decisions, known as circuit splits, cause federal law and, consequently, national policy to vary with geography.

Understanding the causes and consequences of splits is fundamental to “the path of the law” (Holmes 1997). Splits occur regularly and, though they are more likely to be reviewed by the U.S. Supreme Court, are frequently left unresolved (Beim and Rader 2019; Ulmer 1984). The consequences of splits have been well studied; however, less is known about *why* splits occur. While decisions from other circuits are not binding, they can be persuasive (Klein 2002). Decisions by other

circuits often serve as a foundation for courts dealing with new legal issues (Hinkle 2015).

In this article, I propose a new theory of law creation and analyze both the prevalence of and the factors that contributed toward circuit splits. I develop a model of judicial decision making under uncertainty. In the model, two courts with heterogeneous preferences sequentially decide a novel legal issue. The courts face uncertainty about what constitutes a “correct” decision for the case at hand. Before deciding the case, each court receives an imperfect, private signal. The second court observes the decision made by the first court, but not the first court's signal. Using this framework, I examine the effect of the ideological distance between the courts on the probability that a circuit split occurs. I break down the effect of ideological distance into two components: learning from previous decisions and learning from private information.

The first main insight of the model is that as the ideological distance between the two courts increases, the second court is influenced less by the first court's decision. Increasing the ideological distance between the courts increases the set of cases for which the courts disagree about what constitutes the correct decision. Consequently, the decision of the first court is less informative for the second. This increases the probability that a circuit split will occur.

The second theoretic insight is that ideological distance has heterogeneous effects on the probability of a circuit split. The direction of this effect depends on the combination of the ideology of the first court, relative to the second, and the first court's decision. There are two distinct types of preference arrangements. A decision is *bias-compatible* if the first court is more liberal than the second and chose the liberal disposition, or if the first court is more conservative and chose the conservative disposition. In contrast, a decision is *bias-incompatible* if the first court is more liberal than the second and chose the conservative disposition, or if it is

Anthony R. Taboni , Postdoctoral Fellow, Department of Government, University of Texas at Austin, United States, anthony.taboni@austin.utexas.edu.

Handling editor: Monika Nalepa.

Received: July 05, 2024; revised: December 19, 2024; accepted: August 20, 2025.

¹ In a given year, the Supreme Court grants full review to under 80 cases. In the same period, the U.S. Courts of Appeals decide approximately 40,000 cases (U.S. Courts 2023).

more conservative and chose the liberal disposition. The model provides a clear empirical prediction for bias-compatible cases—as the first court becomes more distant from the second, the probability of a circuit split increases. However, for bias-incompatible decisions, increased ideological distance can increase or decrease the probability of a circuit split.

To disentangle the mechanisms that lead to disagreement, I extend the model to multiple courts and examine the relationship between when in the sequence of decisions a court acts and two outcomes: the probability of subsequent agreement and the probability that the court creates a circuit split. If courts do not learn from each other's decisions, these probabilities are constant regardless of when in the development of an issue a court makes a decision.

I use the model to inform and guide an empirical analysis of circuit splits in the U.S. Courts of Appeals. This empirical analysis requires cases where courts are both unconstrained by precedent and faced with uncertainty over their preferred disposition. For the majority of cases heard by the Courts of Appeals, neither of these factors apply. Circuit courts are constrained by both the decisions of the Supreme Court and the decisions of previous panels within their circuit. The Courts of Appeals sit in a hierarchy, supervised by the Supreme Court. If the Supreme Court has already decided an issue, the circuits are formally obligated to follow (Haire, Lindquist, and Songer 2003; Songer, Segal, and Cameron 1994). Similarly, judges in the Courts of Appeals must treat decisions made by previous panels as binding precedent (Cross 2003; Klein 2002).

Accordingly, I collect an original dataset of cases of *first impression*. When a circuit hears a case of first impression, the case presents “a legal issue that has never been decided by the governing jurisdiction” (Legal Information Institute 2025). These cases, in particular, provide an opportunity to study decision making in an uncertain environment and understand what courts can learn from one another. My original data consist of over 1,300 decisions on issues active between 1993 and 2021. This dataset constitutes, to the best of my knowledge, the first comprehensive collection of these cases.

Circuit splits are relatively common and occur in 35% of issues of first impression in my data. I find that, for bias-compatible cases, increased ideological distance leads to an increased probability of a circuit split. For bias-incompatible cases, I find the opposite effect. Increasing the ideological distance between courts instead reduces the probability of a circuit split. To gain insight into the mechanisms behind these effects, I test predictions about the effect of sequence order, and I find that the empirical patterns are inconsistent with a model where courts *do not* learn from the actions of their peers.

My findings build upon an existing literature that examines the importance of information in judicial decision making (see Iaryczower and Shum 2012). I find that decisions are not solely determined by judge's preexisting beliefs. Instead, courts

incorporate both public and private information about cases when deciding new legal issues. My analysis provides both a theoretical and empirical foundation for understanding what courts learn from one another and what factors can lead them to come to different conclusions.

INTERACTIONS AMONG COURTS

Circuit court decisions are consequential for the development of law. How are these decisions made? Previous work has focused primarily how an upper court learns from lower court decisions as opposed to lower courts learning from one another (see Beim 2017; Clark and Kastellec 2013; but see Cameron 1993; Strayhorn 2020). Studies that examine sequential decision making have focused on situations where courts either *ex ante* know which decision they prefer (Strayhorn 2020) or share the same preferences (Talley 1999).²

Despite the formal independence that they have from one another, courts look to the decisions of other circuits for guidance (Klein 2002). Except for cases with regionally specific or *sui generis* case facts, most issues of law that arise in one circuit will be heard by others as well.³ If another circuit has previously considered a legal issue, current judges can gain information from the previous court's decision. Evidence suggests that judges regularly look to what their colleagues in other circuits have done when deciding a case (Klein 2002). Despite having no legally binding value, judges will use the decisions of other circuits as persuasive precedent when explaining their rationale in a case (Hinkle 2015).

These decisions can influence the Supreme Court. When the circuits come to conflicting decisions, an issue is more likely to be reviewed by the Supreme Court (Beim and Rader 2019; Black and Owens 2009; Ulmer 1984). Judges are aware of the increased scrutiny surrounding circuit splits and may alter their behavior in anticipation of review (Barnes Bowie and Songer 2009; Klein 2002). When the Supreme Court grants review, the actions of lower courts can influence the decision ultimately taken by it. When writing an opinion, the Supreme Court often borrows language from the lower court decision it reviewed (Corley, Collins, and Calvin 2011). Even if a court is not directly reviewed, their decisions can have an impact. The proportion of circuits on each side of a split is correlated with the position the Supreme Court ultimately takes in a case (Lindquist and Klein 2006).

In economics, a large body of literature investigates how actors can learn from one another (see Banerjee

² I use courts and judges interchangeably throughout the article. In the U.S. Courts of Appeals, cases are heard by three-judge panels with dispositions chosen by a majority vote. In the model, I abstract away from judges making decisions and instead focus on courts. A court could be a single judge or a multimember panel.

³ An exception is cases considered by the U.S. Federal Circuit. Unlike the other 12 circuit courts whose jurisdiction is defined by geography, the Federal Circuit handles appeals involving patents or contract claims against the U.S. government (Re 2001).

1992; Chamley 2004; Ellison and Fudenberg 1993). What actors can learn from one another is dependent on whether signals are public and observed by everyone or private and only observed by one individual. When signals are public, information grows more and more accurate over time. When signals are private, however, an informational cascade can occur—individuals' actions no longer reflect their private information and no new information about the state of the world is generated (Bikhchandani, Hirshleifer, and Welch 1992). These informational cascades can persist even when individuals have heterogeneous preferences (see Bala and Goyal 2001; Smith and Sørensen 2000).

Observational work on social learning has been limited in its ability to classify decisions as the result of learning from the actions of other actors. As Volden, Ting, and Carpenter (2008) note, without accounting for the effect of private information on beliefs, one is unable to determine if uniformity is a consequence of learning from the actions of others as opposed to actors independently making the same decision. To address this critique, recent work on policy diffusion in courts has focused on the reasoning courts use for their decisions as opposed to the choice of disposition itself (see Hinkle and Nelson 2016; Matthews 2024; but see Klein 2002). An opinion, however, is downstream to the disposition in a case (Parameswaran, Cameron, and Kornhauser 2021).

Uncertainty in Judicial Decision Making

Much of the focus on judicial uncertainty in existing theoretical work has been on the uncertainty faced by an upper court when deciding whether to review a lower court's decision (see Beim, Hirsch, and Kastellec 2014; Cameron, Segal, and Songer 2000; Spitzer and Talley 2000). These approaches center case facts as the source of uncertainty for an upper court. An inability to directly observe the facts of a case translates into uncertainty over whether a lower court decision is compliant.

This uncertainty, however, differs from what courts face when deciding new legal issues. In these cases, courts observe the facts of the case, but face uncertainty into how these facts translate into relevant outcomes. To better understand judicial decision making under uncertainty, I conducted a handful of exploratory interviews with former clerks from the U.S. Circuit Courts of Appeals.⁴ In these semistructured interviews, I asked the clerks about the process of deciding cases involving new legal issues. While uncommon relative to the overall caseload of a court, cases sometimes arose where no

precedent applied. The clerks noted that uncertainty would often arise following some sort of shock to an existing legal rule. This shock could take the form of a new law, a recent Supreme Court case, or facts that required a new understanding of old law. The uncertainty faced by the clerks in these cases was not about what the facts of the case were but instead about how these cases related to existing preferences in light of the shock.

The clerks noted that when these circumstances occurred, they were not completely in the dark about what to do. Instead, they would look to what other courts had written as guideposts to understand the new legal landscape. This did not mean that the clerks' judges always followed their peers; instead, they looked to the actions of other circuits, updated their beliefs, and then took the action they thought was best.

MODEL

How do courts deal with uncertainty, and when should we expect them to come to different outcomes? To model this process, I utilize the "case space" framework of judicial decision making (Cameron and Kornhauser 2017a; Kornhauser 1992; Lax 2011). This approach links the actions taken by courts to dispute resolution. In contrast to spatial models of policymaking, a court does not directly choose a legal rule. Instead, for a given case, they choose a disposition consistent with their preferred legal rule. For many issues, courts are not choosing from a broad set of legal standards. Courts are faced with a choice to either rule for or against a litigant on a given issue. Most circuit splits only involve two sides of an issue (Beim and Rader 2019). The dispositional framework captures this decision problem.

The setup of the model is as follows:

Players. There are two courts, C_1 and C_2 .

Timing. The model has two periods. First, nature draws observable case facts, $\hat{x} \sim U[0, 1]$, and an unobserved shock, ω , drawn according to a known distribution $G(\omega) = U[-\Omega, \Omega]$. Larger values of Ω correspond to higher levels of uncertainty. In the first period, C_1 observes \hat{x} and receives a signal, s_1 , that provides them with information about ω .

If $\hat{x} + \omega \geq y_i$

$$s_i(y_i, \hat{x}, \omega, p) = \begin{cases} \widehat{lib} & \text{with probability } p \\ \widehat{con} & \text{with probability } 1-p. \end{cases}$$

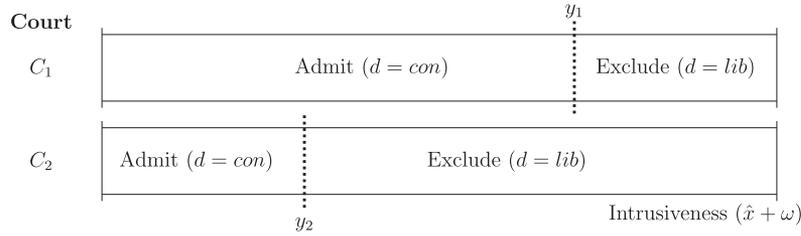
Whereas, if $\hat{x} + \omega < y_i$,

$$s_i(y_i, \hat{x}, \omega, p) = \begin{cases} \widehat{lib} & \text{with probability } 1-p \\ \widehat{con} & \text{with probability } p \end{cases}$$

with $p \in (\frac{1}{2}, 1)$ observed by both courts.⁵ The court then chooses one of two dispositions, $d_1 \in \{\widehat{lib}, \widehat{con}\}$. In the second period, C_2 observes \hat{x} (the same case facts as the first court) and d_1 but not s_1 . The second court

⁴ I interviewed five former clerks who served at some point between 2012 and 2022. These clerks were identified via personal connections of colleagues. Three of the clerks' judges were appointed by Democratic presidents while the other two clerks' judges were appointed by Republican presidents. None of the judges was in senior status when the interviewees clerked. Interviewees were granted anonymity for participating in the study. They were provided with consent forms prior to the interview, and this part of the study was approved by Princeton University's Institutional Review Board (No. 14504).

⁵ Note that $\hat{\cdot}$ denotes a signal as opposed to a disposition.

FIGURE 1. Preferences over Dispositions


Note: The top panel displays the preferences for a conservative court, with a cut point further to the right. The intrusiveness of a search must be relatively high before they are willing to exclude the evidence. In contrast, C_2 , who has a cut point further to the left, is more liberal as they will exclude evidence from searches with low levels of intrusiveness.

receives their own signal, s_2 . This signal is also correct with probability $p \in (\frac{1}{2}, 1)$. The second court then chooses $d_2 \in \{lib, con\}$.

Preferences. I assume that a common shock affects case facts such that utility depends on the sum of observed case facts \hat{x} and an unobserved shock ω .⁶ For court C_i , preferences over dispositions can be characterized by a cut point y_i . If $y_i \leq \hat{x} + \omega$, the court prefers $d = lib$; otherwise, they prefer $d = con$. I assume that $y_i - \Omega > 0$ and that $y_i + \Omega < 1$. There is no court so extreme that a shock would push its cut point out of the observable fact range. Let $r(y_i, \hat{x}, \omega)$ define a court's preferences over dispositions such that:

$$r(y_i, \hat{x}, \omega) = \begin{cases} lib & \text{if } y_i \leq \hat{x} + \omega \\ con & \text{if } y_i > \hat{x} + \omega. \end{cases}$$

I assume that courts suffer a constant loss for choosing a disposition inconsistent with their preferences. Following Cameron, Segal, and Songer (2000), I utilize a constant loss utility function. This utility form does not satisfy the increasing difference in dispositions criteria of Cameron and Kornhauser (2017b). Using a utility function that satisfies this criterion would change the set of case facts that are informative for the second court; however, conditional on the decision being informative, what the second court learns from the first's decision is equivalent as how the second court updates its beliefs does not depend on the form of the utility function.

$$u(d; y_i, \hat{x}, \omega) = \begin{cases} 0 & \text{if } d = r(y_i, \hat{x}, \omega) \\ -\ell & \text{if } d_i \neq r(y_i, \hat{x}, \omega), \end{cases}$$

where $\ell > 0$.⁷

Information. Both courts have prior beliefs that $\omega \sim U[-\Omega, \Omega]$. Cut points for all courts are common knowledge. C_2 does not observe s_1 but does observe d_1 ; that is,

they observe C_1 's disposition but not the signal they received. Conditional on \hat{x} and ω , s_1 and s_2 are independent.

Strategies and equilibrium concept. For the first court, a pure strategy is a mapping from the set of observable case facts and signals to a disposition: $\hat{X} \times \{lib, con\} \rightarrow \{lib, con\}$. For the second court, a pure strategy is a mapping from the set of observable case facts, signals, and possible first court dispositions to a disposition: $\hat{X} \times \{lib, con\} \times \{lib, con\} \rightarrow \{lib, con\}$. I assume that courts update beliefs via Bayes' Rule and take actions that maximize their utility given these beliefs.

Model Illustration

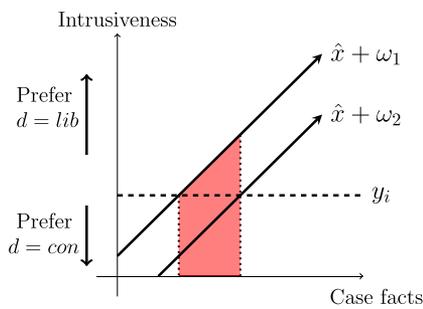
As an example, consider a court reviewing the constitutionality of a search. The court observes the facts of the case—what was searched, the scope of the search, how the search was conducted—and must decide whether to admit ($d = con$) or exclude ($d = lib$) the evidence. The court's preferences over dispositions can be characterized in terms of how intrusive a search was. Figure 1 visualizes these preferences for two different courts. The courts' cut points divide levels of intrusiveness into two regions: levels of intrusiveness that justify exclusion and those that do not. C_2 is more liberal than C_1 . As their cut point is further to the left, a lower level of intrusiveness is required before they prefer to exclude evidence from a search; whereas, C_1 , the more conservative court, requires a higher level of intrusiveness before they would exclude the evidence from a search.

Uncertainty manifests in the mapping from case facts to intrusiveness. Consider the use of a new technology, such as a thermal camera, to conduct a search of a house.⁸ Whether the court wants to exclude the evidence depends on how intrusive the search was; however, the court does not directly observe the intrusiveness of the search. Instead, the court sees that a thermal camera was used to detect the amount of heat emanating from the house. How these observable case facts translate into intrusiveness may be unclear. The

⁶ This characterization of uncertainty is similar to that of Cameron, Segal, and Songer (2000), Gilligan and Krehbiel (1987), and Epstein and O'Halloran (1994)—where outcomes depend on both an observed and an unobserved component. I assume that the information courts have about the unobserved component is imperfect.

⁷ If $\ell = 0$, the court is always indifferent between dispositions.

⁸ See *Kyllo v. United States*, 533 U.S. 27 (2001).

FIGURE 2. Case Space Conceptualization of Judicial Uncertainty

Note: In the shaded region, the court's preferred disposition differs depending on the realization of ω .

thermal camera can “see” through walls, which may increase the intrusiveness, but what it is seeing is heat as opposed to images, which may decrease the intrusiveness.

If the court faces uncertainty over how intrusive the search is, this can lead them to be unsure of which disposition is correct for the case. Figure 2 captures this understanding of uncertainty. The same observed case facts result in two different levels of intrusiveness depending on the size of the shock. For case facts in the shaded region, the court's preferred disposition depends on whether the shock is ω_1 or ω_2 . Courts then have an interest in learning about ω in order to choose the “correct” disposition.

Model Discussion

Before proceeding to the analysis, three features of the model warrant discussion. First, the distribution that generates a court's signal depends on the preferences of the court. When deciding new legal issues, the information courts have available comes primarily from litigants. Litigants are not concerned with providing judges with information for the sake of learning about the law but instead are trying to convince a judge (or judges) why they should rule for the litigant's side. Litigants will target information at judges whose votes they seek to secure (Hazelton and Hinkle 2022). Consequently, the information available to judges is often specific to what that judge should do for the case at hand and not about the world more generally. By conditioning the distribution of signals on the preferences of the court hearing the case, I account for this strategic incentive of litigants.

Second, in line with previous work on social learning, I assume that actors receive private signals before making their decision (see Bikhchandani, Hirshleifer, and Welch 1992). When courts decide cases, their decisions are often accompanied by an opinion outlining the reasons for their decision. This opinion may provide information about the court's signal. Talley (1999) examines this setting under the additional assumption that courts have the same preferences. Opinions, however, are written after the court has decided which disposition to choose.

The opinion that can be written is conditional on which decision the court chooses (Parameswaran, Cameron, and Kornhauser 2021). Consequently, a court's opinion may not always be a perfect accounting of its private information.

Finally, I assume that a court's utility is independent of the actions of other actors. Under the model, courts have no inherent preference for uniformity and face no cost for being part of a circuit split. This assumption transforms the court's choice into a decision-theoretic one. Importantly, the courts' decisions are not completely independent. The second court can still be influenced by the actions of the first court; however, each court's utility depends only on their own actions.

Canonical models of informational cascades also assume that actors' payoffs are independent of each other's actions (see Bikhchandani, Hirshleifer, and Welch 1992). Under my framework, this assumption allows me to easily extend the model and look at the behavior of multiple courts as opposed to just two. Additionally, while courts observe each other's actions, they are not obligated to follow one another. Any cost of a circuit split is informal.⁹ While some judges may internalize a cost to generating a split, this is not universal (Klein 2002). Finally, in the Courts of Appeals, courts face uncertainty over whether a future circuit will even decide the same question and, if so, what the preferences of this court might be. Utilizing a decision-theoretic framework allows me to abstract away from this additional uncertainty and instead focus on what information courts can learn from each other's actions.¹⁰

ANALYSIS

First Court

I begin by analyzing the behavior of the first court. I assume that courts are unconcerned with how their decisions will be perceived by future courts. Consequently, the first court will simply choose the disposition that has a higher probability of being correct.¹¹ The

⁹ The cost of disagreement varies with institutional context. In matters of state law, state Supreme Courts face no explicit costs for inconsistency (Caldeira 1985). By contrast, disagreement between courts and legislatures has the potential to be quite costly (Staton and Vanberg 2008; Vanberg 2001). The extent to which inconsistency influences the behavior of courts depends on the consequences of disagreement.

¹⁰ In Appendix B in the Supplementary Material, I consider a model, where courts face a cost if a circuit split occurs. Adding a cost to generating a circuit split changes behavior in two ways. First, the second court becomes more likely to choose the same decision as the first court. Second, the first court becomes more likely to make the decision it expects the second court to make. If the second court is likely to choose the liberal decision—for example, this makes the first court more likely to as well. Which cases are informative changes under this framework; however, results about what the second court can learn from an informative decision are unchanged.

¹¹ A derivation of these (and all) beliefs can be found in Appendix A in the Supplementary Material.

first court has two sources of information available about the state of the world: their signal and their prior beliefs over ω . While the signal is influential on the court, it is not deterministic. The shock to case facts is limited in value. If the observed case facts are extreme relative to the court's cut point, no shock is large enough to change which disposition they prefer ex ante. For $\hat{x} \in (0, y_i - \Omega) \cup (y_i + \Omega, 1)$, a court's choice is the same irrespective of signal.

Even if a shock is theoretically large enough to change which disposition a court prefers, they will not necessarily follow their signal. Signals are sometimes wrong. The accuracy of the signal affects its influence on a court's decision. If $\hat{x} < y_1 - \Omega(2p-1)$, the case facts are so conservative that the court "ignores" the signal and always chooses $d = con$. If $\hat{x} \geq y_1 + \Omega(2p-1)$, the case facts are so liberal that they will always choose $d = lib$.

Whether the signal will affect which disposition is chosen depends on the observable case facts, \hat{x} . Figure 3 divides the case-space into three regions. For observed case facts in regions *I* and *III*, which disposition the court chooses does not depend on their signal. Only in region *II*, does their disposition depend on the signal. If case facts are in this region, then the court will choose the disposition that matches their signal. The size of region *II* depends on p . When signals are very accurate ($p \sim 1$), courts will follow their signal for even relatively extreme case facts. In contrast, for inaccurate signals ($p \sim \frac{1}{2}$), the court's choice of disposition almost never depends on their signal.

Second Court

Whether the second court learns anything from the first court's actions depends on the relationship between y_1 and \hat{x} . The second court does not directly observe the first's signal. Instead, the second court must rely on both the observable case facts and the first court's disposition to update their beliefs about ω . When case facts are extreme relative to the first court's cut point, a disposition provides no information about ω because the first court chooses the same disposition regardless of signal. As the first court's signal is private, the second court learns nothing about ω . In contrast, if the case facts are more moderate (relative to the court), the first court's disposition is identical to their signal. As

dispositions are public, the second court then learns the first court's signal.

From the perspective of a second court, there are two distinct types of decisions which are as follows.

Definition 1 (Informative). A decision is *informative* when the first court's disposition depends on their signal.

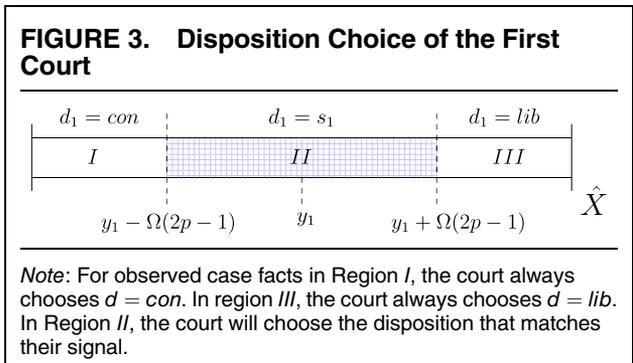
Definition 2 (Uninformative). A decision is *uninformative* when the first court's disposition does not depend on their signal.

As the second court observes \hat{x} and y_1 , they can always derive whether the first court's decision is informative or uninformative. If the first court's decision was uninformative, the second court proceeds as if they were the first decision maker.¹² They observe case facts \hat{x} and a signal s_2 about their preferred disposition for the case. The second court then chooses d_2 . Instead, if the first court's decision was informative, they learn that $s_1 = d_1$.

Preference heterogeneity leads to differences in the influence of the first court's decision on the second. The first court's signal is dependent on their preferences. If the first and second courts have different preferences, this affects what the second court can learn from the first's actions. As an example, consider $y_2 < y_1$ and $d_1 = lib$, that is, the first court is more conservative than the second and chose the liberal disposition. The top panel of Figure 4 displays this preference arrangement. Assume this decision is informative. As $d_1 = lib$, the second court learns that $s_1 = \widehat{lib}$. A signal of $s_1 = \widehat{lib}$ means that, with probability p , $y_1 \leq x$. As the first court is more conservative than the second ($y_2 < y_1$), this implies that $y_2 < y_1 \leq x$. The second court's cut point is also to the left of the case facts.

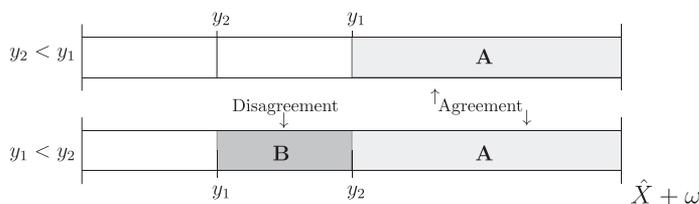
In any case, where a more conservative court prefers the liberal disposition, the second court does as well. This concept is analogous to the "Nixon goes to China" principle discussed in the judicial auditing literature. In models of judicial auditing, an upper court does not review a liberal disposition chosen by a more conservative lower court, as they would always agree with the decision (Cameron, Segal, and Songer 2000). In my model, when a conservative court chooses the liberal disposition, future liberal courts can infer that they would have made the same choice.

When the first court is more liberal than the second ($y_1 < y_2$), as in the bottom panel of Figure 4, this relationship changes. Again assume that $d_1 = lib$. Unlike before, the first court prefers $d = lib$ for a larger set of case facts than the second court. While the two courts might agree on many cases, there is an interval of



¹² If signals were public as in Talley (1999), we no longer need the distinction between informative and uninformative cases. As the second court can directly observe the first court's signal, they no longer need to rely on the first court's disposition to learn about s_1 . Though public, these signals still depend on the preferences of the court deciding the case. Dynamics about what the second court can learn from the first are then equivalent to that of informative cases.

FIGURE 4. Potential Location of $\hat{x} + \omega$ Based on a Signal of $s_1 = \widehat{lib}$



Note: In A, the two courts prefer the same disposition. In B, the courts prefer different dispositions.

case facts between the first and second courts' cut points for which the first court prefers $d_1 = lib$ and the second court prefers $d_2 = con$.

The second court receives their own signal as well. Using both signals, the second court updates their beliefs about ω . Choosing a disposition is then simple. The court compares the probability that the case facts are to the left of their cut point with the probability that they are to the right and chooses the disposition corresponding to whichever probability is higher.

What Contributes to Splits?

A split occurs if the two courts choose different dispositions. I analyze the factors that make a split more likely to occur. Under the model, courts have no innate preference for uniformity in their decisions. They are only concerned with choosing a "correct" disposition according to their ideal rule. The court's beliefs about which disposition is correct are central to their decision. As the novel feature of my model is preference heterogeneity, I focus on how the ideological distance between the courts affects the probability of a split.

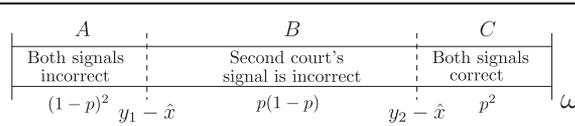
Learning from Previous Decisions

First, I analyze the effect of ideological distance on the second court's beliefs over ω . Assume the first court chose $d_1 = lib$. An increase in the second court's belief that the liberal disposition is correct increases the probability that they will choose the same disposition—the second court becomes more likely to always choose $d_2 = lib$, regardless of signal, and less likely to always choose $d_2 = con$. Together, these effects make it more likely that the second court will choose the same disposition as the first court.

If the first court's decision is uninformative, the second court learns nothing from the disposition. Assume the first case is informative such that the second court learns s_1 . Consider $s_1 = s_2 = \widehat{lib}$ and a second court more conservative than the first ($y_1 < y_2$). How likely are these signals given ω ?

Figure 5 divides the case space into three regions. In Region A, $\hat{x} + \omega$ is to the left of both courts. Within this region, both courts prefer a disposition of $d = con$; however, both courts received signals of $s = \widehat{lib}$. For ω to be in Region A, both of these signals would have to be incorrect. Moving to the right, Region B is the

FIGURE 5. Accuracy of $s_1 = s_2 = \widehat{lib}$ Based on the Value of ω



Note: The figure demonstrates what would have to be true of the signals received by the courts for ω to be in various regions of the space.

interval between the courts' cut points. For values of ω in this interval, the first court prefers $d = lib$ while the second court prefers $d = con$. If ω were in this region, the second court's signal would still be incorrect; however, the first court's signal is now correct. For both signals to be correct, ω would need to be in Region C ($\omega > y_2 - \hat{x}$), where both courts prefer a disposition of $d = lib$. When updating their beliefs about ω , the second court weights their prior beliefs by the probability that neither, one, or both of the signals are correct.

The second court is able to learn something from an informative case regardless of the disposition chosen or the relative position of the previous court. *How much* the court can learn, however, depends on these factors. Importantly, "learning" does not immediately translate to more accurate beliefs. Even if the court is able to perfectly intuit the signals of the other courts, these signals are not perfectly accurate. When signals are imperfect, there is always a chance that incorrect information persists (Bikhchandani, Hirshleifer, and Welch 1992).

Proposition 1. Assume the first court's decision is informative. As the distance between the courts' cut-points increases, the second court is less likely to believe that d_1 is their preferred disposition.

Intuitively, as the ideological distance between the two courts increases, the set of cases for which the two courts prefer different dispositions increases. Consequently, the relative weight the second court places on the first's signal decreases. As the first court's signal becomes less influential on the second court's beliefs, the second court is less likely to take the same action. A corollary of this proposition is that a court learns most

from a perfect ally. Heterogeneous cut points reduce the amount of information a court can learn from a single case.

Learning from Private Information

The two courts might choose the same disposition because the second court observed the decision of the first, updated their beliefs, and chose the same disposition as a result. This process is our common understanding of why uniformity occurs. However, the courts may choose the same disposition even if the second learns nothing from the actions of the first (Volden, Ting, and Carpenter 2008). Outside of previous dispositions, a court has two sources of information available to them when deciding a case: the location of the case facts and a private signal. If both courts are faced with extreme case facts, prior beliefs over ω will outweigh any of the influence signals have on the courts' decisions. Similarly, both courts might have received the same signal and chosen the same disposition as a result. Any combination of these three processes could result in uniformity; however, only the first is the result of learning from another court's actions.

Consider an example from policy diffusion literature—why two cities enact a ban on smoking in restaurants (see Shipan and Volden 2006; 2008). It may be that the second city learned from the first's policy. After the first city enacted a ban, the second observed the effects, updated its beliefs about the utility of a ban, and, as a result of this updating, decided to ban smoking. This explanation fits squarely in our understanding of policy diffusion. However, both cities could adopt a ban in the absence of learning; that is, both could have received private information about the negative consequences of secondhand smoke and enacted a ban as a result. Either of these processes could lead both cities to ban smoking, but we would refer only to the former as policy diffusion.

What looks like courts learning from one another may actually be courts simply taking actions consistent with their private information about the case. If the ideological distance between two courts is related to the private information they receive, it could lead us to incorrectly attribute uniformity to learning from previous decisions (Volden, Ting, and Carpenter 2008). If we could directly observe the information available to the courts, we would be able to account for the influence of private information on uniformity. However, we are limited in our ability to observe the case facts and the courts' signals.¹³ Consequently, I model how courts' private information relates to the relationship between ideological distance and the probability of a circuit split. Combining these results with Proposition 1 then gives the expected effect of ideological distance.

The probability that the courts receive the same signal and the probability that case facts are extreme vary as the ideological distance between the courts increases. To isolate the effect of a court's private information,

I calculate the effect of ideological distance *assuming that the courts do not observe each other's actions*. Formally, I assume that regardless of the actions of the first court, the second court always has a posterior belief equal to their prior, $G(\omega|d_1 = d, y_1) = U[-\Omega, \Omega]$. The assumption that dispositions are private is unrealistic; however, it allows me to identify whether there is an effect of ideological distance that is not related to the second court learning from the first.

To characterize the effect of ideological distance, we need to distinguish between two types of orderings.

Definition 3 (Bias-compatible). A disposition-court pair is *bias-compatible* if the first court is more liberal than the second and the first court chose the liberal disposition. Similarly, a pair is bias-compatible if the first court is more conservative than the second and chose the conservative disposition.

Definition 4 (Bias-incompatible). A disposition-court pair is *bias-incompatible* if the first court is more liberal than the second and chose the conservative disposition, or vice versa.

Whether a decision is bias-compatible or bias-incompatible is relative to the cut point of the second court. The same decision by the same first court could be bias-compatible or bias-incompatible, depending on whether the first court is more conservative or more liberal than the second court.

Ideological distance relates to the probability of a split through two mechanisms: changes in the probability of expected case facts and changes in probability that $s_2 = \widehat{lib}$. For building intuition, I focus on changes in the probability of expected case facts. Propositions are with respect to the combined effect of both mechanisms.

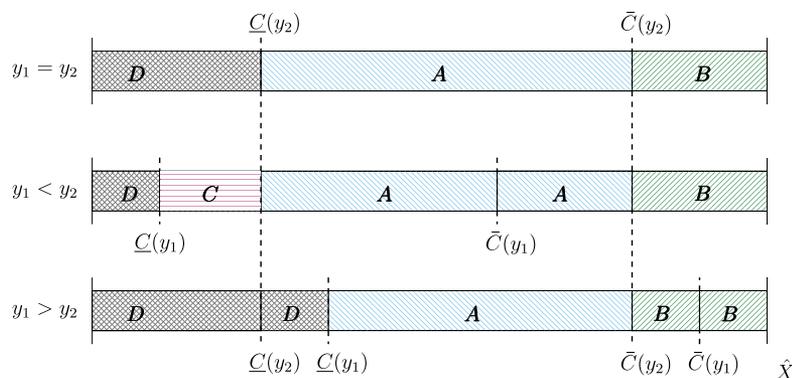
As a comparison point, I first assume the courts have the same cut point. The top panel of Figure 6 plots the potential locations of the observable case facts, \hat{x} , conditional on the first court choosing $d_1 = lib$. The space can be broken down into two regions. If $\hat{x} \in A$, the second court will choose $d_2 = lib$ if and only if she receives a signal of $s_2 = lib$. Instead, if $\hat{x} \in B$, the second court will always choose $d_2 = lib$.

Moving down a panel, I consider the bias-compatible case, where the first court is more liberal than the second. I note first that Regions *A* and *B* remain the same as before; however, there is a third region: *C*. If $\hat{x} \in C$, the second court will choose $d_2 = con$ regardless of which signal they received. As the distance between y_1 and y_2 increases, the size of Region *C* increases. Consequently, the probability that the two courts choose different dispositions increases. This leads to the following proposition.

Proposition 2. Assume dispositions are private. As the ideological distance between the courts increases, the probability of a circuit split increases for bias-compatible decisions.

The effect of ideological distance on the probability of a circuit split is in the opposite direction for bias-

¹³ While it is technically possible to observe case facts, quantifying them is imprecise and difficult to standardize across issues.

FIGURE 6. Possible Locations of \hat{x} for Different Preference Arrangements


Note: In Region A, the second court chooses $d_2 = lib$ if and only if $s_2 = \widehat{lib}$. In Region B, the second court always chooses $d_2 = lib$. In Region C, the second court never chooses $d_2 = lib$. If $y_1 < y_2$, increasing the ideological distance between the two courts increases the size of region C. If $y_2 < y_1$, increasing ideological distance reduces the size of Region A. The probability that \hat{x} is in Region D is 0. $\underline{C}(y) = y + G^{-1}(1-p)$. $\bar{C}(y) = y + G^{-1}(p)$.

incompatible decisions. Returning to the third panel of Figure 6, the second court is now more conservative than the first. As with the special case, where the two courts shared a cut point, there are only two regions: A and B. The size of Region A, however, is smaller than before. While the absolute size of Region B does not increase, its relative size does. The greater the distance between the first and second court, the greater the probability that $\hat{x} \in B$. When $\hat{x} \in B$, the second court always chooses $d_2 = lib$, so legal uniformity is assured. Consequently, increasing the distance between the two courts decreases the probability of a circuit split for bias-incompatible cases. This leads to the following proposition.

Proposition 3. Assume dispositions are private. As the distance between the courts increases, the probability of a circuit split weakly decreases for bias-incompatible decisions.

The Effects of Ideological Distance

As a reminder, increasing ideological distance reduces the influence of the first court's decision on the second court's actions. Combining this effect with the results concerning private information allows me to characterize the effect of ideological distance on uniformity. For bias-compatible decisions, increased ideological distance results in a decreased probability that the courts choose the same disposition.

Proposition 4. For bias-compatible decisions, as ideological distance between the two courts increases, the probability of a circuit split increases.

This result yields a clear empirical prediction for bias-compatible cases: increased ideological distance should yield more circuit splits. The clear empirical prediction, however, prevents one from determining what type of learning is occurring. Both effects are in

the same direction, so a positive effect could result from learning from previous decisions or from private information (Figure 7).

Bias-incompatible cases provide an opportunity to identify an independent effect of ideology on the probability of a disagreement resulting from courts learning from their peers. For these cases, the direction of the effect of ideological distance depends on what type of information the court is learning from. Focusing only on learning from peers, ideological distance has a positive effect on the probability of a circuit split. In contrast, ideological distance has a negative effect on the probability of a circuit split if we look only at a court's private information. Taken together, this results in unclear expectations on the effect of ideological distance on the probability of a circuit split; however, a positive effect can only occur if courts are learning from each other's decisions.

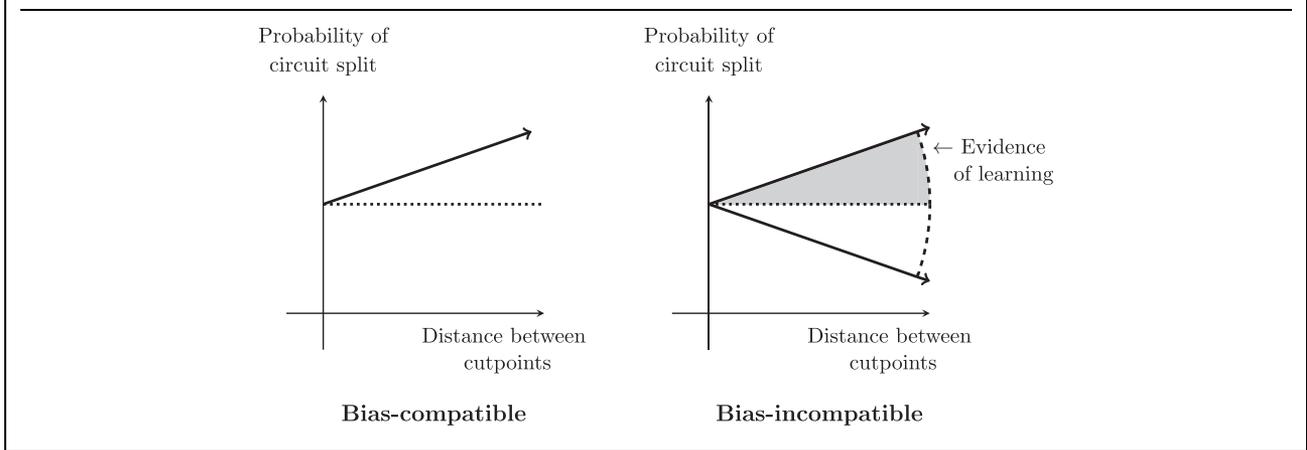
Proposition 5. For bias-incompatible decisions, a positive effect of ideological distance on the probability of a circuit split is evidence of courts learning from each other's decisions.

What If Courts Didn't Learn from Each Other?

Proposition 5 provides an opportunity to distinguish between a model, where courts learn from previous courts' decisions and one where courts independently make decisions solely based on their private information. For bias-incompatible cases, a positive effect of ideological distance on the probability of a circuit split is evidence that courts are learning from each other's decisions. A negative effect is still consistent with the model; however, an effect in this direction does not distinguish between the processes that lead to legal uniformity.

What would decisions look like if courts did not learn from one another? The primary difference between an environment in which courts learn from each other's

FIGURE 7. Expected Effect of Ideological Distance on the Probability of a Circuit Split



decisions and one where decision making is sequestered is that the former allows for information aggregation. When actors are able to learn from each other’s decisions, over time the influence of other actors’ decisions can outweigh one’s own signal (Bikhchandani, Hirshleifer, and Welch 1992). Absent this information aggregation, when in the sequence a decision is made is unrelated to which actions are taken.

Extending the analysis to N courts, as opposed to just two, highlights the lack of a relationship between order and outcomes if courts do not learn from each other. Assume that N courts decide the same case with observed facts \hat{x} . Let d_t correspond to the t^{th} disposition in the sequence. I refer to t as the period of the court.

Assume that the courts do not learn from each other’s decisions. First, I examine the probability that subsequent courts choose the same disposition as the t^{th} court, d_t . As the courts are not updating their beliefs from each other’s actions, each court has the same prior beliefs about ω , regardless of when in the sequence they act. Consequently, the probability that a court with cut point y_t choose d_t is constant in the period it decides. As long as there is no relationship between when a court hears a case and their cut point, the probability that any two courts agree is not affected by either of their periods. The probability that the second court chooses the same disposition as the first is equal to the probability that the tenth court chooses the same disposition as the ninth. This leads to the following proposition.

Proposition 6. Assume courts do not learn from each other’s decisions. The probability that $d_t = d_{t'}$ is constant for all $t, t' > t$.

A similar lack of a relationship exists between the period of a court and when circuit splits occur. Assume that all but one court chose the same disposition—for example, $N-1$ courts chose $d = \textit{lib}$ and one court chose $d = \textit{con}$. I refer to decision chosen by only one court as the *minority decision*. When in the development of an issue, the minority decisions most likely to occur? As above, if courts do not learn from each other’s decisions then prior belief about ω do not vary with a court’s

period. Consequently, the minority decision is equally likely to occur in any period.¹⁴

Proposition 7. Assume courts do not learn from each other’s decisions and only one court chose d . $Pr(d_t = d) = Pr(d_{t'} = d)$ for all t, t' .

Both of the above propositions provide an opportunity to identify if courts are making decisions independent from one another.

CASES OF FIRST IMPRESSION

Testing predictions from the model requires a set of cases where courts are unconstrained by precedential concerns. Accordingly, I focus on cases of first impression in the U.S. Courts of Appeals. Identifying these cases presents a unique challenge. The U.S. Courts of Appeals decide over 40,000 cases a year (U.S. Courts 2023), and judges are not mandated to declare that an issue is one of first impression when writing an opinion. The difficulty of identifying these cases has limited their collection. Existing work has primarily focused either on specific issues or only the first decision on an issue.¹⁵

¹⁴ If courts learn from each other’s decisions, the relationship between sequence order and both of these outcomes becomes more complex. In Appendix A in the Supplementary Material, I examine these relationships via simulation, allowing for courts to learn from each other’s decisions.

¹⁵ The *Seton Hall Circuit Review* collected a subset of cases of first impression from 2005–2018. In their publication, the editors only focused on the occurrence of an issue in a single circuit and did not track the development across circuits. In around 80% of the cases collected by Seton Hall, “first impression” appears in the opinion (Seton Hall Circuit Review 2018). Additionally, Klein (2002) collected a set of cases involving new legal issues from 1984–1991. Klein does not specifically describe these issues as ones of first impression; however, they fit similar criteria. An important distinction between my approach and Klein’s is how cases were identified. Klein relies on secondary sources to identify new legal rules. In contrast, I rely on judges’ characterizations of the issue as novel.

Accordingly, I constructed a new dataset of cases of first impression. I started with cases in the Third Edition of the Federal Reporter, which comprises cases published in the Courts of Appeals from 1993 to 2021. I searched opinions and case synopses for the phrase “first impression,” which yields over 6,000 cases for this period. I narrowed down this initial set by focusing on cases that deal with search and seizure, criminal law, or freedom of speech.¹⁶

I first determined whether the case decided a federal issue of first impression.¹⁷ I excluded cases that announced that a question was one of first impression but declined to decide it, decided on a state-specific issue of first impression, or simply used the phrase “first impression” in an unrelated context. Additionally, I excluded cases that were unpublished or from a period outside the interval of study. For each case, I identified the legal question of first impression by relying on both the synopsis of the case and the text of the opinion. In cases, where the court’s description of the question is unclear, I rely on the discussion of the legal question by other circuits.

Once the question of first impression had been defined, I used citations to identify cases from other circuits that previously decided the issue. For each citation, I first checked that the cited opinion discussed the issue. If it did, I then determined if the circuit had previously decided the question of first impression. Like Beim and Rader (2019), I focus on “precedent-setting cases”—the first time a circuit resolved an issue of first impression. By identifying precedent-setting cases, I was able to determine the order in which circuits decided the legal question. I determined if a court had previously decided the issue by looking for in-circuit precedent cited in their discussion of the issue. If an in-circuit precedent was cited, I read it to determine if it also decided the issue. The process repeated until I found the precedent-setting case for the issue.

¹⁶ These cases were identified in three ways. First, using the text of opinions from the Harvard Case Law Access Project (Harvard Law School 2023), I took the intersection of the set of opinions that used “first impression” and the set of cases identified by Westlaw as citing the Fourth Amendment of the U. S. Constitution. The Case Law Access Project database appeared to be missing opinions from recent years, so I next turned to Westlaw. Westlaw allows cases to be filtered using a key number. Once again starting with opinions that use “first impression,” I subset to cases that involve search and seizure (394) and cases involving freedom of speech (92XVIII). Finally, Westlaw allows cases to be filtered by “practice area.” Using this feature, I filtered cases to the practice area of criminal law. I took a random subset of these criminal law cases.

¹⁷ I rely on a court’s characterization that an issue is novel. In my exploratory interviews with clerks, they noted that looking for the phrase “first impression” in the opinion text would serve as an effective method of identifying cases with uncertainty. Courts may have strategic incentives to claim that an issue is novel, as doing so affords them greater freedom in decision making. This greater latitude these cases provide could lead judges to over claim that an issue is new relative to some “true” standard of novelty. Crucially, for judges to be incentivized to claim an issue is new despite it being settled, there must be some underlying change—for example, a shift in the composition of the Supreme Court, that enables the judge ignore existing legal rules. In these cases, courts will still act unconstrained in a way consistent with the absence of precedent.

I repeated the process for each circuit cited.¹⁸ If in the process of looking through the cited cases, I discovered a citation to an additional circuit, I included the case. Once cases cited by the opinion were exhausted, I repeated the process using cases that cited the opinion to look for subsequent decisions on the issue.

For each case, I coded whether the court chose the liberal or conservative disposition. For criminal or search and seizure cases, I coded the pro-defendant choice as the liberal disposition. For the freedom of speech cases, I coded the pro-speech choice as the liberal disposition.¹⁹ If at least one circuit chose the conservative disposition and at least one circuit chose the liberal disposition, I coded the legal question as a circuit split. In total, my data consist of 239 unique questions of first impression encompassing 1,326 precedent-setting decisions. To the best of my knowledge, this constitutes the most comprehensive collection of these cases to date (see Taboni 2025).

Figure 8 provides some descriptive statistics about the cases. The cases were roughly uniformly distributed across the circuits with two exceptions: the Ninth Circuit heard more than average and the DC Circuit heard fewer than average.²⁰ Panel b of the figure plots the publication years for cases in the data. My starting point is issues active between 1993 and 2021; however, decisions on these issues could occur before or after this period. The oldest decision on one of these issues was in 1953, while the newest was from 2024. Unsurprisingly, the vast majority of cases are from 1990 to 2020.²¹

Prevalence of Circuit Splits

Unlike previous work, I capture both circuit splits and legal uniformity among the courts. How common were circuit splits? Of the 239 legal questions, there were 83 circuit splits (about 35% of questions). Despite ideological variation across panels and circuits, there is substantial uniformity in the Courts of Appeals. On almost two-thirds of issues, circuits were in agreement with one another. For more recent cases, the circuits still have an opportunity to split on a question of first impression. As I include legal issues that started as late as 2021, a split may not yet have occurred. Subsetting to

¹⁸ In cases, where a panel was overruled by the circuit en banc, I counted the en banc case as the precedent-setting case.

¹⁹ These coding decisions are consistent with the procedure of the Supreme Court Database. I excluded issues, where the ideological direction of a case was not clear. For example, what is the proper standard of review for issues involving the speech and debate clause? Additionally, when I took a subset of criminal law cases, I excluded cases, where the issue of first impression was not a criminal law question.

²⁰ The Ninth Circuit decided 170 issues of first impression—the most of any circuit by far. The Ninth Circuit is the largest in the Courts of Appeals, so its increased activity in cases of first impression is expected. In contrast, the DC Circuit decided 34 cases. The DC Circuit only covers the District of Columbia, so the types of cases it hears may differ from the other geographic circuits. For an issue to appear in the data, at least two circuits had to hear it. This may explain why the DC Circuit decided fewer cases.

²¹ The median amount of time between the first and last decision in the data on a question of first impression is 14 years.

FIGURE 8. Descriptive Plots about Cases of First Impression

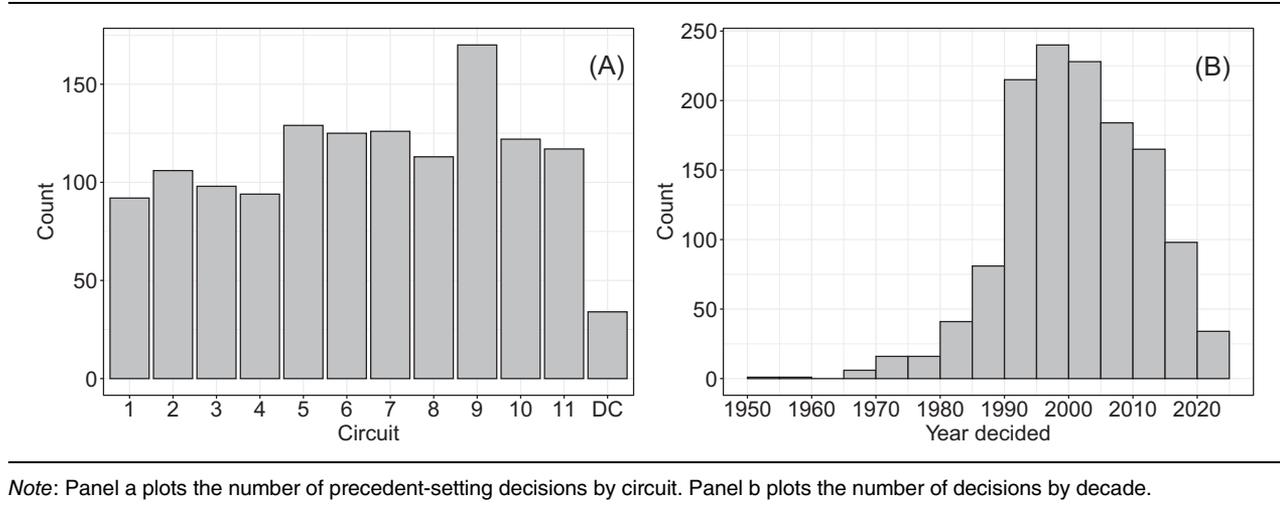
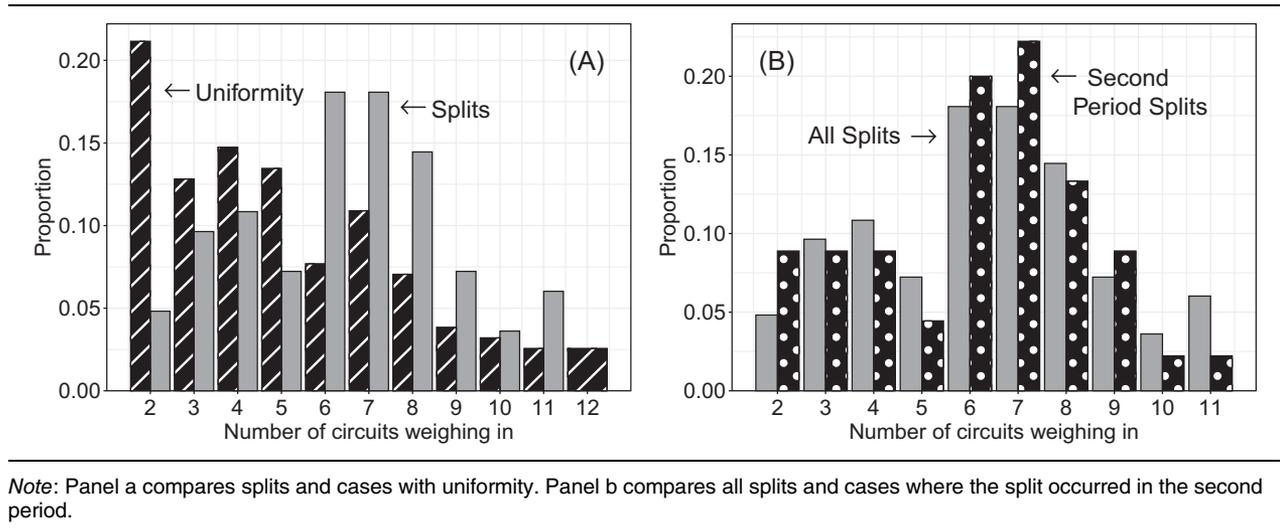


FIGURE 9. Number of Decisions on Issues of First Impression



issues, where the last decision was in or before 2015 (171 issues) yields a split in about 35% of questions, the same as all issues.

Circuit splits occur in a minority of cases of first impression. While the presence of a split distinguishes these cases, so do other factors of the case. Figure 9 (panel a) examines the number of circuits that weighed in on a legal question. Legal questions that divided the circuits were heard by more courts on average than questions, where the circuits agreed. The median number of courts that ruled on a decision that divided the circuits is six; whereas, the median number of courts that rule on a question where the courts agreed is five.

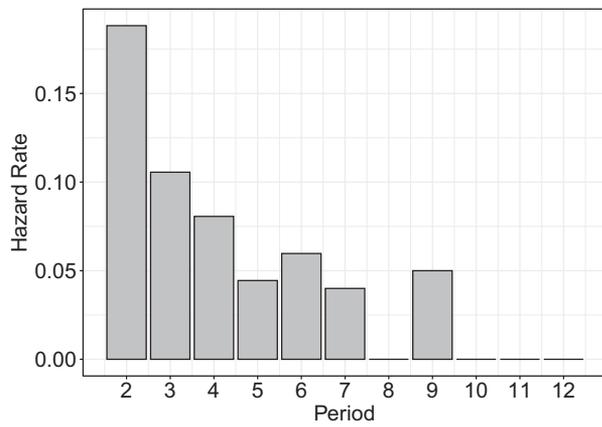
This relationship is *not* caused by the weakly increasing probability of a circuit split in the number of courts that decide an issue. Subsetting to cases, where the split occurred in the second period (the earliest it can happen), the relationship persists. Figure 9 (panel b) compares these splits to all splits.

My data allow me to calculate how likely a circuit split is, conditional on the number of courts that decide an issue. Figure 10 plots the circuit split hazard rate. A circuit split is most likely to happen immediately. In the second period, a split happens in around one out of every five issues—by far the most likely time for a split to occur. After this point, the probability of a split appears to weakly decrease. If an issue makes it to a seventh court, the probability of a split drops below 5%. After period 9, the observed hazard rate drops to zero.²²

Testing Predictions

The model generates predictions on when legal uniformity is more likely to occur. While the model only focuses

²² Very few issues made it to a tenth decision without a split. Out of the 239 issues, only 13 reached this point (5%).

FIGURE 10. Split Hazard Rate

Note: Hazard rate is calculated as the proportion of previously uniform cases in period $t-1$ that result in a split in period t .

on the first two courts that decide an issue, the main theoretical predictions apply to multiple courts as well. When multiple courts have already decided an issue, the next court in the sequence is able to potentially learn from all of the previous decisions. *What* it learns from these decisions, however, is governed by the same structure as the two-court model. The current court learns previous courts' signals from informative decisions, and decisions by more ideologically distant courts are less influential. Similarly, the relationship between two courts' private information and ideological distance does not depend on the other courts that decide an issue. Consequently, it is unaffected by the inclusion of more decisions.²³

To capture an environment with multiple actors deciding the same issue, I rely on an empirical approach commonly used to study policy diffusion (Volden 2006). Studies of policy diffusion seek to capture how the actions of multiple other actors can influence current decision-makers. This approach corresponds well to the process of legal development and has been used to study the transmission of precedent across state (Hinkle and Nelson 2016). Within each issue, I create a dyad between each pair of courts that decided the question. For example, if three courts decide an issue, there are three dyads: (C_1, C_2) , (C_2, C_3) , and (C_1, C_3) . I create a *Disagreement* variable that is equal to 1 if the two courts chose different dispositions and 0 otherwise. To measure ideological distance, I utilize Judicial Common Spaces Scores (Boyd 2015; Epstein et al. 2007; Giles, Hettinger, and Peppers 2001).²⁴ Divergence is measured as the distance between the panel medians.²⁵

²³ An exact test of the model would only look at the first two courts to decide an issue. With the current data, this test is underpowered.

²⁴ When a Supreme Court justice served on a panel, I use their score from their time on the Courts of Appeals if they previously served as a circuit judge. Otherwise, I use the average of their JCS score.

²⁵ If the case was heard en banc, I use the decision from the circuit en banc. For these cases, I use the circuit median to measure ideological distance. Of the 1,326 cases, 33 were decided by a panel en banc.

I begin by examining the relationship between ideological distance and disagreement, pooling together bias-compatible and bias-incompatible cases. Columns 1 and 2 in Table 1 present results from this simple regression. In both models, the effect of ideological distance on disagreement is not significant and close to zero. An interpretation of these results that stopped there would suggest that ideological distance does not contribute to the occurrence of circuit splits.

The formal model, however, suggests that this apparent null result is masking important heterogeneity in the relationship between ideological distance and agreement. To capture this effect, I include a variable for whether the first court's decision is bias-compatible. I code the variable such that it is equal to 1 if the previous court is more liberal than the second and chose the liberal disposition or if they are more conservative than the second and chose the conservative disposition, and 0 otherwise. I interact this variable with the ideological distance between the two courts. I estimate models including the aforementioned variables and fixed effects for each issue of first impression.

Columns 3 and 4 in Table 1 display results from this analysis.²⁶ I begin with the direct effect of a case being bias-compatible. When ideological distance is equal to 0, there is no difference in the probability of a circuit split between bias-compatible and bias-incompatible cases. The coefficient for bias-compatible is insignificant and near zero. Turning next to ideological distance, we need to distinguish between bias-compatible and bias-incompatible orderings. The main effect of ideological distance captures the effect of distance on the probability of a split for *bias-incompatible* cases. When a previous decision is bias-incompatible, increasing the distance between the two courts reduces the probability of a circuit split. As the two courts move further apart, the current court is *more* likely to choose the same disposition as the previous court. When the previous court is more liberal and chose the conservative disposition, or vice versa, ideological distance increases agreement among the circuits.

The relationship between ideological distance and the probability of a split is opposite for bias-compatible cases. The effect of distance on disagreement among *bias-compatible* cases is captured by the sum of the coefficients for ideological distance and the interaction between distance and bias-compatible. Consistent with the predictions from the formal model, I find that, for *bias-compatible* decisions, as ideological distance increases, the current court is less likely to choose the same disposition as the previous court. In other words, a circuit split is more likely to occur when the previous court is ideologically distant and the first court renders a bias-compatible decision.

²⁶ In Appendix C in the Supplementary Material, I estimate additional models as robustness checks. Including circuit fixed effects does not change the substantive takeaways of the main results. The effect of ideological distance depends on whether the previous decision was bias-compatible or bias-incompatible. Results are robust to using CF-scores generated by Bonica and Sen (2021) as opposed to JCS scores.

TABLE 1. OLS Estimates of the Effect of Ideological Distance on Disagreement between Circuits

	Dependent variable:			
	Disagreement			
	1	2	3	4
Ideological distance	-0.03 (0.03)	0.02 (0.03)	-0.13* (0.03)	-0.05* (0.03)
Bias-compatible			0.02 (0.02)	0.02 (0.02)
Distance*Bias-compatible			0.20* (0.06)	0.15* (0.05)
Constant	0.19* (0.02)	-0.01 (0.01)	0.18* (0.02)	-0.04* (0.02)
Issue fixed effects		✓		✓
Adj. R ²	0.00	0.30	0.02	0.31
No. of obs.	3855	3855	3855	3855
N Clusters	239	239	239	239

Note: Robust standard errors are clustered at the issue of first impression. * $p < 0.05$.

The results from the analysis suggest the importance of distinguishing between the two types of cases. Looking only at the effect of ideological distance on agreement suggests that ideology has no effect. This apparent null effect is instead the result of significant but opposite effects of ideological distance based on the ideological ordering of the courts.

Figure 11 (panel a) plots the predicted probability of a circuit split as ideological distance increases, distinguishing between bias-compatible and bias-incompatible cases. Figure 11 (panel b) plots the

difference in probability of a circuit split between bias-compatible and bias-incompatible cases along with the 95% confidence intervals of this difference. When the courts have the same cut point, there is no significant difference in the probability of a circuit split between bias-compatible and bias-incompatible cases. As ideological distance increases, however, different patterns emerge. The probability of a split decreases for bias-incompatible cases, while the probability increases for bias-compatible cases. What is the substantive significance of this difference? The mean liberal panel ($y < 0$) and the mean conservative panel ($y > 0$) have an ideological distance of 0.64. This results in a 10% difference in the probability of a circuit split depending on whether the first court's decision was bias-compatible or bias-incompatible. As the average probability of a split is around only 35%, this difference is notable.

How do the above findings relate to previous work on judicial decision making? Ubiquitous in the judicial auditing literature is the idea that supervising courts rarely review decisions that run counter to a court's known bias. A liberal decision made by a more conservative lower court or a conservative decision made by a more liberal lower court is unlikely to be reviewed (Beim, Hirsch, and Kastellec 2014; Cameron, Segal, and Songer 2000). The results from the model and subsequent analysis demonstrate that a similar principle exists with respect to the probability of a circuit split. A conservative decision made by a more liberal court indicates that conflict is unlikely. These results can be characterized in terms of liberal and conservative dispositions. If the first court chose a liberal disposition, a circuit split is *less* likely as the relative conservativeness of the first court increases. If the first court chose a conservative disposition, a circuit split is *more* likely as the relative conservativeness of the first court increases.

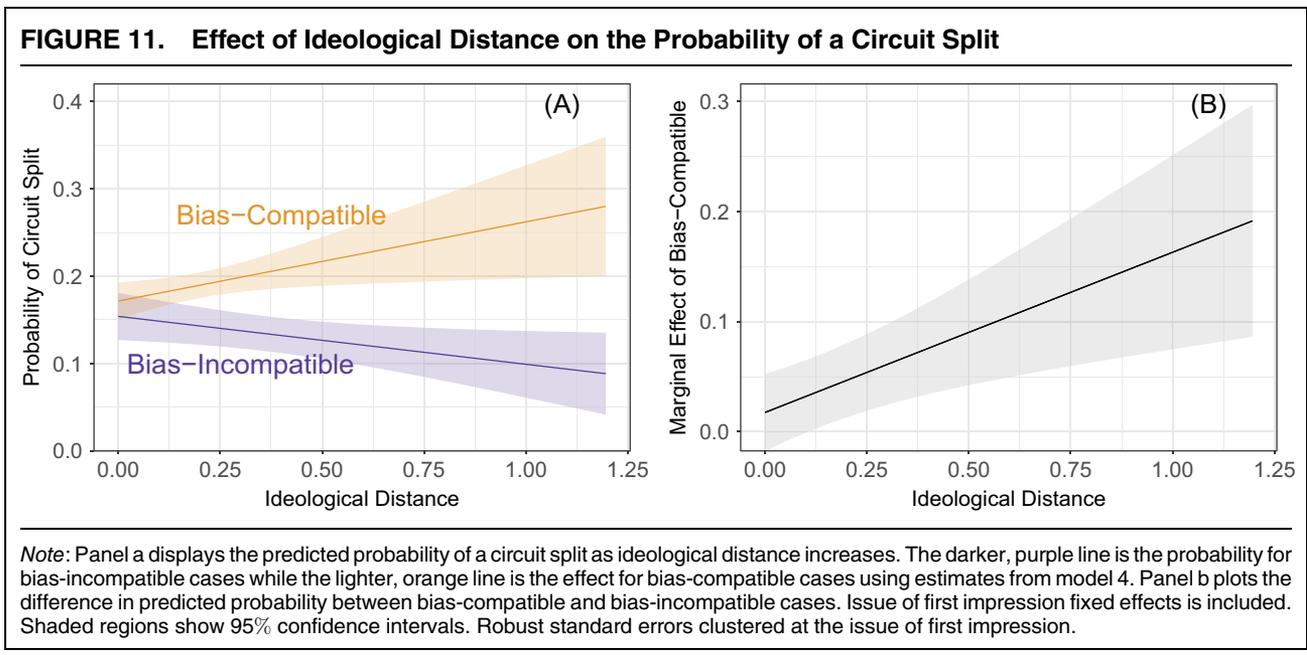
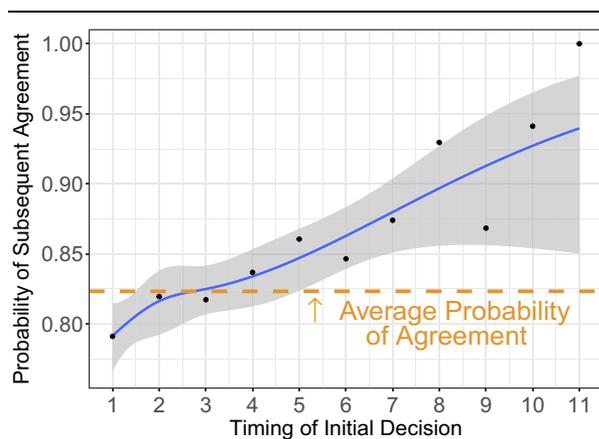


FIGURE 12. Relationship between the Probability of Subsequent Agreement and the Period in Which a Court Acted

Note: The dashed line is the empirical average probability of agreement between two courts. Dots indicate average probabilities of agreement with a court deciding in period t . The solid line is a loess estimate with shaded region depicting 95% confidence interval.

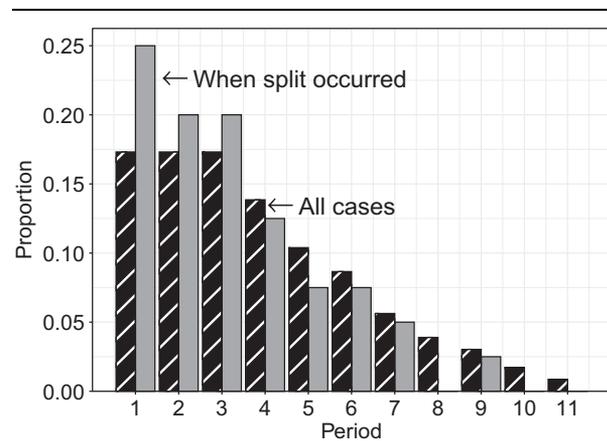
Distinguishing Mechanisms

The above findings provide evidence that ideological distance between courts affects the probability of agreement. Importantly, the effect of ideological distance for bias-compatible matches the prediction of the model. Increasing the distance between two courts increases the probability that a circuit split occurs. For bias-incompatible cases, results are consistent with the model, as the model allows for a positive or a negative effect; however, a negative effect does not allow us to say anything about whether the findings result from courts learning from one another.

If courts were making decisions independently, we should observe no relationship between when in the sequence of decisions a court acts and either the probability of subsequent agreement (Proposition 6) or the probability that a split occurs in that period (Proposition 7). First, I test whether sequence order is related to the probability of subsequent agreement. Figure 12 captures this relationship. In contrast to the flat relationship that a model with independent decision making predicts, we observe a positive relationship between the period in which a court makes a decision and subsequent agreement with that decision.

A similar result arises with respect to when the minority decision in a circuit split is made. If the circuits were making decisions independently, we would expect the distribution of when minority decisions are chosen to look identical to the distribution of cases overall.²⁷ Instead, Figure 13 demonstrates that if all but one court

²⁷ Issues are heard by different numbers of circuits. Consequently, the probability that for a given issue, a decision was made in period n is weakly decreasing in n . I subset to issues decided by at least three circuits to be able to define the minority decision.

FIGURE 13. Period in Which the Minority Decision Occurred

Note: Solid, gray bars plot proportions of period where a court chose *lib* (*con*) conditional on all other courts choosing *con* (*lib*). Black, striped bars plot the proportions of all decisions occurring in the period. Figure subsets to issues involving at least three decisions in order to code a minority decision.

chooses the same disposition, the minority decision is more likely to occur in earlier periods. Both of these findings provide evidence that judicial behavior is not consistent with courts making actions independently based solely on their private information.²⁸

DISCUSSION AND CONCLUSION

“It is one of the happy incidents of the federal system that a single courageous State may, if its citizens choose, serve as a laboratory; and try novel social and economic experiments without risk to the rest of the country.”²⁹ In *New State Ice Co. v. Liebmann*, Justice Brandeis famously highlighted the importance of decentralized decision making in the United States. While Brandeis’s focus was on the policy actions of the states, the decentralized nature of the U.S. Courts of Appeals plays a similar role in the formation of national policy. Decisions by circuits are often the final say on an emerging legal question. From interactions with law enforcement to environmental quality, courts play an important role in policy development. The legal landscape, however, is constantly changing. Technological shocks, changes in precedent, and new laws alter existing legal understandings. These shocks generate uncertainty over courts’ decisions. While the task of deciding important legal issues is daunting, courts are not making these

²⁸ Regression analyses confirm the relationship in the figures. Tables can be found in Appendix C in the Supplementary Material. I note that observed empirical patterns are consistent with simulations where courts learn from each other’s decisions. I caution against interpreting findings as evidence in direct support of the model I present and instead as evidence against an environment with decisions based only on private information.

²⁹ *New State Ice Co. v. Liebmann*, 285 U.S. 262 (1932).

decisions in a vacuum. They can rely on the decisions of their peers to better inform their actions.

My theory sheds light on *how much* courts learn from their peers by studying sequential decision making. First, I find that courts learn most from their ideological allies. When faced with uncertainty, the decision of a like-minded court is most informative. Second, I decompose the relationship between ideological distance and the probability of a circuit split into two mechanisms: learning from previous decisions and learning from private information. Finally, I characterize what we would observe if courts did not learn from each other's decisions.

I test predictions from the model using an original dataset consisting of over 1,300 cases of first impression. These data provide insight into the prevalence of circuit splits and allow me to compare circuit splits with cases where the Courts of Appeals agree. Using the data, I estimate a model of agreement and find that without disaggregating between bias-compatible and bias-incompatible cases, one would erroneously conclude a null effect of ideological distance on the probability of a circuit split. When the cases are disaggregated into these two groups, I find that ideological distance has significant, heterogeneous effects. Ideological distance increases disagreement for bias-compatible orderings and decreases disagreement for bias-incompatible orderings. I find that court's behavior is inconsistent with a model where they do not learn from the actions of their peers.

What do these findings tell us about the "path of the law" (Holmes 1997)? First, I add to the large body of literature on the importance of ideology to judicial decision making. Unsurprisingly, I find that ideology plays a central role in the development of circuit splits. More importantly, the theory provides insight into *why* ideology matters. Differences in ideology influence what courts can learn from each other's decisions. Importantly, I find that ideologically polarized courts need not come to different conclusions. When a court makes a decision inconsistent with their known bias, future courts are more likely to agree. Even when they are ideologically distant, a split is less likely to occur. Additionally, my empirical analysis provides insight into the frequency of disagreement in the Courts of Appeals. While circuit splits are by no means uncommon, I find that they occur in a minority of new legal issues.

There are a number of ways in which the model could be expanded. I assume that the courts face no risk of review or reversal; however, courts face the risk of the Supreme Court reviewing and overturning their decision. Future work should examine how these pressures affect which decisions courts are willing to make under uncertainty.

While my focus has been primarily on the U.S. Courts of Appeals, the model has relevance to other institutions as well. State supreme courts also face uncertainty when deciding new legal questions. When deciding questions of state law, state courts of last

resort are bound only by precedent from within the state. Despite this formal independence, these courts often look to their peers in other states for guidance on new legal questions (Hinkle and Nelson 2016). Similarly, within a circuit, district courts are not bound by the actions of other district courts (Mead 2011). The model captures the dynamics between these courts as well. Uniformity among the circuits may be the norm; however, it is by no means a guarantee. Understanding why circuits come to different conclusions and when to expect circuit splits is central to understanding legal development.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit <https://doi.org/10.1017/S0003055425101160>.

DATA AVAILABILITY STATEMENT

Research documentation and data that support the findings of this study are openly available at the American Political Science Review Dataverse: <https://doi.org/10.7910/DVN/VYDUKV>.

ACKNOWLEDGEMENTS

I thank Deborah Beim, Roel Bos, Charles Cameron, Alina Dunlap, Sean Gailmard, Gleason Judd, John Kastlelec, Eric Manning, Hye Young You, and participants at the 2024 APSA, MPSA, and Emerging Scholars in Public Law conferences for helpful comments and suggestions.

FUNDING STATEMENT

This research was funded by the Center for the Study of Democratic Politics (#108122).

CONFLICT OF INTEREST

The author declares no ethical issues or conflicts of interest interests in this research.

ETHICAL STANDARDS

The author declares the human subjects research in this article was reviewed and approved by the Princeton University Institutional Review Board and certificate numbers are provided in the text. The author affirms that this article adheres to the principles concerning research with human participants laid out in APSA's Principles and Guidance on Human Subject Research (2020).

REFERENCES

- Bala, Venkatesh, and Sanjeev Goyal. 2001. "Conformism and Diversity under Social Learning." *Economic Theory* 17 (1): 101–20.
- Banerjee, Abhijit V. 1992. "A Simple Model of Herd Behavior." *The Quarterly Journal of Economics* 107 (3): 797–817.
- Barnes Bowie, Jennifer, and Donald R. Songer. 2009. "Assessing the Applicability of Strategic Theory to Explain Decision Making on the Courts of Appeals." *Political Research Quarterly* 62 (2): 393–407.
- Beim, Deborah. 2017. "Learning in the Judicial Hierarchy." *The Journal of Politics* 79 (2): 591–604.
- Beim, Deborah, Alexander V. Hirsch, and Jonathan P. Kastellec. 2014. "Whistleblowing and Compliance in the Judicial Hierarchy." *American Journal of Political Science* 58 (4): 904–18.
- Beim, Deborah, and Kelly Rader. 2019. "Legal Uniformity in American Courts." *Journal of Empirical Legal Studies* 16 (3): 448–78.
- Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. 1992. "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades." *Journal of Political Economy* 100 (5): 992–1026.
- Black, Ryan C., and Ryan J. Owens. 2009. "Agenda Setting in the Supreme Court: The Collision of Policy and Jurisprudence." *The Journal of Politics* 71 (3): 1062–75.
- Bonica, Adam, and Maya Sen. 2021. "Estimating Judicial Ideology." *Journal of Economic Perspectives* 35 (1): 97–118.
- Boyd, Christina L. 2015. "Federal District Court Judge Ideology Data."
- Caldeira, Gregory A. 1985. "The Transmission of Legal Precedent: A Study of State Supreme Courts." *American Political Science Review* 79 (1): 178–94.
- Cameron, Charles M. 1993. "New Avenues for Modeling Judicial Politics." Paper presented at the Conference on the Political Economy and Public Law, Rochester, NY.
- Cameron, Charles M., Jeffrey A. Segal, and Donald Songer. 2000. "Strategic Auditing in a Political Hierarchy: An Informational Model of the Supreme Court's Certiorari Decisions." *American Political Science Review* 94 (1): 101–16.
- Cameron, Charles M., and Lewis A. Kornhauser. 2017a. "What Courts Do... and How to Model It." Research Paper.
- Cameron, Charles M., and Lewis A. Kornhauser. 2017b. "What Do Judges Want? How to Model Judicial Preferences." Research Paper.
- Chamley, Christophe. 2004. *Rational Herds: Economic Models of Social Learning*. Cambridge: Cambridge University Press.
- Clark, Tom S., and Jonathan P. Kastellec. 2013. "The Supreme Court and Percolation in the Lower Courts: An Optimal Stopping Model." *The Journal of Politics* 75 (1): 150–68.
- Corley, Pamela C., Paul M. Collins Jr., and Bryan Calvin. 2011. "Lower Court Influence on U.S. Supreme Court Opinion Content." *The Journal of Politics* 73 (1): 31–44.
- Cross, Frank B. 2003. "Decisionmaking in the U.S. Circuit Courts of Appeals." *California Law Review* 91 (6): 1457–515.
- Ellison, Glenn, and Drew Fudenberg. 1993. "Rules of Thumb for Social Learning." *Journal of Political Economy* 101 (4): 612–43.
- Epstein, David, and Sharyn O'Halloran. 1994. "Administrative Procedures, Information, and Agency Discretion." *American Journal of Political Science* 38 (3): 697–722.
- Epstein, Lee, Andrew D. Martin, Jeffrey A. Segal, and Chad Westerland. 2007. "The Judicial Common Space." *The Journal of Law, Economics, & Organization* 23 (2): 303–25.
- Giles, Micheal W., Virginia A. Hettinger, and Todd Peppers. 2001. "Picking Federal Judges: A Note on Policy and Partisan Selection Agendas." *Political Research Quarterly* 54 (3): 623–41.
- Gilligan, Thomas W., and Keith Krehbiel. 1987. "Collective Decisionmaking and Standing committees: An Informational Rationale for Restrictive Amendment Procedures." *The Journal of Law, Economics, and Organization* 3 (2): 287–335.
- Haire, Susan B., Stefanie A. Lindquist, and Donald R. Songer. 2003. "Appellate Court Supervision in the Federal Judiciary: A Hierarchical Perspective." *Law & Society Review* 37 (1): 143–68.
- Harvard Law School. 2023. "Case Law Access Project." Accessed 2023-9-14. <https://case.law/>
- Hazelton, Morgan L. W., and Rachael K. Hinkle. 2022. *Persuading the Supreme Court: The Significance of Briefs in Judicial Decision-Making*. Lawrence: University Press of Kansas.
- Hinkle, Rachael K. 2015. "Legal Constraint in the U.S. Courts of Appeals." *The Journal of Politics* 77 (3): 721–35.
- Hinkle, Rachael K., and Michael J. Nelson. 2016. "The Transmission of Legal Precedent among State Supreme Courts in the Twenty-First Century." *State Politics & Policy Quarterly* 16 (4): 391–410.
- Holmes, Oliver Wendell. 1997. "The Path of the Law." *Harvard Law Review* 110 (5): 991–1009.
- Iaryczow, Matias, and Matthew Shum. 2012. "The Value of Information in the Court: Get It Right, Keep It Tight." *American Economic Review* 102 (1): 202–37.
- Klein, David E. 2002. *Making Law in the United States Courts of Appeals*. Cambridge: Cambridge University Press.
- Kornhauser, Lewis A. 1992. "Modeling Collegial Courts. II. Legal Doctrine." *The Journal of Law, Economics, & Organization* 8 (3): 441–70.
- Lax, Jeffrey R. 2011. "The New Judicial Politics of Legal Doctrine." *Annual Review of Political Science* 14: 131–57.
- Legal Information Institute. 2025. "Case of First Impression." Accessed 04/08/2025. https://www.law.cornell.edu/wex/case_of_first_impression
- Lindquist, Stefanie A., and David E. Klein. 2006. "The Influence of Jurisprudential Considerations on Supreme Court Decisionmaking: A Study of Conflict Cases." *Law & Society Review* 40 (1): 135–62.
- Matthews, Abigail A. 2024. "Learning from and Emulating Their Peers: Policy Diffusion in the Courts." *State Politics & Policy Quarterly* 24 (2): 218–27.
- Mead, Joseph W. 2011. "Stare Decisis in the Inferior Courts of the United States." *Nevada Law Journal* 12 (3): 787–830.
- Parameswaran, Giri, Charles M. Cameron, and Lewis A. Kornhauser. 2021. "Bargaining and Strategic Voting on Appellate Courts." *American Political Science Review* 115 (3): 835–50.
- Re, Joseph R. 2001. "Brief Overview of the Jurisdiction of the US Court of Appeals for the Federal Circuit under Sec. 1295 (a)(1)." *Federal Circuit Bar Journal* 11: 651.
- Seton Hall Circuit Review. 2018. "First Impressions."
- Shipan, Charles R., and Craig Volden. 2006. "Bottom-up Federalism: The Diffusion of Antismoking Policies from U.S. Cities to States." *American Journal of Political Science* 50 (4): 825–43.
- Shipan, Charles R., and Craig Volden. 2008. "The Mechanisms of Policy Diffusion." *American Journal of Political Science* 52 (4): 840–57.
- Smith, Lones, and Peter Sørensen. 2000. "Pathological Outcomes of Observational Learning." *Econometrica* 68 (2): 371–98.
- 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 (3): 673–96.
- Spitzer, Matt, and Eric Talley. 2000. "Judicial Auditing." *The Journal of Legal Studies* 29 (2): 649–83.
- Staton, Jeffrey K., and Georg Vanberg. 2008. "The Value of Vagueness: Delegation, Defiance, and Judicial Opinions." *American Journal of Political Science* 52 (3): 504–19.
- Strayhorn, Joshua A. 2020. "Ideological Competition and Conflict in the Judicial Hierarchy." *American Journal of Political Science* 64 (2): 371–84.
- Taboni, Anthony R. 2025. "Replication Data for: The Path of Law: Legal Uncertainty and Issues of First Impression in the U.S. Courts of Appeals." Harvard Dataverse. Dataset. <https://doi.org/10.7910/DVN/VYDUKV>.
- Talley, Eric. 1999. "Precedential Cascades: An Appraisal." *Southern California Law Review* 73: 87–137.
- Ulmer, S. Sidney. 1984. "The Supreme Court's Certiorari Decisions: Conflict as a Predictive Variable." *American Political Science Review* 78 (4): 901–11.
- U.S. Courts. 2023. "2023 Year-End Report on the Federal Judiciary." <https://www.supremecourt.gov/publicinfo/year-end/2023year-endreport.pdf>
- Vanberg, Georg. 2001. "Legislative-Judicial Relations: A Game-Theoretic Approach to Constitutional Review." *American Journal of Political Science* 45 (2): 346–61.
- Volden, Craig. 2006. "States as Policy Laboratories: Emulating Success in the Children's Health Insurance Program." *American Journal of Political Science* 50 (2): 294–312.
- Volden, Craig, Michael M. Ting, and Daniel P. Carpenter. 2008. "A Formal Model of Learning and Policy Diffusion." *American Political Science Review* 102 (3): 319–32.