Getting a Fair Hearing?
Bias and Error in the Judicial Hierarchy*

Ryan Hübert†

University of California, Davis

July 12, 2018

Abstract

I present a formal model illustrating how institutions that constrain trial judges' influence over legal cases have important, under-appreciated downsides. The model provides three novel insights. First, an unbiased trial judge makes judgments that are systematically biased in favor of powerful litigants. This results from their strategic anticipation of litigant appeals, and it allows them to put less effort into resolving cases. Second, when unbiased judges are assigned randomly to cases, they will produce more errors than a case assignment system allowing trial judges more freedom to select cases. This has implications for judicial selection: diversity on the bench may promote high quality decision-making under flexible case assignment procedures, but not under random case assignment. Third, even when a trial judge is explicitly biased in favor of a litigant, they will often (but not always) make fewer errors in their decision-making than an unbiased judge.

Supplementary material for this article is available
in the Technical Appendix to be made available online.

*All errors and omissions are my own.

†Ryan Hübert is Assistant Professor, Department of Political Science, University of California, Davis. *Email*: `rhubert@ucdavis.edu`. *Website*: `http://www.ryanhubert.com/`.

A key concern for the American legal system is whether litigants can get a fair hearing in court. A set of institutions exist, at least in part, to facilitate this. For example, on the front end of a case, judges are usually randomly assigned in order to prevent them from taking cases in which they may be biased. On the back end of a case, parties who lose in a lower court are generally afforded an opportunity to appeal to an higher court and obtain a reversal of a bad decision. Many of these institutions are expressly designed to minimize the influence of individual judges over legal outcomes. The presumption is that, by constraining judges, these institutions will reduce the number of cases that will be resolved erroneously.

I present a formal model of adjudication illustrating how institutions that constrain trial judges have important, under appreciated downsides. The model provides three novel insights. First, an unbiased trial judge will make systematically biased judgments when subject to appellate review. This occurs because of their strategic anticipation of litigant appeals, and it allows them to put less effort into resolving cases. As a result, they make more errors. Second, when unbiased judges are assigned randomly to cases, they will produce more errors than a case assignment system that allows trial judges freedom to select the cases they hear. Moreover, this has implications for judicial selection: diversity on the bench may promote high quality decision-making under flexible case assignment procedures, but not under random case assignment. Third, even when a trial judge is explicitly biased in favor of a litigant, they will often (but not always) make fewer errors in their decision-making than an unbiased judge.

The common thread that links these three findings is a trial judge's control over case outcomes. Indeed, fundamental features of the adjudication process make it difficult to *completely* constrain judges. Most obviously, by having the authority to make legal judgments (and thus decide which litigant mounts an appeal), judges have a kind agenda-setting power that may be difficult to counteract. They also have discretion to choose how much time, energy and resources to devote to resolving individual cases. This is hard to observe, let alone control. In spite of the conventional wisdom that the American legal system is primarily litigant-driven, the very fact that those litigants are seeking the judgment of a court transfers some degree of agency to presiding judges. The model shows how these factors interact with judge-constraining institutions in counterintuitive—and perverse—ways.

The findings have important implications for the ability of litigants to get a fair hearing in the courts. At a basic level, a court system generating a lot of erroneous judgments will mean that litigants will often get legally incorrect decisions from trial judges. While it may be unreasonable to expect that an adjudication system will have no errors, the model shows how certain institutional choices can exacerbate or mitigate the problem. Moreover, certain kinds of errors are qualitatively different. In the model, errors are systematically biased toward powerful litigants. Thus, over forty years after the publication of Galanter (1974)'s ground-breaking article, I provide a new judge-centered channel through which the "haves" may come out ahead.

Law Application in the Judicial Hierarchy

The bulk of the time, energy and effort put into legal cases happens at the first stages of adjudication—in the trial courts. For example, in the year ending March 31, 2017, there were 344,929 civil and criminal cases terminated in the federal trial courts, as opposed to 59,040 in the lower federal appellate courts (Administrative Office of the U.S. Courts 2017, Tables B-D). In addition to being a much larger presence in the judiciary's docket than appellate courts, trial courts also serve a unique purpose. They do not generally make new law, and they are only tangentially involved in the creation of doctrine. They are instead tasked with developing a deep understanding of individual cases so that they can figure out how to apply the law in those cases. This is often referred to as "law application." A key question is whether they do this well. In other words, if we take existing law as given, how many errors do courts make in resolving cases? These errors have two sources: mistakes and willful non-compliance.

On both counts, existing research presents a relatively optimistic view. Scholars have found little empirical evidence of non-compliance by trial judges (Gruhl 1980; Schanzenbach and Tiller 2007; Randazzo 2008; Boyd and Spriggs II 2009; Epstein, Landes, and Posner 2013; Boyd 2015). These studies largely demonstrate that trial judges strategically alter their decision-making to conform to their appellate court overseers. While compelling, this account underemphasizes the role that litigants play in appeals. Indeed, a trial judge's decision-making does not just occur in the shadow of an appellate court. By making a decision against a litigant, a trial judge is potentially activating that litigant to mount a strong appeal. In the model I present, the trial judge rationally anticipates an appellate court's review posture, but precisely *how* they respond depends on the litigants' behavior.

By emphasizing litigants in the appeal process, this paper fits into a long theoretical literature, largely in law and economics (*e.g.*, Dewatripont and Tirole 1999; Daughety and Reinganum 2000; Cameron and Kornhauser 2005, 2006; Talley 2013). Again, the existing results paint an optimistic picture. Using a model of advocacy, Dewatripont and Tirole (1999) shows that appellate review can improve adjudication by reigning in biased decision-makers. Cameron and Kornhauser (2006) demonstrates how the appeals process minimizes the number of errors in a three-tiered hierarchy. However, both of these models feature judges who are relatively passive recipients of information that is privately known or acquired by litigants. In reality, litigation often involves a more complex informational environment where the litigants and the judge are all working to learn about the "merits" of a case, *i.e.*, how the substantive facts and law map into a decision.

The starting point of the model I analyze below is that each case involves a potentially protracted and costly process to establish the merits (Hornby 2009; Kim et al. 2009). A judge's effort in this process is a key feature. This effort enables the judge to acquire information that will allow her to make a more accurate decision. However, trial judges face dueling incentives. On the one hand, it is costly to exert effort. For example, in some areas of law—such as patent law or employment discrimination law—complexity reigns and judges have to a more active role resolving cases (Posner 2013). Moreover, huge caseloads may impose a limit the amount of time a judge can devote to any given case, increasing the cost of effort. On the other hand, they may be motivated to get the legally correct outcome on a case, either because they are intrinsically motivated to be interested in the case or because the threat of reversal by an appellate court raises the stakes for them. The model incorporates each of these elements.

In order for trial judges to respond to the appeals process, as the compliance literature suggests, appeals must impose a meaningful constraint on trial judges. In practice, appellate courts have a limited set of tools. Most obviously, they can reverse a trial judge's decision. Reversals impose at least two costs. First, they force a trial judge to reopen and reconsider a case, which takes time and resources. For example, reversed decisions are often sent back to the lower court ("remanded") for further proceedings. Second, reversals are often a source of professional embarrassment for trial judges. This may be especially acute for trial judges who have greater opportunity to be promoted to a higher court (Choi, Gulati, and Posner 2012). In the model, the judge faces a cost when she is reversed.[1]

Broadly speaking, the results in this paper make important contributions to at least two bodies of research. First, the model contributes to research on learning in the judicial hierarchy (*e.g.*, Cameron, Segal, and Songer 2000; Kastellec 2007; Lax 2012; Clark and Carrubba 2012; Carrubba and Clark 2012; Beim, Hirsch, and Kastellec 2014; Clark and Kastellec 2013; Beim 2017). In particular, the model demonstrates that law application can be biased even when there is no conflict between upper and lower courts over doctrine.

1. While I do not explore it here, an interesting avenue for future research is the appellate court's incentives to increase or decrease these costs in light of the trial judge's strategic anticipation of reversals.

The novel feature that drives this result is the active role of litigants. Second, the model contributes to existing research on oversight dynamics across political institutions. Some of the closest analogues to the present model come from studies of bureaucracy. For example, researchers have repeatedly demonstrated the importance of bureaucratic effort and expertise for the quality of policy making (*e.g.*, Prendergast 2007; Gailmard and Patty 2007, 2013). Moreover, recent models of bureaucratic incentives, such as Dragu and Polborn (2013), Turner (2016) and Gailmard and Patty (2017), show how review by a supervising body (such as domestic court or an international human rights court) can affect a bureaucrats' effort, and thus outcomes.

In the next section, I describe the baseline model of adjudication. In the subsequent sections, I first demonstrate the limits of appellate review, then I show the limits of random case assignment procedures. Finally, I extend the model to examine whether the main findings apply in the context of explicitly biased judges.

Model

I analyze a stylized civil case between two litigants, P and D, initially heard by a trial judge, T (which I refer to as the "trial judge" or simply "judge"), and then appealed to an appellate court, A.

Cases. In the model, the "merits" of the case are represented by $\omega \in \{P, D\}$, indicating whether the law supports the plaintiff or the defendant. A case's merits are a function of both law and facts, and so ω can be micro-founded using a case space approach (Lax 2011). The merits of the case are initially unknown to all players, although they share a common prior belief that the merits support the plaintiff with probability π. Essentially, this prior belief represents all the publicly available information—albeit imprecise—about the specific case before the court. For example, if $\pi = \frac{1}{4}$, then all the players think that there was a 25%

chance that the merits support the plaintiff. I refer to this as the "strength" of the case.

The fact that there is uncertainty over the case merits does not imply that individual litigants do not have better knowledge over some specific details of the case. Indeed, other models, such as Cameron and Kornhauser (2005, 2006), have emphasized how litigants' private information over facts may shape adjudication. I abstract away from this issue to focus on the *trial judge's* process of making sense of what she has learned about a case. The mere fact of ongoing litigation suggests that there is some degree of uncertainty about how the case should be resolved (Priest and Klein 1984). Moreover, American court procedures favor a relatively open exchange of private information via, for example, liberal discovery rules (Kane 2013). However, this does not mean cases are "easy" since, in the words of U.S. District Judge Hornby: "although the underlying facts may be not all that uncertain, they need an authoritative legal exposition" (Hornby 2009, p. 90). The judge's process of coming to this "authoritative legal exposition" is a core feature of the model.

I adopt an important assumption about the nature of the case merits, which rules out the possibility that players in the game lie about what they learn. Substantively, this could reflect prohibitively high sanctions for perjury combined with a high probability of detection.

Assumption 1 (no fabrication)**.** The case merits cannot be fabricated.

At each point in the game, the player who is making a decision may or may not know for sure whether the merits support the plaintiff. I label player i's belief that the merits support the plaintiff as $\hat{\pi}_i$. Throughout the paper, I refer to both π and $\hat{\pi}_i$ as "beliefs," and it should be apparent from the context that π is the prior belief, whereas $\hat{\pi}_i$ is a generic representation of a player's (prior or posterior) belief at any point in the game.

Payoffs. The appellate court gets higher utility when the outcome of the case reflect the merits than when it does not, and I normalize these utilities to 1 and 0, respectively. Given that appellate review is not discretionary for the federal circuit courts, I assume that any cost of review is a sunk cost and thus plays no role in the court's decision-making. The trial judge gets a benefit $\delta > 0$ from resolving the case in a manner consistent with the case merits. Accordingly, the judge in this baseline model is unbiased. This assumption represents a "best case scenario" for studying how much error the judge makes, since it ensures that she has no ex ante incentive to tilt her decision making in favor of either litigant. In a later section, I relax this assumption and examine how the results apply when a judge has biased preferences.

The judge also bears two costs. The first is a cost for effort on the case, $\frac{1}{2}c_T e^2$, where $c_T > 0$. This cost reflects both the concrete costs associated with case management and as well as the opportunity cost associated with devoting extra time and energy to the present case and not to other cases. The second is a cost she incurs whenever she is reversed, $k > 0$. This cost reflects the concrete costs associated with reversals (*e.g.*, having to reopen a case) as well as her career or reputational concerns (*e.g.*, the professional embarrassment of being reversed).

The litigants prefer to receive a judgment in their favor, whether at the trial level or at the appellate level, and they receive a benefit normalized to one if they do (and zero otherwise). If they lose they can opt to make a costly appeal to reverse the judge's decision. The losing litigant, $L \in \{P, D\}$, may choose to appeal the case, and pays a cost $\frac{1}{2}c_L a_L^2$. There is no cost associated with filing the appeal, but they prepare a brief, b_L, that offers the appellate court a justification for why the case should be reversed. To rule out substantively unusual equilibria, I assume L faces an arbitrarily small reputational cost, $\varepsilon_L > 0$, whenever they provide information to the appellate court that undermines their own appeal (*i.e.*, $b_P = D$ and $b_D = P$).

Because the cost parameters c_P and c_D represent both the opportunity costs of the litigants and the technology available to the litigants for building a strong appeal, they also capture, in a sense, the relative power of the litigants. I define the litigant with lower costs as the "more powerful" litigant, and in the analysis, I assume that the defendant is the more powerful litigant, so that $c_P > c_D$. The main qualitative lessons of the formal results do not depend on this assumption.

> **Assumption 2** (powerful litigant). The defendant is the "more powerful" litigant, so that $c_D < c_P$.

Finally, I make the following assumption about the players' costs, which aids the analysis by ruling out equilibria where the effort of one or more of the players is so high that the judicial system is able to resolve every single case correctly. Substantively, these equilibria would represent uninteresting scenarios whereby the players face no genuine trade-off about how to allocate resources.

> **Assumption 3** (sufficient resource constraints). For the litigants, $c_P > c_D > 1$. For the judge, $c_T > \bar{c}_T$, where $\bar{c}_T$ is defined in the Appendix.

Sequence. The model begins with a judge being assigned to the case. Formally, δ is chosen via an arbitrary procedure, and it is observed by all players. In the baseline analysis in the next section, I take δ as given, but in the subsequent section, I explicitly study the process by which judges are assigned. Then, the judge makes some unobserved effort $e \in [0, 1]$, which helps her learn whether the merits of the case support the plaintiff or defendant. I denote what she learns by m_T. With probability e she learns whether the merits support the plaintiff or not ($m_T = \omega$) and with probability $1 - e$ she learns nothing (which I denote by $m_T = \phi$). To simplify the analysis, I assume that m_T is public information.[2] After learning (or not) the case merits, the judge then makes a judgment, $x \in \{D, P\}$, where $x = D$ is a decision in favor of the defendant and $x = P$ is a decision in favor of the plaintiff.

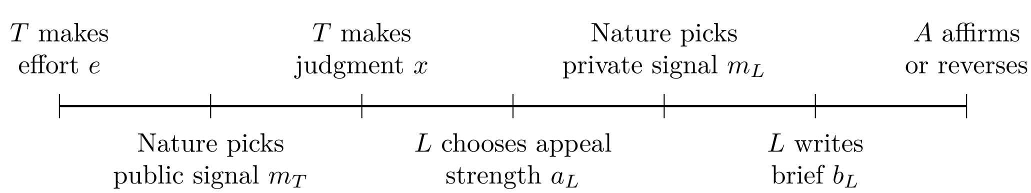
The losing litigant has the option to appeal the judge's decision and accordingly decides how much unobserved effort to put into an appeal, $a_L \in [0, 1]$. While this effort has a

2. The fact that the judge's information is public is a simplification that obviates the need to study multiple equilibria driven by the judge's indifference over her message to the appellate court. It could be replaced by a very small benefit from revealing—or very small cost for concealing—confirmatory information.

mathematically identical format as that of the judge, it is substantively different. Appeals by losing litigants typically revolve around the issue of whether the lower court erred in its application or interpretation of the law. This requires that the judge first be given an opportunity to give authoritative legal exposition to the case at hand. In fact, two important provisions of the Federal Rules of Appellate Procedure highlight why a model of appeals may feature the type of sequential effort described here. The final judgment rule bars litigants from appealing decisions that are not final, thus allowing judges an opportunity to fully consider the case at hand. More importantly, however, is the rule requiring preservation of error. This rule requires litigants properly preserve errors made by the trial judge by (1) making an objection on the record, and (2) giving the trial judge an opportunity to correct the error (Castanias and Klonoff 2008). Taken together, these rules empower a trial judge to make the first attempt to resolve the legal controversies in a case.

As with the judge's effort, the losing litigant learns the merits of the case or not, $m_L \in \{\omega, \phi\}$, where ω is learned with probability a_L. Unlike before, the information learned by the litigant is private, and the litigant can decide whether to communicate the information to the appellate court or conceal it. However, if the litigant receives no information as a result of their effort, then they cannot communicate any information. Finally, the appellate court decides whether to affirm or reverse the district court, $r \in \{0, 1\}$. Figure 1 depicts the model's sequence.

Figure 1: *Model Timeline*



Strategies and Equilibrium. I derive perfect Bayesian equilibria in pure strategies.

The adjudication strategy for the judge is a mapping from their information sets yielding a pair $(e, x) \in [0, 1] \times \{P, D\}$, specifying a level of effort and a judgment. A strategy for the losing litigant is a mapping from their information sets yielding a pair $(a_L, b_L) \in [0, 1] \times \{m_L, \phi\}$, specifying a level of effort and a brief to submit to the appellate court. A strategy for the appellate court is a mapping from its information sets yielding $r \in \{0, 1\}$, which specifies whether the appellate court affirms or reverses the district court's judgment. I adopt the following assumption.

> **Assumption 4** (indifference)**.** When indifferent, the judge rules in favor of the defendant and the appellate court's reversal strategy favors the defendant.

In the next sections, I focus on deriving the main results. Supporting information about the derivations, including proofs, are collected in the Appendix.

Limits of Appellate Review

First, it is straightforward to observe that the appellate court always reverses cases it knows are incorrect. That is, if either the judge or the losing litigant reveals hard information that definitively demonstrates that the district court's judgment is erroneous, then the appellate court reverses. However, if neither the judge nor the litigant offers hard information to the appellate court, then the appellate court will reverse a judgment whenever it has a belief that the judgment is more likely to be incorrect than correct.

A losing litigant who discovers an error by the trial judge during its appeal never has an incentive to conceal that information as it leads to a reversal and a final disposition in its favor. In light of this, the losing litigant has to decide whether to expend resources to mount a strong appeal to discover an error. Clearly, if the judge has already presented a solid legal and factual interpretation of the case to the appellate court—as reflected by revelation of hard information, $m_T = \omega$—then the losing litigant would be wasting resources by exerting effort to mount a costly appeal. In this situation, $a_L^*(m_T = \omega) = 0$. However, if the judge does not present hard information to justify its decision, then the litigant's interim expected utility of mounting a strong appeal is:

$$U_L(a_L, m_T = \phi) = \begin{cases} a_P \pi - \frac{1}{2} c_P a_P^2 & \text{if } L = P \\ a_D(1-\pi) - \frac{1}{2} c_D a_D^2 & \text{if } L = D \end{cases}$$

The optimal effort to put into an appeal in the absence of hard information therefore balances the potential benefit of discovering an error by the judge with the cost of mounting the appeal. Formally, it is derived by maximizing the interim utility function:

$$a_L^*(m_T = \phi) = \begin{cases} \frac{\pi}{c_P} & \text{if } m_T \neq \omega \text{ and } x^* = D \\ \frac{1-\pi}{c_D} & \text{if } m_T \neq \omega \text{ and } x^* = P \end{cases}$$

Next, consider the judge's strategy, which has two components: the judgment and a level of effort in managing and resolving the case. The judge rules in favor of the defendant if the interim expected utility of doing so is greater than or equal to the interim expected utility of ruling in favor of the plaintiff. For simplicity, suppose that the appellate court always affirms judgments in the absence of hard information (I discuss this more below). Then, this condition is:

$$\underbrace{(1-\hat{\pi}_T)\delta + \hat{\pi}_T a_P^*(\delta - k)}_{U_T(x=D)} \geq \underbrace{\hat{\pi}_T \delta + (1-\hat{\pi}_T) a_D^*(\delta - k)}_{U_T(x=P)} \tag{1}$$

Conversely, the judge rules in favor of the plaintiff when the inequality is strictly satisfied in the other direction. If the judge has successfully discovered whether the merits support the plaintiff or not (*i.e.*, learned ω), this condition is easy to check since $\hat{\pi}_T = 0$ or $\hat{\pi}_T = 1$. Specifically, if the judge learns that the merits support the plaintiff, then they rule in favor of the plaintiff, and if not, then they rule in favor of the defendant.

Despite exerting effort, sometimes a judge will not successfully learn the case merits before having to make a ruling. In those situations, the judge will need to make a best guess using the prior belief that the merits support the plaintiff. I refer to those kinds of uninformed judgments as "predispositions," and I formally denote them as x^*_ϕ, where the ϕ in the subscript indicates that the judge makes the judgment without hard information. Using the expressions from equation (1) and substituting π for $\hat{\pi}_T$ and the equilibrium values for a^*_D and a^*_P, the judge rules for the defendant in the absence of information if:

$$(1-2\pi)\delta + (\delta - k)\left(\frac{\pi^2 c_D - (1-\pi)^2 c_P}{c_D c_P}\right) \geq 0 \tag{2}$$

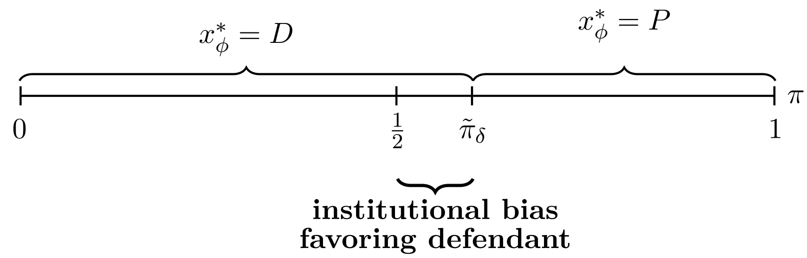
There is a value $\tilde{\pi}_\delta$ (which depends on δ) such that (2) holds with equality. Intuitively, this represents how certain the judge has to be that the merits favor the plaintiff before being willing to rule in favor of the plaintiff in the absence of hard information.[3] That is, a judge who is uninformed rules in favor of the defendant for all $\pi \leq \tilde{\pi}_\delta$ and in favor of the plaintiff for all $\pi > \tilde{\pi}_\delta$.

This analysis yields an important insight. A completely "impartial" trial judge would base their judgment solely on the merits (or what she believes about them), and not on other considerations. Formally, they would rule in favor of the defendant if they believe $\hat{\pi}_T \leq \frac{1}{2}$ and in favor of the plaintiff if they believe $\hat{\pi}_T > \frac{1}{2}$. As I have shown, however, the judge bases its decision on the threshold $\tilde{\pi}_\delta$, which is not generally equal to $\frac{1}{2}$. Figure 2 illustrates.

Notice that there is a "tilt" toward one of the litigants in the judge's decision-making. This represents an *institutional bias* in favor of one of the litigants, despite the fact that the judge does not have preferences that favor either of the litigants. Whether this institutional bias favors the plaintiff or the defendant overall depends on how reversal averse the trial

3. The proof of Lemma A3 explicitly characterizes $\tilde{\pi}_\delta$ as a function of the parameters in equation (2).

Figure 2: *Decision-making by an uninformed judge*



judge is. If $k > \delta$, then this institutional bias favors the more powerful litigant (here, the defendant). In this situation, the trial judge dislikes reversals sufficiently that they minimize their chance of being reversed by stacking the deck in favor of more powerful litigants, who would mount stronger appeals. If $k < \delta$, then the institutional bias favors the weaker litigant (here, the plaintiff). In this situation, the trial judge cares very little about being reversed. She rules against the powerful party in the absence of information because she knows the powerful party will mount a more effective appeal. After all, she values the losing party's effort to figure out the merits.

> **Proposition 1.** Due to appellate review, the trial judge's decision making is not generically impartial. Formally, $\tilde{\pi}_\delta \neq \frac{1}{2}$ for almost all combinations of the model's parameters.

In theory, the trial judge's aversion to reversals allows the appellate court to credibly sanction a trial judge who does a bad job adjudicating. However, this kind of sanction often ends up being "over powered" in the sense that it induces the trial judge to engage in perverse behavior that actually worsens her decisions. More specifically, as long as δ is sufficiently low, then a judge will end up being *less* impartial, and thus *less* compliant with the appellate court, as the cost of being reversed increases.

> **Proposition 2.** If δ is not too high, then a judge is less impartial when reversal costs are higher. Formally, if $\delta < k$, then $\frac{1}{2} \leq \tilde{\pi}_\delta(k) < \tilde{\pi}_\delta(k')$ for $k' > k$.

More generally, regardless of whether the institutional bias ends up favoring the plaintiff or defendant overall, the judge's adjudication strategy always *leans* toward the more powerful litigant. This follows directly from the fact that the judge is reversal averse ($k > 0$). Whatever other incentives the judge has, she also always has an interest in minimizing her probability of being reversed. An effective way to do this is to rule in favor of the powerful litigant more often.

As I mentioned above, the derivation of this optimal judgment relied on a supposition that the appellate court affirms district court judgments in the absence of information. However, because of this institutional bias toward one litigant, the appellate court may infer that the judge made a bad judgment and thus reverse it. Lemma A3 in the Appendix formally derives the condition under which it is optimal for the appellate court to affirm district court judgments. In brief, as long as $\tilde{\pi}_\delta$ is sufficiently low so that the judge does not tilt adjudication *too* far in favor of the defendant (*i.e.*, $\tilde{\pi}_\delta < \overline{\pi}$, where $\overline{\pi}$ is defined in the proof) then the judge's optimal judgment is according to $\tilde{\pi}_\delta$. When this condition fails, the judge's judgment is constrained (although not completely) by appellate review.

The judge also makes a decision about how much effort to invest in adjudicating the case. Like the losing litigant, the judge exerts the level of effort that maximizes her expected utility:

$$U_T(e; x_\phi^*, a_D^*, a_P^*, \cdot) = \begin{cases} e\delta + (1-e)\left[\pi\delta - (1-\pi)a_D^*(k-\delta)\right] - \dfrac{c_T}{2}e^2 & \text{if } x_\phi^* = P \\ e\delta + (1-e)\left[(1-\pi)\delta - \pi a_P^*(k-\delta)\right] - \dfrac{c_T}{2}e^2 & \text{if } x_\phi^* = D \end{cases} \quad (3)$$

As is reflected in equation (3), the judge's interim utility depends on the judgment she makes when only partially informed (*i.e.*, x_ϕ^* analyzed in previous section). This is due to the fact that the trial judge's ruling affects who the appealing litigant is, which in turn affects the effort that the judge makes. Maximizing (3) with respect to e yield's the judge's optimal level of effort:

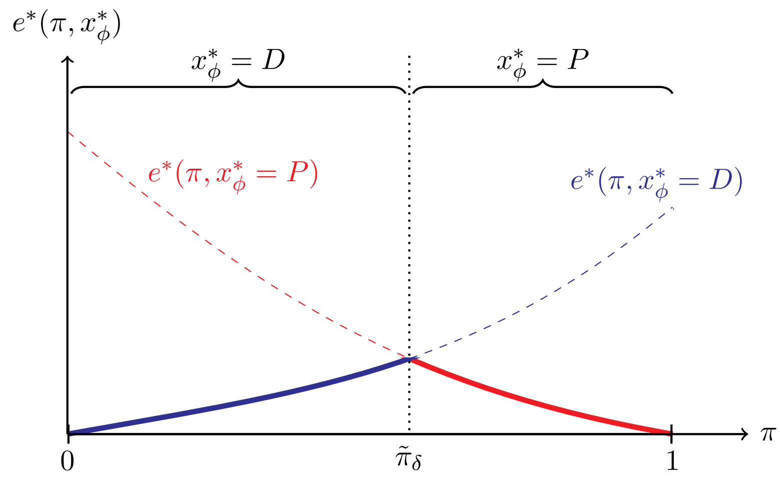
$$e^*(x^*_\phi) = \begin{cases} \frac{1}{c_T c_D}\left[(1-\pi)(c_D-(1-\pi))\delta+(1-\pi)^2 k\right] & \text{if } x^*_\phi = P \\ \frac{1}{c_T c_P}\left[\pi(c_P-\pi)\delta+\pi^2 k\right] & \text{if } x^*_\phi = D \end{cases}$$

Litigant-driven appellate review affects the trial judge's effort in two important ways. First, the threat that the losing litigant will discover a reversible error induces the trial judge to exert more effort. To see this, notice that e^* always increases as k increases. This underscores a standard "agency benefit" of litigant-driven appeals: as long as judges dislike being reversed, it provides appellate courts with an effective way to mitigate a trial judge's incentive to shirk.

However, the agency benefit of litigant-driven appeals is limited by the trial judge's ability to choose *which* litigant mounts the appeal. When the trial judge is sufficiently reversal averse ($k > \delta$), then she tilts her decision-making against the weak litigant. Since weaker litigants are less likely to mount strong appeals and discover reversible errors, this allows the judge to reduce her effort. On the other hand, when the trial judge is not too reversal averse ($k < \delta$), then she tilts her decision-making against the powerful litigant. In this case, the losing litigant can mount a stronger appeal and since the trial judge is not too concerned about being reversed, she "free-rides" on the powerful litigant's effort. This allows her to exert less effort. Figure 3 plots $e^*(x^*_\phi = P)$ and $e^*(x^*_\phi = D)$ as a function of π, for some fixed parameter values

The thick line is the equilibrium level of effort, given that x^*_ϕ changes as π changes. Notice that the judge's equilibrium judgment allows her to minimize the level of effort that she exerts. For example, suppose that the appellate court could require her to rule in favor of the plaintiff whenever she has no information about the merits of the case. Then, she would exert effort according to $e^*(\pi, x^*_\phi = P)$ for all π, and whenever $\pi < \tilde{\pi}_\delta$, her effort would

Figure 3: *Equilibrium effort*



be strictly higher (the dashed line) than her effort when she gets to issue her own preferred judgment (which favors the defendant).

> **Proposition 3.** Suppose that the systemic bias in favor of the powerful litigant is not too great. Then the trial judge's optimal judgment minimizes judicial effort. Formally, suppose $\tilde{\pi}_\delta \leq \overline{\pi}$. Then, $e^*(x^*_\phi = D) \leq e^*(x^*_\phi = P)$ if and only if $x^*_\phi = D$.

The analysis has so far described how a specific trial judge adjudicates a specific case. In the next section, I explore the broader institutional implications of these findings given that a large set of cases must be resolved by a pool of heterogeneous judges. Before proceeding, however, I note that there exists an equilibrium of the game, which is formally described in Proposition A1 in the Appendix.

Case Assignment and Diversity on the Bench

The analysis has so far established two important findings: (1) there is an institutional bias in a trial judge's decision-making even when the judge is not biased, and (2) trial judges minimize their effort through their strategic choice of who wins a case. These two findings have important implications for the way that the aggregate set of cases gets assigned to, and thus resolved by, individual judges. In this section, I explore this issue formally.

In order to assess the effects of the model on decision-making, I must first articulate an appropriate normative benchmark. Many models of adjudication focus on the courts' error correction role (*e.g.*, Cameron and Kornhauser 2006). Indeed, the model in this paper revolves around the process by which trial judges and litigants acquire information about the merits of the case. In the model, "accuracy" is presumed to be an achievable goal. The logic of the model is less well suited for contexts where the law is fluid or not well established. However, accuracy *is* an important concern of appellate courts. For example, in *Strathie v. Department of Transportation* (1983), the Third Circuit found that the district court had failed to address or acknowledge many of the arguments raised by the plaintiff, a school bus driver who sued the Pennsylvania Department of Transportation over a provision prohibiting hearing-aid wearers from driving school buses. The trial judge's decision was not vacated due to any ideological or doctrinal disagreement, but rather due to its seemingly low quality.

In light of this, a reasonable question to ask is: how accurate are case outcomes in light of the equilibrium behaviors described? I refer to this as "quality," which I formally define as follows.

> **Definition 1.** The **quality** of a judge's equilibrium decision-making is represented by the probability that the outcome of a case will be correct. I denote this by $\xi(x^*_\phi, \delta) \equiv \Pr(y^* = \omega | x^*_\phi, \delta, \cdot)$, where $y^* \in \{P, D\}$ is the (post-appeal) outcome of the case generated by the equilibrium.

First, I explicitly characterize $\xi(x^*_\phi, \delta)$:

$$\xi(x^*_\phi, \delta) = \begin{cases} e^*(x^*_\phi, \delta) + [1 - e^*(x^*_\phi, \delta)][1 - \pi(1 - a^*_P)] & \text{if } x^*_\phi = D \\ e^*(x^*_\phi, \delta) + [1 - e^*(x^*_\phi, \delta)][1 - (1 - \pi)(1 - a^*_D)] & \text{if } x^*_\phi = P \end{cases} \tag{4}$$

$\xi(x^*_\phi, \delta)$ gives a measure of the probability that a judge will make an error on a given case.

Now, I use this measure to explore the *aggregate* ramifications of the model in a judicial hierarchy with a distribution of cases entering the trial courts and a distribution of judges available to hear those cases.

Suppose that there is a pool of trial judges, each of whom could be assigned to a case. Each judge differs in the degree to which they are intrinsically motivated by different kinds of cases. Formally, there is heterogeneity among judges in their δs. Then, in a judicial system with K issues that can come before the courts, each judge can be characterized by a vector, $(\delta_1, \delta_2, ..., \delta_K)$, indicating their degree of intrinsic motivation for each issue. I assume judges do not differ on any other dimension, such as their costs, although different forms of heterogeneity is an interesting avenue for future research.

To focus on the core substantive lessons, I examine the simplest situation where there are two kinds judges, two issues and two possible levels of issue motivation, zero and $\delta > 0$. Moreover, I assume that the two kinds of judges are interested in different cases, so that one kind of judge can be described by $(\delta_1 = \delta, \delta_2 = 0)$ while the other can be described by $(\delta_1 = 0, \delta_2 = \delta)$.[4] I refer to the former judges as "issue 1 judges." Similarly, I refer to the latter judges as "issue 2 judges." For example, issue 1 judges may be very interested in civil rights cases but not in securities litigation, while issue 2 judges may be very interested in securities litigation, but not civil rights cases. The fact that the judges differ with respect to their issue interests implies the following result, which is useful in the subsequent analysis.

> **Lemma 1.** An issue 1 judge is a lower quality adjudicator on an issue 2 case than an issue 2 judge, and vice versa. Formally, $\xi(x^*_\phi, \delta) > \xi(x^*_\phi, 0)$.

This result comes from two sources, one obvious and one subtle. First, judges working on issues they do not care about have less intrinsic motivation and thus exert less effort. This is is a straightforward implication of the model. Second, since those judges are *also* less

4. The qualitative lessons can be recovered from a more complicated setup, although with some loss of clarity.

impartial, they strategically choose litigants in a way that *further* reduces their equilibrium effort. This, in turn, reduces the quality of a mismatched judge's adjudication even more than one would expect from simply having a less issue motivated judge hearing a case.

Next, I describe two case assignment procedures, which are at the extremes of a broader set of potential assignment procedures. Suppose the mix of judges given by q, which is the probability that a given judge is an issue 1 judge. Accordingly, $1-q$ is the probability that a given judge is an issue 2 judge. Moreover, suppose an individual case is an issue 1 case with probability p and an issue 2 case with probability $1-p$. First, I assume that each case is randomly assigned to a judge and that random assignment gives equal weight to each judge. Then, using $\xi(x^*_\phi, \delta)$ from above, the expected proportion of correct decisions under random assignment, Δ_r, is given by the following,

$$\Delta_r \equiv p\big(q\xi(x^*_\phi, \delta) + (1-q)\xi(x^*_\phi, 0)\big) + (1-p)\big(q\xi(x^*_\phi, 0) + (1-q)\xi(x^*_\phi, \delta)\big).$$

Second, at the opposite extreme, suppose instead that when cases come into the court system, trial judges volunteer to take them. It is straight forward to observe that, all else equal, an individual issue 1 judge will choose to take issue 1 cases over issue 2 cases and an individual issue 2 judge will choose to take issue 2 cases over issue 1 cases. I set aside the specific micro-foundation of this process in order to focus on core substantive lessons. Accordingly, I assume that the volunteer assignment process satisfies two assumptions. First, cases end up on the dockets of judges who have an interest in them as long as there is room on the dockets of those judges. Second, all judges' dockets must be equally populated by cases. Suffice to say, the aggregate effect of this assignment system is that cases will be assigned to judges based on their interest so long as those judges are not overburdened. Before proceeding, note that the analysis below would be unchanged with a slightly more nuanced procedure where judges *specialize* in an issue and having cases randomly assigned to them within their specialization.

Given that there may be substantial differences in the number of issue 1 (or 2) judges than there are issue 1 (or 2) cases and I assume dockets must be equal across judges, a quantity of interest will be $q - p$. If $q - p > 0$, then there are more issue 1 judges than there are issue 1 cases, and vice versa if $q - p < 0$. The knife-edge case of $q = p$ represents a situation where there are just enough issue 1 judges to take all issue 1 cases and just enough issue 2 judges to take all issue 2 cases. Again, using $\xi(x^*_\phi, \delta)$ from above, the expected proportion of correct decisions under volunteer assignment, Δ_v, is given by the following,

$$\Delta_v \equiv \begin{cases} p\xi(x^*_\phi, \delta) + (1-p)\left(\left(\frac{1-q}{1-p}\right)\xi(x^*_\phi, \delta) + \left(\frac{q-p}{1-p}\right)\xi(x^*_\phi, 0)\right) & \text{if } q - p > 0 \\ p\xi(x^*_\phi, \delta) + (1-p)\xi(x^*_\phi, \delta) & \text{if } q - p = 0 \\ p\left(\left(\frac{q}{p}\right)\xi(x^*_\phi, \delta) + \left(\frac{p-q}{p}\right)\xi(x^*_\phi, 0)\right) + (1-p)\xi(x^*_\phi, \delta) & \text{if } q - p < 0 \end{cases}$$

Let $m \equiv \max\{q - p, p - q\}$ represent the divergence between the mix of judges and the mix of cases. Substantively, m is the proportion of cases that have to be resolved by mismatched judges. Then, some simple algebra reduces this to:

$$\Delta_v \equiv (1-m)\xi(x^*_\phi, \delta) + m\xi(x^*_\phi, 0) \tag{5}$$

Having characterized the quality of adjudication under these two assignment procedures, the following result demonstrates an important downside to random assignment.

> **Proposition 4.** In the baseline model of adjudication, random assignment of judges to cases leads to strictly fewer accurate decisions than voluntary assignment. Formally, $\Delta_r < \Delta_v$.

Proposition 4 compares two extreme assignment procedures. More generally, an assignment procedure that allows judges more flexibility to pick their cases will generate more accurate—and thus higher quality adjudication—than random assignment. There are two obvious institutional implications. First, this finding underscores a benefit of specialized courts, since they allow judges to select into a set cases that interest them. What drives Proposition 4 in the fact that a judge's effort is affected by their level of issue interest. Proposition 4 suggests that one way to improve the quality of adjudication is to allow judges to choose into cases that interest them. This provides a rationale for why specialized courts are often focused on certain areas of law that are complex and require a lot of active involvement from a judge (*e.g.*, bankruptcy, patents, etc.).

A second institutional implication of Proposition 4 involves judicial selection. In light of the predictions of this model, one optimal way to select judges is to appoint those who are highly motivated on every single issue that could come before the courts. Realistically, though, this is not feasible. Judicial appointments almost always require choices between candidates with differing priorities. One potential response is to appoint judges from a wide variety of backgrounds, and thus with a variety of priorities. Indeed, in recent years, the importance of "diversity on the bench" has been front and center in debates about the composition of the judiciary (Alliance for Justice 2016). I represent "diversity on the bench" using the parameter q, the mix of judges on the bench. Formally, I explore how Δ increases or decreases as q changes under the two case assignment mechanisms. First, under random assignment, the quality of adjudication either increases monotonically in q or decreases monotonically in q, depending on the value of p:

$$\frac{\partial \Delta_r}{\partial q} = (2p-1)\underbrace{\left(\xi(x^*_\phi,\delta) - \xi(x^*_\phi,0)\right)}_{+}$$

If there are fewer issue 1 cases than issue 2 cases, *i.e.* $p < \frac{1}{2}$, Δ_r is decreasing as q increases. Then, under random assignment, the best allocation of judges for maximizing the accuracy of adjudication is to have a bench of *entirely* issue 2 judges. In contrast, if $p > \frac{1}{2}$, then the best allocation of judges for maximizing the accuracy of adjudication is to have a bench of entirely issue 1 judges. To the extent that courts care about increasing the accuracy of their decisions, random assignment encourages a *homogeneous* bench. Intuitively, under random assignment a court can "cut its losses" by maximizing the number of correct outcomes in whichever issue has the larger set of cases.

The quality of adjudication under voluntary assignment, in contrast, is non-monotonic in q, meaning that it increases and decreases depending on q:

$$\frac{\partial \Delta_v}{\partial q} = -m'(q) \underbrace{\left(\xi(x^*_\phi, \delta) - \xi(x^*_\phi, 0)\right)}_{+}$$

where $m'(q) = \frac{\partial}{\partial q} \max\{q - p, p - q\}$. In particular, for all $q < p$, $m'(q) < 0$ and Δ_v increases with q, and for all $q > p$, $m'(q) > 0$ and Δ_v decreases with q. Intuitively, under voluntary assignment, more cases are resolved correctly as mix of judges, q, approaches the mix of cases, p. Interestingly, $q = p$ can be considered a formal representation of a particular diversity benchmark: the judges who sit on the bench are motivated by cases in the proportion to which those cases percolate into the legal system. Using this benchmark, increasing diversity (*i.e.*, moving q towards p) always improves the quality of adjudication under voluntary assignment but improves the quality of adjudication under random assignment only when it makes the bench more homogeneous toward whichever issue dominates the set of cases.

Proposition 4 has already established that $\Delta_r < \Delta_v$ for all q. However, increasing diversity on the bench (in the sense of moving q toward p) magnifies the benefit of voluntary assignment over random assignment. To see this point, I examine the quality gap between these two procedures:

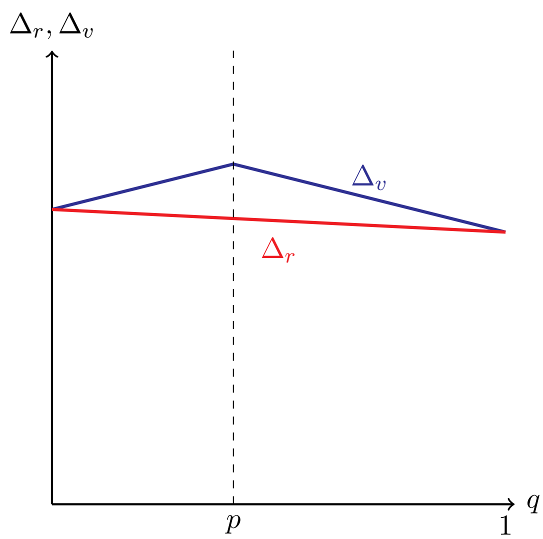
$$\Delta_v - \Delta_r = \begin{cases} (1 + 2q(1-p))\xi(x^*_\phi, \delta) - 2q(1-p)\xi(x^*_\phi, 0) & \text{if } q < p \\ (1 + 2p(1-q))\xi(x^*_\phi, \delta) - 2p(1-q)\xi(x^*_\phi, 0) & \text{if } q > p \end{cases}$$

Taking the derivative with respect to q yields

$$\frac{\partial}{\partial q}[\Delta_v - \Delta_r] = \begin{cases} 2(1-p)(\xi(x^*_\phi, \delta) - \xi(x^*_\phi, 0)) & \text{if } q < p \\ -2p(\xi(x^*_\phi, \delta) - \xi(x^*_\phi, 0)) & \text{if } q > p \end{cases}$$

Thus, as q approaches p from the left ($q < p$), the quality gap increases and voluntary assignment becomes even more beneficial relative to random assignment. And, the same conclusion holds as q approaches p from the right ($q > p$).[5] To see this more clearly, consider the following graph, where I assume $\xi(x^*_\phi, \delta) = 0.5$, $\xi(x^*_\phi, 0) = 0.25$ and $p = 0.4$.

Figure 4: *Quality under case assignment procedures*



The higher blue line represents the quality of adjudication under voluntary assignment, while the lower red line represents the quality of adjudication under random assignment. As is apparent, voluntary assignment is always higher quality than random assignment, and the benefits of diversity are particularly pronounced under that assignment mechanism. One interesting issue that I do not explore here is the effect of these dynamics on the mix of cases coming into the courts. Essentially, I assume p is exogenous and unchanged as q changes. An

5. Note: the derivative is negative, meaning that as q decreases, the quality gap increases.

interesting avenue for future research would be to analyze the impact of increasing diversity ($i.e.$, changing q) on litigants' decisions to file cases in court.

The result in Proposition 4 applies to a model of adjudication as described in the analysis of the previous sections. Random assignment, however, also speaks to a set of concerns about judicial bias, which is not captured in the core model. Specifically, if judges volunteer to take cases, they may take cases in which they are biased in favor of one party or another. Random assignment is supposed to prevent this. Importantly, Proposition 4 does *not* say that random assignment is *generally* worse than volunteer assignment. Indeed, the result holds for the model of adjudication analyzed in the previous section. So, to the extent that the dynamics of that model operate in real-world decision-making and judges have heterogeneous preferences over which cases they are interested in, then Proposition 4 establishes an often overlooked disadvantage of random assignment: it puts too many "dispassionate" judges on cases, reducing the overall quality of adjudication.

Biased Judges

In the analysis thus far, there is no conflict over doctrine between the trial judge and appellate court. This helped focus attention on the way that the appeals process can alter a judge's decision-making, and underscored that a lack of biased preferences is not a necessary condition for institutionally biased outcomes. However, appellate review and random assignment might also serve to mitigate the deleterious impact of judges who are biased in favor or against particular types of litigants. In this section, I extend the baseline analysis to briefly examine this issue. Here, I focus specifically on whether biased judges reduce the quality of adjudication. In a separate paper, I analyze in more detail the implications of preference conflicts between appellate and trial courts and show how trial judges can exercise substantial control over substantive law (Hübert 2017).

I now assume there are two kinds of judges, "biased judges" and "unbiased judges." As before, all players know whether the judge is biased or unbiased. The model of adjudication proceeds as in the baseline model with a few modifications. First, I assume the biased judge prefers that the plaintiff win, regardless of the particular circumstances of the case. If the final disposition favors the plaintiff, then a biased judge receives a benefit $\beta > 0$, and if the final disposition favors the defendant, then a biased judge receives no benefit. The unbiased judge is indifferent over case outcomes, and thus does not favor one party or the other. Second, because the judge may now have an incentive to conceal information, I also relax the assumption that the judge's information is public, and I assume that a judge gets a very small benefit $\varepsilon_T > 0$ from writing a decision that provides the appellate court with incontrovertible information that the judgment is correct.[6] I also focus on equilibria supported by reasonable off-equilibrium path beliefs in order to avoid analyzing those that are driven by implausible beliefs in situations that do not arise.[7] Given that the trial judge may now have an incentive to conceal his or her information, two types of equilibria emerge, which are characterized by Lemma A8 in the Appendix.

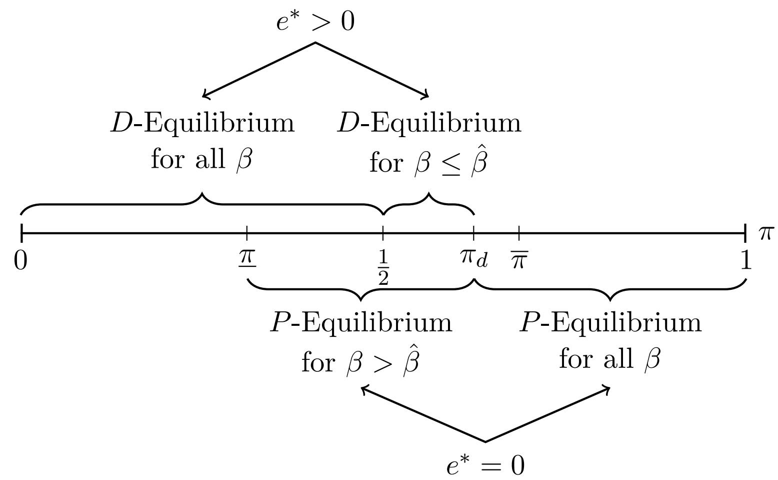
The equilibrium behavior of the appellate court and the litigant are the same across the two types of equilibria, but the trial judge's decision-making varies. Figure 5 depicts the equilibria.

Due to the presence of bias in this revised model, it is now harder to incentivize the trial judge to exert effort when the plaintiff has a very strong case. For all cases where $\pi > \pi_d$, there only exists a P-Equilibrium where the judge rules in favor of the plaintiff in the absence of hard information and makes no effort. Moreover, there also exists a "bad" equilibrium where the trial judge makes no effort and rules for the plaintiff *despite the fact* that the

6. Specifically, I assume $0 < \varepsilon_T < \pi(1-\pi)\beta$. This rules out equilibria where the judge conceals information that supports the judgment that she or he makes.

7. In particular, a deviation from the equilibrium judgment in the absence of information leads the appellate court to infer that the trial judge has no information. The details are in the proof of Lemma A8.

Figure 5: *Equilibria with plaintiff-biased judges.*



available information would support a judgment for the defendant. On the other hand, appellate review still constrains a biased judge. whenever the plaintiff's case is relatively weak (*i.e.*, $\pi < \frac{1}{2}$), then there always exists an equilibrium where the judge rules for the defendant in the absence of information. For all $\pi \in (0, \underline{\pi})$, that is the only equilibrium and the presence of a biased judge does *not* lead to more pro-plaintiff outcomes despite the fact that the judge is biased in favor of the plaintiff.

Would an unbiased judge produce higher quality adjudication? If a biased judge adjudicates according to a P-Equilibrium, then her decisions are clearly lower quality than an unbiased judge. In such equilibria, she never exerts effort, whereas an unbiased judge exerts *some* effort. Since effort increases the quality of decisions, it immediately follows that biased judges make lower quality decisions. However, if a biased judge adjudicates according to a D-Equilibrium, then a biased judge actually produces a higher quality outcome (in the sense of Definition 1) than an unbiased judge. Moreover, since the D-Equilibrium is the only kind of equilibrium for all $\pi < \frac{1}{2}$ or for all $\pi < \pi_d$ and $\beta \leq \hat{\beta}$, in many situations, a biased judge will always make higher quality decisions.

Proposition 5. Suppose that the plaintiff's case is not too strong (*i.e.*,

$\pi < \pi'$, where π' is defined in the proof). Then, exists an equilibrium where adjudication by an unbiased judge is lower quality than adjudication by a plaintiff-biased judge. Moreover, if either the judge is not too biased in favor of the plaintiff or the case is a very weak case (or both), then adjudication by an unbiased judge is always lower quality than adjudication by a plaintiff-biased judge.

While these results are more conditional, they provide some solace for those worried about the effect of biased judges on legal outcomes. Indeed, biased judges are often constrained by appellate review, and when they are, they are higher quality adjudicators than unbiased judges. This also suggests that a random case assignment system may still perform worse than another case assignment system even when some (or all) judges are biased. The optimal system with biased judges, however, will not be the voluntary system described above. Instead, the optimal case assignment system with biased judges will systematically assign cases favoring one litigant to judges who are biased *against* that litigant. To illustrate, consider a hypothetical example. Suppose that all Republican appointees are biased against plaintiffs in employment discrimination cases while all Democratic appointees are biased in favor of plaintiffs. If these cases generally favor defendants (in the sense of a low π), then the optimal assignment system in light of the judges' biases would be to systematically assign Democrats to employment discrimination cases.

Such a system would clearly be impractical, but some more feasible assignment mechanisms could push in this direction. I leave a full analysis to future work, but allowing judges to specialize into issue areas and then be assigned cases randomly may reap some of the benefits of having biased judges adjudicate cases while minimizing the costs.[8] In any event, a necessary feature of a case assignment system that allocates biased judges optimally is sufficient numbers of judges of opposing biases to hear cases. Even in the context of biased judges, this suggests a potential benefit of efforts to diversify the trial court bench.

8. Briefly, the logic goes as follows. A biased judge will find it beneficial to choose to specialize in issue areas where their bias is most activated. Then, random assignment ensures that they receive a proportional share of cases favoring the party they're biased against.

Conclusion

Views about which actors drive outcomes in the American judiciary are bifurcated. On the one hand, most scholars of judicial politics view the influence of individual judges as deeply consequential. For example, most models of judicial politics do not include litigants as active players. On the other hand, in the legal and law and economics traditions, litigants are the primary drivers of outcomes. Indeed, many models in these traditions view judges as passive adjudicators, often sharing a set of "team" preferences (*e.g.*, Dewatripont and Tirole 1999; Cameron and Kornhauser 2005, 2006).

In this paper, I incorporate both strategic litigants *and* strategic judges. I show that institutions designed to constrain judges have perverse consequences. First, judges' decision-making will be systematically biased even when judges are not themselves biased. Moreover, this systematic bias benefits powerful litigants and allows trial judges to minimize the effort they put into resolving cases. Second, random assignment of judges produces more error than other ways to assign cases to judges. Diversity on the bench interacts with case assignment procedures in nuanced ways, but the optimal system combines a diverse bench with a case assignment procedure that incorporates judicial preferences over cases. Third, even when judges are biased in favor of a litigant, they often make more accurate decisions than unbiased judges. These results provide reason to be skeptical that the judicial hierarchy encourages high quality adjudication. They also reinforce the notion that institutions designed to limit the influence of decision-makers can counterintuitively magnify their negative influence.

References

Administrative Office of the U.S. Courts. 2017. *Federal Judicial Caseload Statistics 2017 Tables.* http://www.uscourts.gov/statistics-reports/federal-judicial-caseload-statistics-2017-tables.

Alliance for Justice. 2016. *Broadening the Bench: Professional Diversity and Judicial Nominations.* Accessed July 11, 2018. https://www.afj.org/wp-content/uploads/2014/11/Professional-Diversity-Report.pdf.

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. 2015. "The Hierarchical Influence of Courts of Appeals on District Courts." *Journal of Legal Studies* 44 (1): 113–141.

Boyd, Christina L., and James F. Spriggs II. 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.

Cameron, Charles M., and Lewis A. Kornhauser. 2005. "Decision Rules in a Judicial Hierarchy." *Journal of Institutional and Theoretical Economics* 161 (2): 264–292.

———. 2006. "Appeals Mechanisms, Litigant Selection and the Structure of Judicial Hierarchies." In *Institutional Games and the U.S. Supreme Court,* edited by James R. Rogers, Roy B. Flemming, and Jon R. Bond. University of Virginia Press.

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.

Carrubba, Clifford J., and Tom S. Clark. 2012. "Rule Creation in a Political Hierarchy." *American Political Science Review* 106 (3): 622–643. doi:10.1017/S000305541200024X.

Castanias, Gregory A., and Robert H. Klonoff. 2008. *Federal Appellate Practice and Procedure in a Nutshell.* St. Paul, MN: Thomson-West.

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. doi:10.1017/S0022381611001794.

Clark, Tom S., and Jonathan P. Kastellec. 2013. "The Supreme Court and Percolation in the Lower Courts: An Optimal Stopping Model." *Journal of Politics* 75 (1): 150–168.

Daughety, Andrew F., and Jennifer F. Reinganum. 2000. "On the Economics of Trials: Adversarial Process, Evidence, and Equilibrium Bias." *Journal of Law, Economics, and Organization* 16 (2): 365–394.

Dewatripont, Mathias, and Jean Tirole. 1999. "Advocates." *Journal of Political Economy* 107 (1): 1–39. doi:10.1086/250049.

Dragu, Tiberiu, and Mattias K. Polborn. 2013. "The Administrative Foundation of the Rule of Law." *Journal of Politics* 75 (4): 1038–1050. doi:10.1017/S002238161300100X.

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. 2007. "Slackers and Zealots: Civil Service, Policy Discretion, and Bureaucratic Expertise." *American Journal of Political Science* 51 (4): 873–889.

Gailmard, Sean, and John W. Patty. 2013. *Learning While Governing: Expertise and Accountability in the Executive Branch.* Chicago: University of Chicago Press.

———. 2017. "Participation, Process, & Policy: The Informational Value of Politicized Judicial Review." 37 (3): 233–260.

Galanter, Marc. 1974. "Why The "Haves" Come Out Ahead: Speculations on the Limits of Legal Change." *Law & Society Review* 9 (1): 95–160.

Gruhl, John. 1980. "The Supreme Court's Impact on the Law of Libel: Compliance by Lower Federal Courts." *The Western Political Quarterly* 33 (4): 502–519.

Hornby, D. Brock. 2009. "The Business of the U.S. District Courts." Chap. 6 in *Judges on Judging: Views from the Bench,* 3rd, edited by David M. O'Brien, 88–96. Washington, DC: CQ Press.

Hübert, Ryan. 2017. "Getting Their Way: An Informational Theory of Ex Post Deference in Appellate Review."

Kane, Mary Kay. 2013. *Civil Procedure in a Nutshell.* 6th. St. Paul, MN: Thomson-West.

Kastellec, Jonathan P. 2007. "Panel Composition and Judicial Compliance on the US Courts of Appeals." *Journal of Law, Economics, and Organization* 23 (2): 421–441.

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. doi:`10.1146/annurev.polisci.042108.134842`.

———. 2012. "Political Constraints on Legal Doctrine: How Hierarchy Shapes the Law." *Journal of Politics* (New York) 74 (3): 765–781. doi:`10.1017/S002238161200045X`.

Posner, Richard A. 2013. *Reflections on Judging.* Cambridge, MA: Harvard University Press.

Prendergast, Canice. 2007. "The Motivation and Bias of Bureaucrats." *American Economic Review* 97 (1): 180–196.

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

Randazzo, Kirk A. 2008. "Strategic Anticipation and the Hierarchy of Justice in U.S. District Courts." *American Politics Research* 36 (5): 669–693.

Schanzenbach, Max M., and Emerson H. Tiller. 2007. "Strategic Judging Under the U.S. Sentencing Guidelines: Positive Political Theory and Evidence." *Journal of Law, Economics, and Organization* 23 (1): 24–56. doi:10.1093/jleo/ewm002.

Strathie v. Department of Transportation. 1983. 716 F.2d 227. LexisNexis Academic (August 22, 2015).

Talley, Eric. 2013. "Law, Economics, and the Burden(s) of Proof." In *Research Handbook on the Economic Analysis of Tort Law,* edited by Jennifer H. Arlen, 305–329. Edward Elgar Publishing Co.

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