

# Do judges actually care about the law? Evidence from circuit split data

Scott Baker and Anup Malani<sup>1</sup>

**Abstract.** A central debate among scholars of judicial decisionmaking is whether judges are merely political actors that vote based on their policy preferences or legal actors that try to identify the “correct” legal rule, perhaps looking to prior decisions to determine what that rule may be. We present models that embed these two perspectives. We use data on so-called circuit splits to test these theories. Circuit splits are sequences of conflicting decisions by different appellate courts on the same legal question. We find evidence of streaks in court decisions, inconsistent with a model of purely political, independent judicial decisionmaking. We also find that the first court to decide after an initial split is more likely to side with the court that last addressed the issue, consistent with a model of judicial learning with variable quality records or litigants. We conclude that the data support the view that judges rely on persuasive precedent, consistent with a model of judges as legal rather than purely political actors.

## INTRODUCTION

There are long-running, parallel debates among political scientists and among lawyers about the motivation and behavior of judges when deciding cases. These debates, in their simplest form, pitch those who view judges as mini-legislators that decide cases based on personal policy preferences against and those who view judges as trying to decide cases in conformity with existing law or precedent. We call the former view the political actor model and the latter the legal actor model. In political science, the political actor model is called the attitudinal model and is implicit in the widely-used, ideal-point models used to estimate the policy preference of judges and the positive political theory models used to model strategic behavior among judges. In law, the political actor model is implicit in views of the legal realism and critical legal studies schools, while the legal actor model is implicit in formalism schools.

Both approaches offer an oversimplified view of judicial behavior. Calls have been made for a more nuanced approach (Epstein and Knight 2013), though simplistic models of judges, formally or informally, remain widespread in academic discussions of judicial decisionmaking. They also play a role in judicial politics. On the one hand, candidates for federal judgeships depict themselves as purely legal actors in hearings over their nomination, e.g., using the metaphor of an umpire calling balls and strikes. On the other hand, candidates running for election to state judgeships decide cases to demonstrate to the

---

<sup>1</sup> Baker is Professor of Law and Professor of Economics at Washington University; Malani is the Lee and Brena Freeman Professor of Law and Professor of Medicine at University of Chicago. The authors thank Kaushik Vasudevan, Kevin Jiang, Sarah Wilbanks and Bridget Widdowson for outstanding research assistance. We also thank Charles Barzun, William Hubbard, Richard McAdams, Tom Miles, workshop participants at the Max Planck Institute in Bonn, Northwestern Law School, and ZTH, and participants at the Harvard Conference on Blinding and the L&E Theory Conference at Yale for helpful comments.

voting public that they would have a certain policy preference, e.g., tough on crime (Shephard and Kang 2014).

A precise and direct test of the competing views in a context in which judges have some freedom to demonstrate their preferences has been hard to construct. A finding that judges follow the precedent produced by higher courts is consistent with the idea that lower court judges view themselves as constrained by law but also the idea that lower court judges make decisions with an eye toward avoiding reversal (McNollgast 1995). Likewise a finding that judges follow their own circuit's precedent is consistent with the idea that judges view those decisions as binding or that failure to follow the circuit's own precedent will lead to en banc review and reversal (O'Hara 1993; Rasmusen 1994).

The debate about what motivates judges is important for at least two reasons. First, from a social sciences perspective, it will help explain and predict the behavior of a central institution in American government. Second, from a legal perspective, understanding the motivation of judges affects the legitimacy, whether that term is defined as public approval or as moral authority, of the judiciary. Whether one's concern is positive or normative analysis, the judiciary is one of three branches and has substantial and increasing authority to decide matters, at least for social arrangements such as affirmative action, gay marriage, and privacy regulation.

In this paper we present models that capture the intuition of judges as purely political actors and as purely legal actors that generate contrasting and testable predictions. Our models examine how a sequence of different judges from different federal circuits would rule on the same legal question over time. Each judge's decision is known to future judges and is only persuasive precedent, if precedent at all, for those judges. The context helps provide inferentially powerful tests of judicial preferences. Because the judges are from different circuits, the legal question they tackle is always one of first impression in the circuit. Moreover, because there is uncertainty about the Supreme Court's views, judges are unsure what position will trigger reversal. Finally, because the judge often has conflicting persuasive precedent from which she can draw, she has the cover to decide cases purely on the basis of her political preferences.

We test the predictions of the different models using data on roughly 650 circuit splits identified by U.S. Law Week (USLW) between 1983 and 2014. USLW reports the citations of federal appeals court cases that all decide the same legal question and the order in which they decide the issue. As noted above, because they are different appellate courts, each court's decision is only persuasive precedent for later courts. Because USLW is looking for circuit splits in order to predict which cases might go before the Supreme Court, all our sequences have at least 1 disagreement between courts.

Our political actor model posits that judges care only about deciding cases according to their own private preferences. Under the assumption that the order in which cases reach judges is independent, however, judges' decisions are independent of the history of prior judges' decisions on the same legal question. However, our splits data clearly rejects that prediction. History affects judicial decisions on common legal questions.

Our legal actor model posits that judges care only about getting the right answer to legal questions. Each judge receives a private signal about the correct answer to a legal question, where the signal represents the information provided by litigants. Judges decide cases in a randomly selected sequence and their decisions – though not their private signals – are known to subsequent judges in the sequence. Unlike the political actor model, the legal actor model implies that history matters. Each judge considers not just her own private signal but also the decisions of prior cases. Those cases convey information, albeit imperfect, about the correct answer. However, the models also make predictions about how history matters, depending on whether the quality of information in each case is identical or varies across cases. Our splits data suggest that history matters in a way that suggests that case quality varies, a reasonable finding.

In short, our data favor the view that judges do learn from prior judges and are not purely political actors. Of course these findings have important limitations. First, our analysis really tests whether judicial decisions are independent or not. Although we interpret the political actor model as more consistent with independent decisionmaking, that is not necessarily the case. It is possible judges care about politics and not law, and learn about the political meaning of a case from decisions by prior courts. Alternatively, it is theoretically possible that judges only care about the law but are so confident in their own ability they do not rely on prior decisions. We do not think these possibilities are likely. Moreover, in a future draft we plan to control for the political (Republican/Democrat) alignment of judges to determine whether there is reliance on prior judges beyond wanting to decide the same way prior judges of the same party decided a matter.

A related concern is that, because the political actor model does not take a position on what motivates the judge's preference over case outcomes, it is possible that the model capture "legal" motivations. For example, perhaps judges have a preference not to be reversed by the Supreme Court and their "preference" is their best guess as to what the Supreme Court wants.<sup>2</sup> Thus, even if we were to reject our legal actor model in favor of our political actor model, it is possible that judges are still trying, indirectly, to get the "law" right, where law is defined as what the Supreme Court believes it is.

Second, our models are not the only possible models of judicial behavior. A commonly cited alternative is Judge Richard Posner's labor economics model, which does not include a policy preference or a preference for get the law right (Posner 1993). Another model is that judges care about the embarrassment of being reversed and that they are just trying to predict what the Supreme Court is likely to do. This model does not have obvious implications for the independence of judicial decisions – judges may only be upward looking or may look to decisions of prior judge to predict what the Supreme Court is likely to decide. We do not test this and certain other models.

Third, our models focus on judges' decisions on the assumption that they decide cases by themselves. By contrast our data, in fact, includes the decisions of panels of judges, even as those decisions are

---

<sup>2</sup> Of course, if each judge has the same preference, judges may consider prior judges' decision in order to obtain information about what the Supreme Court wants.

announced by one judge. We ignore how the political or legal actor models play out when there may be strategic interactions between multiple judges with heterogeneous preferences on a panel.

That said, our paper also has some important merits. Methodologically it is more informative than most empirical papers on judicial behavior because it conducts a horse race between multiple models that only differ in their assumptions about judicial preferences. Moreover, to the best of our knowledge, it is the first paper to use data on circuit splits, which, as noted, allows us to separate out the effects of binding precedent from the informational content of persuasive precedent on judicial decisionmaking.

Finally, it should be acknowledged that many legal scholars, particularly those outside the law and economics and political science traditions, may ask whether our extensive apparatus is necessary to demonstrate what to them is intuitively obvious. These scholars will argue that one need only to read any randomly selected judicial opinion to see that history matters: judges constantly cite prior caselaw. Yet this observation shortchanges the legal realism challenge to formalism. Judges who cite prior cases may only be citing cases that validate their political views to provide cover or legitimacy for their behavior. The fact that judges are political does not mean that observers of judges approve of such behavior, and astute political judges may take actions that disguise their political behavior to observers. Our analysis provides a theoretically more powerful test of whether judges are purely political as it makes predictions not about cheap talk (citations), but about costly behavior (decisions).

The remainder of our paper has 4 parts. Section I examines the different theories of judicial behavior in prior literature. Section II presents our political and legal actor models of judicial behavior and the predictions they spawn. Section III presents our data and empirical tests. The last section sets forth our conclusion about the preferences of judges and discusses the merits and limitations of our study.

## **I. LITERATURE REVIEW**

### **A. POLITICAL ACTOR MODEL IN POLITICAL SCIENCE AND LAW**

The political actor model of judging manifests itself in two theoretical models in political science. One is the attitudinal model, which contends that ideology or politics, rather than law, is the primary driver of judicial behavior. Pritchett (1941, 1948) introduced this model, which was subsequently popularized through the work of Spaeth and Segal (2002).

Under this model, judges are assumed to have an ideal point in some policy or case space. Judges, then, decide cases to maximize the degree to which the law reflects those ideals (Gennaioli & Shleifer 2007; Lax 2011). The framework is transplanted from work done on voters and legislators. Empirically, what that means is that the same techniques for identifying the underlying policy preferences carry over: one simply looks at judicial voting patterns (Martin and Quinn 2002).

The second model falls under the rubric of positive political theories of judging. These models look at strategic interactions between judges or between judges and other political actors that care only about their personal policy preferences (Marks 1988; Guy and Spiller 1992; Jacobi 2006; Tally and Spitzer

2013). In many of these models, judges decide cases to the degree they can without provoking a response by other institutional actors.

Although the political actor model is better formalized in political science, it has a long history in legal scholarship too. Legal realists, such as Walter Cook, Leon Green, Karl Llewellyn, and Herman Oliphant suggested that one couldn't identify the true driver of a judicial decision by looking at the stated reasons and doctrines articulated in the decision (Leiter 2003). Instead, the reasons given were often post hoc rationalizations for a result reached on other grounds. The legal realists looked for patterns in the cases and sought to identify the underlying factors that better explained the decisions (Leiter 2003 at 8-9). In some sense, the legal realists were not that different from early law and economics scholars: both tried to translate abstract legal concepts (which they felt did little work in the actual cases) into workable ways to better catalog and characterize the case outcomes. The critical legal studies movement did much of the same. Like the legal realists, these scholars felt like that the legal materials – the language of the statute, the prior precedent – rarely constrained the judge. Citations to these materials justified results reached on other, less legitimate grounds such as a desire to maintain the status quo.

## B. LEGAL ACTOR MODEL IN LAW

The legal actor model posits that judges feel constrained by the legal materials and doctrine (Tiller and Cross 2006). This assumption is uncommon in the political science literature (Knight and Epstein 2013), but arises in several strands the legal literature.<sup>3</sup> “Doctrine-as-constraint” is not the only aspect of a legal model of judging. Even when the legal materials “run out” the assumption is that judges try and make the decision that best fits with the legal materials, not necessarily the decision that comports closest with their personal policy preferences. Importantly, the legal actor model assumes judges start with the legal materials. These materials aren't simply rolled out as after-the fact-dressing. Instead, the process follows a pattern. The appellate judges receive a case. They listen to the lawyers and look at the record. They then consult the precedent from the higher court and from other circuits. They think about whether the function of the legal rule would be served better by deciding the case one way rather than another. Given all the materials and information, they make the best decision they can.

Baker and Mezzetti (2012) formalize aspects of the legal model and show how the efficient use of judicial resources can explain patterns of behavior often identified as evidence of the political model (like, for example, the exact same case being decided differently at different points in time).

## C. PRIOR EMPIRICAL WORK ON JUDICIAL PREFERENCES

Prior empirical work on judicial behavior is vast. Epstein, Landes, and Posner (2013) contain a useful summary. We mention a few of the findings here, before suggesting why our data is especially well-suited to investigate the motivation of judges.

A large number of studies have studied the effect of politics on judicial outcomes. They have found evidence of political party-associated panel effects on judicial voting patterns. For example, a

---

<sup>3</sup> [Insert note on legal process school and originalism.]

Republican-appointed judge is more likely to cast a conservative vote if he sits with two other Republican-appointed judges (Sunstein et al. 2006). Studies have looked at the influence of ideology (typically proxied by the party of the appointing president) on judicial votes (e.g., Chen, Levonyan, and Yeh 2014). They have also looked at the degree to which circuit judges cite decisions by judges appointed by a different party (Choi and Gulati 2008).<sup>4</sup> At the same time, other studies have offered suggestive evidence that courts care about law, including evidence that own-circuit precedents influence federal district courts (Chen, Frankfurter and Yeh 2014) and that the Solicitor General's impact on Supreme Court decisionmaking is unaffected by the political complexion of the Supreme Court (Epstein and Knight 2013).

Our data and analysis is a bit different. Unlike many studies, we do not need to rely on the coding of decisions as "liberal" or "conservative" to test competing theories. We are just looking for patterns: whether say, two circuit court decisions going in one direction make the next circuit more likely to rule that same way. As noted, all our dispositions are ones of first impression in the circuit. They aren't bound by existing precedent. The judge has a lot of freedom to decide and multiple sources of law to draw from.<sup>5</sup> Moreover, our analysis attempts to run a horse race between two models that really just differ on one assumption: that the preferences or private signal of judges in a sequence of decisions on a common legal question are independent. Thus, our empirical analysis is a bit more theoretically driven and discriminating than most prior studies of judicial decisionmaking.

That said, we are not the first to consider the empirical fact of a circuit split in explaining judicial behavior. Prior studies have used the fact of splits to estimate the probability of Supreme Court review (Bowie and Songer 2009) and have examine how the nature of split affects Supreme Court substantive decisions (Lindquist and Klein 2006). However, we believe we are the first to use the sequence of case decisions and to use split data to distinguish between political actor and legal actor theories of judicial behavior.

## II. THEORIES AND PREDICTIONS

This section presents three different models of judging in federal appeals courts in the context of persuasive precedent from other circuit courts. The first part presents a political actor model of appellate judges that corresponds to a legal realist or attitudinal model judge. The motivating assumption is that judges decide cases to maximize the degree to which the law in their circuit accords with their policy preferences. We do not take a position what those preferences are or why they are

---

<sup>4</sup> Although unrelated to political or legal motivation, studies have also examined the impact of increasing caseloads in one area of law (immigration) on the affirmance of trial court decisions in other areas of law (Huang 2011) and even the degree to which judge might be hungry or not affects the case disposition (Danziger et al. 2011).

<sup>5</sup> It is possible that circuit court judges make decisions with an eye to avoiding reversal by the Supreme Court, so with less freedom than we suspect. However, Klein and Hume (2003) found little influence of Supreme Court predicted preferences in search and seizure cases on circuit court decisionmaking and concluded that risk of reversal did not have a big effect on circuit court decisionmaking. This is consistent with the obvious empirical point that the probability of Supreme Court review of an appellate decision is rare.

what they are. They could be chosen because they maximize career opportunities, minimize effort, or satisfy (altruistic) political views.

The second section presents a legal actor model that presumes a “right” or “correct” answer to the legal issue facing the circuits. Judges, then, maximize the degree to which their decision accords with the correct one. From the litigation, each judge observes a private signal correlated with the correct resolution of the case. These signals correspond to the information presented by the litigants and in the case record. Each judge’s signal is presumed to be independent but equally correlated with the correct decision, i.e., of equal quality. Each judge also observes the decisions reached by the other circuits. In deciding what to do, a judge combines the information extracted from their own litigation with the information contained in the prior decisions by other circuits, which has persuasive appeal. The model is basically the information cascade model from the economics literature (Banerjee 1992; Bikhchandani et al. 1992). A stark and somewhat highly unlikely implication of this model is that any time two circuits in a row decide the case the same way all subsequent circuits herd on that ruling.

The final section presents a more realistic legal model where the information contained in the record varies across cases. In this model, judges observe prior decisions by the other circuits but do not observe the degree of confidence the judges had in those decisions – the strength of the record before that court. Instead, judges make inferences about the strength of the reasoning in prior decisions based on the order in which those decisions were made. Based on these inferences, the judge decides how much to weight each one of the prior opinions (i.e., should she listen more closely to the second decision or the first decision). This model yields a more likely prediction that herding, while possible, is not always sudden.

#### A. POLITICAL ACTOR MODEL

The issue space is the unit interval. Without loss of generality, assume the most conservative position is 0 and the most liberal position is 1. The legal issue can be resolved in one of two ways: A or B, where  $0 < A < B < 1$ .

Take, as an example, a case involving the Credit Repair Organization Act. The Act provides a statutory private right to sue for its violation. The courts were asked to decide whether the right to sue in litigation could be waived in favor of arbitration. The Third and Eleven Circuits said yes. The Ninth Circuit said no. The judges had one of two choices. They could hold that the statutory right of action could or could not be waived via arbitration. Those choices correspond to A and B in the model.

We assume judges have a policy preference  $\alpha$ . This preference is distributed over the issue space -- the unit interval -- according to the distribution  $F(\cdot)$ . The idea is that some judges are more conservative or more liberal than other judges. The overall makeup of the judiciary is captured by the location of the “mass” of judges at each point on the interval. If the judiciary is relatively conservative there will be more judges located close to zero than if the judiciary is relatively liberal.

If the judge decides A, her loss is the absolute value of the distance between the issue A and her preferred policy preference  $\alpha$ . If the judge decides B, the loss is the absolute value of the distance between the issue B and her policy preference  $\alpha$ .

For example, suppose the judge has a policy preference of .1, i.e., she is a pretty conservative judge. Further suppose that resolving the issue as A institutes into the law a conservative position of .3, whereas resolving the case as B institutes in the law a liberal position of .6. This judge's loss from deciding the case as A equals .2. Her loss from deciding the case as B equals .5. As a result, this judge prefers deciding the case as A rather than B.

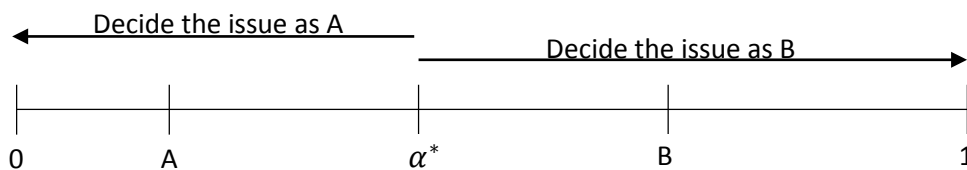
The judicial decision is constrained by the way the issue is presented and the possible ways it can be resolved. In the above example, the judge prefers to institute a policy of .1. Perhaps, to return to our example, he believes that the Credit Repair Organization Act should never have been enacted. That said, he is unable to invalidate the entire act. Instead, he is constrained to one of two choices when it comes to interpreting the statute: either the right to sue can be exercised through arbitration or not.

Mathematically, we can represent the judge's loss from deciding the case as A as  $-\lvert\alpha - A\rvert$ . His loss from deciding the case as B is  $-\lvert\alpha - B\rvert$ . Clearly, all judges with policy preferences below A will issue a decision of A and all judges with policy preferences above B will issue a decision of B. The marginal judge – the one indifferent between deciding the issue as A and as B – is determined by  $\alpha - A = B - \alpha$  or

$$\alpha^* = \frac{B + A}{2}$$

The probability a case is decided as A is just the probability that the preference of the judge deciding the case lies below  $\alpha^*$ . Given the assumption about the distribution of judges in population, this probability is just  $F(\alpha^*)$ . The probability the judge decides the issue as B soaks up the remainder and equals  $1 - F(\alpha^*)$ . Figure 1 represents this analysis visually.

*Figure 1. Mapping of preferences onto decisions (A or B).*



The model predicts that the probability of an A decision will be determined by the mean value of A and B. Moreover, if the bulk of the distribution of judicial preferences in a circuit lies below  $\alpha^*$  – i.e., most of the appellate judges are conservative – an A decision will be more likely than a B decision. Finally, as the mean of the two possible legal outcomes moves to the left more judges will be inclined to adopt the less liberal position. Take an example where A moves from .4 to .6, while B remains at .7. Following this shift, more judges will be inclined to decide the issue as A.



For this model of judge to generate a sequence of decisions, we start with a given legal question and then select judges at random from the pool of all judges. This is equivalent to drawing  $\alpha$ 's from  $F$ . We also draw a random number between 2 and 13 (the number of circuits) to be our sequence length and stop picking judges after we hit this sequence length. The order in judges are picked from the pool is the order in which the answer the given legal question. Because the draw of judges is random, the draw of judicial preferences is i.i.d.

The political model now generates some predictions about the impact of horizontal persuasive authority – the decisions of sister circuits – on judicial decision-making. The attitudinal model or legal realist judge, as we model it here, is not persuaded or influenced by the decisions of the other circuits. One reason is that decision order is assumed random. Another is that judges only care about their own preferences and neither directly or indirectly the preferences of judges deciding case before them.<sup>6</sup> Thus, the probability a judge decides a case in a specific way is independent of the decisions made previously by other circuits. This yields the following predictions, the second of which will contrast with a prediction of one of the legal actor models.

*Prediction 1a: In the political actor model, the history of prior case dispositions should have no effect on the judge's resolution of the issue.*

*Prediction 1b: Among other things, Prediction 1a trivially implies that if the history of prior cases is balanced, i.e., it contains an equal number of A and B decisions, then then the probability of an A decision in a case is  $F(\alpha^*)$ .*

## B. LEGAL ACTOR MODEL (CONSTANT QUALITY CASES)

Our legal actor model for circuit splits replicates the classic herding models (Banerjee 1992, Bikhchandani et al. 1992) and follow closely ones already in the law and economics literature (Daughety and Reinganum 1999, Talley 1999).

As before, judges decide cases as either A or B and in sequence, the order in which judges hear cases in a sequence is random, and sequence length is randomly selected. However, as between A and B, there is now a "correct" rule  $R$ , i.e., one that describes the true (legal) state of the world. Judges start with an uninformative prior about the correct rule (i.e.,  $Pr(R = A) = Pr(R = B) = (1/2)$ ).

Before issuing a decision  $D$ , each judge observes a private informative signal  $s$  from the record in the case before them about the correct outcome. The private signal, which is independently and identically distributed across judges, takes on values  $s = a$  or  $b$ . The probability the signal correctly identifies the correct outcome is  $\pi = Pr(s = a|R = A) = Pr(s = b|R = B) \in (.5, 1]$ .<sup>7</sup> Judges also observe decisions

---

<sup>6</sup> We assume that a judge's  $\alpha$  is not defined by the  $\alpha$  of the judges before her. For example, the 4<sup>th</sup> Circuit does not systematically learn from earlier 9<sup>th</sup> Circuit decisions about what the, say, conservative view on a legal question is. Nor do judges have a direct preference for being followers or contrarians. They simply have a preference  $\alpha$  such that, if the sequencing of decisions is independent, then the decisions are independent.

<sup>7</sup> The private signal in this legal actor model is analogous to the draw of  $\alpha$  in the political actor model.

rendered by prior judges. However, they do not observe the private signals – the quality of the records – received by these prior judges.<sup>8</sup>

Judges reap a payoff 1 from a correct decision and zero from an incorrect decision. Given these payoffs, a judge will decide  $D = A$  if she believes that the true state is more likely to be A than B and  $D = B$  otherwise. Even though sequence order is random, these beliefs, of course, will be based on the information extracted from the prior judicial decisions and the value and precision of the judge's own private signal. If a judges beliefs suggest that the correct answer could equally be  $R = A$  or  $B$ , the judge will go with her own private signal.

Studying a sequence of three circuit court decisions is sufficient to generate the key predictions of this model. Given that judges start with uninformative priors, the first judge follows his own private signal. If the signal is  $s = a$ , she decides  $D = A$ . If the signal is  $s = b$ , she decides  $D = B$ . The second judge observes the decision of the first judge and can infer from that the private signal that judge obtained. Suppose that the first judge decides  $D = A$ . The second judge issues a decision agreeing with the first judge, i.e.,  $D = A$ , if she receives a private signal  $s = a$ . She splits and issues a decision  $D = B$  upon receiving signal  $s = b$ .

Now consider the third judge. First, suppose that judge 3 observes the prior two circuits decide the case as A, i.e., sees a history  $H = AA$ . This judge knows that both judge 1 and judge 2 received private signals with a value  $s = a$ . As a result, even if the third judge received a private signal  $s = b$ , she will rationally ignore her own conflicting private signal and go with the “herd”, deciding  $D = A$  as well.<sup>9</sup> The reason is that the information contained in two  $s = a$  signals outweighs the information contained in one  $s = b$  signal.

Finally, consider the fourth judge. She learns nothing from the decision rendered by the third judge. The third judge would have decided  $D = A$  whether she received an  $s = a$  or  $b$  private signal. So the fourth judge sits in the same position as the third judge in terms of the information available in the case law. She therefore makes the same choice, no matter the value of her own private signal. She also decides  $D = A$ .

In this model, the impact of two prior identical decisions is stark. The moment any judge sees this, she ignores her own private signal and follows precedent, yielding the following predictions.

---

<sup>8</sup> Of course readers may get some sense of the confidence a judge has in her opinion by reading her opinion. However, all our model requires that the opinion does not fully convey the confidence that the judge had in the record. That is a more reasonable assumption and enough to justify our simplified modeling assumption.

<sup>9</sup> The analysis can be confirmed by an application of Bayes rule. Suppose that the first two judges decided  $D = A$  and the third judge received signal  $s = b$ . In that case, Bayes rule implies that

$$\Pr(R = A|s = b, H = AA) = \frac{\Pr(s = b, H = AA|R = A) \Pr(R = A)}{\Pr(s = b, H = AA|A) \Pr(R = A) + \Pr(s = b, H = AA|R = B) \Pr(R = B)}$$

Because the signals are conditionally independent and the prior is uninformative, this expression can be written as  $(1 - \pi)\pi^2 / [(1 - \pi)\pi^2 + (1 - \pi)^2\pi]$ . By contrast  $\Pr(R = B|s = b, H = AA) = (1 - \pi)^2\pi / [(1 - \pi)\pi^2 + (1 - \pi)^2\pi]$ . Since  $> 1 - \pi$ , it follows that  $\Pr(R = A|s = b, H = AA) > \Pr(R = B|s = b, H = AA)$ .

*Prediction 2a: In the legal actor model with constant quality signals, history generally matters. For example, the probability of observing an outcome  $D = A$  ( $D = B$ ) following two  $A$ 's ( $B$ 's), i.e.,  $H = AA$  ( $H = BB$ ), is one.*

*Prediction 2c: In the legal actor model with constant quality signals, the probability of that the  $n$ -th case in a sequence is the first to split with prior decisions is 0 for  $n > 2$ .*

While history can have a big impact in this legal actor model, there are certain circumstances when history does not matter. Let's go back to the third judge but now suppose that she observes a split in the first two cases, e.g.,  $H = AB$  or  $BA$ . The judge knows that both judge 1 and 2 received opposite private signals. But she also knows that the quality of those signals are identical. Thus the third judge is in the same position as judge 1 and will decide based solely on her private signal. Indeed, this logic suggests applies to any future judge facing a balanced history of prior cases with no set of consecutive identical decisions. This yields the following prediction.

*Prediction 2b: In the legal actor model with constant quality signals, the probability that a case with a history that is balanced and include no consecutive identical decisions will choose  $A$  is the probability that the private signal is  $s = a$ .<sup>10</sup>*

To test this prediction, we need to know the probability that a judge receives a private signal  $s = a$ . Unfortunately this is not observed and our best estimate of it, the number of  $A$  decisions in a sequence divided by the length of the sequence, is rather imprecise. However, given that the order in which judges hear a case is random, the decision that comes first in a sequence clarifies which signal has expected probability 0.5 or greater. Thus, any signal that differs from the first decision should on average have probability 0.5 or lower. Thus, we can test the modified version of the preceding prediction.

*Prediction 2b: In the legal actor model with constant quality signals, the probability that a case with a history that is balanced and include no consecutive identical decisions will choose the last decision in the history is at most 0.5.*

### C. LEGAL ACTOR MODEL (VARIABLE QUALITY CASES)

It seems unrealistic to assume that all judges receive the same quality signal from underlying litigation. Some sets of fact are more informative than others. Lawyers vary in quality. So some records provide a better source of information for judges than others. To capture these differences, we make just one change to the model of the preceding section. We assume that the precision of the private signal can take one of two values. With probability  $p$  the private signal is barely informative, having a value  $\pi = \pi_L = 0.5 + \varepsilon$ , where  $\varepsilon > 0$  but "small." With probability  $1 - p$ , the signal is highly informative, taking a value  $\pi = \pi_H > \pi_L$ .

---

<sup>10</sup> We rule out consecutive decisions because, with herding, two consecutive decisions have less average information content than a single decision. This the sequence ABAB generates Prediction 2b, but sequence ABBA does not.

As in the analysis above, judges observe decisions rendered by prior judges. They do not observe the private signal received by past judges or the precision of those signals. An important contrast is that, whereas in the prior model judges know the quality of each judge's private signal, in the present model judges know the quality of their signal but not those of other judges.

Consider the decision-making of the first three different circuit courts to hear the same issue. The first judge receives her signal. Because the signal is informative (even if only slightly), she decides the case according to her signal. If the signal says  $s = a$ , she decides  $D = A$ , and vice versa.

The second judge to consider the issue receives a new private signal. If that signal lines up with the decision rendered by the first judge, he decides the case the same way. If, however, the signal differs from the decision reached by the first judge, the second judge might or might not create a split.

Suppose that the first judge decided the issue as  $D_1 = A$  and the second judge obtained a conflicting private signal  $s = b$ . Further suppose that the quality of that signal was high,  $\pi_H$ . The second judge will follow her own conflicting high quality signal rather than deferring to the judgment of the first judge. The second judge knows that the first judge received a  $s = a$  signal. That much is revealed by the first judge's decision. The second judge is unsure, however, whether that signal was of good or bad quality – based on a good or bad record or strong or flimsy reasons. The expected quality is  $E[\pi] = p\pi_L + (1 - p)\pi_H$ . By contrast, she knows for sure that her own conflicting signal has good quality,  $\pi_H > E[\pi]$ . As a result, she will follow her own conflicting signal because it is higher expected quality, i.e., more likely informative about the correct rule. On the other hand, if the second judge's conflicting signal  $s = b$  was of low quality,  $\pi_L < E[\pi]$ , she would ignore it, preferring to follow the first judge because the first judge's signal was of higher expected quality. There is some chance, after all, that the first judge made her decision based on a good record.

Now let's turn to the third judge. First consider what happens if the first two judges agree, i.e., the third judge observes a history of  $H = AA$ . In the legal actor model with constant quality cases, as in the standard herding model, the third judge follows, no matter his signal. In the legal actor model with variable quality cases, however, the third judge might or might not follow. If the quality of the high quality signal is strong enough, she will create a split upon seeing such a signal.

Suppose it is, then the probability of judge 3 being the first to split given that judge 1 decided A is

$$\begin{aligned} \Pr(\text{judge 3 is first to split}) &= (1 - \Pr(s_2 = b, \pi_2 = \pi_H)) \Pr(s_3 = b, \pi_3 = \pi_H) \\ &= (1 - \Pr(s = b) (1 - p)) \Pr(s = b) (1 - p) \\ &< \Pr(s = b) (1 - p) = \Pr(s_2 = b, \pi_2 = \pi_H) = \Pr(\text{judge 2 is first to split}) \end{aligned}$$

which is less than the probability that judge 2 is the first to split. If the quality of judge 3's signal is not high enough to cause judge 3 to split, then the probability she is the first to do so is 0. Thus, we know that, whatever the quality of the high quality signal, the probability that just 3 is the first to split is smaller than the probability that judge 2 is the first to split. The converse of this result is that judge 3 is more likely to follow than judge 2.

Because this logic generalizes to subsequent judges, we obtain the following predictions.

*Proposition 3a. In the legal actor model with variable quality cases, history matters. For example, the probability of following precedent increases with the number of circuits that follow precedent before.*

Now consider the decision of a third judge following a split, e.g., following  $H = AB$ . In creating the split, the second judge signaled that she must have made the decision based on a strong record or strong reasons. Otherwise, she wouldn't have been so bold. The first judge, by contrast, could have made her decision on a weak or a strong record. That information will be useful to subsequent judges. Before considering her own private signal, the third judge is more likely to believe that judge 2 got the decision right as opposed to judge 1.

Now suppose that the third judge draws the  $s = a$  signal, but that signal is not terribly informative. Maybe the lawyers did not present that compelling case for deciding the case as A. What should the third judge do? Should he follow his own signal rendered from a lousy record or defer to the decision-making of the second judge -- a decision he knows was made on better information? The following initial result identifies when a judge following a split will discard the information contained in his own private signal and defer to the judge who created the split.

*Remark (1) Following any split, a judge who receives a weak private signal that conflicts with the decision of the second judge will ignore his own private signal and defer to the decision made by the second judge. (2) Following any split, a judge who receives a weak private signal that conflicts with the decision of the first judge will follow his own private signal.*

If the third judge draws a highly quality signal, he always follows that signal. Combining with the results in remark 1, this insight allows us to compute the probability of observing the sequence of decisions  $ABB$ . That is to say, the decision sequence where the first two judges split and the third judge follows the second judge's position. We can also compute the probability of observing the sequence  $ABA$  -- the path where the first two judges split and the third judge follows the first judge's position. Because of the second judge must be really confident to create a split, the third judge will always find that decision more persuasive. The first path is more likely to occur. Formally, we have:

*Prediction 3b: In the legal actor model with variable quality private signals, it is possible for a case with a balanced history that includes no consecutive decisions to follow the last case in the history with probability greater than 0.5. For example, a case with history  $H = AB$ , will be more likely to be decided B than A if  $p > 2 - (1/\pi_H)$ .*

This proposition stands in stark contrast to the political actor model or the legal actor model with constant quality cases, where decision following a balanced histories yields at best equal probabilities of  $D = A$  or  $B$ .

### III. EMPIRICAL TESTS

Table 1 summarizes the predictions our three theories make about the decisionmaking of a sequence of circuit courts deciding a common legal question. Although the theories sometimes generate similar predictions, there are enough that differ that we can discriminate to some extent between the theories.

*Table 1. Predictions from three models of decisionmaking by sequence of circuit courts.*

Subject of predictions	Theory and description		
	Political actor: Independent decisions	Legal actor: Persuasive precedent - constant quality cases	Legal actor: Persuasive precedent - variable quality cases
Dependence upon history	Pred. 1a. No dependence on history	Pred. 2a, 2c: Probability of following precedent is one after two consecutive cases with the same result.	Pred. 3a. Probability of following precedent rises with number of prior cases following precedent.
Decision in case with balanced history	Pred. 1b: Probability a case with a balanced history follows the very last decision is 0.5 or less.	Pred. 2b: Case with certain balanced histories (AB, BA, ABAB, BABA, etc.) are equally likely to follow or not follow the very last decision.	Pred. 3b: A case with a balanced history is more likely to follow the very last decision.

## A. DATA

We employ data on sequences of circuit court decisions that tackle a common question. These data are drawn from sequences identified in the publication U.S. Law Week (USLW), which seeks sequences with at least one circuit split. (Circuit splits are a thought to be a good predictor of whether the Supreme Court takes up a legal question.). We conducted our data gathering by working backwards from the January 2015 issues of USLW and gathering the last 1000 splits recorded in that periodical. So far gathered 977.<sup>11</sup> Because USLW only reports on splits, we do not have any sequences in which all courts make the same decision about a legal question.

We employ two different methodologies for gathering our data. For most sequences featured in USLW, both the latest case in the sequence is reported in a Circuit Split roundup at the beginning of the USLW and the entire sequence is reported in a separate article. But for 264 sequences, USLW did not report the entire sequence in a separate article. For the latter sequences, we had law student research

<sup>11</sup> Specifically, we gathered 1001 splits but one of them started with a Supreme Court decision and was followed by a divergent circuit court case. We dropped that observation. We have 23 sequences for which we are still recording decisions.

assistants manually identify the cases in the sequence and code the decisions in those cases. As a result, our data varies in how cases and decisions in a sequence are identified.<sup>12</sup>

Our initial data take the form of sequences. E.g., AAB, ABBAB, AAAABC, etc. A sequence is defined as a series of circuit court cases addressing the same legal question. The first decision in a sequence is always assigned the label A. The first *different* decision in the sequence is assigned the label B.<sup>13</sup> If the sequence yields a third, fourth, or fifth different decision, that will be called C, D and E, respectively. Because our models only allow two states of the world, we artificially terminate each sequence just before any C decision appears. Thus AAAABC, for example, would become AAAAB.

We gathered a number of covariates for each of the cases in our data. We obtain these from Google Scholar searches on each case. From Google Scholar, we are able to scrape for each cases the circuit and date of decision. For most cases we are also able to obtain the names of each judge on a panel and are in the process of scraping whether the decision was unanimous or with dissent. We use the judges' names to extract biographical data, such as appointing President (and his party), from the Federal Judicial Center (FJC) database available at <http://www.fjc.gov/history/home.nsf/page/export.html>. This matching is incomplete, in some cases because our records from Google Scholar are imperfect and in other cases because data on the deciding judges are not in the FJC data. For example, some cases are decided by state supreme courts and those judges are not the in FJC data, which only contain data on federal judges. Where the matching is complete, our regression sample size may be lower than the total number of sequences or cases in the data.

Table 2 provides a statistical description of the case sequences in our data. To check if the manner in which we code decisions had an effect on the data we provide summary statistics separately by whether decisions were coded by USLW or our research assistant.<sup>14</sup> Our own hand-coding of sequences identified generates longer sequences – by number – but shorter sequences – by duration -- and a smaller fraction that are truncated because of a C decision. In any case, only the difference in the fraction truncated is statistically significant. To test whether differences in data gathering methodology has any effect on outcomes, we run each empirical test once on the whole sample and once on just the subsample of decisions coded by USLW. We report any instances in which the two samples give different results.

Sequences can (and do) have different lengths. This can happen for multiple reasons. First, as noted, we chose to terminate sequences once we observed a third possible resolution of the legal issue. Second, we chose to stop observing a sequence the date it was reported in USLW. The reason for doing so is that

---

<sup>12</sup> USLW is available online until mid-2007. Prior to that it is only available in paper form. However, the format of the online and paper versions are similar and so the data should not vary depending whether USLW was online or paper-only. We verified this by comparing summary statistics for the online and paper-only data subsamples. Sequence lengths are slightly shorter (3.736 v. 4.0002) and the fraction of splits truncated because of C decision is higher (0.0116 v. 0.015) for splits reported in online versions, with both difference significant at 95% confidence level. Differences in other covariates are not significant.

<sup>13</sup> For certain statistical tests, it is helpful to randomly assign A or B with equal probability to the first decision. When we assign B to the first decision, the first different decision is assigned A. We will clarify when we do that.

<sup>14</sup> Appendix Table 5 and Table 6 do the same for the distribution of cases across circuits and time.

we wanted the method for coding a decision to be consistent within sequences. Third, the sequence may be terminated because the Supreme Court took up the legal question addressed in the sequence and resolved it, though we do not have the Supreme Court decision in the sequence.<sup>15</sup>

Although these truncations may potentially cause selection in the distributions of A's and B's over positions in a sequence, we find little evidence of this. First, Table 3 replicates Table 2 except for sequences truncated because they had a C decision. We find, surprisingly, that truncated sequences are significantly longer (by number of slots), perhaps because they have to have at least 3 decisions to get a C decision in the first place, but also significantly shorter (by duration). (Note that the table only tests for whether one cause of truncation – C decisions – impacts the data.) Second, we regressed an indicator for A decisions on a sequence length and found that the coefficients on sequence length (coef=0.019, se=1.21, 0.772) and length squared (-0.000, 0.14) were small and insignificant. Finally, as a precaution, we either include sequence fixed or random effects or include sequence length and length squared as coefficients in our regression analysis where possible.

## B. SPECIFIC TESTS

### 1. *Random order of circuits*

A critical assumption of our political and two legal actor models is that the judges – hence courts or circuits – are chosen in random order. If this assumption were violated, then it could be that judges do not consider outcomes in prior cases but decisions are not independently distributed over time because, e.g., litigants seek out courts where outcomes are likely to be similar to those in prior cases. Likewise, it is possible that judges do consider outcomes in prior cases but that decisions look independently distributed because, e.g., litigants seek out courts where outcomes are likely to differ from prior cases. In other words, random court ordering makes it more likely that correlation in case outcomes identifies judicial learning rather than other mechanisms that may influence case outcomes.

If court order in each sequence were random, then the probability of observing court  $k$  in slot  $m$  in a sequence would be the same as observing court  $k$  across all slots in a sequence. So, to test our assumption of random order, we conduct a non-parametric, bootstrap-derived test of whether the probability of observing a decision from court  $k$  in slot  $m$  is not significantly different than the probability of observing court  $k$  in any slot.<sup>16</sup> To implement the test, for each actual sequence of cases in the actual data, we draw a new sequence of cases from the actual cases in that sequence with replacement, i.e.,

---

<sup>15</sup> There are two exceptions, both fully reported by USLW. One sequence started with a Supreme Court case and had 1 divergent circuit court cases; we dropped this observation. Another sequence ended with a Supreme Court case; we kept the sequence but dropped the Supreme Court case because the latter resolved the circuit split. Many other sequences with splits may have ultimately been

<sup>16</sup> We also conducted a parametric test where we examined whether the fraction  $f_{km}$  of cases from court  $k$  in slot  $m$  in the actual data were significantly different than the fraction  $f_k$  of cases from court  $k$  in any slot in the actual data. We assumed that the fraction of cases from court  $k$  in slot  $m$  was a random binomial variable with mean  $f_{km}$  and variance  $f_{km}(1 - f_{km})$  and tested that this was different than the fraction of cases from court  $k$  which was the constant  $f_k$ . Our analysis found 6 court-slot combinations that were significant. Even then, because we are running a test for each court-slot combination, i.e., 126 tests, up to 6.3 can be significant even if the truth is that court order is random. So we can still not reject the hypothesis that court order is non-random.



we randomly reorder each sequence. We do this 5000 times for each sequence, resulting in 5000 drawn sequences for each actual sequence. Each sequence of decisions – actual or drawn – implies a sequence of courts that hear the legal question the actual sequence addressed. Thus for each actual sequence of courts we have 5000 alternative drawn sequences of courts. We assemble the n-th draw on each sequence into a data set representing the n-th draw on the set of all sequences, resulting in 5000 draws on the entire set of sequences. Then, for each court and slot combination, we look at the fraction of draws on the set of all sequences that have fewer cases from that court in that slot. If the fraction is below 0.025 or above 0.975, then we can reject random order with 95% confidence in a two-sided test.

Table 4 reports the fractional of draws that have fewer cases from court k in slot m than the actual data. No court-slot combinations in the actual data are significantly more or less likely than from randomly drawn data.<sup>17</sup> Thus we cannot reject random ordering with 95% confidence.

## *2. Independence of decisions*

The political model asserts that a judge's decision does not depend on – i.e., is independent of – the outcomes of prior cases on the same subject. The legal actor models each predict that history has a positive influence on a judge's decision, though in particular ways. Our tests examine these different assumptions and predictions.

We begin by testing the independence assumption of the political actor model without testing the specific ways in which history matters for decisions. Our initial test – a check for the number of runs in the data – is unconditional on covariates, i.e., looks at whether there is correlation across decisions within a sequence, but does not condition on other variables such as sequence length or party of the authoring judge. If the actual data were comprised of correlated draws on A or B, the number of runs it contains would be lower or higher than the number were the actual data comprised of independent draws. (Here we define a run as a set of adjacent decisions that are identical; e.g., the sequence AAAB has 1 run, AAA.)

Because we have a large number of small sequences of variable length, we cannot employ, e.g., the Wald-Wolfowitz approximation (Wald and Wolfowitz 1940) for the distribution of runs in a sequence. Even exact distributions are not helpful because many of the sequences are so small that there are less than 20 combinations and so we can never achieve 95% confidence that the data are not independent. Our solution is to sum up the runs across sequences, but employ a bootstrap technique to calculate an empirical distribution for the total number of runs across a large number of small sequences that vary in size and probability of A. Specifically, we randomly reorder the A and B's in each sequence in the actual data. Because we only reorder sequences, we ensure that the number of A's in each sequence is always the same and the probability of A varies across sequences. We then count the total number of runs across the reordered sequences. We repeat this 5000 times to generate the empirical distribution of

---

<sup>17</sup> The closest combination to significance is a state court in the 8<sup>th</sup> slot, which is more likely in the actual data than in 96% of draws, implying only 8% significance in a 2-sided test. Moreover, with a multiple testing adjustment, we still cannot reject that the state court in the 8<sup>th</sup> slot is random.

runs. The mean and variance of runs in the empirical distribution is 737.87 and 173.98. We are able to reject that the actual data, with 772 runs, are independent with a p-value of 0.0046.

### *3. Cases with a Balanced History*

Our third test examines Prediction 3b, which helps discriminate between a legal actor model with constant quality cases on one with variable quality cases. Recall the former, like the political actor model, predicts that the probability that a case with a balanced history without consecutive decisions, i.e., an equal number of A and B decisions preceding it and no repeats, follows the immediately prior case is at most  $1/2$ . If, for example, the sequence started out with an A, in expectation the probability of signal  $s = b$  is less than or equal to 0.5. By contrast, the political actor model with variable quality cases predicts that a case with a balanced history can be more likely to follow the decision in the case that immediately proceeds it.

Table 5 presents unconditional, exact hypothesis tests of Prediction 3b for cases with balanced histories of length 2, 4, and 6. We find that the probability of following is greater than 0.5 for each of these balanced history lengths. However, only cases with balanced histories of length 2 have a probability of following that is significantly different than  $1/2$ . Those with balanced histories of length 4 are marginally insignificant at the 90% level. The insignificance of results for cases with balanced history lengths greater than 2 may be due to small sample size: the number of cases with balanced history lengths of 4 and 6 are much smaller (84 and 26, respectively) than the number of cases with balanced histories of length 2 (333).

## **CONCLUSION**

Our paper conducts a horse race between models of purely political judicial decisionmaking, as implied by the attitudinal model or the legal realism school, and models of purely legal decisionmaking, as implied by the formalism school.

Our model of a purely political judging assumes that each judge rules only on the basis of their preference – the prior history of decisions is irrelevant – and the distribution of preferences is independent across judges. The model generates a prediction that the history of decision by prior judges has no effect on the decision of a present judge. Our data reject that prediction. The history of cases is positively correlated with a judge's decision.

We offer two purely legal actor models of judging. Both are based on models of herding or information cascades. These models assumes each judge cares only about identifying the correct legal decision, i.e., the true state of the world and each judge gets a private signal about whether the true state is A or B. This private signal reflects information that judges receive from the record and arguments before the judge, information that is hard to fully convey to subsequent courts. The fact that judges all seek to identify the true state of the world and observe prior decisions means that history matters, i.e., judges decide cases based on their private signal and also the public decisions – though not the private signals –

of prior judges. The difference between our two legal actor models is in the distribution of the quality of the private signal – the probability that the signal indicates the answer is A when the true state of the world is A – across judges. This distribution of private signals will affect the way in which the history affects a judge's decision.

In one legal actor model, the quality of the private signal is assumed to be constant across judges. On the one hand, this implies strong herding, i.e., the probability of the first split vanishes after identical decisions by the first two judges addressing a common legal question. On the other hand, if the prior decision of judges are balanced, e.g., the first judge says A and the second says B, the probability that the third judge says B is roughly the same as the probability that she gets a private signal that says B is the true state, which probability is at most 0.5 in our example. Our data reject both predictions.

In the other legal actor model, we permit the quality of private signals to vary across judges according to some known distribution of signal quality. This implies weak herding, i.e., the probability of the first split declines faster with time than implied by the political model, but slower than implied by herding model with constant quality private signals. Moreover, the variable quality private signal model implies that judges can be more likely to follow the last case than earlier cases in a line of balanced cases, i.e., the probability the sequence AB is followed by a B can be greater than merely the probability of receiving a private signal of B, i.e., more than 0.5. Our data cannot reject either of these predictions.

Given our empirical results, we reject the purely political actor model as well as purely legal actor model with constant quality litigants and judges, i.e., constant quality private signals. However, we conclude that the data support – in the sense of not rejecting – the predictions of the purely legal actor model with variable quality litigants and judges.

## REFERENCES

- Baker, Scott and Claudio Mezzetti. 2012. "A Theory of Rational Jurisprudence." *Journal of Political Economy*, 120(3), pages 513 – 551
- Banerjee, Abhijit. 1992. "A Simple Model of Herd Behavior." *Quarterly Journal of Economics*, vol. 107, pp. 797-817.
- Bikhchandani, Sushil, Hirshleifer, David and Ivo Welch. 1992. "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades." *Journal of Political Economy*, vol. 100, pp. 992-1026
- Choi, Stephen and Mitu Gulati. 2008. "Bias in Judicial Citations: A Window into the Behavior of Judges?" *Journal of Legal Studies*, vol. 27, pp. \_\_\_\_-\_\_\_\_.
- Daughety, Andrew and Jennifer Reinganum. 1999. "Stampede to Judgment: Persuasive Influence and Herding Behavior by Courts." *American Law and Economics Review*, vol. 1, pp. 158-89
- Shai Danziger, Jonathan Levav, and Liora Avnaim-Pesso. 2011. "Extraneous factors in judicial decisions, Proceedings of the National Academy of Sciences of the United States of America"
- Epstein, Lee and Jack Knight. 2013. "Reconsidering Judicial Preferences." *Annual Review of Political Science*, vol. 16, pp. 19.1-19.21.
- Epstein, Lee, William M. Landes, and Richard Posner. 2013. *The Behavior of Federal Judges: A Theoretical and Empirical Study of Rational Choice*. Cambridge, Mass.: Harvard University Press.
- Gely, Rafael and Pablo Spiller. 1992. "A Rational Choice Theory of Supreme Court Decisions with Applications to the State Farm and Grove City Cases." *Journal of Law, Economics, & Organization*, vol. 6, pp. 263-\_\_\_\_.
- Gennaioli Nicola & Andrei Shleifer. 2007. "The Evolution of the Common Law." *Journal of Political Economy*, vol. 115, pp. 43-68.
- Huang, Bert I. 2011. "Lightened Scrutiny." *Harvard Law Review*, vol. 124, pp. 1109-1147.
- Jacobi, Tonja. 2006. "The Impact of Positive Political Theory on Old Questions of Constitutional Law and Separations of Powers." *Northwestern Law Review*, vol. 100, pp. 259-278.
- Lax, Jeffrey R. 2011. "The New Judicial Politics of Legal Doctrine." *Annual Review of Political Science*, vol. 11, pp. 131-157.
- Leiter, Brian. 2003. "American Legal Realism." *The Blackwell Guide to Philosophy of Law and Legal Theory*; W. Edmundson and M. Golding, eds. Oxford: Blackwell.
- Marks, Brian A., *A Model of Judicial Influence on Congressional Policymaking: Grove City College v. Bell* (Hoover Institution, Working Paper No. P-88-7, 1988).

Martin, Andrew D. and Kevin M. Quinn. 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953-1999." *Political Analysis*, vol 10, pp. 134-153.

McNollgast. 1995. "Politics and the Courts: A Positive Theory of Judicial Doctrine and the Rule of Law." *Southern California Law Review*, vol. 68, pp. 1631-1683.

O'Hara, Erin. 1993. "Social constraint or implicit collusion-toward a game theoretic analysis of stare decisis." *Seton Hall Law Review*, vol. 24, pp. 736-\_\_\_\_.

Posner, Richard A. 1993. "What do judges and justices maximize? (The same thing everybody else does)." *Supreme Court Econ. Rev.*, vol. 3, 1-41.

Pritchett, C. Herman. 1948. *The Roosevelt Court*.

Pritchett, C. Herman. 1941. "Divisions of Opinion Among Justices of the U.S. Supreme Court, 1939-1941." *Amer. Pol. Sci. Rev.*, vol. 35: pp. 890-\_\_\_\_.

Rasmusen, Eric. 1994. "Judicial legitimacy: An interpretation as a repeated game." *Journal of Law, Economics, and Organization*, vol. 10, pp. 63-83.

Segal, Jeffrey A. and Harold J. Spaeth. 2002. *The Supreme Court and the Attitudinal Model Revisited*.

Shepherd, Joana, and Michael S. Kang. 2014. "Skewed Justice: Citizens United, Television Advertising and State Supreme Court Justices' Decisions in Criminal Cases." Available at <http://skewedjustice.org/#summary>.

Spitzer, Matthew and Eric, Talley. 2013. "Left, Right, and Center: Strategic Information Acquisition and Diversity in Judicial Panels." *Journal of Law, Economics, and Organization*, vol. 29, pp. 638-680.

Sunstein, Cass, David Schkade, Lisa M. Ellman, and Andres Sawicki. 2006. *Are Judges Political*. Brookings Institute: Washington D.C.

Talley, Eric. 1999. "Precedential Cascades: An Appraisal." *Southern California Law Review*, vol. 73, pp. 87-\_\_\_\_.

Tiller, Emerson and Frank Cross. 2006. "What is Legal Doctrine?" *Northwestern Law Review*, vol. 100, pp. 517-534.

Wald, A. and J. Wolfowitz (1940). "On a Test Whether Two Samples are from the Same Population." *The Annals of Mathematical Statistics* 11(2): 147-162.

## FIGURES AND TABLES

*Table 2. Summary statistics on various covariates for sequences of federal appellate cases, by how decisions in the sequence were coded.*

	USLW coded decisions	Hand- coded decisions	Differ- ence	P-value	All
Sequence length, all decisions (no. of decisions)	3.875 (1.878)	4.114 (2.075)	0.239	0.240	3.970 (1.945)
Sequence length, A and B decisions (no.)	3.704 1.800	4.083 (2.069)	0.379	0.032	3.848 (1.896)
Sequences truncated because C decisions (fraction)	0.094 (0.292)	0.015 (0.122)	0.079	0.000*	0.072 (0.259)
Duration (years)	12.310 (10.87)	10.290 (8.05)	2.020	0.184	11.750 (10.21)
"A" decisions (fraction of cases)	0.554 (0.161)	0.536 (0.182)	0.018	0.889	0.549 (0.167)
Opinion by Republican judges (fraction)	0.591 (0.306)	0.575 (0.309)	0.016	0.449	0.586 (0.307)
Observations (No. of sequences)	713	264			977

Note: All sequences are identified by USLW, but decisions may be coded by USLW or a research assistant (hand-coded). Table provides summary statistics by how decisions are coded. Cells contain means with standard deviations in parentheses. P-values are calculated using Wilcoxon rank sum test for continuous variables and Fisher exact test for binary variables. \* indicates significance with 95% confidence in a 2-sided test.

Table 3. Summary statistics on various covariates for sequences of federal appellate cases, by whether the sequence was truncated because of a first C decision..

	Not truncated	Trun- cated because C deci- sions	Differ- ence	P-value	All
Sequence length, all decisions (no. of decisions)	3.875 (1.915)	5.188 (1.92)	1.313	0*	3.970 (1.945)
Sequence length, A and B decisions (no.)	3.875 (1.915)	3.507 (1.605)	0.368	0.039	3.848 (1.896)
Duration (years)	11.556 (10.19)	9.464 (10.29)	2.092	0.011*	11.760 (10.2)
"A" decisions (fraction of cases)	0.548 (0.17)	0.560 (0.127)	0.012	0.688	0.549 (0.167)
Opinion by Republican judges (fraction)	0.589 0.304	0.554 (0.343)	0.035	0.333	0.586 (0.307)
Observations (No. of sequences)	906	71			977

Note: Table provides summary statistics by whether the sequence of decisions was ended because a court made a C decision, i.e., the first decision different than two prior decisions. Cells contain means with standard deviations in parentheses. P-values are calculated using Wilcoxon rank sum test for continuous variables and Fisher exact test for binary variables. \* indicates significance with 95% confidence in a 2-sided test.

Table 4. Non-parametric test of whether courts are chosen in random order.

Court	Slot (cells contain probabilities)								
	1	2	3	4	5	6	7	8	9
Circuit 1	0.27	0.26	0.56	0.57	0.61	0.67	0.78	0.40	0.21
Circuit 2	0.69	0.69	0.39	0.24	0.40	0.41	0.37	0.30	0.21
Circuit 3	0.55	0.54	0.45	0.43	0.44	0.48	0.43	0.39	0.11
Circuit 4	0.21	0.21	0.50	0.67	0.68	0.64	0.75	0.65	0.39
Circuit 5	0.60	0.60	0.49	0.42	0.33	0.44	0.38	0.40	0.44
Circuit 6	0.42	0.43	0.53	0.50	0.59	0.44	0.36	0.55	0.45
Circuit 7	0.49	0.48	0.57	0.48	0.46	0.37	0.41	0.27	0.45
Circuit 8	0.37	0.36	0.60	0.55	0.56	0.39	0.54	0.59	0.45
Circuit 9	0.83	0.85	0.46	0.26	0.14	0.15	0.19	0.25	0.11
Circuit 10	0.27	0.26	0.39	0.56	0.71	0.71	0.71	0.54	0.29
Circuit 11	0.42	0.42	0.45	0.53	0.42	0.49	0.60	0.60	0.44
DC Circuit	0.43	0.42	0.50	0.46	0.49	0.54	0.57	0.62	0.43
Fed Circuit	0.52	0.51	0.52	0.19	0.64	0.45	0.49	0.31	0.12
State Court	0.17	0.18	0.20	0.47	0.47	0.79	0.73	0.96	0.11

Note: Table reports the fraction of draws in which the number of times a given Circuit is in a given slot (across all sequences) is less than the number of times in that Circuit is in that slot in the actual data (across all sequences). We made 5000 draws. In each draw, we reorder the courts in each actual sequence at random. Draws do not reallocate courts across sequences.



Table 5. Probability of following prior case given a balanced decision history, by length of that history.

	Balanced history length		
	2	4	6
Follow last case (fraction)	0.598	0.595	0.576
Std error	0.027	0.054	0.098
P-value v. 1/2	0.000	0.101	0.557
Obs.	333	84	26

Notes. Table calculates the probability a case follows the decision in the case immediately prior to it, given that the history of cases prior to the case is balanced. If the balanced history length is 2/4/6, then the first 2/4/6 cases in the sequence have equal numbers of A and B decisions, and we are examining whether the case in position 3/5/7 follows that in position 2/4/6.

## APPENDIX

### *Proof of the Remark*

Before proceeding to prove the remark, we first show that the second judge in a sequence only creates a split if

- (a) he receives a private signal that conflicts with the decision of the first judge and
- (b) the signal is of high quality.

Assume that the first judge receives signal  $s_1 = a$  and thus decides the case as  $D_1 = A$ . The judge in the next circuit to consider the issue receives a new private signal. If that signal lines up with the decision rendered by the first judge, he decides the case the same way. If, however, the signal differs, the second judge might create a split. Specifically, if the second judge receives a signal  $b$  at a precision level  $\pi_H$ , she will decide  $D_2 = B$ . To see why, Bayes rule implies that

$$\Pr(R_2 = A | s_2 = b, D_1 = A) = \frac{(1 - \pi_H)(p\pi_L + (1 - p)\pi_H)}{(1 - \pi_H)(p\pi_L + (1 - p)\pi_H) + \pi_H(p(1 - \pi_L) + (1 - p)(1 - \pi_H))}$$

or

$$\Pr(R_2 = A | s_2 = b, D_1 = A) = \frac{(1 - \pi_H)E(\pi)}{(1 - \pi_H)E(\pi) + \pi_H(1 - E(\pi))}$$

Likewise, and by similar reasoning,

$$\Pr(R_2 = B | s_2 = b, D_1 = A) = \frac{\pi_H(1 - E(\pi))}{(1 - \pi_H)E(\pi) + \pi_H(1 - E(\pi))}$$

The second expression exceeds the first since  $\pi_H > E(\pi)$ . Thus, the second judge creates a split, deciding  $D_2 = B$  when he receives a high quality signal whose value is  $b$ . We can do the same analysis and show that the second judge ignores a conflicting signal  $s_2 = b$  and decides the case as  $D_2 = A$  if the signal is of low quality. As a result, the split perfectly informs the third judge that the second judge based his decision on a high quality signal.

Turning to the proof of the remark part (1), consider a sequence of decisions --  $D_1 = A$  and  $D_2 = B$  -- presented to the third judge (i.e., he observes a split). Suppose further than he receives a signal "a" at a precision level  $\pi_L$ . Bayes rule implies

$$\Pr(R_3 = B | s_3 = a, D_2 = B, D_1 = A) = \frac{(1 - \pi_L)\pi_H(1 - E(\pi))}{(1 - \pi_L)\pi_H(1 - E(\pi)) + \pi_L(1 - \pi_H)E(\pi)}$$

and

$$\Pr(R_3 = A | s_3 = a, D_2 = B, D_1 = A) = \frac{\pi_L(1 - \pi_H)E(\pi)}{(1 - \pi_L)\pi_H(1 - E(\pi)) + \pi_L(1 - \pi_H)E(\pi)}$$

The third judge ignores his own private signal  $s_3 = a$  (a signal consistent with the decision of the first judge) if

$$(1 - \pi_L)\pi_H(1 - E(\pi)) > \pi_L(1 - \pi_H)E(\pi)$$

or

$$\frac{\pi_H}{1 - \pi_H} > \frac{\pi_L}{(1 - \pi_L)} \frac{E(\pi)}{(1 - E(\pi))}$$

For small enough value of  $\varepsilon$ , the inequality always holds (at  $\varepsilon = 0$ , the inequality reduces to

$$\frac{\pi_H}{1 - \pi_H} > \frac{E(\pi)}{(1 - E(\pi))}, \text{ which always holds since } \pi_H > E(\pi).$$

To prove part (2), consider the following sequence. The first judge decides  $D_1 = A$ ; the second judge decides  $D_2 = B$ ; and the third judge receives a private signal  $s_3 = b$  with low quality. The third judge's posteriors are

$$\Pr(R_3 = B | s_3 = b, D_2 = B, D_1 = A) = \frac{\pi_L \pi_H (1 - E(\pi))}{\pi_L \pi_H (1 - E(\pi)) + (1 - \pi_L)(1 - \pi_H)E(\pi)}$$

and

$$\Pr(R_3 = A | s_3 = b, D_2 = B, D_1 = A) = \frac{(1 - \pi_L)(1 - \pi_H)E(\pi)}{\pi_L \pi_H (1 - E(\pi)) + (1 - \pi_L)(1 - \pi_H)E(\pi)}$$

The first expression exceeds the second expression (and, as a result, the third judge decides B) if

$$\pi_L \pi_H (1 - E(\pi)) > (1 - \pi_L)(1 - \pi_H)E(\pi)$$

This inequality must hold because

$$\frac{\pi_H}{(1 - \pi_H)} \frac{\pi_L}{(1 - \pi_L)} > \frac{E(\pi)}{(1 - E(\pi))}$$

Completing the proof of part (2) of the result.

### *Proof of Prediction 3c*

Conditional on  $B$  being the correct resolution, an empirical observer will see the sequence  $\{D_1 = A, D_2 = B, D_3 = B\}$  in two scenerios. In the first, the first judge sees a private signal  $s_1 = a$ , the second judge sees a private high quality signal  $s_2 = b$  and the third judge observes a signal  $s_3 = b$  (no matter the signal strength). In the second, the first judge sees  $s_1 = a$ , the second judge sees a high quality signal  $s_2 = b$  and the third judge observes a low quality private signal  $s_3 = a$ . Combined, then, the probability of this sequence (assuming  $B$  is the correct answer) is

$$[(1 - E(\pi))(1 - p)\pi_H E(\pi) + [1 - E(\pi)](1 - p)\pi_H p(1 - \pi_L)]$$

An empirical observer will see the sequence  $\{D_1 = A, D_2 = B, D_3 = A\}$  if the first judge sees a private signal  $s_1 = a$ , the second judge sees a high quality signal  $s_2 = b$ , and the third judge observes a high quality signal  $s_3 = a$ . Conditional on  $B$  being the correct answer, the probability of this sequence is

$$[1 - E(\pi)](1 - p)\pi_H(1 - p)(1 - \pi_H)$$

The first expression always exceeds the second expression if

$$\left(\frac{p}{1-p}\right) > 1 - 2\pi_H$$

a condition that always holds since  $\pi_H > 1/2$ .

This proves that the sequence  $\{D_1 = A, D_2 = B, D_3 = B\}$  is more likely than the sequence  $\{D_1 = A, D_2 = B, D_3 = A\}$  if  $B$  is the correct answer.

Conditional on  $A$  being the right answer, an observer will see the sequence  $\{D_1 = A, D_2 = B, D_3 = B\}$  with probability

$$(1 - p)(1 - \pi_h)E(\pi)(1 - E(\pi)) + (1 - p)(1 - \pi_h)E(\pi)p\pi_l$$

An observer will see the sequence  $\{D_1 = A, D_2 = B, D_3 = A\}$  with probability

$$(1 - p)\pi_H(1 - p)(1 - \pi_H)E(\pi)$$

Simple algebra shows that the first sequence is more likely given the condition on  $p$  in the prediction. Given that the sequence  $\{D_1 = A, D_2 = B, D_3 = B\}$  is more likely than the sequence  $\{D_1 = A, D_2 = B, D_3 = A\}$  irrespective of whether  $A$  or  $B$  is the right answer, it is more likely to be observed in the data.

## APPENDIX FIGURES AND TABLES

Table 6. *Distribution of cases in sequences across circuit courts, by how decisions in the sequence were coded.*

Court	No USLW citation	With USLW citation	All
1	5.68	5.20	5.34
2	7.92	9.61	9.12
3	6.34	8.09	7.58
4	8.57	7.86	8.07
5	8.67	10.36	9.87
6	8.67	8.24	8.36
7	10.90	10.29	10.47
8	8.95	7.14	7.66
9	12.40	14.01	13.54
10	8.85	6.19	6.96
11	8.29	7.93	8.04
Fed	0.00	0.95	0.67
DC	3.91	3.19	3.40
State	0.84	0.95	0.92
Total	100	100	100

Note: Table lists percent of cases in data from each circuit or state court

Table 7. Distribution of sequences by date of last observed case in sequence.

	USLW- coded decisions (percent)	Hand- coded decisions (percent)	All
1933	0.00	0.15	0.11
1984	0.00	0.08	0.05
1989	0.19	0.19	0.19
1990	0.00	0.19	0.13
1991	0.00	0.42	0.29
1992	0.00	0.15	0.11
1993	0.19	0.00	0.05
1994	0.00	0.11	0.08
1995	0.56	0.11	0.24
1996	0.00	0.23	0.16
1997	0.00	0.15	0.11
1998	0.00	0.72	0.51
1999	1.76	0.15	0.62
2000	0.46	0.30	0.35
2001	29.32	3.47	10.96
2002	27.75	4.08	10.94
2003	31.54	4.98	12.68
2004	8.23	1.55	3.48
2005	0.00	1.06	0.75
2006	0.00	6.72	4.77
2007	0.00	8.38	5.95
2008	0.00	6.38	4.53
2009	0.00	6.98	4.96
2010	0.00	8.30	5.90
2011	0.00	11.85	8.42
2012	0.00	7.21	5.12
2013	0.00	9.66	6.86
2014	0.00	11.74	8.34
2015	0.00	4.72	3.35
Total	100.00	100.00	100.00

Note: All sequences are identified by USLW, but decisions may be coded by USLW or a research assistant (hand-coded). Table lists percent of decisions in data from each year.