Getting Their Way: Bias and Deference to Trial Courts*

Ryan Hübert†

April 7, 2019

Abstract

How much do trial judges influence the law in the United States? I analyze a model of adjudication by a trial judge who engages in fact-finding before deciding a case, but whose decision may be reversed. The model makes three broad points. First, it provides an informational rationale for ex post deference to biased trial judges that does not require an ex ante commitment by an appellate court to a standard of review. Second, it shows how procedural discretion can bring biased trial judges' rulings closer to appellate doctrine despite enabling trial judges to "get their way" more often. Third, de facto law as represented by trial judges' case-by-case adjudication will differ substantially from de jure law. As long as there are not too many extremist trial judges, de facto law will reflect the predispositions of trial judges, not legal doctrine.

Keywords: judicial politics, judicial hierarchy, trial courts, appellate review, fact-finding, formal theory

*I thank Sean Gailmard, Sean Farhang, Kevin Quinn, Eric Schickler, Ryan Copus, Tiberiu Dragu, James Hicks, John Kastellec, Jonathan Nash, Carlo Prato, participants at the 2017 Emory Conference Institutions and Law-Making and anonymous reviewers for helpful feedback. All errors and omissions are my own.

†Ryan Hübert is Assistant Professor of Political Science, University of California, Davis, 469 Kerr Hall, One Shields Avenue, Davis, CA 95616 (rhubert@ucdavis.edu).

How much do trial judges influence the law in the United States? The conventional wisdom pulls in two directions. On the one hand, accounts by judges and practitioners portray trial judges as distinctly important. For example, U.S. District Judge Henry N. Graven famously remarked that "the people of this district either get justice here with me or they don't get it at all" (quoted in Carp and Rowland 1996, p. 1). Justice Scalia once quipped that "There are a lot of non-reviewable ways in which [a trial judge] can make the case come out 'right'" (American Constitution Society and Federalist Society 2006). On the other hand, key institutional features limit trial judges' impact on law. A primary purpose of trial courts is to serve as the judiciary's workhorse fact-finders on specific cases, not to craft new law. As a result, their decisions do not constitute binding legal precedent and their rulings are much more routinely subjected to appeal.

This article presents a model of trial court decision-making that helps reconcile this puzzle. The model demonstrates that trial courts' institutionally limited role can paradoxically enable them to exert substantial de facto control over the law even in contexts where the de jure law established by appellate courts is clear and settled. This occurs because trial judges' closeness to case facts ensures that they have the ability to determine how *accurately* legal rules are implemented. If appellate courts prefer that trial judges correctly apply legal rules to case facts and trial judges are uniquely able to establish each case's factual record, then appellate courts may find it worthwhile to defer to trial judges' decisions even if they are not likely to be compliant with settled law. In the aggregate, this means that case-by-case adjudication will generate outcomes that differ substantially from the appellate court's established doctrine.

With the model, I make two arguments that emphasize the important role that trial courts play in judicial politics. First, appellate courts sometimes have a strategic incentive to defer to trial court decisions even when those decisions appear at odds with settled law. In the model, deference occurs ex post and therefore requires no pre-commitment by the appellate

court to defer. This finding provides a fully endogenous explanation for deference to trial courts that does not rely on the presumption that appellate courts hold themselves to their pre-specified standards of review. To the extent that appellate courts face a commitment problem in adhering to these standards, the model identifies a strategic rationale for why they may do so *even when* confronted with a set of judges who routinely issue judgments that are non-compliant.

The second argument is that trial judges' procedural discretion, which enables biased trial judges to "get their way" more often, may counterintuitively bring adjudication closer in line with the appellate court's legal doctrine. In the model, "procedure" affects how easy it is for a trial judge to communicate what it knows to the appellate court. As a result, procedural discretion allows trial judges to influence what an appellate court believes about how compliant her decision is. While the appellate court cannot prevent itself from making negative inferences about certain judgments made by a trial judge, it is sometimes *overly* concerned that those judgments are non-compliant. This skepticism undermines a trial judge's incentives to work hard, lowering the accuracy and quality of decisions overall. The appellate court cannot credibly prevent itself from being overly skeptical, but it can solve this problem by letting the trial court make obfuscatory procedural choices that have the effect of actually reducing the appellate court's skepticism.

These two arguments underscore the degree to which the incentives of trial judges influence legal outcomes in society, despite the fact that trial judges are nominally the lowest ranked decision-makers in the judicial hierarchy. If trial judges are routinely making non-compliant decisions that are affirmed by appellate courts, as shown in the model, then the law as written on paper may differ from the law as experienced by individuals in society. In the last section of this article, I show how trial judges may influence law through their adjudication by pulling de facto law away from the appellate court's de jure law. This has potentially large implications for how we understand judicial politics. For example, conflict

between appellate courts over the establishment of a new legal rule will occur in the shadow of trial courts' distortion of de facto law (see Lax 2012). The model therefore makes explicit a set of concerns about lower court compliance that have preoccupied judges sitting on higher courts (as well as scholars, e.g., Boyd and Spriggs 2009; Westerland et al. 2010). Specifically, non-compliance by trial courts may be both blatant *and* unchecked.

1 Fact-Finding in the Judicial Hierarchy

There has been a lot of theoretical research on how the judicial hierarchy affects law in the United States (for a recent survey, see Kastellec 2017). Much of it has focused either explicitly or implicitly on the two upper tiers of the three-tiered hierarchy (but see Cameron and Kornhauser 2006; Baker and Kornhauser 2015). There is good reason to suspect that hierarchical dynamics are particularly influential in trial courts. As Choi, Gulati, and Posner (2012) points out, "district judges are constrained: Unlike appellate court judges, whose opinions are subject only to discretionary (and occasional) review by the Supreme Court, district court decisions are subject to mandatory (and routine) review by circuit courts" (p. 520). A major difficulty that arises when trying to apply existing theories of judicial hierarchy to the trial court context is that trial courts are *very* different institutions than appellate courts (Kim et al. 2009; Boyd 2017).

Three distinct institutional features of trial courts stick out most prominently. First, and most obviously, trial court determinations do not constitute binding legal precedent.[1] At a minimum, this means that any new legal rule a trial judge articulates has only limited intrinsic value. For the vast majority of cases, a trial judge exerts influence over law via individual decisions about who wins cases. As a result, theories that hinge on a lower court's

[1]For example, see *Fishman & Tobin, Inc., v. Tropical Shipping & Construction Co., Ltd.*, 240 F.3d 956 (11th Cir. 2001), where the court said "While the decisions of their fellow [district] judges are persuasive, they are not binding authority. ... As a result, the district court cannot be said to be bound by a decision of one of its brother or sister judges" (p. 965).

ability to establish a binding legal rule (e.g., Cameron, Segal, and Songer 2000; Carrubba and Clark 2012; Clark and Carrubba 2012) are less applicable when the lower court is a trial court.

Second, trial courts are crucial for establishing the facts of legal cases. As a reflection of this, most of a trial court's day-to-day work is focused on actively *building* the factual records of the cases they hear. This takes time and effort. As a recent, prominent example of this effort, consider *Perry v. Schwarzenegger*, 704 F. Supp. 2d 921 (N.D. Cal. 2010), a case about the constitutionality of California's ban on same-sex marriage. U.S. District Judge Vaughn Walker was widely seen as exerting a lot of effort establish the facts of the case (Dolan 2013). In November 2009, for example, he personally reviewed some of the defendants' documents to determine whether they should be provided to the plaintiff as part of discovery. In his order, he outlined what he reviewed, indicated that future disputes should be brought to him (instead of a magistrate judge) and made clear he was "ready to assist the parties should further disputes arise."[2] When he eventually ruled for the plaintiffs after a bench trial, fifty-five pages of Judge Walker's 138 page opinion succinctly articulated his specific findings of fact.

As the example highlights, establishing the facts of a case entails two distinct tasks. First, a trial judge has to figure out for herself the relevant case facts from the mountains of information made available by the litigants. To do so, she may have to review documents, consider the admissibility and relevancy of evidence, hear and consider expert and witness testimony, meet with attorneys to resolve disputes, and preside over public hearings. However, the fact that a trial judge's efforts help her learn about the relevant case facts does not automatically imply that what she learned becomes readily available to appellate courts reviewing her decision. Indeed, as the Supreme Court acknowledged in *Pierce v. Underwood*, 487 U.S. 552 (1988), after "...settlement conferences and other pretrial activities, the district

[2]The public record of this order is available at `https://ecf.cand.uscourts.gov/doc1/03516202322`.

court may have insights not conveyed by the record, into such matters as whether particular evidence was worthy of being relied upon..." (p. 560). A trial judge's second task is to record what she learned in order to justify her decisions. This can be more or less costly, depending on whether the trial judge's knowledge of the case facts is based on concrete and specific pieces of information—such as specific documents provided in discovery—or harder to convey information—such as "a balancing of relatively subjective factors" under Rule 403 of the Federal Rules of Evidence (Rothstein, Raeder, and Crump 2012, p. 7).

Existing theories have demonstrated that a lower court's private information about a case's facts creates an incentive for an upper court to defer (e.g., Cameron, Segal, and Songer 2000; Clark and Carrubba 2012; Carrubba and Clark 2012; Lax 2012; Beim, Hirsch, and Kastellec 2014; Beim 2017). However, the particular way that fact-finding is conceptualized (and therefore modeled) matters for understanding both the scope and form of this deference (see Stephenson 2011). Counterintuitively, deference is *less* common when lower courts have to work hard to find facts and when they must communicate what they learned to a reviewing court, as is the case with trial courts. Costlier fact-finding may cause the upper court to infer the lower court is uninformed; easier record-building may cause the appellate court to infer a lower court is concealing unfavorable information. Either way, these negative inferences sometimes make a reviewing court *too* skeptical of the trial judge, which in turn causes a trial judge to give up on making *any* fact-finding effort.

A third distinctive feature of trial courts is that their decisions are subject to more routine review by appellate courts. Intermediate appellate courts are *required* to hear (properly filed) appeals from trial courts and therefore cannot avoid making these overly negative inferences. One potential solution to this problem is for appellate courts to articulate formalized standards of review that guarantee a certain degree of deference to trial court determinations. Indeed, the U.S. court system, a trial judge's factual determinations are reviewed under a "clearly erroneous" standard, which accords a lot of deference to trial judge. Unfortunately,

this solution only goes so far. Scholars have extensively demonstrated that judges disagree with one another about how to decide similar cases (for recent examples, see Boyd, Epstein,and Martin 2010; Epstein, Landes, and Posner 2013; Kastellec 2013; Boyd 2016). To the extent this disagreement exists between trial and appellate judges, how can an appellate court credibly commit itself to applying a deferential standard of review when it knows a trial judge would prefer to use a very different legal rule to resolve cases?

In addition to the literature on how judicial hierarchy shapes the law, this article contributes to at least two other recent strands of research. First, it contributes to a burgeoning formal theory literature on the incentives that institutions provide for lower-level government officials to improve (or not) the quality of policy or legal implementation (e.g., Bueno de Mesquita and Stephenson 2007; Gailmard and Patty 2013; Dragu and Board 2015; Patty and Turner 2016; Turner 2016, 2018). For example, Turner (2018) presents a model of judicial review of agency rule-making that demonstrates how ex post review can undermine a policy-maker's incentive to work hard to make good policy.

Second, this article adds to recent models focused on the role of opinion writing in managing hierarchical relationships between courts, and between courts and legislatures (e.g., Staton and Vanberg 2008; Clark and Carrubba 2012; Beim, Hirsch, and Kastellec 2014). Clark and Carrubba (2012) conceives of opinions as consisting of legal rules and persuasive quality. One distinctive feature of trial court opinions (and other written documents, such as orders on motions) is that they help record what the trial judge learned about the facts of a case. As I show below, if opinions are informative about the case facts, then their "persuasiveness" is not unambiguously beneficial for upper and lower courts. In Staton and Vanberg (2008), a court may decide to issue a vague opinion to head off non-compliance by a legislature. Similarly, in the model here, completely informative opinions reveal non-compliance to a reviewing court, so a trial judge may have an incentive to issue an uninformative opinion in order to mask their non-compliance.

The model in this article is most closely related to two recent articles. In Gennaioli and Shleifer (2008), deference to a trial court occurs because it is implicitly assumed that an appellate court can commit to defer to a trial judge's fact finding. Since an appellate court's review occurs ex post, what is missing from that account is an explanation for why the appellate court defers when it fully understands that the trial judge has an incentive to engage in fact distortion. Baker and Kornhauser (2015) presents a model demonstrating that an appellate court has an incentive to defer ex post after a trial judge engages in fact-finding. In that model, a trial judge costlessly receives private information on two dimensions, only one of which is visible to an appellate court. The model here departs from Baker and Kornhauser (2015) by assuming that a trial judge's fact-finding and fact-recording efforts are costly, which helps rationalize appellate courts' grant of *procedural* discretion.

2 Model

I analyze a simplified model of litigation where a trial judge (T, which I refer to as "she") hears a case, which is then reviewed by an appellate court (A, which I refer to as as "it").[3] The trial judge engages in fact-finding, then makes a judgment that is accompanied by an opinion that may include her findings of fact. The appellate court reviews the case and decides whether to affirm or reverse.

Cases and Legal Rules. I model legal cases using a case space approach (Lax 2011). The fact pattern of a specific case is represented by a number in a unidimensional space, $\omega \in \mathbb{R}$. Fact patterns are distributed in society according to a distribution with a cumulative distribution function F_ω. There is common knowledge that fact patterns come from this distribution, but neither the trial judge nor the appellate court knows the specific value of ω

[3]Because I am interested in the direct relationship between appellate and trial courts, I do not study litigant behavior in this model. However, other research incorporates strategic litigants, and this is a fruitful avenue for additional research (for example, see Priest and Klein 1984; Dewatripont and Tirole 1999; Cameron and Kornhauser 2006; Beim 2017).

when a case is initially filed in court. Legal rules are represented as cut points $\ell \in \mathbb{R}$ where fact patterns to the left of the legal rule favor the defendant and fact patterns to the right favor the plaintiff.

There is an existing legal rule, $\ell_A \in \mathbb{R}$, which the appellate court has previously established as legal doctrine. This is settled law. The appellate court prefers that the trial judge make decisions about individual cases using this legal rule. However, the trial judge is biased and prefers to use a different legal rule to resolve cases. I assume the existing legal rule is strictly lower than the trial judge's own preferred legal rule, $\ell_A < \ell_T$, but the main substantive results of this article would be unaltered if this were reversed.

> **Definition 1.** A disposition x is **consistent** with a legal rule ℓ_i if and only if it favors the defendant when $\omega \leq \ell_i$ and favors the plaintiff otherwise, and is **compliant (with doctrine)** when it is consistent with ℓ_A. Moreover, let $\mathbf{1}(\ell_i, x)$ be an indicator function taking a value of 1 if x is consistent with ℓ_i.

The prior probability that a particular case will favor the defendant under legal rule ℓ_i is given by the cumulative distribution function $F_\omega(\ell_i)$ evaluated at that rule. I will denote $F_\omega(\ell_i)$ by the Greek letter δ_i indicating that it is the prior probability that a case's fact pattern favors the defendant, according to legal rule ℓ_i. Substantively, this prior probability represents a *predisposition* in favor of either the defendant or the plaintiff. For example, in a civil rights case, a conservative judge may have a predisposition favoring the defendant (e.g., $\delta_i = \Pr[\omega \leq \ell_i] = 0.6$). That said, once this judge learns the facts of the specific case, she may be convinced they are extreme enough to rule for the plaintiff (e.g., $\omega > \ell_i$).

Formally, I assume that the existing settled law has a built-in predisposition that favors plaintiffs, $\delta_A < \frac{1}{2}$, whereas the trial judge's preferred legal rule has a predisposition that favors defendants, $\delta_T > \frac{1}{2}$. Because I am interested in an appellate court's incentive to defer to a non-compliant trial judge, this provides the most interesting case to study. Figure 1 illustrates the case space and the courts' predispositions given a hypothetical distribution of cases.

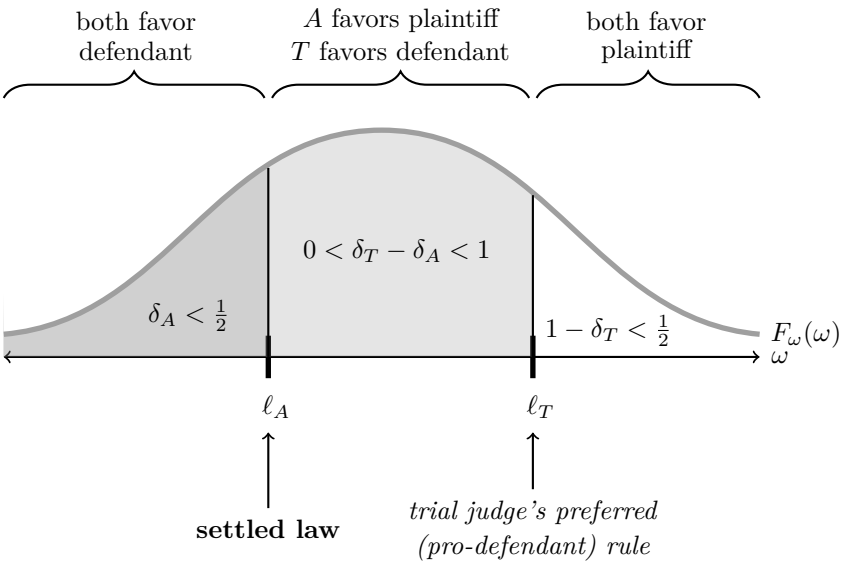

Figure 1: An Example Case Space When Courts Have Conflicting Predispositions

Even though fact patterns are represented in a very stylized way, the model explicitly considers a trial court's difficulty both learning and communicating these fact patterns. First, I assume that if the trial judge decides to engage in fact-finding, she pays a fact-finding cost c in order to learn the fact pattern. (Note: I will allow the trial judge to shirk and not engage in fact-finding.) Second, I assume that a trial judge's knowledge about a fact pattern can be based on information that is more or less difficult to convey to the appellate court. For example, her knowledge of ω could be based on a set of physical documents submitted into the record as evidence (easy to convey), or it could be based on her subjective assessment of the credibility of various witnesses (difficult to convey). Formally, the trial judge faces a cost $\kappa \in \{\kappa_L, \kappa_H\}$ (where $0 \leq \kappa_L < \kappa_H$) when she attempts to convey her knowledge of the fact pattern to the appellate court.

Most concretely, κ represents how costly it is for the trial judge to write a fully informative opinion articulating her findings of fact. A more general interpretation is that it represents how costly it is for the trial judge to make sure that the case record is fully informative to

the appellate court. She privately learns this cost if and when she learns the case facts. Both the appellate court and the trial judge have a prior belief that $\kappa = \kappa_L$ with probability η and $\kappa = \kappa_H$ with probability $1 - \eta$. If the trial judge decides to write an informative opinion, I assume that she cannot include false information about ω. She can either reveal or conceal anything she knows about ω.

> **Assumption 1.** The trial judge cannot provide false information to the appellate court but she can conceal information.

Sequence. The game proceeds as follows. First, $\omega \in \mathbb{R}$ and $\kappa \in \{\kappa_L, \kappa_H\}$ are chosen according to their respective distributions. The trial judge then privately decides whether to exert effort $e \in \{0, 1\}$ to learn the case facts.[4] If she does, then she learns the fact pattern ω with probability ε and learns nothing with probability $1 - \varepsilon$. Throughout the analysis, I represent a lack of information about the fact pattern using the symbol ϕ. If she learns ω, she also learns κ. Next, the trial judge issues a judgment, which consists of two features: a disposition favoring the plaintiff or defendant, $x \in \{p, d\}$, and an opinion. The trial judge decides whether to pay cost κ to make her opinion fully informative, and I denote this decision by $o \in \{\omega, \phi\}$. With some abuse of language, I refer to the pair (x, o) as the court's "judgment" and to the specific decision in favor of the plaintiff or defendant as the "disposition." Finally, the appellate court decides whether to affirm or reverse the decision, $\rho \in \{0, 1\}$, where $\rho = 1$ is a reversal and $\rho = 0$ is an affirmance.

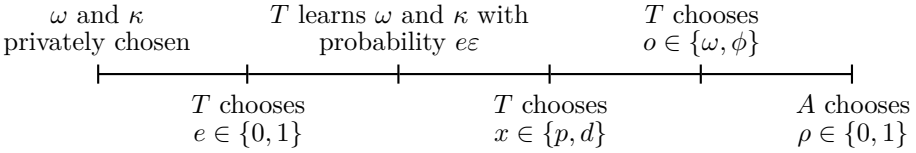

Figure 2: Game Sequence

[4]The fact that effort is privately chosen is not consequential. The results remain unchanged with observable effort given the structure of this game and the requirement that players have correct beliefs about equilibrium strategies.

Given this set up, the model most clearly represents cases resolved solely by trial judges (such as bench trials or cases terminated with pre-trial motions). However, a small number of cases are ultimately decided by juries.[5] Many of the same lessons of this model apply in that context because, even though the jury is the ultimate decision-maker, a trial judge still spends considerable effort before and during the trial on fact-finding. She decides, for example, what information is discoverable, what evidence is admissible and whether a case can be terminated with a pre-trial motion. Moreover, a trial judge can decide how informative a jury's fact-finding is by determining whether it must issue a general verdict on the case or a special verdict that definitively resolves a specific set of relevant factual questions. The main difference is that the trial judge no longer has complete control over the case disposition. Because the strategic relationship between a judge and a jury is not obvious, and I do not explore it here. It is an interesting avenue for future research.

I derive perfect Bayesian equilibria in pure strategies. Formally, each of the players' strategies are functions mapping information sets into their actions. I will use e, x, o and ρ to indicate the players' strategies as well as specific choices. For example, $\rho(x, o)$ indicates the appellate court's review strategy (as a function of the trial judge's disposition and opinion), whereas ρ indicates a specific decision about whether to affirm or reverse, $\rho \in \{0, 1\}$.

Preferences. The appellate court prefers all dispositions to be compliant with doctrine ℓ_A. Therefore, the appellate court's preferences can be represented by the following simple utility function: $u_A = \mathbf{1}(\ell_A, x)$. The trial judge gets two benefits: $\alpha > 0$ when the case's outcome is consistent with her preferred legal rule, and $q > 0$ when she writes an informative opinion (where $\kappa_L < q < \kappa_H$, so that it is only *sometimes* optimal to write an informative opinion). The latter benefit may reflect a variety of motives, such as an intrinsic preference for writing a high quality opinion (as in Clark and Carrubba 2012) or a preference for

[5] In the year ending March 31, 2018, only 0.7% of civil cases (of those terminated after court action) and 2.2% of criminal cases in the U.S. District Courts ended with a jury trial (see Tables C-4 and D-4 of Administrative Office of the U.S. Courts 2018).

demonstrating her competence in her job. For example, retired U.S. District Judge Shira Scheindlin once said: "What I really like to do is write opinions ... There you get to do what you think is right, what you believe in. You're pushing the margins of the envelope, being willing to be creative" (as quoted in Toobin 2013).

The trial judge faces three costs. First, she suffers a cost $k > 0$ whenever she is reversed by the appellate court. This could represent a judge's career concerns,[6] or the extra costs imposed by having to reopen a case and handle the issues raised by the appellate court's reversal. Indeed, trial judges dislike reversals for reasons beyond their impact on case outcomes (see George and Yoon 2003; Choi, Gulati, and Posner 2012). For example, a judge's reversal rates are routinely considered by Senators deciding whether to confirm a judge who is being promoted. The other two costs were discussed above: she incurs a cost $c > 0$ if she exerts effort to find the facts of a case and a cost $\kappa \in \{\kappa_L, \kappa_H\}$ to write an opinion that is informative. The trial judge's preferences are represented by the following utility function:

$$u_T = \alpha \left[\mathbf{1}(\ell_T, x)(1-\rho) + (1 - \mathbf{1}(\ell_T, x))\rho\right] + (q - \kappa)o - k\rho - ce.$$

I adopt the following assumption about the trial judge's preferences, which ensures that the trial judge is sufficiently motivated by case outcomes.

> **Assumption 2.** The trial judge is sufficiently motivated by case outcomes. Formally, $\alpha \geq \frac{2c}{\varepsilon}$ and $\alpha \geq q - \kappa_L$.

Practically speaking, Assumption 2 rules out equilibria in which (1) some judges do not engage in fact-finding even when they receive full deference from the appellate court and (2) judges care more about writing opinions than getting their preferred outcomes on cases. While each of these situations are potentially interesting avenues for future analysis, they do not speak to this article's core question about why appellate courts purposefully defer to

[6] Baker and Kornhauser (2015) micro-founds a judge's career concerns in a dynamic model where a trial judge with a private type builds a reputation for being non-compliant.

trial judges they know to be biased and sufficiently motivated to "get their way."

The analysis of the players' equilibrium strategies proceeds via backward induction: I first consider the appellate court's optimal review strategy, and then I analyze the trial judge's optimal judgment and effort decisions. I present the relevant main results in intuitive language below, and refer readers to the Supplemental Information for formal results and proofs.

3 Rationalizing Deference

The appellate court considers both the trial judge's disposition x and accompanying opinion o before making a decision whether to affirm or reverse. As a result, I express the appellate court's reversal strategy as a function of both the disposition and the accompanying opinion, $\rho(x, o)$. Its decision about whether to reverse will depend on whether it believes the trial judge's decision is compliant with doctrine. Since the appellate court is strictly better off when the judgment is compliant with doctrine, it will affirm a judgment if and only if it believes the judgment is more likely than not to be compliant. When it is indifferent, I assume it affirms. I next consider how the appellate court evaluates the trial judge's decisions when they are accompanied by informative and uninformative opinions, respectively.

Informative Opinions. Because the trial judge cannot fabricate information, an informative opinion ($o = \omega$) means that the appellate court knows for sure whether the judgment is compliant. Then, the appellate court's reversal decision is easy: it reverses whenever it sees that the trial judge's disposition x is not compliant with doctrine.

> **Result 1.** The appellate court never defers when it is fully informed. Formally, $\rho(x, o = \omega) = 1 - \mathbf{1}(\ell_A, x)$.

This first result highlights the appellate court's commitment problem. Deference is never possible when the trial judge fully reveals what it knows. The appellate court simply uses

the information to force compliance with its existing doctrine.

Uninformative Opinions. If the opinion is not fully informative ($o = \phi$), then the appellate court's inference is more complicated. There are four possible kinds of equilibria, each corresponding to a different reversal strategy that the appellate court could use when the opinion is uninformative:

$$\underbrace{\rho(p,\phi) = \rho(d,\phi) = 0}_{\text{always affirm}} \qquad \underbrace{\rho(p,\phi) = 0, \rho(d,\phi) = 1}_{\text{only reverse pro-defendant}}$$

$$\underbrace{\rho(p,\phi) = \rho(d,\phi) = 1}_{\text{always reverse}} \qquad \underbrace{\rho(p,\phi) = 1, \rho(d,\phi) = 0}_{\text{only reverse pro-plaintiff}}$$

Three preliminary observations will streamline the analysis. First, Lemma A5 in the Supplemental Information shows that there is no equilibrium where $\rho(p,\phi) = 1$ and $\rho(d,\phi) = 0$. Second, there always exists a non-deferential equilibrium where $\rho(p,\phi) = 0$ and $\rho(d,\phi) = 1$. As Proposition A1 in the Supplemental Information describes, in this equilibrium, the appellate court reverses any decision that appears non-compliant (such as a judgment favoring the defendant) unless there is an informative opinion justifying the decision. In turn, the trial judge issues pro-plaintiff decisions unless it is able to write an informative opinion justifying a pro-defendant decision. Moreover, the trial judge exerts fact-finding effort only when she is sufficiently confident that if she does so, she will be able to write it into an informative opinion. Formally, she exerts effort if $\eta \geq \hat{\eta}$, where $\hat{\eta}$ is defined in the proof of Lemma A10. I refer to this as the "**non-deferential equilibrium**." Third, as Corollary A1 in the Supplemental Information demonstrates, if the trial judge exerts no effort to find the facts of the case ($e = 0$), then the *only* possible equilibrium of the model is this non-deferential equilibrium since the appellate court knows the trial judge cannot learn the case facts.

In the remaining two candidate equilibria, the appellate court defers: $\rho(p,\phi) = \rho(d,\phi) = 0$ and $\rho(p,\phi) = \rho(d,\phi) = 1$. In these equilibria, the trial judge is able to get its preferred

disposition as an outcome of the game. I will not analyze the always-reverse equilibrium since it is substantively implausible and it provides similar qualitative lessons as the never-reverse equilibrium. Interested readers may refer to the Supplemental Information for a full characterization of the always-reverse equilibrium. In the remainder of this section, I demonstrate that there exists an always-affirm equilibrium and characterize when it occurs. I refer to it as the "**deferential equilibrium**."

3.1 Deferential Equilibrium

Recall that a trial judge's strategy consists of a decision to engage in fact-finding, a disposition and an opinion: e, x and o. When the appellate court uses a deferential review strategy, the trial judge is able to get her preferred disposition. It follows immediately that $x = d$ if she learns $\omega \leq \ell_T$ or if she does not learn ω, and $x = p$ if she learns $\omega > \ell_T$. Moreover, as long as it is not too costly, the trial judge will write an informative opinion only if she learns that the case's fact pattern ω is not in the range where the two courts disagree. If she writes an informative opinion in that scenario, then the trial judge's disposition would be revealed to be non-compliant, and the appellate court would reverse her decision, as in Result 1.

Next consider the trial judge's decision about whether to exert fact-finding effort. Since she never receives deference from the appellate court if she does not exert effort (as discussed above) then she must find it optimal to exert effort in a deferential equilibrium. This requires that her expected utility from exerting effort under deference is weakly greater than her expected utility from not exerting effort (and not receiving deference). Formally:

$$\alpha(\delta_T + \varepsilon(1 - \delta_T)) + \varepsilon(\delta_A + 1 - \delta_T)\eta(q - \kappa_L) - c \geq \alpha(1 - \delta_T)$$

By Assumption 2, this always holds and the trial judge finds it optimal to exert effort when the appellate court defers to her decision.

Given the trial judge's strategy in a deferential equilibrium, I conclude the analysis by characterizing appellate court's belief. Specifically, in order for a deferential equilibrium to hold together, the appellate court must be convinced that the trial judge's decision is more likely to be compliant than non-compliant. This is easy if the trial judge (who is biased in favor of the defendant) rules for the plaintiff. In that situation, the appellate court infers that the trial judge must have discovered such extreme facts that even the trial judge with her pro-defendant bias is convinced the plaintiff should win (i.e., $\omega > \ell_T$). However, if the trial judge issues a pro-defendant decision, the appellate court's belief is more equivocal. On the one hand, it is possible the trial judge learned that $\omega \leq \ell_A$ (compliant) but was just unable to write an informative opinion because it was too costly to do so. On the other hand, it is possible that the trial judge learned that $\omega > \ell_A$ (or didn't learn ω, both non-compliant) and is attempting to contravene appellate doctrine.

Formally, the appellate court's belief about whether the trial judge's decision in favor of the defendant is compliant can be calculated using Bayes' rule:

$$\Pr[\omega \leq \ell_A | x = d, o = \phi] = \frac{\delta_A(1 - \eta\varepsilon)}{\delta_A(1 - \eta\varepsilon) + (\delta_T - \delta_A) + (1 - \delta_T)(1 - \varepsilon)} \tag{1}$$

Deference to the trial judge is optimal if and only if the appellate court is more convinced than not that the decision is compliant—that is, the expression (1) is weakly greater than $\frac{1}{2}$. Some algebra reduces this condition to:

$$\eta \leq \frac{1}{\delta_A}\left(1 - \delta_T - \frac{1}{\varepsilon}(1 - 2\delta_A)\right) \equiv \tilde{\eta} \tag{2}$$

Since the appellate court does not have access to an informative opinion, its review is focused on the trial judge's fact-finding process rather than the actual fact pattern of the case. First, did the trial judge undertake enough fact-finding? This requires that $e = 1$ and ε is sufficiently high. For example, condition (2) never holds if $\varepsilon < \frac{1-2\delta_A}{1-\delta_T}$. Second, is

the case sufficiently complicated that it is reasonable that she did not write an informative opinion? Recall that η represents how easy it is for the trial court to write an informative opinion. Consider the extreme example where $\eta = 1$. Then, any time there is not an informative opinion, the appellate court knows that either the trial judge is uninformed *or* she is concealing information. Either way, this makes the appellate court very skeptical of the judgment and it reverses. On the other extreme where $\eta = 0$, it is never feasible for the trial judge to write an informative opinion. The appellate court thus never suspects her of concealing information and is much less likely to reverse. This logic is captured in the following result, which characterizes the existence of a deferential equilibrium.

> **Result 2.** A deferential equilibrium exists if $e = 1$ and $\eta \leq \tilde{\eta}$, and is characterized by Proposition A2 on page A13 of the Supplemental Information. Moreover, a deferential equilibrium never exists if $\delta_T > \overline{\delta} \equiv 1 - \frac{1}{\varepsilon}(1 - 2\delta_A)$.

Result 2 underscores the difficulty of establishing ex post deference. When the preferences of the trial judge and the appellate court are too far apart (i.e., $\delta_T > \overline{\delta}$) or if it is too easy for the trial judge to communicate what she learns about the case facts (i.e., $\eta > \tilde{\eta}$), then the appellate court cannot credibly defer. However, Result 2 also demonstrates that deference is still possible despite the fact that the appellate court knows the trial judge is getting its preferred dispositions. It is willing to trade off some non-compliant decisions in order to get more informed decision-making overall. Moreover, deference in this model does not occur as the result of standards of review that mandate deference. From a strategic perspective, it is unsatisfying to explain deference to lower courts by emphasizing the appellate court's standards of review. In the absence of some other benefit obtained by deferring to trial judges, appellate courts would not establish those standards if they allowed trial judges to routinely issue non-compliant decisions.

3.2 What Are Standards of Review For?

The analysis demonstrates that appellate courts have an incentive to defer ex post without the need for formalized standards of review. It also rationalizes the two main standards of review that appellate courts apply to trial courts' judgments. Result 1 establishes that appellate courts will never defer to a trial judge when it knows that the incorrect legal rule was applied to a fact pattern. Substantively, this mirrors the federal standards of review for legal issues—"*de novo*" review—which are the completely non-deferential and hinge on an analysis of which legal rule was applied.

In contrast, an appellate court's review of the trial judge's decision in Result 2 hinges on its assessment of the trial judge's fact-finding. Indeed, it establishes that the appellate court has an incentive to defer to a trial judge's decision whenever it is convinced that the trial judge's *fact-finding process* is sufficient: the trial judge did sufficient fact-finding ($e = 1$ and ε is sufficiently high) and had little incentive to conceal the results of that fact-finding (η is sufficiently low). In the federal courts, a trial judge's fact-finding is reviewed under a "clearly erroneous" standard of review, which as *Miller v. Mercy Hospital, Inc.*, 720 F.2d 356 (4th Cir. 1983) emphasizes, is "properly focused upon fact-finding processes rather than fact-finding results" (p. 361).

A remaining question is why appellate courts articulate these formalized ex ante standards of review if they already have a strategic incentive to defer ex post? A full exploration of this issue is beyond the scope of this article, but the model suggests one possible reason why the appellate court may find it beneficial to enshrine standards of review into the formal rules of appellate process. In some situations, the model generates multiple equilibria that each provide reasonable predictions about how an appellate court and a trial court interact with each other. In one equilibrium, the appellate court defers to the trial judge, while in another equilibrium, it does not. When both equilibria are possible at the same time, the equilibrium in which the appellate court does not defer to the trial judge is Pareto dominated

by the equilibrium in which the appellate court does defer to the trial judge. In other words, the non-deferential equilibrium makes both courts worse off than the deferential equilibrium.

> **Result 3.** When deference and non-deference are both possible, each court is better off when the appellate court defers to the trial judge (see Proposition A3 on page A15 of the Supplemental Information).

Non-deference is therefore a "fragile" outcome whenever a deferential equilibrium is possible since both courts would be better off coordinating on the deferential equilibrium. As a result, if the appellate court announces in advance that it intends to play a deferential equilibrium whenever multiple equilibria are possible, this announcement will be credible and the appellate court will follow through and defer to the trial judge. In this way, formalized standards of review can help the courts coordinate on better outcomes.

In this formulation, however, formalized standards of review cannot constrain an appellate court's review whenever a deferential equilibrium does not exist. In that situation, only non-deference is possible, and the appellate court never has an incentive to defer ex post even if it announces a formalized standard of review. So, while formalized standards of review can help the courts coordinate on better outcomes, there are important limits to their usefulness in light of the appellate court's ex post incentives. Moreover, even though formalized standards of review can eliminate some inefficient equilibria, the analysis in the next section demonstrates that other inefficient equilibria may still persist, and the appellate court may still defer too little.

4 Procedural Discretion

Because the trial judge can choose whether to write an informative opinion, the composition of the opinion will affect what the appellate court believes about the trial judge's motives. As I discuss in the previous section, if it is relatively easy for the trial judge to write an

informative opinion (i.e., η is high), then whenever the opinion is actually uninformative, the appellate court will be suspicious that the trial judge is concealing unfavorable information in order to issue a non-compliant decision. As a result, deference will counterintuitively be *more* common when the appellate court is *less* often able to scrutinize the trial judge's decision. While it makes sense that the appellate court might be suspicious that the trial judge wants to conceal unfavorable case facts, it can sometimes be *too* suspicious.

For example, consider Figure 3, which plots when deference is a possible equilibrium outcome for a particular set of parameter values satisfying the model's assumptions: $\delta_A = 0.45$, $\varepsilon = 0.8$, $\alpha = 3$, $c = 1$, $q = 2$, and $\kappa_L = 1$. Suppose that the appellate court reviews a case from the gray region in Figure 3.

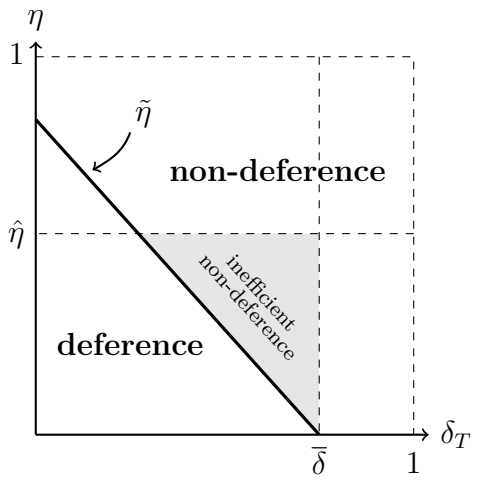

Figure 3: When Expected Opinion-Writing Cost Is High (Low η), Non-Deference Is Inefficient for Some Trial Judges

In this region, η is sufficiently high that the appellate court does not defer to the trial judge (see Result 2) but low enough that the trial judge does not find it worthwhile to exert effort (see Lemma A10 on page A12 of the Supplemental Information). However, if the trial judge had received deference instead, she would have exerted effort and the appellate court would have been better off. In other words, the appellate court's suspicion of the

trial judge ends up undermining the trial judge's incentive to make a higher quality (more informed) decision. To see that the appellate court is worse off, note that the appellate court's expected utility in a non-deferential equilibrium with no effort by the trial judge is $1 - \delta_A$. In a deferential equilibrium, its expected utility would instead be $\delta_A + \varepsilon(1 - \delta_T)$. The latter is always larger than the former in the gray region.

> **Result 4.** Non-deference to a biased trial judge is inefficient if $\eta \in (\tilde{\eta}, \hat{\eta})$ and $\delta_T \leq \overline{\delta}$ (see Proposition A4 on page A15 of the Supplemental Information).

This inefficiency results from dual commitment problems on the part of the trial judge and the appellate court. The trial judge would be better off it could credibly commit to writing fewer informative opinions. The appellate court would be better off if it could credibly commit to defer to the trial judge more often. Of course, it is not credible for either court to make these commitments in the baseline model. However, a potential solution to these interlocking commitment problems is for the trial judge to take advantage of its *procedural* latitude to change the context in which she adjudicates the case. Doing so unambiguously makes the trial judge better off if she is granted additional opportunities to get her way. This may also make the *appellate court* better off.

To demonstrate how this works, I extend the baseline model to allow the trial judge to make procedural decisions that affect the context in which the case gets decided. In the extended model, the trial judge may first choose procedures that reduce her ability to write an informative opinion. For example, she could limit discovery of physical documents and rely heavily on in-person witness testimony, which would make it more difficult for her to convey her findings of fact to the appellate court. Before proceeding, it is important to note that this analysis focuses on one specific effect of procedure—i.e., the trial judge's ability to effectively communicate information to a reviewing court. This may not be the *only* effect of procedural discretion on adjudication. Indeed, others have modeled various aspects of procedure to study a wide range of questions (see, for example, Cooter and Ulen 2012;

Talley 2013). However, to the extent that procedural decisions do shape the trial judge's ability to convey her information to a reviewing court, the results in this article provide a new way to understand the effect of procedure on trial court decision-making.

Formally, the extended game begins with the trial judge making a procedural choice, $\pi \in [0, \eta]$, where η can be considered the court's "baseline procedure." I assume some inherent limitation on her ability to write an informative opinion, so she can only use procedure to reduce η. Then, game proceeds as before until the appellate court's review. Upon review, the appellate court may now choose to review on the merits (i.e., either affirm or reverse the decision) or remand the case to the lower court to remedy a procedural defect. If the case is remanded on procedural grounds, the trial judge pays a cost m and must rehear the case under the baseline procedure η. As in the baseline model analyzed above, the appellate court's review is ex post, and so any deference (to a trial judge's decision about the merits of the case or her procedural decisions) must emerge endogenously and ex post.

To simplify the analysis, I assume that the appellate court defers whenever both deference and non-deference are rational. As I discuss in the previous section, this is reasonable if we assume that the appellate court uses formalized standards of review to help the courts coordinate on deference, which is a Pareto improvement over non-deference (see Result 3). The appellate court's decision about whether to remand the case or review on the merits is straight-forward.[7] The appellate court remands if the trial judge's choice of procedure obscures the case facts "too much." Formally, if the non-deferential equilibrium is not inefficient under the baseline procedure η (e.g., the case is in the white region of Figure 3), then the appellate court is worse off allowing the trial judge to pick a different procedure $\pi < \eta$, and will accordingly remand.

The trial judge finds it optimal to exercise procedural discretion by setting $\pi < \eta$ as long

[7]Note that because the trial judge's choice of procedure occurs before she fully learns the case facts, it does not signal anything to the appellate court about whether the trial judge's decision is compliant.

as doing so changes the appellate court's review posture from one that is non deferential to one that is deferential. That is, she uses her procedural discretion as long as it helps her shield her decision from the appellate court's scrutiny. Importantly though, in equilibrium, the trial judge's use of procedural discretion is upheld by the appellate court because it makes the appellate court better off. The next result characterizes when the trial judge exercises procedural discretion.

> **Result 5.** In the extended model, the trial judge exercises procedural discretion by choosing a procedure $\pi < \eta$ if she has an intermediate level of bias and η is not too high. This appellate court is better off when trial judges have this discretion (see Proposition A5 on page A17 of the Supplemental Information).

There are two main substantive take-aways from the extended model. First, it is obvious to see that procedural discretion sometimes allows a trial judge to "get their way" even more often than in the baseline model. More subtly, trial judges exercise their procedural discretion *only* when it enables them to do this. Neither the most moderate judges (who always receive deference) nor the most extreme judges (who never receive deference) find it optimal to exercise discretion since it limits their ability to write informative opinions. For judges with intermediate biases, inducing the appellate court to defer to them by lowering η is beneficial. Second, from the appellate court's perspective, it is better to give procedural latitude (i.e., discretion) to the trial judge and reign in clear abuses ex post than to articulate a single procedure that should apply across all cases and all judges. The extended model therefore provides a rationale for the third commonly used standard of review, "abuse of discretion," which is applied to procedural decisions. The ex post application of a standard like this allows for heterogeneity in the way different trial judges use procedure, and this turns out to make the appellate court better off.

5 De Facto Law in Trial Courts

The previous sections present a model of a single case and demonstrate how trial judges are often able to get their way. They get their way because appellate courts sometimes purposefully defer to them, and because they choose procedures to receive this deference more often. This case-specific behavior has wider implications for how all cases in the legal system are resolved, and thus how law operates in society. Because trial judges often get their way, the model suggests that an appellate court's doctrine, ℓ_A, may be significantly distorted by the the trial judges' case-by-case adjudication. In other words, there may be a significant divergence between de jure law and de facto law. In this section, I discuss the implications of the previous analysis for an adjudication system that hears many cases by many judges.

Two caveats are in order. First, to the extent that de facto and de jure law differ, litigants and appellate courts are likely to respond to this in the way they file and review cases. That is an interesting avenue for future research, which I do not explore here. Suffice to say that the trial judge's de facto power over legal rules will influence rule-making in many under-appreciated ways. Second, I still assume that the appellate panel and the trial judge have opposing predispositions. As a result, the analysis here illustrates a worst case scenario. If, for example, the legal system operates similar to "team models" where trial and appellate judges have the same preferences over law (e.g., Cameron and Kornhauser 2006) then it is immediate to see that there will be minimal divergence between de facto and de jure law. However, that conclusion would require *all* trial judges to prefer the same legal rules as their supervising appellate courts, which is implausible in light of well-documented preference heterogeneity among judges.

To explore this issue formally, I now consider a how a large set of cases is resolved by a trial court. Accordingly, I will treat the proportion of all cases that favor the defendant as

an aggregate measure of how law is being applied. If all cases were decided without error and according to the appellate court's doctrine, then the proportion of decisions favoring the defendant would be δ_A. This is the proportion of pro-defendant outcomes that would be generated by a perfect application of de jure law.

Now suppose that adjudication occurs as in the extended model from the previous section, and a large set of cases are assigned to the trial judge. Conditional on a specific procedure π, the proportion of pro-defendant outcomes in an equilibrium of the model will depend on the trial judge's bias. If the trial judge is sufficiently moderate, then she will receive deference from the appellate court (possibly by using procedural discretion). For this kind of judge, the proportion of decisions favoring a defendant will be $1-\varepsilon(1-\delta_T)$. If, however, the trial judge is an extremist, then she never receives deference and the proportion of decisions favoring a defendant will be $e\delta_A\varepsilon\eta$.

> **Result 6.** If the trial judge is moderately pro-defendant, then de facto law will favor defendants more that de jure law, formally $1-\varepsilon(1-\delta_T) > \delta_A$. If the trial judge is extremely pro-defendant, then de facto law will favor plaintiffs more than de jure law, formally $e\delta_A\varepsilon\eta < \delta_A$ (see Proposition A6 on page A18 of the Supplemental Information).

This result is specific to a trial judge, and the broader implications for the judiciary as a whole will depend on both the distribution of preferences among all the trial judges, as well as the process by which cases get assigned to judges. Even without making additional assumptions, however, the result demonstrates that the presence of moderately biased judges will pull de facto law toward those judges' preferred outcomes. For example, consider the numerical example plotted in Figure 3, where perfect application of de jure law would be $\delta_A = 0.45$. Moreover, assume that there is a uniform distribution of biased trial judges on the interval $[\frac{1}{2}, 1]$ and that the baseline procedure is $\eta = \frac{1}{2}$. Then, the proportion of pro-

defendant outcomes under de facto law (with biased judges) can be calculated as follows:

$$\underbrace{\int_{0.5}^{0.875} 2 \cdot (1 - 0.8(1 - \delta_T))\ d\delta_T}_{\text{moderates}} + \underbrace{\int_{0.875}^{1} 2 \cdot 0\ d\delta_T}_{\text{extremists}} = 0.5625$$

With this simple numerical example, $0.5625 - 0.45 = 11.25\%$ more cases get pro-defendant outcomes under de facto law, relative to de jure law.

The result also suggests that if there are large numbers of extremely biased judges, de facto law may be substantially more *pro-plaintiff* than de jure law. This results from the fact that extremely biased judges never receive deference and thus avoid reversals by routinely ruling for plaintiffs in line with the appellate court's predisposition. In this situation, the incentives created by appellate review create a perverse outcome that can make de facto law much more pro-plaintiff than the appellate court prefers.

A subtle, but important, empirical implication emerges from this discussion. Many studies of district judge decision-making find that the effect of a judge's party (as defined by party of appointing president) is smaller in the district courts than in appellate courts (e.g., Zorn and Bowie 2010; Epstein, Landes, and Posner 2013). This result suggests that a set of judges who are biased relative to their reviewing appellate court (e.g., Democratic district judges in a Republican circuit) will have heterogeneous effects on de facto law: some will produce outcomes that are more pro-plaintiff than de jure law, some will produce outcomes that are more pro-defendant than de jure law. As a result, aggregating individual judge effects will tend to *understate* the degree to which trial judges affect outcomes. Researchers may actually find substantial effects of partisanship on outcomes among the subsets of judges who, in the model, receive deference and who exercise procedural discretion. To see this more concretely, again consider the numerical example. Among all defendant-biased judges, 56.25% of cases end up favoring the defendant. Restricting attention only to the most moderate judges, this figure increases dramatically to 75%.

6 Conclusion

Despite institutional limitations on their power, there is reason to believe that trial judges exert substantial control over law in the United States. In this article, I have presented a model of legal decision-making that rationalizes appellate courts' deference to trial judge decisions. This deference results from trial judges' closeness to case facts, which enables them to make more accurate judgments. However, the appellate courts' skepticism of (potentially biased) trial judges runs deep. They are sometimes so skeptical of trial judges that they end up undermining those judges' incentives to exert effort in fact-finding. This is where procedural discretion can help. By allowing a trial judges to use case management procedures to obfuscate their fact-finding and shield themselves from reversals, appellate courts can counterintuitively bring the trial judge closer in line with doctrine.

Taken as a whole, these dynamics mean that de facto law will differ substantially from the appellate court's de jure law. However, exactly *how* it differs is complicated. Moderate judges will pull de facto law toward their preferred legal rules whereas extremist judges will do the opposite. This suggests that empirical research on the effects of partisanship on judging in the trial courts may mask important heterogeneity among judges. Indeed, the model demonstrates how they may influence legal outcomes more than is immediately apparent.

References

Administrative Office of the U.S. Courts. 2018. "Federal Judicial Caseload Statistics 2018 Tables." http://www.uscourts.gov/federal-judicial-caseload-statistics-2018-tables.

American Constitution Society and Federalist Society. 2006. "A Conversation on the Constitution: Perspectives from Active Liberty and A Matter of Interpretation." Video recording. http://youtu.be/qjAYuMLDGyI.

Baker, Scott, and Lewis A. Kornhauser. 2015. "A Theory of Judicial Deference." *Unpublished manuscript* (November 2, 2015). https://perma.cc/C4QG-LJ46.

Beim, Deborah. 2017. "Learning in the Judicial Hierarchy." *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): 904–918.

Boyd, Christina L. 2016. "Representation on the Courts? The Effects of Trial Judges' Sex and Race." *Political Research Quarterly* 69 (4): 788–799.

Boyd, Christina L. 2017. "Gatekeeping and Filtering in Trial Courts." In *The Oxford Handbook of U.S. Judicial Behavior*, eds. Lee Epstein and Stefanie A. Lindquist. Oxford University Press.

Boyd, Christina L., Lee Epstein, and Andrew D. Martin. 2010. "Untangling the Causal Effects of Sex on Judging." *American Journal of Political Science* 54 (2): 389–411.

Boyd, Christina L., and James F. Spriggs. 2009. "An Examination of Strategic Anticipation of Appellate Court Preferences by Federal District Court Judges." *Washington University Journal of Law* and Policy 29:37–81.

Bueno de Mesquita, Ethan, and Matthew C. Stephenson. 2007. "Regulatory Quality Under Imperfect Oversight." *American Political Science Review* 101 (3): 605–620.

Cameron, Charles M., and Lewis A. Kornhauser. 2006. "Appeals Mechanisms, Litigant Selection and the Structure of Judicial Hierarchies." In *Institutional Games and the U.S. Supreme Court*, eds. James R. Rogers, Roy B. Flemming, and Jon R. Bond. University of Virginia Press, 173–204.

Cameron, Charles M., Jeffrey A. Segal, and Donald R. Songer. 2000. "Strategic Auditing in a Political Hierarchy: An Informational Model of the Supreme Court's Certiorari Decisions." *American Political Science Review* 94 (1): 101–116.

Carp, Robert A., and C.K. Rowland. 1996. *Politics and Judgment in Federal District Courts*. Lawrence, KS: University Press of Kansas.

Carrubba, Clifford J., and Tom S. Clark. 2012. "Rule Creation in a Political Hierarchy." *American Political Science Review* 106 (3): 622–643.

Choi, Stephen J., Mitu Gulati, and Eric A. Posner. 2012. "What Do Federal District Judges Want? An Analysis of Publications, Citations, and Reversals." *Journal of Law, Economics, and Organization* 28 (3): 518–549.

Clark, Tom S., and Clifford J. Carrubba. 2012. "A Theory of Opinion Writing in a Political Hierarchy." *Journal of Politics* 74 (2): 584–603.

Cooter, Robert, and Thomas Ulen. 2012. *Law and Economics*. 6th edition. Boston, MA: Pearson Education, Inc.

Dewatripont, Mathias, and Jean Tirole. 1999. "Advocates." *Journal of Political Economy* 107 (1): 1–39.

Dolan, Maura. 2013. "Prop. 8: 'I Have Done My Part,' Retired Judge Vaughn Walker Says." *Los Angeles Times*. `http://articles.latimes.com/2013/mar/26/local/la-me-ln-prop-8-judge-vaughn-walker-20130326`.

Dragu, Tiberiu, and Oliver Board. 2015. "On Judicial Review in a Separation of Powers System." *Political Science Research and Methods* 3 (3): 473–492.

Epstein, Lee, William M. Landes, and Richard A. Posner. 2013. *The Behavior of Federal Judges: A Theoretical and Empirical Study of Rational Choice*. Cambridge, MA: Harvard University Press.

Gailmard, Sean, and John W. Patty. 2013. "Formal Models of Bureaucracy." *Annual Review of Political Science* 15:353–377.

Gennaioli, Nicola, and Andrei Shleifer. 2008. "Judicial Fact Discretion." *Journal of Legal Studies* 37 (1): 1–35.

George, Tracey E., and Albert H. Yoon. 2003. "The Federal Court System: A Principal-Agent Perspective." *St. Louis University Law Journal* 47 (819): 819–834.

Kastellec, Jonathan P. 2013. "Racial Diversity and Judicial Influence on Appellate Courts." *American Journal of Political Science* 57 (1): 167–183.

Kastellec, Jonathan P. 2017. "The Judicial Hierarchy." In *Oxford Research Encyclopedia of Politics*. Oxford: Oxford University Press.

Kim, Pauline T., Margo Schlanger, Christina L. Boyd, and Andrew D. Martin. 2009. "How Should We Study District Judge Decision-Making?" *Journal of Law and Policy* 29 (83): 83–112.

Lax, Jeffrey R. 2011. "The New Judicial Politics of Legal Doctrine." *Annual Review of Political Science* 14:131–157.

Lax, Jeffrey R. 2012. "Political Constraints on Legal Doctrine: How Hierarchy Shapes the Law." *Journal of Politics* 74(3): 765–781.

Patty, John W., and Ian R. Turner. 2016. "Ex Post Review and Expert Policymaking: When Does Oversight Reduce Accountability?" *Unpublished manuscript* (February 9, 2017). doi:10.2139/ssrn.2781750.

Priest, George L., and Benjamin Klein. 1984. "The Selection of Disputes for Litigation." *Journal of Legal Studies* 13:1–55.

Rothstein, Paul F., Myrna S. Raeder, and David Crump. 2012. *Evidence in a Nutshell*. 6th edition. St. Paul, MN: Thomson Reuters.

Staton, Jeffrey K., and Georg Vanberg. 2008. "The Value of Vagueness: Delegation, Defiance, and Judicial Opinions." *American Journal of Political Science* 52 (3): 504–519.

Stephenson, Matthew C. 2011. "Information Acquisition and Institutional Design." *Harvard Law Review* 124:1422–1483.

Talley, Eric. 2013. "Law, Economics, and the Burden(s) of Proof." In *Research Handbook on the Economics of Torts*, ed. Jennifer H. Arlen, 305–329. Northampton, MA: Edward Elgar Publishing Co.

Toobin, Jeffrey. 2013. "Rights and Wrongs: A Judge Takes on Stop-and-Frisk." *New Yorker*, May 27.

Turner, Ian R. 2016. "Working Smart and Hard? Agency Effort, Judicial Review, and Policy Precision." *Journal of Theoretical Politics* 29 (1): 69–96.

Turner, Ian R. 2018. "Reviewing Procedure vs. Judging Substance: The Scope of Review and Bureaucratic Policymaking." *Unpublished manuscript* (May 28, 2018). doi:10.2139/ssrn.3019371.

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

Zorn, Christopher, and Jennifer Barnes Bowie. 2010. "Ideological Influences on Decision Making in the Federal Judicial Hierarchy: An Empirical Assessment." *Journal of Politics* 72

(4): 1212–1221.