Title
Learning through Litigation: How Judicial Process Shapes Decision Making in the U.S. District Courts

Author
Hubert, William Ryan

Publication Date
2016

Peer reviewed|Thesis/dissertation

Learning through Litigation:
How Judicial Process Shapes Decision Making in the U.S. District Courts

By

William Ryan Hubert

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Political Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Sean Gailmard, Chair
Professor Sean Farhang
Professor Kevin Quinn
Professor Eric Schickler

Fall 2016

Abstract

Learning through Litigation:
How Judicial Process Shapes Decision Making in the U.S. District Courts

by

William Ryan Hubert

Doctor of Philosophy in Political Science

University of California, Berkeley

Professor Sean Gailmard, Chair

The U.S. District Courts resolve the vast majority of cases in the U.S. federal legal system, but we know much less about them than we do about the U.S. Courts of Appeals or the U.S. Supreme Court. The disproportionate focus on appellate courts has reinforced the perception that courts are primarily policymaking institutions. However, district courts rarely make new law, and are mostly at the front line of law *enforcement*. In the aggregate, the quality of their decisions on individual cases determines the degree to which laws are applied fairly and consistently. However, their work is not trivial, and the quality of outcomes depends on whether district judges have sufficient incentives to produce high quality decisions. Indeed, district courts are the primary fact-finders in legal disputes. Judges often have to cajole litigants, witnesses and attorneys into being forthcoming about what they know. They have to rule on a flurry of motions about what information can be discovered or admitted as evidence. District judges are therefore intricately involved in creating the case records, which means that they must engage in some costly effort to figure out the facts of cases they oversee.

A district judge's work on any given case constitutes a process of "learning through litigation." The importance of this process for the enforcement of laws is the point of departure for this dissertation. In a series of three papers, I analyze the agency problem confronting appellate courts: how can district judges be induced to undertake maximum effort to learn the facts and the law pertaining to the legal cases before them when this effort is costly, but imperfectly observed by principals in the appellate courts? Moreover, what are the consequences for litigants (and society more broadly) of district court institutions and practices that serve to induce effort and compliance by district judges?

To address these questions, I analyze three original models of district court decision mak-

ing. The models present three main takeaways: (1) due to their informational advantage, deferring to district judges is often optimal for appellate courts, even when those judges are biased; (2) dispassionate judges are often harder to motivate to exert effort to resolve cases than either activist or biased judges; (3) endogenous appeals by litigants can provide perverse incentives to unbiased "legalist" judges to tilt judgments toward powerful litigants. Moreover, the models I study in these papers offer observationally equivalent explanations for commonly observed empirical phenomena, such as compliance with legal doctrine and in-group bias. Taken together, the models underscore the important role that courts play in enforcing laws effectively and accurately. Indeed, if there is significant variation in district judges' effort to resolve cases, then laws will be enforced inconsistently. Perhaps more insidiously, the results suggest that low quality decision making by district judges may affect different kinds of litigants differently.

Introduction

> *In this age of instant reporting and tweets and blogs, there's a temptation to latch on to any bit of information, sometimes to jump to conclusions. But when a tragedy like this happens, with public safety at risk and the stakes so high, it's important that we do this right. That's why we have investigations. That's why we relentlessly gather the facts.* ***That's why we have courts****.*
>
> – President Barack Obama[1]

Courts are institutions for learning. This observation is so obvious to most people as to barely evoke a reaction. Trials are dramatized in the popular media as forums for the presentation and interpretation of evidence. Businesses carefully structure contracts in order to avoid litigation and the costly discovery associated with it. Landmark Supreme Court cases periodically "update" American governance by overturning seemingly out-of-date laws when new circumstances are brought to light.

In a formal sense, "learning" is the idea that new information is acquired and used to update one's belief about some uncertain quantity. Indeed, when modeling this process, we suppose that an actor begins with some prior belief about the world, receives some new (hopefully clarifying) information, then forms a new (posterior) belief. The new belief either reinforces the prior belief or undermines it, thus allowing the actor to be more or less informed than they were before. Information is indispensable for government officials who are tasked with making important decisions. Bureaucrats need to know how proposed regulations will affect particular industries. Members of Congress need to know how their vote on a piece of legislation will affect their electoral prospects. Prosecutors need to know whether a suspect committed a crime when deciding whether to charge them.

[1]Speaking April 19, 2013 in response to the Boston Marathon bombings, transcript at `http://www.whitehouse.gov/the-press-office/2013/04/19/statement-president`. Emphasis added.

The asymmetric allocation of information generates important dilemmas for officials, as well as opportunities for institutional designers. In the context of bureaucratic politics, a recent book by Gailmard and Patty (2013a) underscores this point. Using a series of models, the authors demonstrate how institutions can promote and/or hinder learning by decision makers. In these models, there are two main issues confronting officials: (1) some actors are privately informed about relevant issues and the officials' task is to elicit that private information from those actors; and (2) some actors may acquire expertise that is helpful for policymaking and the officials' task is to induce them to do so. As the authors persuasively demonstrate, such generic informational problems offer powerful explanations for the rise of various institutional structures, such as the closeness of the Securities and Exchange Commission to the interests it regulates (to help elicit information) or the development of a civil service with strong job protections (to induce actors to acquire expertise).

Similar information asymmetries exist in the context of legal disputes, and scholars have long been interested in the ways that the institutions of the U.S. federal courts resolve these problems (or don't). A typical court case involves some degree of uncertainty, either about the facts of a case, the law that applies to the case, or the preferences of the presiding judge(s). To the extent that political scientists have focused on leaning in the courts, the bulk of that attention has been devoted to studying how appellate courts craft new legal rules. For example, Klein (2002) shows that Courts of Appeals judges learn from judges in other circuits when making law in uncharted areas. Carrubba and Clark (2012a, 2012b) analyze models of rule-creation where lower courts take advantage of their private information over case facts and their ability to write high quality opinions in order to pull legal rules toward their ideal points. Clark and Kastellec (2013) presents a model of Supreme Court *certiorari* where the court can learn from a sequence of lower court decisions, and thus decide to resolve a new legal issue at the "optimal" time. Beim (2016) examines how the U.S. Supreme Court can learn about optimal doctrine from observing the (sometimes conflicting) decisions of two circuits on the same legal issue.

Indeed, studying the appellate courts has many advantages. Appellate courts are most directly involved in policymaking since their decisions resolve legal questions and have precedential value. Their rulings therefore affect a wide range of people all at once. Moreover, data on appellate courts (especially federal courts) are relatively easy to collect. This data has been used to vastly expand our understanding of judicial behavior, ranging from the way that race affects judicial decisions to the frequency that judges and justices depart from precedent. In recent decades, we have learned so much about appellate judges that political scientists have become mostly convinced that they are simply "politicians in black robes."

There are drawbacks to the disproportionate focus on the appellate courts, however. First and foremost, district courts resolve the vast majority of legal disputes that come before the federal judicial system. For example, for the year ending March 31, 2015, there were 264,124 cases terminated in the U.S. District Courts, but only 54,756 appeals were terminated in the U.S. Courts of Appeals. Moreover, only 60.9% of cases terminated upon appeal were terminated on the merits, and of those, only 8.6% (2,868 cases total) were reversed or remanded. All-in-all, only 5.2% of cases terminated at the appellate level (on the merits or not) involved a reversal or a remand, which indicates that almost all legal disputes are resolved in the district courts.[2]

Second, the work of district courts is very different than the work of the appellate courts. District court decisions do not create precedent, and thus have less of an impact on legal doctrine than decisions of appellate courts. However, district courts are at the front line of law *enforcement*. The main concern, then, is the degree to which they are able to make decisions that adequately and accurately reflect the state of existing law. This work is not trivial. District courts are the primary fact-finders in legal disputes. They often have

[2]Statistics available at: `http://www.uscourts.gov/statistics-reports/federal-judicial-caseload-statistics-2015-tables`.

to cajole litigants, witnesses and attorneys into being forthcoming about what they know. They have to rule on a flurry of motions about what information can be discovered or admitted as evidence. Thus, district judges are intricately involved in *creating* the case record, which means that they must engage in some costly effort to figure out the facts of the case. Moreover, since district courts have original jurisdiction, they must also determine which precedents to apply and how to apply them. This too is arduous.

District courts and appellate courts also have different decision making processes. At the trial court level, a single judge presides over each case. At the appellate level, on the other hand, there is usually a panel of judges who make decisions by majority vote. The fact that appellate judges sit on panels introduces a separate set of concerns about preference aggregation (Kornhauser and Sager 1986; Lax 2007; Landa and Lax 2009) and intracourt bargaining (Lax and Cameron 2007; Carrubba et al. 2012). District judge behavior is shaped less by their fellow judges in the district, and more by litigants, attorneys and prosecutors, as well as by their place at the bottom of the judicial hierarchy. Variation in the quality of the attorneys representing litigants or prosecutors pursuing criminal charges will make a district judge's work to resolve a case more or less difficult. Moreover, the supervising appellate court's standard of review can also affect the district judge's behavior, especially if the district judge is sensitive to being reversed upon appeal.

Some recent work has focused attention on the key factors in district courts' decision making: litigants and hierarchical review. Cameron and Kornhauser (2006) offer a team model of adjudication where decisions to appeal are seen as costly signals by litigants that allow appellate courts to infer whether an error was committed by a lower court. The inference is that a truly guilty party may not find it worthwhile to pursue several costly appeals whereas an innocent party will. Talley (2013) analyzes the role that litigants play in signaling the strength of their cases through their collection and presentation of evidence. To the extent that appellate review matters in these settings, it is only to correct errors made by district judges who passively receive (noisy) signals from litigants. However, the prospect of review, and possible reversal, also affects the decisions that a judge makes in the initial stages of litigation. As a result, it is not obvious why district judges should be modeled as passive recipients of information instead of active participants in the process of acquiring information.

Indeed, few models of learning in the judicial hierarchy fully model the district judge's work on cases. As a matter of fact, a district judge has two main decision problems. First, they decide how much effort to exert in resolving cases correctly. For example, in litigation over the prescription drug Seroquel, the district court was intricately involved in the management of the case. At one point during the litigation, the court sanctioned the defendant for noncooperation in discovery, generating a twenty-eight page order full of factual details and analysis (*In re: Seroquel Products Liability Litigation* 2007). If we treat case facts as exogenously given, then we implicitly assume away one of the district judge's main strategic tradeoffs. And, because this process is costly for the judge, institutions can alter the judge's incentives. An appellate court that reviews a district judge's decisions more closely may, somewhat perversely, undermine that judge's incentive to work hard to resolve cases.

A district judge's second decision problem involves their judgments on cases, conditional on any information they acquire during the course of litigation. Many scholars have examined district judges' compliance with doctrine, and have found evidence that these judges are indeed compliant (Randazzo 2008; Boyd and Spriggs 2009; Epstein, Landes, and Posner 2013; but see Schanzenbach and Tiller 2006). Yet, whether a judge is compliant is also a function of their success in figuring out what the "right" decision is. For example, suppose that a liberal district judge oversees an employment discrimination case where a woman alleges that her employer wrongfully terminated her based on her gender. If that district judge is reviewed by a generally conservative appellate court, then he or she may exert little effort to resolve the case, anticipating that the appellate court expects to see a pro-defendant outcome. Moreover, if the district judge is averse to being reversed, then they will simply rule for the defendant. However, what if the facts of the case are so egregious that even the appellate court would have found for the plaintiff? If the district judge does not devote much effort on the case, they may never learn this. Yet, the district judge's judgment will *appear* compliant, despite it being actually non-compliant.

A district judge's work on a case constitutes a process that I refer to as "learning through litigation." The importance of this process for the enforcement of laws is the point of departure for this dissertation. In a series of three papers, I analyze the agency problem confronting appellate courts: how can district judges be induced to undertake maximum effort to learn the facts and the law pertaining to the legal cases before them when this effort is costly, but imperfectly observed by principals in the appellate courts? Moreover, what are the consequences for litigants (and society more broadly) of district court institutions and practices that serve to induce effort and compliance by district judges?

To address these questions, I provide three original models of district court decision making. The first, presented in a paper entitled "'Getting Their Way': An Informational Theory of *Ex Post* Deference in Appellate Review" presents a theory of adjudication that offers an informational rationale for deference to district judges. This deference is "ex post" because—contrary to common understandings of standards of review—it does not require an appellate court to pre-commit to affirming district judge decisions. Deference emerges because, in equilibrium, district judges exert some effort to resolve cases, making them more informed about a case than the reviewing appellate court. This informational advantage creates an incentive for appellate courts to defer to district judges, even when those judges are biased. I show that the degree to which this deference is optimal for the appellate court depends on the informational environment. When district judges are unable to convey information to the appellate court, deference occurs more often. Somewhat surprisingly, deference occurs precisely because it is more difficult to hold district judges accountable. Conversely, whenever information can be conveyed, the appellate court defers less often. This increased accountability perversely leads some otherwise hard working district judges to shirk, making the appellate court *worse off* despite being able to scrutinize district judges more closely.

The empirical implications of these findings are not intuitive. For example, the most biased judges and the least activist judges[3] work harder in informational contexts where their decisions can be scrutinized more closely. I evaluate these findings in light of the fact that district judges in the federal court retain discretion over when to end a civil case. They may decide to issue a summary judgment before trial, or continue the proceedings through trial. In the paper, I discuss how this decision amounts to choosing between informational environments. Indeed, the model predicts that both the most biased and the least activist judges will end cases later in the process. Moreover, such predictions may provide an alternative explanation for empirical findings that portray district courts to be "compliant." Indeed, the model predicts that appellate courts do not defer to the kinds of judges who issue judgments on cases at later stages of litigation. Because commonly used datasets contain only cases with appeals that have *published opinions*, these datasets may over-represent cases that are ended later.[4] The theory in the paper thus provides an explanation for seemingly contradictory observations of district judges' ability to "get their way" and empirical studies showing them compliant.

The second paper, titled "The Downsides of Dispassionate Judges: A Theory of Costly Law Enforcement in District Courts," looks more closely at the role that district judges' preferences play in the quality of outcomes. The model features a career-concerned district

[3] I distinguish between bias and activism in the paper. Bias reflects the degree to which a judge wishes to depart from appellate doctrine in the absence of appellate review, whereas activism reflects the degree to which they are willing to pursue their preferred outcome in the face of appellate review.

[4] An interesting avenue for future empirical research would be to explore this issue in more detail.

judge who may exert some effort to resolve a case before issuing a judgment. The judgment can then be appealed by a losing litigant, who can exert their own effort upon appeal. The model demonstrates how (1) district judges tilt their decision making to favor powerful litigants, and (2) relatively dispassionate judges do this even more. Due to this, outcomes are *less accurate* as judges become more dispassionate. As with activist judges in the previous paper, dispassionate judges are not necessarily biased. Rather, they vary with the degree to which they are willing to pursue their goals given that they face the prospect of appellate review. While the main results of the paper feature a judge who is not biased, I extend the model to show that even judges who are explicitly *biased* can generate more accurate decision making than dispassionate judges.

The main result of the model—that dispassionate judges generate less accurate outcomes—provides an alternative explanation for the *appearance* of bias in district court decision making. For example, litigants occasionally petition for the removal of judges from cases because of their backgrounds. Indeed, scholars have found convincing empirical evidence that judges of different backgrounds rule differently on cases affecting their group (Boyd, Epstein, and Martin 2010; Kastellec 2013; Boyd 2015). However, to the extent that these judges *care more* about cases affecting people like them, the model predicts that their decision making will differ because it will be *more accurate*. Thus, rather than viewing such empirical findings as evidence of bias, the model suggests that they could also be taken as evidence of *shirking* on the part of judges who feel little interest in such cases.

The third paper, titled "Endogenous Appeals," builds on the findings of the previous paper to examine the conventional wisdom that endogenous appeals generate more efficient outcomes (*e.g.*, Dewatripont and Tirole 1999; Cameron and Kornhauser 2006). I argue that litigants' endogenous appeals may undermine the accuracy of decision making because those litigants may have the ability to highlight a district judge's errors. Specifically, if litigants are unevenly matched, or if cases are "close calls" and both litigants are relatively bad at mounting appeals, then the equilibrium is inefficient. Moreover, the paper uses three normative criteria to show how banning endogenous appeals can actually improve social welfare. By more fully modeling the district judge's strategic behavior, the model therefore provides an important set of caveats for scholars whose research highlights the advantages of endogenous appeals. Indeed, accountability via endogenous appeals can have perverse effects whenever district judges play an important role in determining the quality of litigation outcomes.

The models presented in these three papers present three main takeaways: (1) due to their informational advantage, deferring to district judges is often optimal for appellate courts, even when those judges are biased; (2) dispassionate judges are harder to motivate to exert effort to resolve cases; (3) endogenous appeals by litigants can provide perverse incentives to unbiased "legalist" judges to tilt judgments toward powerful litigants. Moreover, the models I study in these papers offer observationally equivalent explanations for commonly observed empirical phenomena, such as compliance and in-group bias. Taken together, the models underscore the important role that courts play in enforcing laws effectively and accurately. Indeed, if there is significant variation in district judges' effort to resolve cases, then laws will be enforced inconsistently. Perhaps more insidiously, the results suggest that low quality decision making by district judges may affect different kinds of litigants differently.

The imbalance between our knowledge of appellate courts and our knowledge of district courts has reinforced the perception that courts are primarily policymaking institutions. Many of the conclusions of the appellate court literature are not themselves incorrect, but the relative lack of focus of the lowest levels of the judicial hierarchy is akin to scholars of the bureaucracy restricting their attention to the political appointees at the top of the federal departments while ignoring the civil and foreign service employees. This dissertation therefore speaks to a broad set of normative and policy concerns about the ability of courts to adequately and fairly enforce laws. This is especially relevant for laws affecting historically disadvantaged groups, such as employment discrimination laws, most of which have provisions granting a private right of action in federal courts.

Paper I

“Getting Their Way”: An Informational Theory of *Ex Post* Deference in Appellate Review

> *As I say, if I were a district judge… you know, a district judge can… well you know, those of you in here know. There are a lot of non-reviewable ways in which he can make the case come out right.*
>
> – Former U.S. Supreme Court Justice Antonin Scalia[5]

District courts play an outsized role in the adjudication of federal court cases, but they also sit at the bottom of the federal judicial hierarchy where they are overseen by appellate courts. The question of whether appellate courts are sufficiently able to reign in district judges is very much unresolved. Many observers and commentators are of two minds: on the one hand, the district courts are afforded a great deal of autonomy, while on the other hand, they are greatly constrained by appellate review. Why this apparent contradiction? One common explanation advanced by legal scholars is that discretion is exercised by lower courts when authority runs out. When district judges find no clear guidance from precedent or face new situations requiring flexibility, discretion is essential (Kim 2007). But district courts exercise discretion in other situations, too. For example, suppose that precedent is well established on a particular issue, but that a district court openly deviates from this precedent in some way. If an appellate court defers to the district court's judgment, even when it considers it to be a bad decision, this too is discretion.

In this paper, I present a model of adjudication under appellate review that provides an informational rationale for deference to district court decisions. The theory rests on the observation that district judges play an important role in the process of adjudicating cases

[5]This statement was made during a panel on the U.S. Constitution sponsored by the American Constitution Society and the Federalist Society (American Constitution Society and Federalist Society 2006).

that come into the federal courts. In particular, district judges have to exert some effort to figure out which litigant should ultimately prevail. This information is valuable because it allows the district court to make a better decision. However, this information, if acquired, may be private to the district court and thus may impose an agency cost on the principal. Indeed, the informational environment in which adjudication occurs plays a crucial role in both determining an appellate court's optimal review and incentivizing district judges to exert effort to resolve cases. Moreover, an appellate court's "standard of review" need not be an *ex ante* commitment to let certain decisions stand, as is commonly understood (see, for example, ch. 8 of Castanias and Klonoff 2008). Deference may emerge *ex post* whenever an appellate court is restrained in its ability to learn about the facts of a case.

Such observations also offer a possible explanation for why empirical studies routinely demonstrate a high degree of adherence to appellate doctrine while judges themselves routinely assert (and occasionally express frustration) that district judges are often able to "get their way." In the absence of its own source of information, an appellate court's belief about what the district judge learned about a case will structure its review posture. A highly activist judge—that is a judge for whom the outcome of the case is more salient—will work harder to resolve cases and will make more accurate decisions, which induces appellate courts to defer more often to these judges. However, the incentive to defer may be undermined if that judge is too biased. Deference, therefore, depends on the informational context in which adjudication occurs and the district judge's level of activism and bias.

The structure of civil procedure in the U.S. federal courts allows district judges some leeway to determine the environment in which they make their final judgments. Specifically, district judges can choose whether to end cases early (such as with a summary judgment) or late (such as post-trial). Setting aside the issue of litigant settlement,[6] I argue that this

[6]An interesting avenue for future research would be to incorporate strategic litigants into the two informational environments I study here.

decision allows district judges to *choose* the kind of appellate review that their decision receives. Moreover, since certain kinds of judges will make this decision differently, it raises questions about empirical findings demonstrating a high degree of compliance with appellate courts. In brief, the results in this paper suggest that district judges who are the most compliant will be the ones who are more likely to show up in commonly used datasets on federal court adjudication.

The paper is structured as follows. First, I describe the civil litigation process in the U.S. federal courts and discuss the role that information plays in this process. Next, I set up the model and derive benchmark behavior of a district judge in the absence of appellate review. Then, I analyze the model of adjudication in two informational environments: simple review and endogenous information. In the "simple review" environment, an appellate court cannot observe the true state of the world when reviewing the district judge's decision, regardless of what the district judge learned. In the "endogenous information" environment, the appellate court may be able to observe the state of the world if the district judge's effort yields a "smoking gun" that it is able to communicate to the appellate court. After analyzing the model in these two informational environments, I explore how the process of civil procedure combined with the incentives generated by these two informational environments may affect *when* a district judge issues her final judgment on a case.

1 Litigation as Information Acquisition in Federal Courts

Perhaps the most distinctive feature of judging in the district courts (and trial courts, more generally) is that the judge is deeply involved in figuring out what exactly happened to generate the conflict(s) that resulted in litigation. Given that many cases only end up in litigation if they are sufficiently complicated, this often involves a long and potentially drawn out back-and-forth between the judge, litigants, witnesses and experts. As a result, the federal courts have adopted sets of rules to govern this process, including rules for the production of evidence (the Federal Rules of Evidence), rules governing the resolution of civil actions (the Federal Rules of Civil Procedure) and rules governing criminal prosecutions (the Federal Rules of Criminal Procedure).

The first of these rules, the Federal Rules of Civil Procedure, were adopted in 1938 in order to harmonize the procedures used on civil cases in the federal courts (Wright and Leipold 2008). These rules, which have been amended over the years to reflect new legal realities and shifting norms, now provide a fairly standardized process for litigants seeking relief in federal courts. As Rule 1 states, this process serves an important function: "to secure the just, speedy, and inexpensive determination of every action and proceeding." Implicit in this phrase is the idea that resolving cases involves a certain degree of hard work to discover what the "just" outcome should be. In more abstract terms, the process governing the work of the federal district courts can be viewed as an extended period of information acquisition. This process is managed by the district judge, who has ultimate authority over the final disposition of the case.

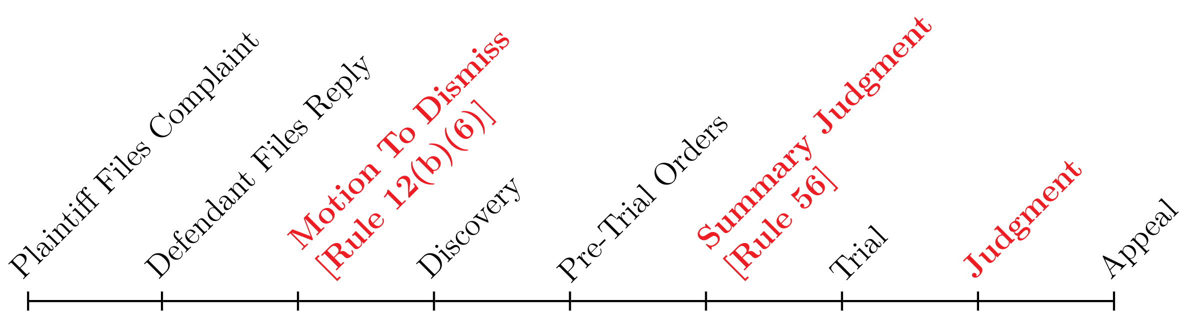


Figure 1: Federal Civil Procedure

Figure 1 depicts the basic trajectory of a federal civil case. A case begins with a plaintiff filing a complaint in federal court, to which the defendant files a reply. At this point, the judge may rule on a defendant's motion to dismiss in accordance with Rule 12(b)(6). If the motion is granted, the case ends. If not, the case continues to discovery, where the litigants begin an often lengthy process of exchanging relevant information under the partial supervision of the court. At the completion of discovery, the judge issues a pre-trial order that will structure the trial proceedings. Either litigant can then opt to seek a summary judgment, which ends the litigation without a trial. If summary judgment is not granted, then a trial occurs, followed by a judgment and possibly, an appeal.

The bold red items in Figure 1 correspond to points in the process where the district judge can terminate a case. Roughly speaking, they represent decision-points in litigation where different kinds of information is available to the district judge when he or she makes her judgment. For example, when ruling on a motion to dismiss, the district judge evaluates the pleadings to determine whether each claim is "plausible on its face" (*Bell Atlantic Corp. v. Twombly* 2007; *Ashcroft v. Iqbal* 2009). Evaluating this requires effort on the part of the judge to determine, based on the limited information in the pleadings, whether there is a case to go forward. Moreover, after litigants engage in discovery, the judge may examine the information provided and decide whether to grant summary judgment. Finally, if the case proceeds to trial, the judge considers all the information provided pre-trial and during trial, as well as the jury's verdict, before issuing a judgment.

The key difference between these decision-points, however, is the information environment in which the decision is being made. In the early stages of litigation, much of what happens is informal and *ad hoc*. For example, the judge may leave a pre-trial conference feeling confident that one side should prevail, and this may be an important basis for her to grant summary judgment. However, the later stages of litigation are significantly more formal. Trial proceedings are conducted on the record and in public. A judgment post-trial will therefore usually occur upon consideration of information that is more structured, more transparent and easier to scrutinize from afar.

In this paper, I argue that the *kind* of information available to a district judge affects not only the way they make decisions, but also their incentives to work hard on cases as well as appellate courts' review posture. In particular, I distinguish between *soft* information and *hard* information. The key distinction between these types of information is whether it can be credibly communicated to another actor who may not share one's preferences. For example, suppose that a district judge has information that would indicate that the defendant should win in a particular case. If that information is soft (such as information that was learned during private pre-trial conferences), then that information cannot be credibly communicated to an appellate court who reviews the case upon appeal. However, if that information is hard (such as a "smoking gun" revealed during testimony by a key witness), then that information can be easily and credibly communicated to the appellate court in a written opinion, for example.

I analyze a model of adjudication under appellate review in two separate informational environments. The first is a **simple review** environment and its distinguishing feature is that the district judge can only acquire soft information. The second is an **endogenous information** environment and its distinguishing feature is that the district judge can acquire hard information if she works sufficiently hard at resolving the case. The degree to which a district judge is able to "get their way" depends the institutional context. As Stephenson (2011) points out, the structure of information available to courts has a significant implications for the design of legal institutions. For this reason, this paper's focus on costly information acquisition sets it apart from other work on judicial hierarchy where an upper court's main problem is eliciting the lower court's (freely acquired) private information (*e.g.*, Cameron, Segal, and Songer 2000; Carrubba and Clark 2012a, 2012b).

The district court's effort to adjudicate cases is also affected by her own personal motivations—that is, her levels of activism, ideological extremity and career concern—which play into her decision making in distinct ways across the various informational contexts. The model I study features a district judge who is relatively pro-defendant and an overseeing appellate court that is relatively pro-plaintiff. In other words, the district judge is *biased* relative to the appellate court. This paper is accordingly distinct from so-called "team models" of adjudication, which see judges in the hierarchy as perfectly aligned in their preferences (*e.g.*, Cameron and Kornhauser 2006). However, the district judge also has career concerns, and suffers a cost whenever she is reversed by the superior court.

The model also features judges that are more or less "activist." In popular usage, judicial activism is regularly understood to indicate that a judge makes decisions that are not consistent with law. For example, Ninth Circuit Judge Diarmuid O'Scannlain has argued that "[j]udicial activism implies that a judge has elevated personal values over text and precedent in order to reach a certain result" (O'Scannlain 2002, p. 132). As a result, many take activism to be synonymous with bias, however the two are distinct. A biased judge prefers to adjudicate a case using a different legal rule than what is prescribed. For example, if existing law says the plaintiff should prevail in a specific situation, a biased judge would prefer that the defendant should prevail. Judges can be more or less biased, and I mention above, the model I study features a biased district judge. However, activism implies more than bias; it implies that the judge be *active* in achieving their goals. Indeed, activist judges make some effort to *implement* their preferred outcomes.

This distinction is especially salient in the lower courts, where routine review by a superior court meaningfully constrains a judge's ability to actively pursue particular outcomes. In the model, a district judge's level of activism is a key determinant of the effort they exert on the case. Conditional on a specific level of bias, a judge can be more or less of an activist. In this sense, "activist" as it is conceptualized in this paper is similar to the "zealots" in Gailmard and Patty (2007). As a result, it should be be clear at the outset that, in this setting, more activist judges will *work harder to resolve cases because they get a higher payoff from resolving the case correctly—i.e., according to their preferred (and possibly biased) legal rule.* However, as I show below, activism can sometimes work against a district judge, by inducing an appellate court to believe that the district judge is concealing information. As I will show, this perversely induces the most activist judges (*i.e.*, the hardest working ones) to avoid increased scrutiny from the appellate court and the least activist (*i.e.*, the "laziest" ones) to welcome increased scrutiny.

2 Model

I study a simplified model of litigation that revolves around a district judges's effort to acquire information on a case that is subject to appellate review. The model is similar in spirit to recent models of bureaucratic policymaking (Gailmard and Patty 2013b) and electoral accountability (Ashworth and Shotts 2011).

Players and Setting

There are two players, a district court (D, which I refer to interchangeably as the "district court," the "district judge" and "she") and an appellate court (A, which I refer to as as "it"), who are responsible for adjudicating a case. This model analyzes only the strategic relationship between the two courts, and does not explicitly consider strategic behavior of litigants.[7] As in the U.S. federal court system, the district court is responsible for finding the facts of the case and applying the relevant law. The appellate court reviews the decision of the district court and decides whether or not to reverse it.

In line with recent formal work on courts, this model adopts a "case space" approach to characterize adjudication in the two courts (see Lax 2011). In particular, a legal case is assumed to be a conflict on a single dimension of the real number line, where points in the case space, $\omega \in \mathbb{R}$, represent individual cases. Cases map into a binary decision,

[7]Some recent papers that incorporate litigants into formal models of the judiciary are Cameron and Kornhauser (2006), Emons and Fluet (2007), Talley (2013) and Hübert (2015). Moreover, there is a long tradition of studying litigant behavior in law and economics, beginning with Priest and Klein (1984).

$x \in \{\text{defendant}, \text{plaintiff}\}$, according to a legal rule, $\widehat{\omega} \in \mathbb{R}$. A legal rule $\widehat{\omega}$ simply partitions the case space and serves as the decision rule for mapping cases into decisions. As an example, suppose the relevant legal rule for a particular case were $\widehat{\omega} = 0$. Then for all fact patterns less than zero ($\omega < 0$), one decision would be made—say in favor of the defendant—whereas for all fact patterns greater than zero ($\omega > 0$) the other decision would be made—say in favor of the plaintiff.

The two courts prefer to apply different legal rules to cases. I assume that the appellate court's ideal point is strictly lower than the district judge's ideal point: $\widehat{\omega}_A < \widehat{\omega}_D$. Substantively, this implies that the district court is relatively more pro-defendant than the appellate court, although the main substantive results of this paper are unaltered if this were reversed. Figure 2 illustrates the case space.

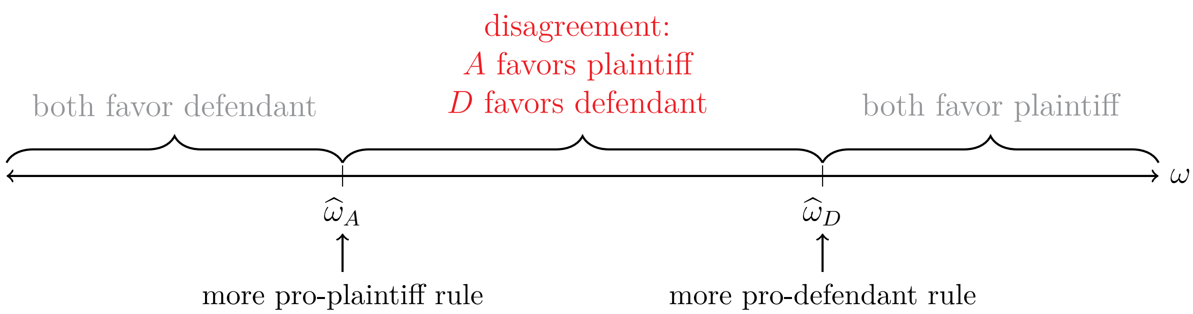


Figure 2: Example Case Space

Notice that sometimes a case is so extreme that both courts agree on the appropriate judgment. However, sometimes the case facts are between the two courts' legal rules, $\omega \in (\widehat{\omega}_A, \widehat{\omega}_D)$. Then, the courts disagree about which judgment they would make. The appellate court would find in favor of the plaintiff whereas the district court would find in favor of the defendant. Substantively, the appellate court's ideal point represents the established doctrine (or precedent) for the legal issue at hand. The district court's attempt to adjudicate according to its own ideal point therefore constitutes a deviation from precedent.

Payoffs and Information

The district judge gets a benefit $\beta > 0$ when the outcome of the adjudicative process yields a result consistent with her preferred legal rule, and pays a cost $k > 0$ when her judgment is reversed upon appeal. The former represents the strength of the district judge's activism: as β increases, she gets a bigger benefit from adjudicating according to her own legal rule. Where relevant, I refer to the post-appeal outcome simply as the "outcome," whereas I refer to the pre-appeal decision by the district judge as the "judgment." The district judge also incurs a cost for the effort she exerts in adjudicating the case. Formally, for an effort level $e \in [0, 1]$, she pays a cost $c(e)$, where $c(\cdot)$ is a strictly convex, continuously differentiable cost function satisfying the following conditions:[8]

$$c(0), c'(0), c''(0) = 0 \qquad \lim_{e \to 1} c'(e) = \infty \qquad c'(e), c''(e) \geq 0, c'''(e) > 0$$

Substantively, a cost function in this general convex form eliminates the relatively uninteresting scenario where the district court exerts maximal effort.[9]

This paper explicitly models the district courts' effort in adjudicating cases, which takes the form of costly effort to discover the case facts. Since information is endogenously acquired, the model explores a general agency problem confronting appellate courts, who rely on district court fact-finding in order to make good decisions.[10] As is the case for most litigation, neither the district court nor the appellate court initially know the specific facts at

[8] Formally, these assumptions guarantee that the marginal cost is also strictly convex, which is more restrictive that the standard assumption that the cost function is itself strictly convex.

[9] Many scholars of the U.S. courts argue that the lower courts' dockets are dominated by "easy" cases Epstein, Landes, and Posner (see, for example, 2013). Those would be the types of cases might not be consistent with the assumptions made on the cost function. In this paper, I make no claim about the existence or prevalence of those types of cases. However, it is clear that *some* cases are more difficult to resolve, and the assumptions on the cost function allow us to focus specifically on the more complex cases that are more interesting.

[10] For examples of similar models, see Stephenson (2006), Bueno de Mesquita and Stephenson (2007), and Gailmard and Patty (2007, 2013a, 2013b). For an excellent discussion of information acquisition in judicial institutions, see Stephenson (2011).

issue for the litigants. However, both courts have a sense about what kinds of fact patterns usually come before the courts. Formally, this means that ω is unknown to all players at the beginning of the game, but it is commonly known to be distributed over the real number line according to the unbounded cumulative distribution function $F(\omega)$, where $f(\omega) > 0$ for all ω.[11]

The probability that a particular case favors the defendant is represented formally by the cumulative distribution function evaluated at the legal rule used to make a judgment. Using the appellate court's ideal point, this probability is $F(\widehat{\omega}_A) \equiv F_A$, and using the district court's ideal point, this probability is $F(\widehat{\omega}_D) \equiv F_D$. The courts have predispositions toward one party or the other, which correspond to the decisions they would make without any specific information about the case under consideration. For example, in the absence of specific information about a case, a court j rules in favor of the plaintiff if $F(\widehat{\omega}_j) < \frac{1}{2}$ in favor of the defendant if $F(\widehat{\omega}_j) > \frac{1}{2}$. In order to focus on interesting situations where the two courts have differing predispositions, I assume that $F_A < \frac{1}{2} < F_D$.

> **Assumption 1** (conflicting predispositions)**.** From an *ex ante* perspective, the appellate court's preferred legal rule is pro-plaintiff, while the district court's preferred legal rule is pro-plaintiff. Formally, $F_A < \frac{1}{2} < F_D$.

Substantively, an empirical proxy for a judge's predisposition could be her political party. A Democratic judge may be inclined to be pro-plaintiff before considering the specific facts of a case, whereas a Republican judge may be inclined to be pro-defendant before considering the specific facts of a case. That said, once the judges know the *actual* facts of the case, those facts may be extreme enough that the two judges actually agree about who should prevail in the case. The judges' predispositions toward one side or the other are therefore not necessarily determinative once a judge learns the particulars of an individual case.

The quality of information available to the two courts is a function of the district judge's

[11] This assumption rules out "perfectly extremist" courts because, loosely speaking, for every possible ideal point, $\widehat{\omega} \in \mathbb{R}$, there is positive density above and below $\widehat{\omega}$.

effort in acquiring factual information relevant for the case. Formally, Nature sends a signal $s_D \in \{\omega, \phi\}$ which privately reveals the case facts to the district judge with probability equal to her effort, $e \in [0, 1]$. With probability $1 - e$, Nature sends the district judge an uninformative signal, $s_D = \phi$. "Effort" should not be taken to imply that judges who choose a low level of effort are abdicating their duty to fairly hear cases. In fact, effort can simply represent how much attention a judge devotes to this particular case. The model presented here is a model of only *one* case, not the entire docket confronting a district judge. Low effort on this one case may indicate that the judge focuses her attention on other cases where she thinks she will be more effectual. This interpretation of effort is especially relevant in light of the exploding dockets of the federal courts.

The information acquired by district court may be soft or hard information, which affects her ability to communicate it to the appellate court. Moreover, the information she acquires about the case is her private information; she can choose whether to reveal it to the appellate court. However, this choice is complicated if her information is soft since it is not verifiable by the appellate court. The fact that her information could be soft and that she could conceal it induces a signaling game between the district judge and the appellate court.

Sequence, Strategies and Equilibrium

The purpose of this paper is to characterize and explore the optimal review strategies of an appellate court overseeing a district court whose effort affects the quality of adjudication. In the next section, I begin by studying a baseline situation where a district judge makes a judgment that is not subject to appellate review. I then proceed to study two environments where the appellate court can reverse a district court's judgment. Both of these environments adopts a different informational structure. First, the district judge's information is soft and cannot be credibly communicated to the appellate court ("simple review"). And second, the district judge's effort endogenously "hardens" her otherwise soft information so that it can be communicated in a verifiable way to the appellate court ("endogenous information"). In each of these environments, I restrict attention to pure strategies.[12]

The basic sequence of the simple review and endogenous information environments are as follows. First, Nature chooses ω according to $F(\omega)$ and does not reveal this information to any player. Then, the district court can decide to invest some effort, e, into discovering the case facts. Nature sends a stochastic hard signal, s_D, to the district court: with probability related to her effort, she learns ω. In the simple review environment, if the district judge learns ω, it is soft information. In the endogenous information environment, it may be hard or soft information. The district court then decides whether to communicate this information to the appellate court (if possible), and makes its judgment, $x \in \{\text{defendant}, \text{plaintiff}\}$. Finally, the appeals court decides whether to reverse or affirm the district court's decision.

The equilibrium concept used to analyze this collection of extensive games with incomplete information is perfect Bayesian equilibrium, which requires that (1) the players' strategies be sequentially rational, and (2) the players' beliefs be consistent with those strategies and updated using Bayes' Rule on the equilibrium path. An equilibrium of each game consists of a profile of sequentially rational strategies and consistent beliefs for each player. The analysis proceeds using backward induction.

2.1 Baseline: No Appellate Review

I first analyze the baseline case where the district judge does not face review. In the federal courts, there are some restrictions on what cases can be appealed. For example, litigants may not appeal judgments that are not final (28 U.S. Code §1291) or appeal errors by the district judge that were not properly preserved by registering objection during the trial (*United States v. Gagnon* 1985). If the district judge is not reviewed, she will make a

[12]Given the sequential set up of the model, studying mixed strategies adds little of substantive interest. Indeed, in this model, mixed strategies may be optimal only in knife-edge regions of the parameter space.

judgment consistent with her belief about ω, which could be degenerate. Thus, if she is able to discover the case facts, the district judge rules in favor of the defendant if $\omega < \widehat{\omega}_D$ and in favor of the plaintiff if $\omega > \widehat{\omega}_D$. If she does not discover the case facts, she rules in favor of the defendant since her prior belief is that the case facts are more likely to support the defendant's claim (formally, $F_D > \frac{1}{2}$).

Whether she is informed or not depends on how much effort she exerts in adjudication. She optimally selects her effort to maximize her *ex ante* utility:

$$U_{NR} = e\beta + (1-e)F_D\beta - c(e)$$

Therefore, her optimal effort is implicitly defined by

$$c'(e^*_{NR}) = \beta(1 - F_D) \tag{1}$$

By the assumptions on $c(\cdot)$, her optimal effort is interior and with probability $1 - e^*_{NR}$ she does learn the case facts.

Lemma 1. Without review, the district court exerts strictly positive effort according to (1), which increases as the judge becomes more of an activist.

Lemma 1 establishes an important utility benchmark for the district court. Because this environment features no review by an appellate court, the district judge is able to decide the case according to her own legal rule. As a result, the effort she exerts in discovering ω is her utility maximizing level of effort for her unconstrained decision problem. In the following sections, I introduce appellate review and study its effect on the district judge's equilibrium effort. In such environments, her optimal choice of effort is now a *constrained* decision problem; any departure from e^*_{NR} will entail lower utility for the district judge.

3 Simple Review Regime (Soft Information)

I now introduce a simple form of appellate review into the model and explore its effect on the district court's effort and judgment. Under the simple review regime, the appellate court reverses or affirms the district judge's decision, and the district judge's information is "soft" information that cannot be conveyed to the appellate court except as a cheap talk signal. In the context of the federal courts, simple review regimes are present when the information that led to a particular decision cannot be verifiably conveyed to the appellate court, such as determinations about excluding evidence after "a balancing of relatively subjective factors" consistent with Rule 403 of the Federal Rules of Evidence (Rothstein, Raeder, and Crump 2012, p. 7). Where appropriate, I denote elements of the simple review regime using the subscript R. The analysis of this game proceeds in two steps. First, I characterize the players' optimal strategies in the subgame beginning with the district judge's judgment. Second, I characterize the district judge's optimal effort.

Before proceeding with the analysis of the simple review regime, it is possible to make a preliminary observation. Because the information available to the district judge is soft, she cannot credibly communicate it to the appellate court. The coarseness of the informational problem ensures that the two courts either completely agree about the appropriate judgment (*i.e.*, when $\omega \notin Z$) or completely disagree about the appropriate judgment (*i.e.*, when $\omega \in Z$). The district court's cheap talk signal is thus completely uninformative to the appellate court, and her judgment therefore serves as the only informative signal about ω. As a result, the district court's cheap talk message about the state of the world is completely uninformative and in the analysis below, I ignore it.

Judgment and Review. The appellate court only observes the district court's judgment, and will condition its reversal on that judgment. First, note that no equilibrium features the appellate court affirming pro-defendant decisions, but reversing pro-plaintiff decisions. Such a review strategy induces the district judge to always rule in favor of the defendant, which makes the appellate court worse off than simply affirming any judgment made by the district judge. Moreover, the only kind of equilibrium where the appellate court reverses pro-plaintiff decisions involves the appellate court reversing *any* decision. To see why, notice that it is only optimal for the appellate court to reverse a pro-plaintiff decision if it believes that $\omega < \omega_A$. Such a strategy can only be supported if pro-defendant decisions are also reversed.[13] However, equilibria of this kind—where the appellate court reverses any judgment—yield the same outcome as equilibria where the appellate court affirms any judgment, except that the district judge faces a reversal cost in the former. Such equilibria are Pareto dominated by equilibria where the appellate court always affirms, and are "fragile" in the sense that introducing an arbitrarily small cost of reversal for the appellate court eliminates them. I will accordingly rule out these inefficient equilibria.

There are therefore two candidate review strategies to consider: (1) the appellate court affirms any judgment, and (2) the appellate court affirms pro-plaintiff judgments and reverses pro-defendant judgments. Because the first strategy allows the district court to implement its preferred outcomes, I say that the appellate court is "deferential" and because the second strategy implements the appellate court's *ex ante* preferred outcome, I say that the appellate court is "non-deferential." The district judge's optimal judgment under these two reversal strategies is straight-forward:

$$x^*(s_D, r) = \begin{cases} \omega & \text{if } D \text{ learns } \omega \text{ and } A \text{ defers} \\ \text{defendant} & \text{if } D \text{ does not learn } \omega \text{ and } A \text{ defers} \\ \text{plaintiff} & \text{otherwise} \end{cases}$$

[13]To see this, note that the district judge faces a cost of reversal, and would receive a benefit β from deciding in favor of the defendant when $\omega < \omega_A$. If the appellate court affirmed pro-defendant decisions, then it would not be optimal for the district judge to rule in favor of the plaintiff when $\omega < \omega_A$.

Now, we characterize the conditions under which it is optimal for the appellate court to defer to the district judge. Suppose that the appellate court defers and observes a ruling in favor of the defendant. In equilibrium, its posterior belief that the judgment is compliant is

$$\Pr(\omega < \widehat{\omega}_A | x = \text{defendant}, \text{full deference}) = \frac{F_A}{F_A + e_R^*(F_D - F_A) + (1 - e_R^*)(1 - F_A)}$$

Therefore, deferring is optimal for the appellate court if this posterior is greater than one half (in other words, its posterior belief is that the defendant should have prevailed). This reduces to the following condition:

$$e_R^* \geq \min\left\{\frac{1 - 2F_A}{1 - F_D}, 1\right\} \equiv \tilde{e}_R \tag{2}$$

The following lemma characterizes the two types of equilibria that emerge in the judgment-review subgame.

Lemma 2. Under simple review, D's equilibrium judgment and A's equilibrium reversal strategy are given by the following:

- If condition (2) holds, then any (efficient) equilibrium involves the district court ruling in favor of the plaintiff when it knows that $\omega > \widehat{\omega}_D$ and ruling in favor of the defendant otherwise. The appellate court always affirms the district court's judgment, and has beliefs formed by Bayes' rule.
- If condition (2) fails, then any equilibrium involves the district court ruling in favor of the plaintiff, and the appellate court affirming a judgment for the plaintiff and reversing a judgment for the defendant. The appellate court's equilibrium belief is formed by Bayes' rule, and its off equilibrium path belief is $\Pr(\omega < \widehat{\omega}_A | x = \text{defendant}) < \frac{1}{2}$.

If the courts' ideal points are sufficiently far apart so that $\frac{F_A}{F_D} < \frac{1}{2}$, then condition (2) fails for all $e_R^* \in [0, 1]$.

Given that $e \in [0, 1]$, condition (2) implies that the appellate court never defers to the district court (regardless of the district court's effort) whenever the two courts are sufficiently far apart.

Corollary 1. If $\frac{F_A}{F_D} \leq \frac{1}{2}$, then any equilibrium of the simple review game involves no deference, and equilibrium judgments favoring the plaintiff.

Figure 3 depicts the appellate court's optimal deference under simple review, as a function of the district judge's bias F_D and equilibrium effort. As is apparent, the existence of

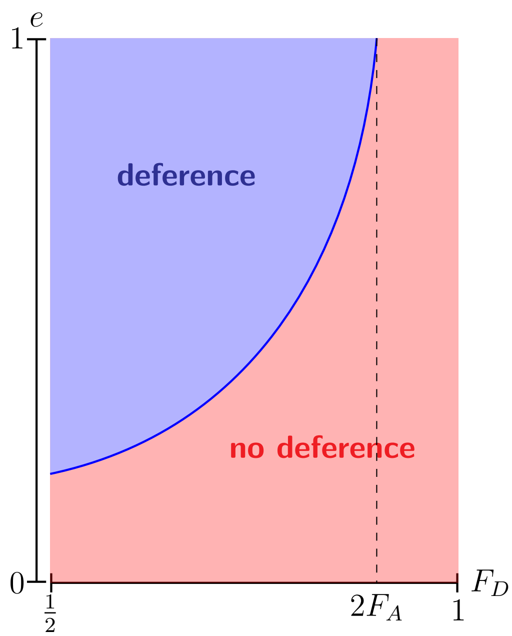


Figure 3: Deference Under Simple Review

an equilibrium with deference depends on the district judge exerting sufficiently high effort. Despite the fact that the appellate court cannot commit to deference *ex ante*, it may optimally defer to the district court *ex post*. Next, I characterize the necessary conditions for the existence of a deferential equilibrium.

District Court Effort. The district court must decide how much effort to invest in discovering the facts of the case, if any. Effort increases the probability that she will learn ω. Given the equilibrium judgment and review strategies described above, the district judge must weigh the cost of effort with the potential benefit of effort. Of course, if the two courts are sufficiently far apart—*i.e.*, $\frac{F_A}{F_D} < \frac{1}{2}$—then condition (2) will always hold and the appellate court will never defer to the district court's judgment. As a result, the district judge's best response is to set $e_R^* = 0$. On the other hand, if the courts are sufficiently close together—*i.e.*, $\frac{F_A}{F_D} > \frac{1}{2}$—then there exist equilibria with strictly positive effort whenever the district judge's optimal level of effort is high enough to cause condition (2) to fail. Define $\tilde{\beta}$ to be the value of β (conditional on F_D) that satisfies equation (1) at the level of effort that binds condition (2): $c'(\tilde{e}_R) = \tilde{\beta}(1 - F_D)$. Then:

Lemma 3. Under simple review, the district court exerts strictly positive effort according to equation (1) if and only if the courts' ideal points are sufficiently close—*i.e.*, $\frac{F_A}{F_D} > \frac{1}{2}$—and the district court is sufficiently activist—*i.e.*, $\beta > \tilde{\beta}$.

Lemma 2 and Lemma 3 together characterize the possible equilibria of a simple review model. In particular, there are two qualitatively distinct review postures. Under non-deference, the district judge exerts no effort and always rules for the plaintiff. Under deference, the district judge exerts positive effort and rules in favor of the defendant unless she learns that ω is so extreme as to warrant a pro-plaintiff decision (even by her own preferred legal rule). Figure 4 illustrates the the two possible review postures in the simple review environment. These two possible review patterns allow us to make an interesting observa-

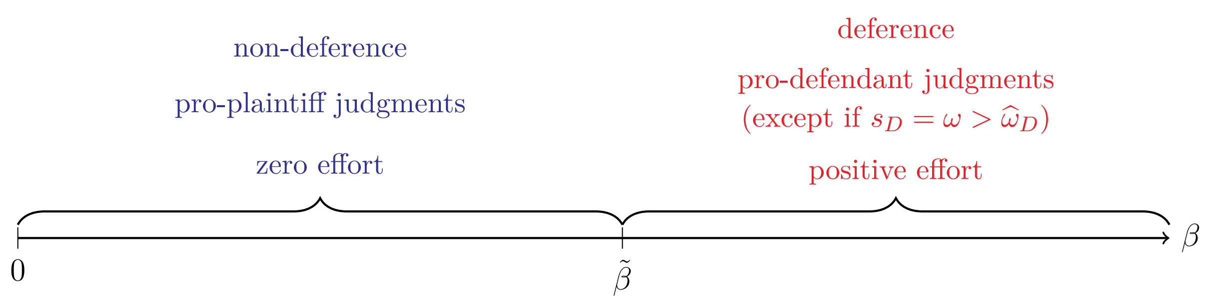


Figure 4: Simple Review When Courts Sufficiently Close Together

tion about adjudication in the real world. Notice that as the district judge becomes more of an activist, her equilibrium effort under deference increases. As long as the two courts are not too far apart, it is optimal for an appellate court to completely defer to a more activist district court even though that court will subvert the appellate court's established doctrine. Now, it is possible to characterize the first main finding.

> **Proposition 1.** If appellate review occurs in a simple review environment, then appellate courts defer to district court judgments if and only if the two courts' ideal points are sufficiently close and the district judge is sufficiently activist.

Proposition 1 characterizes sufficient conditions for purposeful *ex post* deference may arise in the federal courts. First, deference requires that the district judge and appellate panel not be too far apart in their preferred legal rules. Equivalently, the district court cannot be too biased relative to the appellate court. Second, the district judge must be enough of an activist so that the effort she exerts is sufficiently useful to the appellate court.

4 Endogenous Information Review Regime

This section explores a different informational environment where the district court's effort to resolve the case is correlated with her ability to communicate any hard information she receives to the appellate court. As a result, I refer to this setting as appellate review with "endogenous information." The key difference is that now, any signal sent to the appellate court can be verified, which allows the district judge to pass her information on to the appellate court. In U.S. federal courts, there are examples of information hardening. The extent to which district judges are able to write long, detailed orders justifying the decisions they make and spelling out the various factual and legal issues at play will depend on the information available. The basic assumption in this informational environment is that as the judge makes more effort to discover the state of the world, it becomes easier for the appellate court to learn whatever the district judge learned.

Formally, the district judge's effort now has two effects. As before, with probability e she discovers the state of the world ω. However, if she learns ω, then with probability e

she receives a "hard" signal that she can communicate to the appellate court, and with probability $1 - e$ she does not receive this hard signal. Substantively, the hard signal can be conceptualized as a "smoking gun," while the soft signal might be conceptualized as a "gut feeling." Both provide the district court with certainty about the state of the world, but one is easy to communicate to the reviewing appellate panel, while the other is not. Figure 5 depicts this stochastic process.

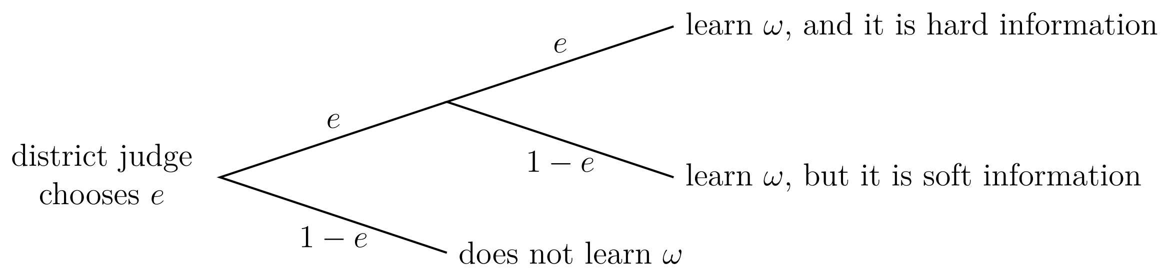


Figure 5: Endogenous Information

It is important to underscore the two main assumptions embedded in this particular stochastic process. First, effort increases both the likelihood that the district court discovers information *and* that the information she discovers is hard information. Second, the district judge does not control whether she receives hard or soft information. This particular process is chosen for convenience, but it could be replaced with a more complex procedure where the district judge first decides whether to discover ω, then whether to harden any information she receives. As long as the district judge's effort is not observed by the appellate court, the results are qualitatively similar.

In addition to making a judgment on the case, the district judge now has an additional action available to her, since she may or may not have hard information. Now, after she makes a judgment, she also sends a message m to the appellate court about the case. If the information the district judge receives is soft information, then her message can only be uninformative: $m = \phi$. If, however, her information is hard, then she chooses her message from $M = \{\omega, \phi\}$. Therefore, with hard information, the district judge can choose whether to pass information about the case to appellate court, or to conceal that information. As before, the appellate court decides whether to reverse or affirm after the district court has issued its judgment. When appropriate, I denote the review regime with endogenous information by the subscript $\bar{R}$.

Judgment and Review. It is straightforward to observe that if the appellate court receives hard information from the district judge, then it is a best response for the appellate court to affirm the district judge's judgment if and only if it is consistent with the appellate court's legal rule. Given this, the district judge is weakly better off releasing hard information to the appellate court whenever $\omega < \widehat{\omega}_A$ or $\omega > \widehat{\omega}_D$. To simplify the analysis, I assume she does so.[14]

If the appellate court does not receive hard information from the district judge—either because the district judge did not learn ω, received a soft signal, or decided to conceal a hard signal—then, it forms a posterior belief about the ω given the district judge's decision. After observing a pro-plaintiff decision, the appellate court believes it more likely to be correct than not, since the district judge is pro-defendant and would only rule in favor of the plaintiff if it found information supporting such a decision.[15] If the appellate court observes a pro-defendant decision and no hard information to support it, then its inferential problem is more complicated. Suppose that the appellate court defers to the district judge's decision in the absence of hard information. Then the appellate court's posterior belief that

[14]This assumption ensures that D doesn't play weakly dominated strategies. If this assumption were violated, then the appellate court's equilibrium belief under deference would be equivalent to its equilibrium belief under deference in the setting with simple review. This assumption could be relaxed if D instead receives some arbitrarily small payoff from revealing any hard information it acquires when $\omega \notin Z$.

[15]Formally, A's posterior belief that a pro-plaintiff decision is correct is in the interval

$$\Pr(\omega > \widehat{\omega}_A | x = \text{plaintiff}, m = \phi) \in \left[1 - F_A, \frac{(1 - F_A) - e^2(1 - F_D)}{(1 - e^2 F_A) - e^2(1 - F_D)}\right]$$

which is always greater than $\frac{1}{2}$.

a pro-defendant judgment is correct is

$$\Pr(\omega < \widehat{\omega}_A | x = \text{defendant}, m = \phi) = \frac{F_A(1-e^2)}{1 - e(1 - F_D + eF_A)}$$

For deference to be optimal, this belief has to be greater than or equal to $\frac{1}{2}$. That reduces to:

$$e(1 - F_D - eF_A) \geq 1 - 2F_A \tag{3}$$

First note that the condition does not hold for a sufficiently high level of effort e, unlike in the simple review environment, where high effort was more likely to generate deference. Here, the higher e, the more likely the appellate court is to believe that the district court is concealing hard information that demonstrates that her judgment goes against the appellate court's preferences. Moreover, as with the simple review environment, this condition does not hold when the courts are sufficiently far apart. Formally, in order for condition (3) to hold—and thus for deference to be possible—the following must hold

$$\underline{e} \leq e^*_R \leq \overline{e} \qquad\qquad \underline{F_A} \leq F_A \leq \frac{1}{2}$$

where $\underline{e}$, $\overline{e}$ and $\underline{F_A}$ are defined in Definition 4 in the appendix. The second condition ensures that deference holds for a smaller region of the parameter space, since $\underline{F_A} > \frac{1}{2}F_D$. Substantively, with endogenous information, the appellate court now needs to be more uncertain about the state of the world in order to justify deference because it can learn more from the district judge. The first condition reveals that there is now a non-monotonic relationship between the district court's effort on the case and the appellate court's deference. Specifically, deferring to the district judge is only optimal for the appellate court if the district court's equilibrium effort is not too low or not too high. In the case of the former, the appellate court infers the district judge is most likely just "guessing" and in the case of the latter, the appellate court infers that the district judge is concealing information. Figure 6 depicts the appellate court's optimal deference under endogenous information, as a function of the district judge's bias and effort.

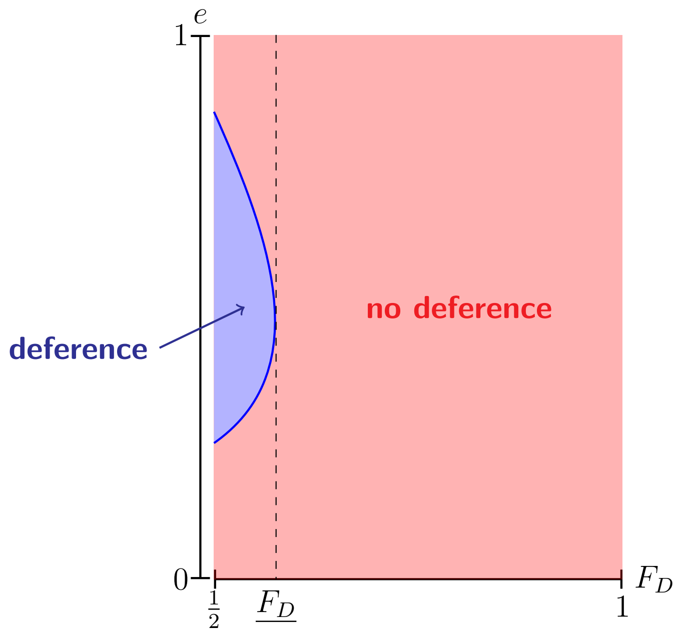


Figure 6: Deference Under Endogenous Information

Given the review strategies above, I now outline the district judge's optimal judgment in the environment with endogenous information. First, suppose that we are in the region of the parameter space where deference in the absence of information is a best response. Then, the district judge's optimal judgment is as follows.

$$x^*(s_D, \text{deference}) = \begin{cases} \text{plaintiff} & \text{if } D \text{ learns that } \omega > \omega_D \\ \text{defendant} & \text{otherwise} \end{cases}$$

Next, consider the region of the parameter space where deference in the absence of information is not optimal for the appellate court. Then, conditional on the district court ruling for the defendant and providing no hard information, the appellate court believes the decision to be incorrect and reverses. As a result, an uninformed district court would be better off ruling in favor of the plaintiff when uninformed.

$$x^*(s_D, \text{no deference}) = \begin{cases} \text{defendant} & \text{if } D \text{ receives hard info that } \omega < \omega_A \\ \text{plaintiff} & \text{otherwise} \end{cases}$$

In contrast to the simple review environment, the appellate court is now able to fine-tune its review posture by being able to occasionally affirm pro-defendant decisions even under a non-deferential review posture. This is only occurs when the district judge provides it with hard information that the pro-defendant decision is justified under the appellate court's established doctrine (*i.e.*, when $\omega < \widehat{\omega}_A$). This observation is significant because non-deference in the endogenous information environment now gives the district court an incentive to exert strictly positive effort adjudicating the case, unlike non-deference in the simple review environment.

The analysis of the equilibrium judgment and review strategies raises two important questions about the effort that the district judge exerts in the endogenous information environment. First, given that deference depends on the equilibrium effort of the district judge (see Figure 6), when does equilibrium effort satisfy the conditions guaranteeing deference? Second, given that the district judge now has an incentive to exert effort even under non-deference, how does effort under non-deference compare to effort under deference?

District Court Effort. First, consider the district judge's optimal effort under a deferential review posture. Because the appellate court never reverses the district judge, her optimal effort does not depend on whether she receives hard or soft information. Her optimization problem looks identical to the optimization problem under simple review. That is, she exerts effort according to equation (1).

Next, suppose that the appellate court does not defer in the absence of information.

Then, her *ex ante* expected utility is

$$U_D(e, \text{no deference}) = (1 - F_D)\beta + e^2 F_A \beta - c(e)$$

Given the judgment and review strategies discussed above, the district judge's *ex ante* utility function represents the following. First, she always gets benefit β when she learns that $\omega > \widehat{\omega}_D$ because this corresponds to a situation where both courts are in agreement that the plaintiff should prevail. Moreover, the appellate court affirms judgments favoring the plaintiff since the district court is more pro-defendant than the appellate court is. Second, the only other circumstance where the district judge receives benefit β, given the judgment and review strategies above, is when she is able to provide hard information to the appellate court that confirms that $\omega < \widehat{\omega}_A$. In this case, both courts agree that the defendant should prevail. This occurs with probability $e^2 F_A$. Finally, the district judge pays a cost for her effort.

Conditional on this review posture, her optimal effort $e^*_{\bar{R}} \in [0, 1]$ is given by the solution to the following first order condition, if such a solution exists:

$$2e^*_{\bar{R}} F_A \beta = c'(e^*_{\bar{R}})$$

Invoking the assumptions on $c(e)$, there are two solutions to this condition, one at zero and the other strictly between zero and one. The second order condition $2F_A\beta - c''(e)$ is positive at zero and negative at the interior solution. Therefore, the utility maximizing level of effort is the interior solution.

The fact that strictly positive effort is optimal for the district judge under *non-deferential* review offers a contrast with the simple review environment studied above. Now, the district judge has an incentive to exert *some* effort in equilibrium just in case she is able to discover hard information that $\omega < \widehat{\omega}_A$. In other words, the ability to discover hard information incentivizes her to exert some effort *even when the appellate court is not deferential.*

How does the now strictly positive effort under non-deference compare to the district judge's effort under deference? Recall that a district judge makes effort according to equation (1) whenever the appellate court defers to her. Given the assumptions on the cost function $c(e)$,[16] Then the district judge exerts (weakly) more effort under non-deference than under deference if:

$$\beta(1 - F_D) \geq c'\left(\frac{1 - F_D}{2F_A}\right) \tag{4}$$

Holding constant F_A and F_D, as β increases, the condition is more likely to hold, indicating that a more activist judge exerts less effort under deference than under non-deference. Equation (4) therefore suggests that there is an threshold level of β that partitions the set of possible district judges into two groups: strong activists and weak activists. Let $\hat{\beta}$ be the level of β such that equation (4) holds with equality if such a level exists, and ∞ otherwise. Then I adopt the following definition.

Definition 1. If $\beta \geq \hat{\beta}$, we say that the district judge is a **strong activist**. If $\beta < \hat{\beta}$, we say that the district judge is an **weak activist**.

Since the level of effort they would make under deference reflects their optimal level of effort in their simple decision problem, then non-deference in an setting with endogenous information induces strong activists to exert *more* effort than they would otherwise exert. Figure 7 depicts the equilibrium effort under deference and non-deference when this condition holds (a strong activist; left panel) and when this condition does not hold (a weak activist; right panel). Using this definition, we can make the following observation.

Lemma 4. In the environment with endogenous information, strong activists exert more effort under non-deference than under deference. Weak activists

[16] And specifically, that $c'''(e) > 0$ and $c''(0) = 0$, which ensures that the marginal cost of effort is strictly convex.

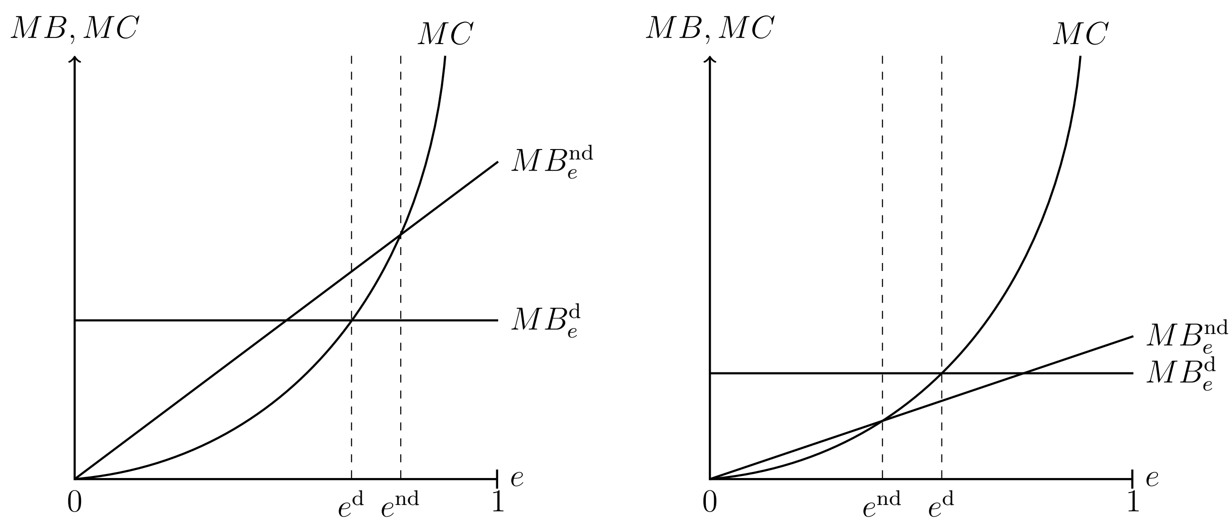


Figure 7: Optimal Effort with Endogenous Information, Strong v. Weak Activists

exert less effort under non-deference than under deference.

Lemma 4 is simple, but substantively important. Because a judge's level of effort under deference is its utility maximizing level of effort in its unconstrained decision problem, appellate review in an environment of endogenous information induces strong activists to exert *more* effort than they would otherwise exert. Moreover, now there is a non-monotonic relationship between the judge's level of activism, β, and deference, whenever deference occurs in equilibrium. Let $\underline{\beta}$ and $\overline{\beta}$ be defined as in Definition 4 in the appendix. Figure 8 depicts the appellate court's optimal review posture as a function of the district judge's level of activism.

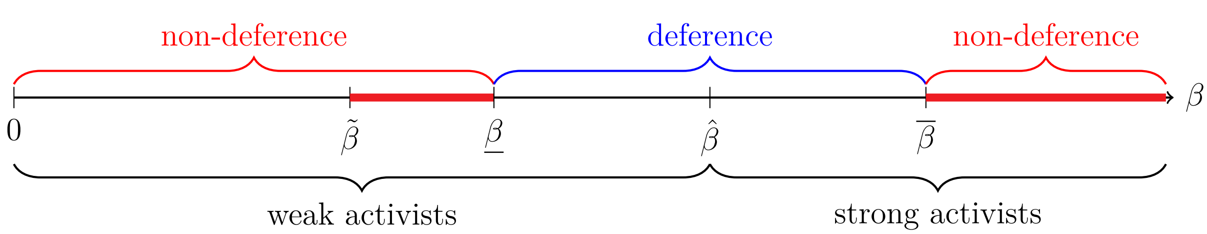


Figure 8: Endogenous Information Review When Courts Sufficiently Close Together

The thick red intervals in Figure 8 represent the levels of activism where the appellate court would defer to the district judge under simple review, but does not defer to the district judge under endogenous information. As a visual comparison of Figure 3 and Figure 6 suggests, deference occurs less often under endogenous information than simple review.

> **Proposition 2.** Deference is optimal in a smaller range of the parameter space in an environment of endogenous information than in an environment of simple review.

Proposition 2 helps elucidate the source of *ex post* deference in this model. Somewhat counterintuitively, the appellate court defers more often when it is less able to scrutinize the district judge's decision, as in the simple review environment. But, when the informational environment makes it possible for the district court to communicate her private information, the appellate court becomes more skeptical about any *lack* of communication. The appellate panel then must weigh the district judge's incentive to acquire information with her incentive to conceal unfavorable information; this reduces the situations in which the appellate panel is willing to defer.

As is the case in most models of asymmetric information, the appellate court is made better off in situations when more information is available, and the endogenous information environment provides this opportunity. Indeed, the appellate court is better off in the endogenous information environment than with the simple review whenever equilibrium effort is higher relative to equilibrium effort in the simple review environment. Lemma 19 (in the appendix) shows that a district judge facing non-deferential review under the simple review environment exerts *more* effort in the endogenous information environment than in the simple review environment, *even though she still faces non-deferential review.* However, it is not generally true that effort by the district judge will be higher under endogenous information than under simple review. For example, consider a judge with a level of activism β such that $\tilde{\beta} < \beta < \underline{\beta}$. Under simple review, these judges face deferential review, whereas under endogenous information, they face non-deferential review. According to Definition 1, these judges are weak activists, and by Lemma 4, they exert strictly higher effort under deference than under non-deference. Since equilibrium effort under deference is the same in either regime, it follows that they exert strictly lower effort under endogenous information than under simple review. This may make the appellate court worse off.

Proposition 3. The appellate court is worse off under endogenous information than simple review when $\tilde{\beta} < \beta < \min\{\underline{\beta}, \hat{\beta}\}$ and $e^*_R - e^*_{\bar{R}} > \frac{F_D - F_A}{1 - F_D}$.

Given the relatively general assumptions on $c(e)$, we cannot pin down whether latter condition in Proposition 3 holds. However, the normative implication of Proposition 3 is interesting in light of the conventional finding that more information makes players better off. Because the appellate court cannot prevent itself from believing that the district court may be concealing information, an inefficiency emerges where some kinds of judges actually decrease their effort in the endogenous information environment. Thus, despite the fact that endogenous information generally makes it easier for the appellate court to acquire information and more easily scrutinize the district judge's decision, it also sometimes leads to the destruction of information by way of lower effort.

5 Optimal Timing of Final Judgments

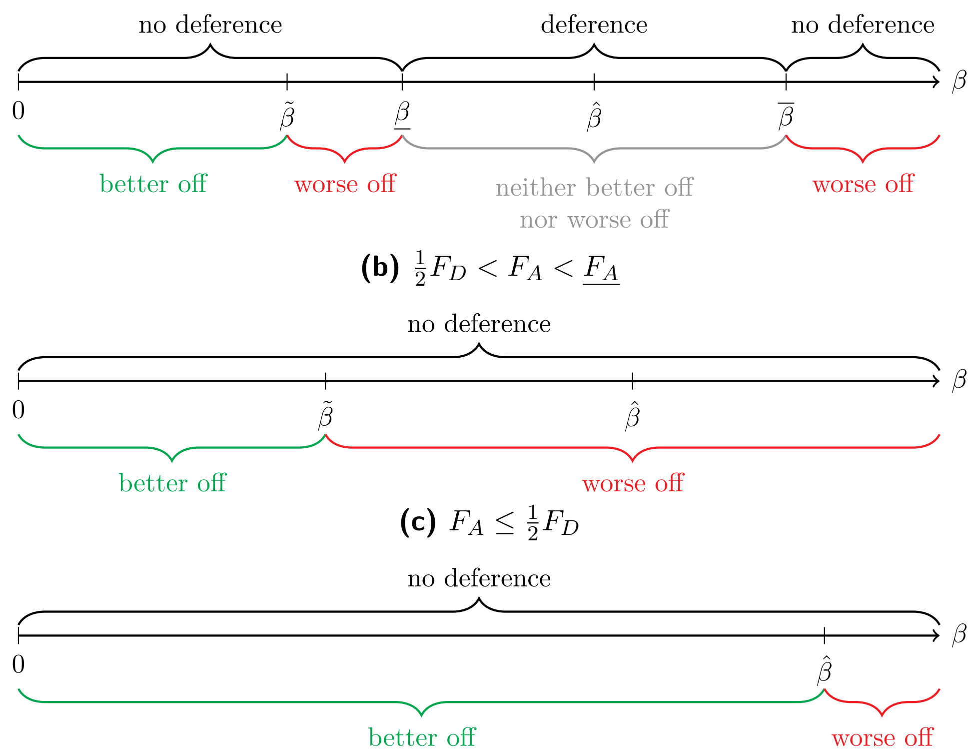
Given the results from the simple review and endogenous information environments, I now explore how these environments can affect a district judge's decision about when to issue her final judgment. To motivate this "meta-game" in substantive terms, assume that a judge can decide whether to end a case by ruling on a motion for summary judgment, or end a case by conducting a trial (thus first denying the summary judgment motion). This description of procedure is purposefully simplified to capture the main choice confronting a district judge.

If she ends the case "early," she makes her decision in a simple review environment, but if she ends her case "late," she makes her decision in an environment with endogenous information. Formally, I describe the district court's decision as t, where $t \in \{\text{pre-trial}, \text{trial}\}$. She makes this decision first, and then the game proceeds as in Section 3 if she chooses $y = \text{pre-trial}$ and as in Section 4 if she chooses $y = \text{trial}$.[17]

With Lemma 4 and Proposition 2 in hand, I now explore whether the district judge is better or worse off under simple review or endogenous information. Figure 9 depicts how the district judge's utility changes in the three regions of the parameter space, whenever moving from simple review to endogenous information. First, note that any district judge

Figure 9: District Court Utility

(a) $F_A > \underline{F_A}$



[17] A more fully specified sequence would involve the district judge playing a two stage game after selecting $y = \text{trial}$: play the simple review game, then the endogenous information game. However, the qualitative results here do not depend on the simplification I make in text. In particular, I do not focus on the region of the parameter space where deference occurs under both simple review and endogenous information. In that region, the decision over y makes less sense as an "either-or" choice.

who faces deference both under both simple review and endogenous information is neither better nor worse off in either environment. This is consistent with the idea that deference—in either environment—allows district judges to achieve their preferred outcomes. Next, note that each of the three panels demonstrates that for sufficiently high level of activism, the district court is (weakly) worse off under endogenous information than under simple review. If $F_A < \underline{F_A}$ (or equivalently, if $F_D > \underline{F_D}$), then more activist district judges are strictly worse off in the endogenous information environment than in simple review. This follows from the fact that she gets her preferred outcome less often, and is accordingly induced to alter her level of effort to convince the appellate court.

It is intuitive that the district court is sometimes worse off under endogenous information than simple review, since it allows the appellate court to exert more control over the district court's decision making. It is less intuitive that it would make the *less activist* judges better off while making the *more activist* judges worse off. Recall that activism is not equivalent to bias, but instead reflects the degree of salience that the case has for the judge. Judges with a relatively low level of activism find it to be an improvement for the appellate court to scrutinize their decisions more closely, despite working less hard than higher activism judges. In contrast, for some higher activism judges, increased scrutiny in the endogenous information environment is welfare reducing because they are no longer freely able to use their effort toward their preferred outcome.

When do district judges opt to end the case during the pre-trial, instead of taking the case through the trial process? The next result characterizes the conditions under which ending early is optimal for the district judge.

Proposition 4. First, suppose that $F_A \leq \underline{F_A}$. If $\beta > \min\{\tilde{\beta}, \hat{\beta}\}$, then the district judge ends the case during pre-trial. Next suppose that $F_A > \underline{F_A}$. If $\tilde{\beta} < \beta < \underline{\beta}$ or if $\beta > \overline{\beta}$, then the district judge ends the case during pre-trial.

Proposition 4 provides a key prediction that is useful for understanding the relationship between appellate review, district court decision making and the characteristics of district judges. To the extent that different steps in the litigation process involve different informational environments, then this result raises important questions about existing empirical findings. Indeed, if earlier stages of litigation are soft information environments, while later stages are endogenous information environments, then the results predict that more activist judges will opt to deciding cases at the earlier stages of litigation than less activist judges do.

6 Empirical Implications

What are the empirical implications of these results? Many legal scholars and practitioners think that district judges are often able to get their preferred dispositions. Famously, federal district judge Henry N. Graven proclaimed that "The people of this district either get justice here [in the district court] or they don't get it at all" (as quoted in Carp and Rowland 1996, p. 1). As evidence of this claim, Judge Graven goes on to explain that he has never been overruled. The low number of reversals of district court decisions potentially undermines the ability of appellate courts to reign in lower courts. Proposition 1 indicates that *ex post* discretion is optimal for appellate courts who care about the district court's effort, but have limited tools to scrutinize district court decisions. As a result, a first order empirical implication of the model studied in this paper is that a low reversal rate of district court decisions will not necessarily indicate compliance with appellate doctrine, since appellate courts sometimes defer. Moreover, it is not obvious when such deference occurs. In this model, deference is correlated with a district judge's bias and her level of activism. While political scientists have many proxies for the former, it is less clear how to measure the latter. Moreover, conditional on a level of bias, there will be heterogeneity among judges of different levels of activism.

Yet, empirical studies have repeatedly claimed to demonstrate the constraining power of appellate review. Most of these recent studies do not focus on reversals, but rather use the ideological direction of district judge decisions as a dependent variable in their analyses. The basic idea is that as the district judge becomes more biased, she may be more likely to depart from the appellate court's preferences. Indeed, such studies generally show this *not* to be the case. Schanzenbach and Tiller (2006) and Fischman and Schanzenbach (2011) find that district judges issue criminal sentences in a manner consistent with their ideologies, but that they are meaningfully constrained by the prospect of appellate review. Randazzo (2008) constructs a strategic choice probit model to explicitly incorporate strategic considerations in the estimation of the effect of reversals on district court behavior. He finds that district court judges are more likely to issue ideological decisions when they anticipate that they will not be reversed. Epstein, Landes, and Posner (2013) consider the effect of discretion on district court decisions and find that "[c]ourts of appeals apparently exert sufficient control over district judges" (p. 241).

One possible explanation for why judges seem compliant in these studies yet are widely seen as being able to "get their way" may come from an examination of their data sources. Many empirical analyses of U.S. District Court decision making draw on a few key data sets. Most notably are the Songer and Sunstein databases, one or both of which are used in Randazzo (2008) and Epstein, Landes, and Posner (2013). These databases contain a sample of published U.S. Courts of Appeals decisions, which the authors link to the preceding district court decisions. Are the district court decisions that resulted in appeals with published opinions representative of cases more broadly?

Ideally, we would be able to determine the proportion of cases decided at pre-trial that yield published decisions and the proportion of cases decided at trial that yield published decisions. However, in the absence of detailed data on case outcomes in the district courts, we cannot know this for sure. That said, cases with published opinions are are widely recognized to be ones that have particularly interesting or novel legal issues in them. Moreover, such cases will usually not feature ambiguity over important pieces of factual information, so that the appellate court may address a "clean" legal issue. This observation would imply that district court judgments that end up generating published appeals will be more likely to have been ended at later stages. And, as discussed above, cases decided at later stages will be more likely to resemble the endogenous information environment studied in this section.

Proposition 4 states that the kinds of judges who end cases early—for example, in pre-trial—will be those who face deference in the simple review environment but non-deference in the endogenous information environment. What kinds of judges are those? The endogenous information environment is one where there is increased scrutiny of the district court from the appellate court. Counterintuitively, the kinds of judges who will subject themselves to increased scrutiny will be the very biased ones, and the least activist ones. Moreover, conditional on a level of bias (that is not too high), it will be the activists who will be more likely to avoid scrutiny because, ironically, they work too hard for the appellate court to defer to them under this increased scrutiny. In other words, the ones who do not work as hard welcome the added scrutiny. Crucially, those judges will also be the ones who are always subjected to a non-deferential standard of review and will appear the most compliant. Thus, observing a high degree of compliance in these datasets will not be surprising.

Finally, it is not a trivial issue that district courts may choose to end cases early. While there may be other reasons for ending a case in pre-trial other than the once advanced in this paper, it is an empirically common phenomenon. In the year ending March 31, 2015, there were 30,102 civil cases terminated with court action during or after pre-trial in the U.S. District Courts. Of these cases, 27,169 (90%) were terminated before trial and 2,933 (10%) were terminated during or after trial.[18] Some number of the cases terminated during

[18]During this year, there were 263,874 cases officially "terminated" in the U.S. District Courts. Of these, 52,561 were terminated without court action and 181,211 were terminated before pre-trial, indicating that the parties likely settled. See `http://www.uscourts.gov/statistics/table/c-4/`

pre-trial were presumably terminated with little involvement of the judge, but many involve rulings on dismissal and/or summary judgment motions. To fully understand the degree to which district courts are compliant with appellate preferences, we would need to study this subset of cases more deeply. The results in this paper suggest that deference (and thus *ex post* appearance of non-compliance) will be more common in the cases that may be more likely to be excluded from traditional datasets.

7 Conclusion

Despite the limitations on district courts' explicit rule-making, their impact is widely felt within the judiciary, as well as in society more broadly. Americans are increasingly coming into close contact with the district courts, as case loads steadily increase. The "rights revolution" coupled with increasing political conflict between Congress and the President, have brought more activists and claimants into the courts to seek relief (see Kagan 1991; Epp 1998; Farhang 2010). As the courts find themselves involved in a broader range of issue areas, the legal claims being made and the factual circumstances being examined are increasingly complex. Appellate courts (and Congress) *rely* on district courts to adequately manage and resolve the collection of factual and legal issues that arise during the ordinary course of litigation. Otherwise, adjudication can lead to too many incorrect decisions.

This paper demonstrates how deferential standards of review can emerge *ex post* when appellate courts review decisions by (potentially biased) district judges whose effort in adjudicating a case is consequential. In particular, deference depends on the informational context in which district courts make their judgments. When it is harder for district judges to credibly communicate case-relevant information to the appellate court as in the simple review environment, deference is more wide-spread. The review posture adopted by appellate

`federal-judicial-caseload-statistics/2015/03/31` for more information.

courts can also have an effect on district court's incentives to engage in high quality fact-finding and case management. The models explored here also show how these dynamics are complicated by the informational environment. Specifically because there is less deference in the endogenous information environment than in the simple review environment, some judges alter their equilibrium effort. This has ambiguous effects on the appellate court. In fact, the appellate court is sometimes *worse off* in the endogenous information environment, despite being able to scrutinize the district judge's effort more closely.

Finally, whenever district judges are able to select the information environment in which they make their judgment by ending cases early (simple review) or late (endogenous information), more activist judges who work the hardest are more likely to end cases early to avoid scrutiny. In contrast, the less activist judges who work the least, as well as the most biased judges, will be more likely to end cases later to increase scrutiny of their decisions. This provides a possible explanation for why quantitative studies routinely find that district judges are compliant with appellate doctrine despite a widespread impression that they are able to "get their way." To the extent that commonly used datasets include samples of cases ended later in the adjudicative process, those samples may disproportionately feature kinds of judges who are more biased and less activist, and thus get less deference. Those judges will appear more compliant than judges in a random sample of cases. This paper suggests that future empirical research on district courts should more fully consider the possible selection issues introduced by the fact that the Federal Rules of Civil Procedure give district judges discretion to end cases at various points in the course of litigation.

Paper II

The Downsides of Dispassionate Judges: A Theory of Costly Law Enforcement in District Courts

In 2009, U.S. District Judge Vaughn Walker, who was presiding over a constitutional challenge to California's ban on same-sex marriage, surprised the parties to the case by deciding to hold a full trial (Dolan 2013). Ordinarily, litigants in constitutional cases usually stipulate to facts, obviating the need for lengthy and costly trials. Walker used his bench trial to establish a detailed and authoritative case record and justified his decision by pointing out that the "parties disputed the factual premises underlying plaintiffs' claims" (704 F. Supp. 2d 921, at 928). His deep involvement in—and high degree of effort on—the case was widely noted in the press, and even received some praise (Lithwick 2010). In the end, Walker acknowledged that the process enabled him to make a better decision since it allowed him to learn a lot about the issues before the court (Totenberg 2013).

However, not too long after Walker's ruling overturning California's ban, prominent supporters of the ban called on his decision to be thrown out. Among those calling for his disqualification were Tony Perkins, the president of the Family Research Council, and the American Family Association, which issued a press release claiming that Judge Walker's decision striking down the ban was "compromised by his own sexual proclivity" since Walker is gay (Stein 2010). The defendants in the case moved to vacate the judgment, arguing that he was unable to be impartial given that he belongs to a group whose legal rights were affected by the result. The motion was eventually denied, but as Ifill (2010) points out, concerns about Walker's impartiality echoed similar concerns that have been raised periodically about judges overseeing civil rights cases.

Federal judges are required to disqualify themselves from overseeing cases for which their "impartiality might reasonably be questioned" (28 U.S.C. 455). As McKoski (2014) explains, this standard has historically been used by litigants to "attack a judge's partiality by claiming that the judge's race, sex, ethnicity, religion, or sexual orientation creates an 'appearance of partiality"' (p. 416). Indeed, empirical studies have found that having minorities and women overseeing cases leads to more favorable outcomes for those groups in discrimination and harassment cases (*e.g.*, Boyd, Epstein, and Martin 2010; Farhang and Wawro 2004; Kastellec 2013; Boyd 2015). In order to alleviate concerns about their impartiality, most judges and nominees insist that they are *dispassionate*.[19] That is, in addition to having no direct interest (*e.g.*, financial or familial) in the cases they oversee, judges also claim to be personally and emotionally detached from them. This is such an entrenched norm in the legal system that a set of expectations and institutions has evolved to support it. For example, Canon 3 of the Code of Conduct for United States Judges prohibits judges from making "public comment on the merits of a matter pending or impending in *any* court" (emphasis added) presumably to preserve the appearance of detachment. Also, in most federal courts, cases are randomly assigned at least in part to reduce judges' ability to take cases they feel strongly about (Macfarlane 2014).

Despite these concerns, however, judges who are *not* detached from their cases can also provide important advantages relative to dispassionate judges. If a district judge's level of effort in resolving a case is costly, then a dispassionate judge who gets little benefit from presiding over a case will exert less effort in doing so unless their superiors provide other incentives. Moreover, if incentivized by concerns that are distinct from the task itself, such as personal career concerns, then that judge may face perverse incentives that undermine both the impartiality and the accuracy of their decisions. Judge Walker's role in the California

[19]Examples of such statements are plentiful. During the 2009 confirmation hearing for district judge nominee Louis B. Butler, Jr., Senator Russ Feingold lauded Butler's view that judges should "apply the law without bias in a neutral, detached, impartial, and independent manner." During his 2012 confirmation hearing to be appointed as a district judge, Jesus G. Bernal assured the committee that he was "ready to transition from being an advocate to being a more objective, dispassionate decisionmaker, which [he] believe[s] is the role of the judge."

marriage case therefore underscores the dilemma that motivates the paper: on the one hand, most observers noticed that Walker worked unusually hard while presiding over the case, but on the other hand, the personal benefit that opponents presumed he would get from the outcome led them to raise concerns about his impartiality. This paper suggests that these two things—high effort and a personal benefit from resolving a case—may actually be closely related to each other.

The paper analyzes a formal model that shows how this principal-agent problem affects adjudication of employment discrimination claims in federal district courts. I focus on this area of law because employment discrimination suits make up a significant proportion of the district courts' dockets, and more importantly, private suits are the major way that Americans seek relief from illegal job discrimination (Rutherglen 2007; Farhang 2010). I model the district judge as the agent and the appellate court as the principal, and accordingly treat the appellate court's preference to match outcomes with its own doctrine as the normative benchmark. A district judge's effort on a given case improves the outcome by improving the information available when making a judgment. In employment discrimination cases (and many other areas of complex civil litigation), judges are often deeply involved in determining how the facts of the case should be interpreted in light of existing law and their effort is thus consequential (Chin 2012). However, a judge's effort has significant opportunity costs: heavy caseloads and the complexity of cases ensure that more effort on one case reduces the judge's capacity to exert effort on other cases (or do other things). Moreover, since federal judges have lifetime tenure and pay protections, the threat of appellate reversal is one of the only tools available for an appellate court to incentivize district judges. In the model, district judges' concerns over their professional reputations are captured by a cost that is faced when a decision is reversed.

The model shows that a dispassionate judge exerts less effort and thus produces a worse outcome on a case—both in terms of impartiality and accuracy of the decision—than a judge who gets an indirect personal benefit, such as an intrinsic or emotional benefit, from working on the case (but who is not necessarily biased). The result provides a new, and somewhat surprising, explanation for outcomes in employment discrimination cases. Appellate review combined with judges' career concerns, induce an "institutional bias" against less powerful litigants, *even in the absence of any explicit bias against those litigants.* By dint of their increased personal experience with discrimination, if certain kinds of judges (such as minorities, women, former civil rights attorneys, political liberals, *etc.*) get larger personal benefits from working on employment discrimination cases, this institutional bias will be partially mitigated. This offers an alternative way to understand recent empirical findings, such as those in Boyd (2015), that women and minority judges rule disproportionately in favor of woman and minority litigants on certain cases. Thus, the formal results in this paper suggest that the differences in average outcomes between different kinds of judges could reflect a *correction* of the underlying institutional bias against less powerful litigants, instead of in-group bias. Finally, to address potential concerns that judges who get indirect personal benefits from resolving these types of cases are simply biased, I extend the model to show that a dispassionate judge is still a worse adjudicator across a wide range of substantively important contexts than a judge biased in favor of plaintiffs.

1 Judicial Effort and Law Enforcement in District Courts

Recent formal theory of American federal courts has focused disproportionately on the appellate courts, and specifically on the evolution of policy and doctrine in the appellate courts (for a recent survey see, Lax 2011).[20] However, the bulk of the time, energy and effort put

[20] But see Spitzer and Talley (2013) and Clark and Kastellec (2013) for models of costly information acquisition in appellate courts, and Lax (2007) and Carrubba and Clark (2012a) for models of costly opinion writing.

into cases happens at the first stages of adjudication—in the district courts.[21] District courts are less involved in the creation of doctrine, and are instead tasked with developing a deep understanding of individual cases, as well as figuring out what law is most relevant for each of those cases. Congress often explicitly delegates statutory enforcement authority to the federal courts, and this mode of enforcement consumes a large share of the federal docket (Farhang 2010). Bureaucracy scholars have long recognized the importance of bureaucratic effort and expertise in the enforcement of laws (*e.g.*, Gailmard and Patty 2007, 2013a; Prendergast 2007). Moreover, recent models of bureaucratic incentives, such as Turner (2015) and Dragu and Polborn (2013), 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. Much like rank-and-file bureaucrats, a district judge's influence comes from the court's original jurisdiction. Each case involves a potentially protracted and costly process to establish both the relevant facts as well as the correct legal interpretation of those facts (Hornby 2009; Kim et al. 2009). Formal models of district courts have focused heavily on the informational role that *litigants* play (*e.g.*, Daughety and Reinganum 2000; Cameron and Kornhauser 2005, 2006; Talley 2013). These models are often presented as signaling games, where judges make judgments after observing costly signals from the litigants (but see Kornhauser 1995, for a judge-centered model).

In some respects, these models reflect features of the American legal system, such as the importance of adversarial litigants (Kagan 1991; Epp 1998). However, in some areas of law, such as patent law or employment discrimination law, complexity reigns and judges play a more active role resolving cases (Posner 2013). Employment discrimination cases are particularly costly to adjudicate for several reasons. It is often difficult to produce definitive evidence of discrimination, as employers have taken steps to eliminate the worst

[21]The U.S. federal legal system has three tiers, and the district courts comprise the lowest tier. They are the federal trial courts.

forms of blatant discrimination and are sometimes able to claim exemptions for "bona fide occupational qualifications" (Rutherglen 2007). The courts have given employers further leeway to articulate nondiscriminatory reasons for their practices—as in *McDonnell Douglas Corp. v. Green* (1973)—a fact that forces courts to grapple with complicated evidentiary issues. And, in cases that make a class claim of employment discrimination, statistical methodologies are necessarily required in order for the plaintiff to satisfy its burden of proof, which the Supreme Court has established in a series of cases.[22] The different kind of work done to enforce statutes, such as those dealing with discrimination, requires a theoretical framework that emphasizes the district judge's role in day-to-day adjudication.

The model in this paper incorporates a district judge's costly effort in acquiring enough information to correctly resolve a case. The more effort a judge exerts, the more likely they will become informed enough about the case to make a good decision. In reality, this effort is important because judges often make determinations about whether the plaintiff has established its burdens and what evidence is admissible.[23] In many cases, litigants only have limited knowledge about whether there was a violation of the law, and the judge's role is crucial in framing the legal and factual issues at hand. In the model, after the district judge's effort and judgment, the losing party can appeal and make its own effort to correct an error made by the district judge. The appellant's effort is only useful when the district court has not already crafted the appropriate interpretation of the case and made the corresponding decision. In those circumstances, the appellant's argument constitutes a statement to the appellate court of the variety: "had the district court not erred in X, then it would be apparent that I should prevail."

Effort is costly and resources available to judges and litigants are limited. Society may

[22]See *Hazelwood School District v. United States* (1977), *Wards Cove Packing Co. v. Atonio* (1989) and *Bazemore v. Friday* (1986).

[23]This even applies to jury trials. The judge determines whether a plaintiff has satisfied its burden of production and thus whether it has established its *prima facie* case.

wish for a judge to "give it their all" in every case, but as a practical matter, this is unfeasible. This is not because judges are lazy or unserious about their work, but rather because they face a host of opportunity costs associated with increasing their effort on any given case. Due to the importance of judicial effort in this model, it is important to ask: what motivates judges to exert a high level of effort? While their motivations surely vary, the model in this paper assumes that federal judges (like most people) care about their professional lives, and generally want to advance in their careers. In the model, such motivations are captured by the fact that the district judge incurs a cost when he or she is reversed by an appellate panel. Reversal costs are an important source of motivation for judges, and they capture a wide array of substantive considerations, from a judge's professional reputation to the direct costs associated with reconsidering a case (Choi, Gulati, and Posner 2011; Sen 2015).

Judges may also be motivated by other factors related to the substance of a case. Some judges may get a personal benefit from resolving certain kinds of cases, such as employment discrimination cases. The theory presented in this paper is agnostic about the source of this benefit, but it could come from personal experience with the issues before the court, professional background, or even political ideology. Professional backgrounds (for example, former civil rights or corporate defense litigators) may lead a judge to get a larger benefit from resolving particular kinds of cases, as has been suggested by some in arguing for increased professional diversity on the bench (Alliance for Justice 2014). The weight of recent research also suggests that personal characteristics of judges matter on these kinds of cases, even after controlling for ideological differences between judges (see, for example, Boyd 2015, for evidence from district courts). It is possible that for some women and minorities, personal experience with discrimination or other issues facing their group contributes to a deeper interest in working on cases involving those groups. Judge Harry Edwards, a black judge sitting on the D.C. Circuit has pointed out that "[b]ecause of the long history of racial discrimination and segregation in American society, it is safe to assume that a disproportionate number of blacks grow up with a heightened awareness of the problems that pertain to these areas of the law" (quoted in Kastellec 2013, p. 169).

The model below studies two kinds of judges: one that gets a personal benefit from resolving an employment discrimination case, and one that does not. This personal benefit is *indirect* in the sense that the judge has no material stake in the outcome case, such as a financial or familial interest. Judges who get this benefit are thus "issue motivated," and I refer to them as such throughout the paper. In contrast, some judges do not get this kind of benefit, and are instead personally detached from the case they oversee. These judges are thus "dispassionate." Before proceeding with the analysis, it is important to emphasize that the difference between these types of judges is modeled as a stark, all-or-nothing distinction only for expositional ease. The main qualitative results can be obtained with a more flexible model that allows all judges to have some degree of indirect personal benefit. What is important for the results, however, is that there are some judges—the issue motivated—who get *higher* benefits than others—the dispassionate.

2 The Model

The formal model builds on the frameworks of two recent working papers, Ashworth and Shotts (2011) and Gailmard and Patty (2013b). It depicts a stylized employment discrimination case between two litigants, P and D, initially heard by a district judge, J (which I refer to as the "district judge" or simply "judge"), and then appealed to an appellate court, A. Throughout the paper, I discuss the model as if it were a model of a trial, but the strategic interaction is more general. In fact, the model can be applied in a variety of situations where a district judge makes an appealable decision, such as rulings on motions to dismiss or rulings on motions for summary judgment.

Information

Adjudication is modeled as a process of costly information acquisition by the judge. The judge is trying to determine whether the plaintiff experienced legally prohibited discrimination, and this is initially unknown to all the players. Relatively liberal discovery rules often remove much uncertainty about factual information, so the judge's effort on the case matters precisely because "although the underlying facts may be not all that uncertain, they need an authoritative legal exposition" (Hornby 2009, p. 90).[24] In this model, the judge's authoritative legal exposition is achieved by exerting costly effort to figure out whether the plaintiff suffered discrimination.

The case that comes before the court is represented by a state of the world, $\omega \in \{P, D\}$, where $\omega = P$ indicates that the plaintiff experienced legally prohibited discrimination whereas $\omega = D$ indicates otherwise.[25] Before the case is heard, the players share a common prior belief that the plaintiff experienced legally prohibited discrimination with probability π. Essentially, this prior belief represents all the publicly available information—albeit imprecise—about how the law applies to the specific case before the court. For example, if $\pi = 0.25$, then all the players think that there was a 25% chance that the plaintiff actually suffered legally prohibited discrimination. For this reason, I refer to this as the "strength" of the case.

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

[24]Because there are strong constitutional protections for criminal defendants, asymmetry of information between the litigants and the judge may be *the* key issue shaping the adjudication of criminal cases. This model is not generally applicable to criminal cases.

[25]The model's informational structure is simple but can be micro-founded using the case space approach that has become popular among scholars of the courts (see Lax 2011).

Assumption 2 (hard information)**.** The state of the world, ω, is verifiable information that cannot be falsified.

At each point in the game, the player who is making a decision may or may not know for sure whether the plaintiff experienced illegal discrimination or not. I label the player's belief that the plaintiff did experience illegal discrimination as μ_k (where k is the index for the player). Throughout the paper, I refer to both π and μ_k as "beliefs," and it should be apparent from the context that π is the prior belief, whereas μ_k is a generic representation of a player's (prior or posterior) belief at any point in the game.

Sequence

The model focuses on adjudication by the judge, so it begins after the litigants have completed discovery and presented their arguments to the judge for adjudication.[26] First, the judge makes some unobserved effort $e \in [0, 1]$ to interpret the information, after which Nature sends a signal $s_J \in \{\omega, \phi\}$ to the players, either revealing whether the plaintiff experienced illegal discrimination (ω) or revealing nothing (which I label ϕ). The signal is stochastic, with $\Pr(s_J = \omega) = e$ and $\Pr(s_J = \phi) = 1 - e$. Substantively, this signal is a stylized (and coarse) representation of the judge's increasing certainty during the course of litigation that one side or the other should prevail.[27] After receiving the signal, 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 party, denoted by $L \in \{P, D\}$, 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

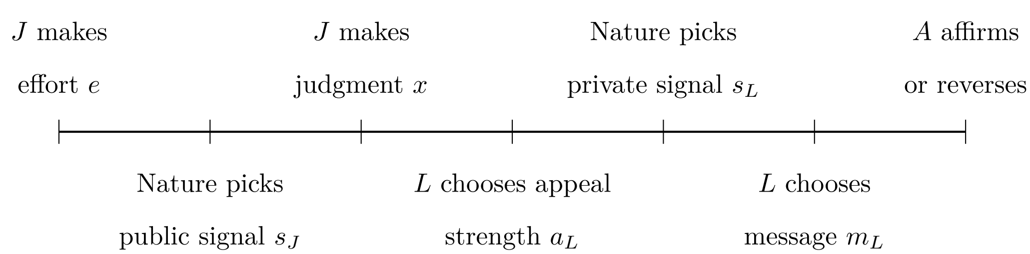
[26]Other models, such as Cameron and Kornhauser (2005, 2006), explicitly study litigants' private information over facts. This is clearly an important issue, but this paper's focus is on the judge's effort in resolving a case, not the asymmetric information between the judge and litigants.

[27]The fact that the signal is public is a simplification that obviates the need to study multiple equilibria driven by the judge's indifference over his or 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.

this effort has a 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 district judge by (1) making an objection on the record, and (2) giving the district judge an opportunity to correct the error. Taken together, these rules empower a district judge to make the first attempt to resolve the legal controversies in a case.

As with the judge's effort, Nature then sends a signal $s_L \in \{\omega, \phi\}$, where ω is revealed to the appellant with probability a_L. Unlike before, the signal received by the appellant is private, and the appellant can decide whether to communicate the information to the appellate court or conceal it. However, if the appellate 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, $r = \{0, 1\}$.

Figure 10: Model Timeline



Payoffs

The appellate court gets higher utility from a decision that matches the state of the world than one that 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. There are two types of judge, $t \in \{0, 1\}$.[28] If $t = 1$, then the judge is issue motivated and gets a benefit $\beta > 0$ from resolving the case in a manner consistent with the state of the world. However, if $t = 0$, then the judge is dispassionate gets no such benefit. The judge also bears two costs. The first is a cost when reversed, $k > 0$ (where $\beta \neq k$), which encompasses the judge's concerns about doing a good job. The second is a (quadratic) cost for effort, $\frac{1}{2}c_J e^2$, where $c_J > 0$. 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 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. If they lose they can opt to make a costly appeal to reverse the judge's decision. The losing litigant may choose to appeal the case, and pays a cost, $\frac{1}{2}c_L a_L^2$, where $L \in \{P, D\}$. There is no cost associated with filing the appeal, but to rule out strange 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.*, $m_P = D$ and $m_D = P$).

An important issue in this paper is whether the litigants' behavior has an effect on district court decision making. 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

[28] There is no uncertainty about judge type in this model, but this is an interesting avenue for future research.

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 lessons of the formal results do not depend on this assumption, but the substantive interpretation does.

Assumption 3 (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 maximal. These equilibria represent uninteresting scenarios whereby the players face no genuine tradeoff about how to allocate resources.

Assumption 4 (interior effort)**.** For the litigants, $c_P > c_D > 1$. For the judge, $c_J > \bar{c}_J$, where $\bar{c}_J$ is defined in Appendix B.

Strategies and Equilibrium

The appropriate equilibrium concept is perfect Bayesian equilibrium, which specifies a profile of sequentially rational strategies and beliefs updated using Bayes' rule on the equilibrium path. The adjudication strategy for the judge is a pair, $(e, x) \in [0, 1] \times \{D, P\}$, specifying a level of effort and a judgment. A strategy for the losing litigant is $(a_L, m_L) \in [0, 1] \times \{s_L, \phi\}$, specifying a level of effort and a message to send to the appellate court. A strategy for the appellate court is $r \in \{0, 1\}$, which specifies whether the appellate court affirms or reverses the district court's judgment. I focus on pure strategies, and adopt the following assumption.

Assumption 5 (indifference)**.** When indifferent, the judge rules in favor of the defendant and the appellate court's reversal strategy favors the defendant.

In the next section, I derive the sequentially rational strategies and characterize the equilibrium beliefs, and in the following section, I present the main results. Many of the results are written in natural language in the main text, with the corresponding formal language in the appendix. The proofs are collected in Appendix A.

3 Equilibrium Analysis

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 appellant 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 appellant offers hard information to the appellate court, then the appellate court's best response is to reverse a judgment whenever it has a belief that the judgment is more likely to be incorrect. In a perfect Bayesian equilibrium, this belief is derived using Bayes' rule where possible.

A losing litigant who discovers an error by the district 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, $L \in \{P, D\}$ 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, $s_J = \omega$—then the appellant would be wasting resources by exerting effort to mount a costly appeal. In this situation, $a_L^*(s_J = \omega) = 0$. However, if the judge does not present hard information to justify its decision, then the appellant's interim expected utility of mounting a strong appeal is:

$$U_L(a_L | x^* \neq L, s_J = \phi) = \begin{cases} a_P \pi - \dfrac{1}{2} c_P a_P^2 & \text{if } L = P \\ a_D(1-\pi) - \dfrac{1}{2} c_D a_D^2 & \text{if } L = D \end{cases}$$

The optimal level of 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:

$$a_L^*(s_J = \phi) = \begin{cases} \dfrac{\pi}{c_P} & \text{if } s_J \neq \omega \text{ and } x^* = D \\ \dfrac{1-\pi}{c_D} & \text{if } s_J \neq \omega \text{ and } x^* = P \end{cases} \tag{5}$$

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.[29] For simplicity, suppose that the appellate court always affirms judgments in the absence of hard information (I discuss this below). Then, this condition is:

$$\underbrace{(1-\mu_J)\beta t + \mu_J a_P^*(\beta t - k)}_{U_J(x=D)} \geq \underbrace{\mu_J \beta t + (1-\mu_J) a_D^*(\beta t - k)}_{U_J(x=P)} \tag{6}$$

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 plaintiff experienced legally prohibited discrimination (*i.e.*, learned ω), this condition is easy to check since $\mu_J = 0$ or $\mu_J = 1$. Specifically, if the judge learns that the plaintiff has experienced illegal discrimination, 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 whether illegal discrimination has occurred before having to make a ruling. In those situations, the judge will need to make a best guess using the prior belief that illegal discrimination occurred. 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

[29]This condition invokes Assumption 5 so that the judge rules in favor of the defendant when indifferent.

hard information. Then, to characterize the optimal judgment when uninformed, first define the following function:

$$\chi(\pi, t, \cdot) \equiv U_J(x = P) - U_J(x = D)$$

Using the expressions from equation (6) and substituting π for μ_J and the equilibrium values for a_D^* and a_P^*, this becomes:

$$\chi(\pi, t, \cdot) = (2\pi - 1)\beta t + (\beta t - k)\left(\frac{(1-\pi)^2 c_P - \pi^2 c_D}{c_D c_P}\right) \tag{7}$$

This expresses the benefit to the judge from ruling in favor of the plaintiff relative to ruling in favor of the defendant, and thus allows us to derive the equilibrium judgment. It also allows us to define a variable $\tilde{\pi}_t$, which is the level of π where where $\chi(\pi, t, \cdot)$ equals zero for a judge of type t. Intuitively, this represents how certain the type-t judge has to be that the plaintiff experienced illegal discrimination before being willing to rule in favor of the plaintiff in the absence of hard information.[30] That is, a type-t judge who is uninformed rules in favor of the defendant for all $\pi \leq \tilde{\pi}_t$ and in favor of the plaintiff for all $\pi > \tilde{\pi}_t$, as illustrated in Figure 11.

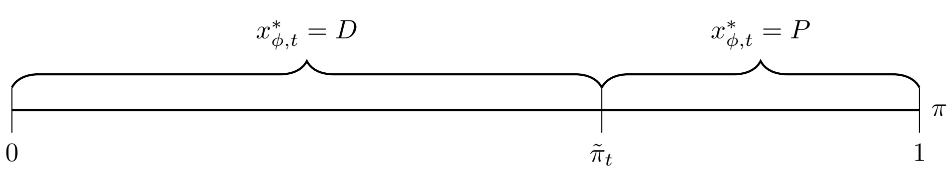


Figure 11: Decision making by an uninformed type-t judge

Notice that there is a "tilt" toward one of the litigants in the judge's decision making, which depends on the parameters of the function χ. For example, in Figure 11, the judge

[30]Appendix C explicitly characterizes $\tilde{\pi}_t$ as a function of the parameters in equation (7).

rules in favor of the defendant for a wider range of cases (as represented by the case strength, π).

As I mention 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 "tilt" toward one litigant, the appellate court may infer that the judge made a bad judgment and thus reverse it. Lemma 21 in Appendix A formally derives the condition under which it is optimal for the appellate court to affirm district court judgments. In brief, as long as $\tilde{\pi}_t$ is sufficiently low so that the judge does not tilt adjudication too far in favor of the defendant (*i.e.*, $\tilde{\pi}_t < \dot{\pi}$, where $\dot{\pi}$ is defined in the proof) then a type-t judge's optimal judgment is according to $\tilde{\pi}_t$. When this condition fails, the judge's judgment is constrained (although not completely) by appellate review.

The equilibrium judgment when informed (x^*_ω) and uninformed (x^*_ϕ) is therefore:

$$x^*_\omega = \omega \qquad\qquad x^*_\phi = \begin{cases} D & \text{if } \pi \leq \min\{\tilde{\pi}_t, \dot{\pi}\} \\ P & \text{if } \pi > \tilde{\pi}_t \end{cases} \tag{8}$$

The judge also makes a decision about how much effort to invest in adjudicating the case. Given the reversal strategy of the appellate court, the appellant's equilibrium effort and argument, and the judge's own equilibrium judgment strategy, J's interim expected utility is given by:

$$U_J(e; x^*_\phi, a^*_D, a^*_P, \cdot) = \begin{cases} e\beta t + (1-e)\left[\pi\beta t - (1-\pi)a^*_D(k-\beta t)\right] - \dfrac{c_J}{2}e^2 & \text{if } x^*_\phi = P \\ e\beta t + (1-e)\left[(1-\pi)\beta t - \pi a^*_P(k-\beta t)\right] - \dfrac{c_J}{2}e^2 & \text{if } x^*_\phi = D \end{cases} \tag{9}$$

As is reflected in equation (9), 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 J's ruling affects who the appellant is, which in turn affects the effort that the judge makes. By maximizing (9) with respect to e, we can derive the judge's optimal effort for both predispositions.

$$e^* = \begin{cases} \dfrac{1}{c_J c_D} \left[(1-\pi)(c_D - (1-\pi))\beta t + (1-\pi)^2 k\right] & \text{if } x^*_\phi = P \\ \dfrac{1}{c_J c_P} \left[\pi(c_P - \pi)\beta t + \pi^2 k\right] & \text{if } x^*_\phi = D \end{cases} \tag{10}$$

The analysis above demonstrates the existence of an equilibrium, which is given in Proposition 5.

Proposition 5. There exists a perfect Bayesian equilibrium of the game where:

- the appellate court reverses decisions it knows for sure are incorrect and affirms otherwise;
- the appellant makes an appeal of strength $a^*_L > 0$ when the judge does not discover hard information (where a^*_L is given by equation 5), and $a^*_L = 0$ otherwise;
- the appellant reveals its private information (if it has any) if and only if it can definitively show that the judge made an error (*i.e.*, if it has hard information that $\omega = L$);
- the judge makes effort $e^* > 0$ to acquire information about who should prevail (where e^* is given by equation 10) and issues a judgment x^*_ω when informed and x^*_ϕ when uninformed (where x^*_ω and x^*_ϕ are given by equation 8).

All beliefs are updated by Bayes' rule on the equilibrium path, and all non-singleton information sets are reached with positive probability.

Using the foregoing analysis, it is possible to examine how this equilibrium affects outcomes for litigants.

4 Assessing Outcomes: Issue Motivated versus Dispassionate Judges

To analyze whether adjudication is better with issue motivated or dispassionate judges, there are really two separate issues to consider. First, which type of judge leads to adjudication that is **more impartial**? And second, which type of judge leads to adjudication that is **higher quality**? To answer these questions, I first explicate normative benchmarks, and then I present some of the paper's main results about the type of outcomes that occur under appellate review of district court judgments.

The following definition formalizes the idea that a judge behaves in an impartial way whenever their judgment is based solely on features of a case, and not on other considerations.

> **Definition 2** (Normative Benchmark: Impartiality)**.** An "impartial" judge rules in favor of the plaintiff if and only if the judge believes it is more likely than not that the plaintiff experienced illegal discrimination. Moreover, a type-t judge is "more impartial" than a type-t' judge if and only if $\tilde{\pi}_t$ is closer to $\frac{1}{2}$ than $\tilde{\pi}_{t'}$.

It is easy to see the implications of this definition for a judge who happens to know for sure whether the plaintiff experienced illegal discrimination. In that situation, the judge rules in a manner consistent with the information available. However, sometimes the judge will not have full information, and the implications of Definition 2 may be less clear. In such situations, a judge who is impartial in their adjudication would issue a judgment reflecting the prior belief, π. For the range of possible values of π, we can illustrate decision making by an impartial judge graphically in Figure 12.

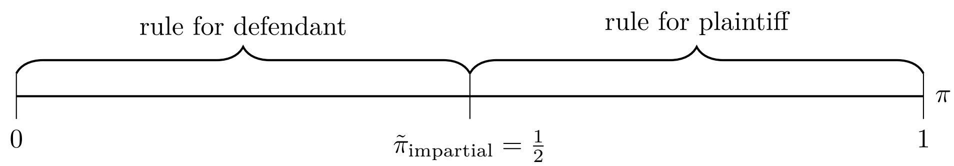


Figure 12: Decision making by a completely impartial judge.

Next, denote an equilibrium outcome of the adjudication model described above as $y^* \in \{D, P\}$. This represents the (final) outcome of the case, after taking into consideration the judge's ruling and the appellate court's reversal decision. Then, the following definition formalizes the quality of a judge's adjudication as the *ex ante* probability that y^* is correct.

> **Definition 3** (Normative Benchmark: Quality)**.** The "quality" of a judge's adjudication is represented by the equilibrium probability of a correct outcome (*i.e.*, $y^* = \omega$). For a type-t judge, denote this by $\xi_t(x_\phi^*) \equiv \Pr(y^* = \omega | x_\phi^*, \cdot)$. Moreover, the adjudication of a type-t judge is "higher quality" than that of a type-t' judge if and only if $\xi_t(x_\phi^*) > \xi_{t'}(x_\phi^*)$.

In the next two subsections, I look at each of these normative benchmarks in turn.

Impartiality of Adjudication

Given the equilibrium judgment made by the judge in equation (8), we can now proceed to examine the impartiality of the two types of judges.

> **Proposition 6.** If the litigants' costs are not too unequally matched, and the judge is sufficiently reversal averse, then dispassionate judges are weakly less impartial in their decision making than issue motivated judges, in the sense of Definition 2.

The result depends both on the judge being sufficiently reversal averse and the litigants' costs not being too different. I address each requirement in turn. The minimum k required for the result depends on the litigant's effort cost, c_D or c_P. Inspection of equation (15) in the proof of Proposition 6 indicates that as long as the costs faced by the litigants are sufficiently high, then the condition is satisfied for an arbitrarily low threshold k. Substantively, in contexts where judges are not reversal averse, the logic of Proposition 6 may break down since dispassionate judges' incentives are severely weakened. However, recent research suggests such scenarios to be rare (Choi, Gulati, and Posner 2011; Epstein, Landes, and Posner 2013). The result also depends on the litigants' costs not being too different from one another (relative to the litigants' common value benefit from winning). This requirement is more substantively constraining, as one can easily imagine scenarios where the defendant and the plaintiffs costs are vastly unequal. If the costs are too unequal, then Proposition 6 fails and the dispassionate judge may be no less impartial than an issue motivated judge.[31]

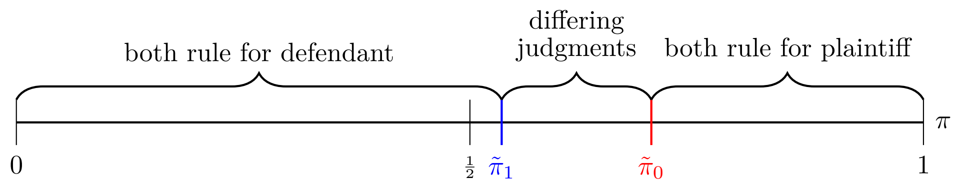


Figure 13: Decision making by dispassionate and issue motivated judges.

Figure 13 illustrates the basic logic of Proposition 6. As is apparent from the figure, a dispassionate judge adjudicates in favor of the more powerful litigant—the defendant—for a wider range of prior beliefs. Surprisingly, then, as long as the cost of reversal is sufficiently high, a dispassionate judge is actually *less* impartial. An important substantive takeaway can be drawn from this result. To the extent that plaintiffs in discrimination cases are less powerful, decision making by dispassionate judges will be tilted against them. This is a kind of "institutional bias" against employment discrimination plaintiffs because it emerges *without any explicit bias against plaintiffs.* Thus claims that focus on the hostility of judges toward employment discrimination plaintiffs need not be true in order for observationally

[31]There is reason to believe cases with vastly unequal litigants are relatively rare since those cases are likely to settle. Moreover, Appendix D shows that, even if Proposition 6 does not hold, a dispassionate judge can still be shown to be less impartial using an alternative conception of impartiality based on the *ex ante* probability that the judge rules in favor of the plaintiff.

equivalent behavior to emerge.

An immediate implication of Proposition 6 is that for some cases (as represented by their strength, π), dispassionate and issue motivated judges will make different judgments. This suggests that plaintiffs will make choices about whether to file discrimination suits based on the likelihood that they get dispassionate judges. To the extent that increasing diversity increases the proportion of judges who are issue motivated when it comes to employment discrimination cases, then substantive representation will be increased by appointing judges from more diverse backgrounds. Importantly, as Proposition 6 establishes, this is not necessarily accompanied by a loss of impartial decision making in this area of law (although this may worsen outcomes on *other* types of cases).

Quality of Adjudication

Appellate review provides important incentives, and without such review, the judge's incentives are weakened. Comparative statics on the judge's equilibrium effort, as described in Lemma 22 in Appendix A, show how the judicial hierarchy can induce higher effort in case management by district judges. Appeals raise the prospect of reversals, which district judges try to avoid by exerting higher effort to get the case right in the first place. However, effort also varies by the type of the judge, as the following result shows.

> **Lemma 5.** A dispassionate judge exerts strictly less effort than an issue motivated judge.

The logic of Lemma 5 is quite straightforward: an issue motivated judge will exert more effort than a dispassionate judge because issue motivated judges get some indirect personal benefit from resolving the case. Therefore, information is less valuable for a dispassionate judge, who just wants to lower the probability of being reversed upon appeal. Now, I explore whether the quality of adjudication is higher under a dispassionate judge or an issue motivated judge. First, I explicitly characterize $\xi(x_\phi^*|t)$. Define $e^*(x_\phi^*, t)$ to be the equilibrium level of effort conditional on predisposition x^*_ϕ and judge type t. Adjudication is correct with the following *ex ante* probability

$$\xi(x^*_\phi|t) = \begin{cases} e^*(x^*_\phi, t) + [1 - e^*(x^*_\phi, t)][1 - \pi(1 - a^*_P)] & \text{if } x^*_\phi = D \\ e^*(x^*_\phi, t) + [1 - e^*(x^*_\phi, t)][1 - (1 - \pi)(1 - a^*_D)] & \text{if } x^*_\phi = P \end{cases} \tag{11}$$

Inspection of equation (11), reveals that adjudication is strictly worse under a dispassionate judge whenever the two types of judges would make the same judgment when uninformed. However, there may be circumstances where this is not the case (as illustrated by Figure 13). Then, the comparison between a dispassionate judge's adjudication and an issue motivated judge's adjudication is not as straight-forward. However, in the interval between $\tilde{\pi}_0$ and $\tilde{\pi}_1$, which is the interval where the types have different predispositions, the proof of Proposition 7 shows that adjudication by a dispassionate judge is still lower quality than adjudication by an issue motivated judge.

> **Proposition 7** (Quality of Adjudication with Review)**.** Adjudication by a dispassionate judge is lower quality than adjudication by an issue motivated judge.

Substantively, Proposition 7 says that issue motivated judges end up making accurate decisions more often than dispassionate judges. In the model, "accuracy" is presumed to be an achievable goal, which is what makes this a model of law enforcement rather than law creation or evolution. 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 district judge's decision was not vacated due to any ideological or doctrinal disagreement, but rather due to its seemingly low quality.

5 Extension: Biased Judges

The foregoing analysis demonstrated advantages of having issue motivated judges over dispassionate judges. In that analysis, there is no conflict over doctrine between the judge and appellate court, which helps focus attention on the way that the appeals process can alter a judge's decision making. Many observers of the courts have also expressed concerns that judges are *biased* in favor or against particular types of litigants. In this section, I extend the analysis to examine the conditions under which an explicitly biased judge can produce better outcomes than a dispassionate judge. Now, the judge can be either a pro-plaintiff biased judge (who I refer to as the "biased judge" and the "plaintiff-biased judge" interchangeably) or a dispassionate judge. As before, all players know whether the judge is biased or dispassionate. 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 $\delta > 0$, and if the final disposition favors the defendant, then a biased judge receives no benefit. 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 $\varepsilon_J > 0$ from writing a decision that provides the appellate court with incontrovertible information that the judgment is correct.[32] I also focus on equilibria supported by reasonable off-equilibrium path beliefs in order to avoid analyzing those that are driven by strange beliefs in situations that do not arise.[33] Given that the district judge may now have an incentive to conceal his or her information, two types of equilibria emerge.

[32]Specifically, I assume $0 < \varepsilon_J < \pi(1-\pi)\delta$. This rules out equilibria where the judge conceals information that supports the judgment that she or he makes.

[33]In particular, a deviation from the equilibrium judgment in the absence of information leads the appellate court to infer that the district judge has no information. The details are in the proof of Lemma 6.

Lemma 6. With a biased district judge, there exist two types of equilibria. In these equilibria:

- the appellate court affirms equilibrium judgments in the absence of specific hard information, and corrects errors otherwise;
- the appellant exerts positive effort according to equation (5);
- and the district judge's equilibrium behavior is as follows:
  - P**-Equilibrium**: for all $\pi \in (\hat{\pi}, \pi_d)$ and $\delta > \hat{\delta}$ or for all $\pi \in (\pi_d, 1)$ and $\delta > 0$, the judge rules in favor of the plaintiff in the absence of information and the judge exerts no effort, where $\hat{\pi} < \frac{1}{2} < \pi_d$, and $\hat{\pi}, \pi_d$ are defined in the proof.
  - D**-Equilibrium**: for all $\pi \in (0, \frac{1}{2})$ and $\delta > 0$ or for all $\pi \in (\frac{1}{2}, \pi_d)$ and $\delta \leq \hat{\delta}$, the judge rules in favor of the defendant in the absence of information and the judge exerts positive effort according to equation (19) in the proof.

Both equilibria are supported by off equilibrium path beliefs and actions described in the proof.

The equilibrium behavior of the appellate court and the appellant are the same across the two types of equilibria, but the district judge's decision making varies. Figure 14 illustrates how the district judge alters its judgment and effort across the parameter space.

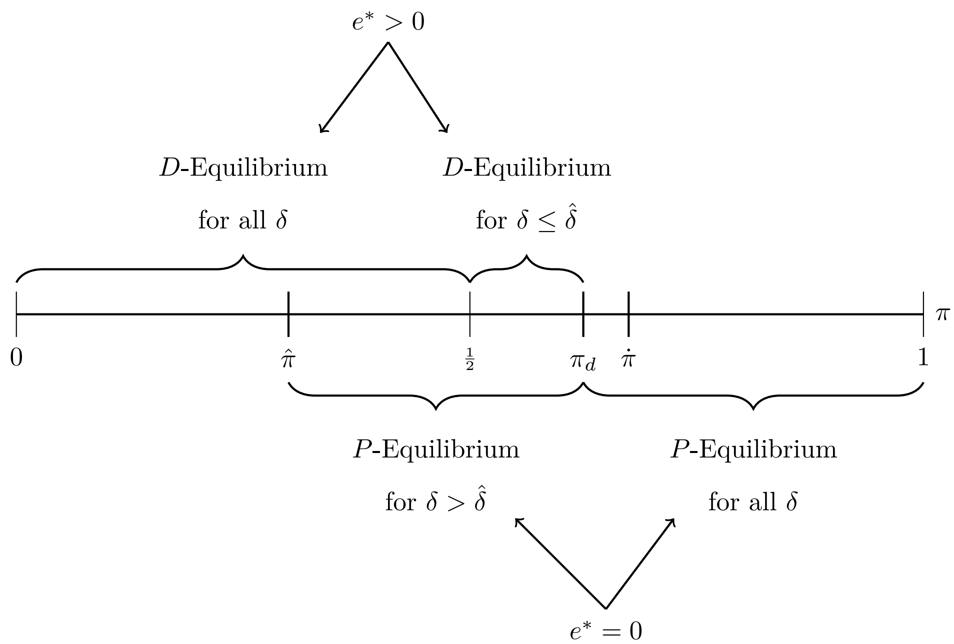


Figure 14: Equilibria with plaintiff-biased judges.

Due to the addition of bias in this revised model, it is now harder to incentivize the district judge to exert effort when the plaintiff has a very strong case. Therefore, 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. On the other hand, whenever the plaintiff's case is relatively weak (*i.e.*, $\pi < \frac{1}{2}$), then there always exists an equilibrium involving a pro-defendant outcome when the appellate court does not have definitive information. However, for a smaller range of cases, there exists an equilibrium involving a pro-plaintiff outcome when the appellate court does not have definitive information. This latter type of equilibrium can be viewed as a "bad" equilibrium in the sense that it entails no effort by the district judge and a pro-plaintiff decision when the prior probability actually supports the defendant. For all $\pi \in (0, \hat{\pi})$, that bad equilibrium does not exist, 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. Moreover, as the proof of Lemma 6 shows, the equilibrium effort of a biased judge is strictly positive in this range.

Would there be a benefit if the biased judge were instead dispassionate? For a D-Equilibrium, the biased judge produces a higher quality outcome (in the sense of Definition 3) than a dispassionate judge. Moreover, for all $\pi < \frac{1}{2}$ or for all $\pi < \pi_d$ and $\delta \leq \hat{\delta}$ the only equilibrium of this revised model is a D-Equilibrium.

> **Proposition 8.** Suppose that the plaintiff's case is not too strong (*i.e.*, $\pi < \bar{\pi}$, where $\bar{\pi}$ is defined in the proof). Then, exists an equilibrium where adjudication by a dispassionate 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 a dispassionate judge is always lower quality than adjudication by a plaintiff-biased judge.

Because of the possibility of multiple types of equilibria, the results with a biased judge are more nuanced than those in the previous sections. In particular, it is possible for an equilibrium to emerge that features worse adjudication by a biased judge relative to a dispassionate judge. However, the results are particularly interesting the context of employment discrimination cases, where it is difficult for plaintiffs to prevail (Clermont and Schwab 2004).[34] This difficulty arises from the fact that claims of illegal discrimination are difficult to prove due to the burdens that a plaintiff must satisfy in order to establish a *prima facie* case and prevent dismissal. Therefore, in the specific situations that cause concern in employment discrimination law—a plaintiff-biased judge presiding over a case that is weak for the plaintiff and which features a relatively powerful defendant—this revised model shows that there *always* exists an equilibrium where the district judge rules in favor of the defendant when uninformed, despite having a pro-plaintiff bias. In this equilibrium, the biased judge is a better quality adjudicator than the dispassionate judge. Moreover, this is the *only*

[34]In a dataset collected by the author, employment discrimination cases accounted for around 11% of the 51,696 civil appeals filed in the Ninth Circuit between 1995 and 2013. Around 11% of those cases were won by the plaintiff at the district court level and 89% were won by the defendant. Moreover, cases won by plaintiffs were much more likely to be reversed: 43% versus 21% for cases with defendant winners. These statistics provide additional suggestive evidence that it is difficult for plaintiffs to prevail in such cases.

equilibrium as long as the judge is not too biased or the case is not too uncertain.

It is important to note that the kind of bias studied here is stark: the judge gets personal utility when the plaintiff wins, regardless of other factors such as specific features of the plaintiff's case. Yet, despite this extreme form of bias, it is still possible to recover equilibria where biased judges are better adjudicators than dispassionate judges. This suggests that the main results of the previous sections underscore a more fundamental set of tradeoffs induced by the process of appellate review; these are tradeoffs that operate even in the presence of biased judges.

6 Discussion and Conclusion

District courts are important institutions for law enforcement, and the model in this paper offers a framework for studying adjudication in lower courts that departs from previous research on the judiciary's role in lawmaking. Enforcement of laws is not always straight-forward, and this is particularly true in areas of law involving complex legal issues or a complicated interplay between factual determinations and law application. In hard-to-adjudicate cases, the effort a judge puts into resolving a case can be consequential; and this is a foundational assumption for the model I analyze. While some scholars have emphasized district judges' role in acquiring information about cases (*e.g.*, Epstein, Landes, and Posner 2013; Posner 2008, 2013), I present a model that formalizes this logic and I study its implications in the context of appellate review. A major contribution of the model is its attention to judges' active role in the *process* of adjudication, beyond just determining outcomes.

Employment discrimination is the prototypical example of a complex body of law where case-by-case enforcement in the courts occupies a significant portion of the federal docket (Farhang 2010; Rutherglen 2007). Yet, in these kind of complicated cases, questions often arise about judicial temperament and the ability of litigants to get both a fair and an adequate hearing in court. As in the California marriage case, some have worried (or suspected) that a judge's membership in a historically disadvantaged group may lead that judge to adjudicate in a manner that is unfairly beneficial to that group. For example, in *Blank v. Sullivan & Cromwell* (1975), the defendants in a sex discrimination suit petitioned for the recusal of U.S. District Judge Constance Baker Motley because she was a woman. In her denial of the defendant's motion, Judge Motley pointed out that "if background or sex or race of each judge were, by definition, sufficient grounds for removal, no judge on this court could hear this case, or many others" (9).

While Judge Motley's words are certainly correct (after all, all judges have a gender identity), they do not directly speak to related, underlying concerns that judges who approach cases with an especially keen interest present the possibility of partiality. For example, Judge Motley had also been prominent civil rights lawyer. In a recent high profile example from the Southern District of New York over New York City's stop-and-frisk practices, Judge Shira Scheindlin was removed from the case by the Second Circuit over concerns about impartiality. The removal was prompted by Scheindlin's media statements on stop-and-frisk, as well as her acceptance of multiple cases on the issue using the district's "related case" rule (thus skipping random assignment). In its rulings, the Second Circuit emphasized that there were "no findings of misconduct, actual bias or actual partiality," but removed her since the *appearance* of impartiality had been compromised by her publicly known interest in the cases (Weiser and Goldstein 2013).

To avoid disrupting the *appearance* of impartiality, judges must remain dispassionate, detached arbiters in the cases they oversee. This precludes them from, among other things, appearing to express excessive interest in particular cases, as Judge Scheindlin did. Yet, Scheindlin's example underscores the tradeoff between dispassion and effort illustrated by the model in this paper. The *New Yorker* article that was cited in the Second Circuit's order removing Scheindlin also noted that "Scheindlin is renowned for her work ethic" (Toobin

2013). In fact, prior to this controversy, Judge Scheindlin was perhaps most famous for a series of rulings she made in 2003 and 2004 in a sex discrimination case, *Zubulake v. UBS Warburg* (2004). She has been widely praised for rolling up her sleeves and working through the complicated issues surrounding the pre-trial discovery process in that case, which ended up revolutionizing electronic discovery in the federal courts (Li 2014). So, by replacing Judge Scheindlin with a more dispassionate jurist, would justice be better served on the stop-and-frisk cases?

It is entirely possible that Judge Scheindlin was both biased against the defendants (the city) and determined to rule against them regardless of the facts at hand. But this account does not fully consider the fact that judges sitting at the lowest tier of the judicial hierarchy face the prospect of appellate review and make decisions in anticipation of that review. The model analyzed in this paper demonstrates how judges' aversion to being reversed can tilt their judgments toward more powerful litigants, and that judges who are dispassionate do this even more. Moreover, since a judge's effort in adjudication is consequential for coming to the correct outcome, the model shows that dispassionate judges are worse adjudicators because they exert less effort to resolve cases. These results stand in stark contrast with a commonly held view that an ideal judge is one who is dispassionate and emotionally detached from the cases. In fact, dispassionate judges are both less impartial and more often wrong.

These findings suggest that caution is warranted when interpreting recent empirical evidence that judges from historically disadvantaged groups rule in a manner that, on average, benefits their group in salient legal areas (*e.g.*, Boyd 2015). In the model, having a degree of issue motivation can induce a judge to exert more effort and to be more impartial, and this is, broadly speaking, beneficial to less powerful litigants. To the extent that judges with personal experience with discrimination or harassment (as in Judge Edwards's quote) have more issue motivation *on average*, then the model suggests that differential decision making could also be the result of higher quality decision making on average by those judges, rather than in-group bias. To put it differently, the empirical patterns recently identified in the literature may simply reflect an average tendency of white or male judges to rule disproportionately in favor of the more powerful litigants who are being accused of discrimination. Interestingly, the model shows how this latter pattern can arise *in the absence of explicit hostility toward any litigants* and represents a sort of "institutional bias" against less powerful litigants.

Of course, the model presented here cannot adjudicate between all competing explanations for the differential decision making. The foregoing discussion simply points out that there is potential observational equivalence between a world where minority and women judges rule unfairly in favor of members of their group and one where they behave as high quality adjudicators. Yet it is seemingly undeniable that a judge's effort is consequential for the quality of decision making in district courts, and to the extent that this plays a role in adjudication, the dynamics of the model will be present. And, as an extension of the model, I examine whether having an explicitly plaintiff-biased district judge will undermine the results in the situations most interesting for employment discrimination law (*i.e.*, when plaintiffs have low *ex ante* probability of prevailing and the defendant is the more powerful litigant). As long as a judge is not too biased, the results show that even explicitly biased judges are higher quality adjudicators in the sense that they exert more effort to resolve cases and come to the right answer more often.

Beyond the issue of diversity on the bench, these results also have broader implications for the optimal staffing of the judiciary. The model raises questions about the desirability of having "generalist" judges serving in district courts. For any given judge, it is inevitable that there will be some areas of law for which the judge has strong preferences and other areas of law in which the judge has little interest. Quite apart from the usual scholarly concern about generalist judges' lack of technical expertise, this model suggests that their day-to-day work product may be inferior than that of specialist judges who select into the particular areas of law that interest them. In the area of employment discrimination, Congress specifically opted to delegate enforcement to a body of generalist judges instead of a set of bureaucrats motivated by civil rights issues. On the issue of whether this is an optimal design for the enforcement regime, this model suggests that there could be substantial, unconsidered downsides in terms of the quality and fairness of enforcement.

Paper III

Endogenous Appeals

What role to litigants play in the resolution of legal disputes? One view is that they have private information about the state of the world and signal (via costly actions) that their case has merit (Priest and Klein 1984; Cameron and Kornhauser 2006; Cooter and Ulen 2012). Another, less common view is that they *acquire* information that helps judges make decisions. For example, Dewatripont and Tirole (1999) study a model of information acquisition in to demonstrate that "advocacy" (*i.e.*, having two competing investigators) can be a more effective way for an organization to acquire important information about a decision than "nonpartisanship" (*i.e.*, having a single investigator). In an extension of their model, they argue that advocacy can also generate efficient endogenous appeals in a context where a decision maker is biased.

The main lesson of this research is that litigants (and appeals, more broadly) are important for helping judges and courts make more informed decisions (see for example, Shavell 2006). While this may be true, there is an overlooked downside of litigant-driven appeals. The model in this paper shows that a decision maker may purposefully make outcomes worse in order to prevent having their decisions overturned. The model features judges who are "legalists" and thus wish to adjudicate cases correctly and and in an unbiased manner. Even so, I show how endogenous appeals by losing litigants can incentivize trial judge to ignore her own information about which litigant should prevail on the merits of a case. Such inefficient equilibria should be common when litigants are relatively good at mounting strong appeals, thus undercutting the widely acknowledged benefits of advocacy.

1 Model

I analyze a model of decision making over civil cases adjudicated in the federal court system.[35] In contrast to other models of courts that focus on judicial policy making (alternatively described as doctrinal politics or rule making), the model in this paper analyzes a process of adjudication where the main goal is to resolve individual cases correctly (alternatively described as law application). A case features information acquisition by a trial judge, J (or "she") to figure out how best to resolve a case,[36] effort by the losing litigant, L (or "he"), to mount an appeal of the trial judge's decision and a decision by the appellate court, A (or "it"), about whether to review, and possibly reverse, the trial judge's decision.

Cases. A case has two parties, a plaintiff P and a defendant D. Moreover, a case is defined by two components: the merits of the case and the quality of the litigants. The first component simply identifies which party—P or D—should prevail on the merits, and is formally described by a state of the world, $\omega \in \{P, D\}$. This state of the world is initially unknown to all parties, who share a common prior belief that $\Pr(\omega = P) = \pi$ and $\Pr(\omega = D) = 1 - \pi$, where ω is independently drawn. The second component of each case captures the quality of each litigant's advocacy and is formally represented by a vector of cost parameters, $c = (c_P, c_D)$. These parameters reflect the litigants' costs for mounting a strong appeal after a loss in the trial court. The litigants' costs are a simple way to capture resource differences between the plaintiff and defendant. For example, if $c_D > c_P$, then the plaintiff is the "more powerful" or "higher resource" litigant since his costs are lower. These costs are commonly known by all players.

Sequence. A case begins with Nature independently choosing c and ω, publicly revealing

[35] Given the significant differences between civil and criminal procedure in the United States, I explicitly model civil cases. However, the underlying strategic trade-off highlighted in the analysis has resonance with criminal cases too. An interesting avenue for future research would be to extend the model to the context of criminal law.

[36] In this model, the judge's information acquisition is non-strategic and the judge's efficacy in doing so reflects her exogenous and commonly known competence rather than hidden actions.

the litigant cost vector but not the merits of each case. Nature chooses ω according to the distribution above: $\Pr(\omega = P) = \pi$ and $\Pr(\omega = D) = 1 - \pi$. Nature chooses c according to a generic distribution $F(c)$, but this distribution plays no role in the analysis since all players immediately learn c at the start of the game. The trial judge then deliberates over the merits of the case. In this model, deliberation represents the judge's process of information acquisition over the (unknown) state of the world. It can therefore take many forms, including pretrial case management (discovery conferences/orders, evidentiary hearings, supervision of jury selection, *etc.*) as well as rulings on dismissal or summary judgment motions and drafting of jury charges.

The trial judge is limited in her ability to fully establish the merits of the case, and is not always successful in doing so. For example, she could be constrained by large case load and have limited time to devote to each individual case. Alternatively, she could be a less competent (or lower quality) judge. This limitation are captured by the parameter $\delta \in (0, 1)$, which is the probability that judge fully discovers the merits of the case (*i.e.*, learns ω). Formally, I represent the judge's deliberation as a signal from Nature to the judge, $s_J \in \{\omega, \phi\}$, where ϕ denotes that the signal is uninformative about the merits of the case. Thus, the probability that the trial judge gets an informative signal is $\Pr(s_J = \omega) = \delta$.

In order to eliminate an analytically uninteresting signaling subgame between the trial court and the appellate court, I assume that the trial judge's signal s_J is public and thus observed by all the players. The reason for this assumption is simple. Given that the trial judge and the appellate panel are both legalists, there is no disagreement over outcomes if $s_J = \omega$. Therefore, a signaling subgame could feature multiple, substantively strange equilibria where the trial judge conceals information supporting her decision because she's indifferent about whether or not to reveal it.

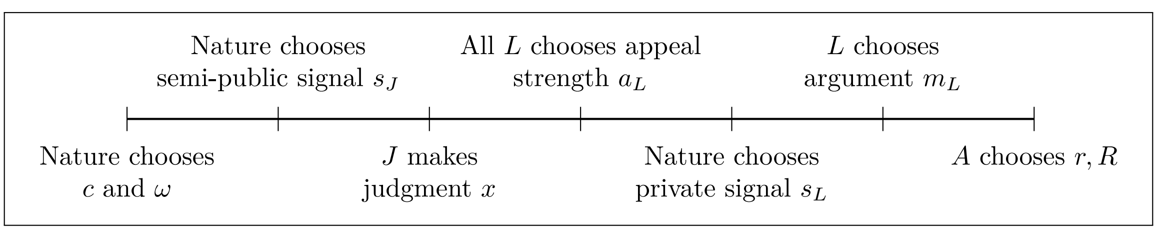
After receiving signal s_J, the trial judge makes a ruling on the case, finding in favor of the plaintiff or the defendant, $x \in \{P, D\}$. The losing litigant $L \neq x$, then chooses whether to mount a strong appeal appeal, $a_L \in \{0,1\}$, where $a_L = 1$ indicates a strong appeal and $a_L = 0$ indicates a weak appeal. Substantively, this represents the losing litigant's effort to preserve a trial court's errors (as required), present novel legal arguments to the appellate court justifying reversal, or in rare circumstances, discover new evidence warranting relief from the lower court's decision. After deciding how much effort to exert on appeal, the litigant receives a private signal, $s_L \in \{\omega, \phi\}$. The signal depends on the strength of the litigant's appeal:

$$\Pr(s_L = \omega) = \begin{cases} \alpha_1 & \text{if } a_L = 1 \\ \alpha_0 & \text{if } a_L = 0 \end{cases}$$

where $1 > \alpha_1 > \alpha_0 > 0$. If L successfully learns ω, he can decide whether to reveal this information to the appellate court or to conceal it, sending a message $m_L \in \{0,1\}$. This message can be thought of as the appellant's briefs and/or oral arguments presented to the appellate court.

Finally, the appellate court observes the judgment x and decides (1) whether to review the case, $r \in \{0,1\}$, which is costly, and (2) if so, whether to reverse the trial court's decision, $R \in \{0,1\}$. If $r = 1$, the appellate court learns s_J and m_J, and may base its reversal decision on this additional information. Figure 15 depicts the time line of the game.

Figure 15: Model Timeline



Payoffs. The trial judge and the appellate court are better off when cases are resolved on their merits so that the final outcome is ω. The two courts also face costs. The trial judge faces a cost c_J when it exerts effort to resolve the merits of a case and κ when it is reversed. The appellate court faces a quadratic cost $\frac{c_A}{2}r^2$ to review each case. The latter cost reflects both the court's opportunity cost for reviewing each case as well as direct costs (*e.g.*, scheduling hearings, conducting deliberations, reviewing materials, writing opinions, etc.). The payoffs of J and A are:

$$u_J = R|x-\omega| + (1-R)\big(1-|x-\omega|\big) - r\kappa$$
$$u_A = R|x-\omega| + (1-R)\big(1-|x-\omega|\big) - \frac{c_A}{2}r^2$$

Implicit in these payoffs is an assumption about judicial preferences that purposefully departs from conventional wisdom among scholars of American courts. In particular, I depart from attitudinalist conceptions of judges (Segal and Spaeth 1993) and assume that both the trial judge and the judges on the appellate panel are idealized "legalist" jurists. They believe (correctly) that there is a correct legal answer about who should prevail, but are *ex ante* uncertain about what this answer is.

> **Assumption 6** (legalist preferences)**.** The trial judge and the appellate court are better off if the ultimate outcome of the case matches the state of the world.

I make an additional and crucial assumption that ensures that the trial judge is sufficiently reversal averse as to create an interesting conflict between her and the appellate court. If her reversal aversion κ is too low, then the trial judge (who is unbiased in this model) simply free-rides on the appellant's effort and maximizes the number of correct decisions made by the courts. Some judges are certainly not very reversal averse, at least not relative to their preferences over other aspects of adjudication such as outcomes or effort aversion. However, there is ample empirical evidence that trial judges widely engage reversal avoidance, and strategic anticipation of appellate review more generally (Schanzenbach and Tiller 2006; Randazzo 2008; Boyd and Spriggs 2009). Thus, while it is plausible that some judges behave this efficient way, the focus of this paper is on studying the potential for *inefficiency* induced by endogenous appeals.

Assumption 7 (reversal averse)**.** J is sufficiently reversal averse: $\kappa > 1$.

Finally, each litigant $L \in \{P, D\}$ receives a benefit b from winning the case, and pays a cost c_L for mounting a strong appeal. In addition, I assume that the losing litigant cares about his reputation. In particular, I assume that he gets a benefit from whenever his appeal communicates to the appellate panel that he *should have won* and pays a cost whenever his appeal communicates to the appellate panel that he *should have lost*. Substantively, these costs capture *advocacy*: the attorneys representing the party get a benefit when their advocacy in favor of their client is substantive and visible to outsiders, and suffer a cost when they appear to publicly advocate for the other side.[37]

Assumption 8 (advocacy)**.** Let $\varepsilon > 0$. Then L receives an additional advocacy payoff of:

- ε if $m_L = \omega = L$; and
- $-\varepsilon$ if $m_L = \omega \neq L$.

Information. The merits of the case are initially unknown to the players, and I make assumptions on the players' (common) prior belief. Specifically, I assume that each case is somewhat favorable to the plaintiff *ex ante*. The assumption that $\pi > \frac{1}{2}$ does not affect the main takeaways of the model, as one could easily reverse this and obtain results that are mirror images of those derived below.

2 Analysis

I begin the analysis by examining the appeal process. First notice that in this model, if the appellate court does not review a case, this automatically results in an affirmance of the trial

[37]Formally, this assumption eliminates a type of equilibrium driven by L's indifference whenever he finds out that the state of the world does not support his cause.

court's decision. This mirrors the appellate process in federal judiciary, where a large share of filed appeals are not reviewed on the merits by appellate courts.[38] This effectively affirms a lower court's decision.

If, however, the appellate court reviews the case, it observes s_J and m_L, forms a posterior belief and decides whether to affirm or reverse. Its sequentially rational reversal strategy is straightforward: reverse if and only if it believes that the decision is more likely to be incorrect than correct.

Lemma 7. Let α_L^* be L's equilibrium appeal strength and w_L^* be A's equilibrium belief that $\omega = L$. Moreover, assume that A does not reverse when indifferent.[39] A's equilibrium reversal strategy is

$$R(\alpha_L^*) = \begin{cases} 1 & \text{if } w_L^* > \dfrac{1}{2} \text{ and } \alpha_L^* < \underline{\alpha} \text{ and } A \text{ reviews} \\ 0 & \text{otherwise} \end{cases}$$

where $\underline{\alpha}$ is the function defined in the proof.

The appellate court's review strategy will be a function of the equilibrium judgment and the beliefs induced by the trial judge's public signal s_J. Then, denote A's review strategy when L is the appellant (*i.e.*, when L loses in the trial court, $x^* \neq L$) by r_L. Then, given her reversal strategy, the appellate court selects a review probability that maximizes her expected net benefit from review.

Lemma 8. A's optimal review strategy is:

$$r_L(\alpha_L^*) = \begin{cases} \min\left\{\dfrac{w_L^*}{c_A}, 1\right\} & \text{if } w_L^* > \dfrac{1}{2} \text{ and } \alpha_L^* < \underline{\alpha} \\ \min\left\{\dfrac{w_L^*\alpha_L^*}{c_A}, 1\right\} & \text{otherwise} \end{cases} \tag{12}$$

where w_L^* is A's equilibrium belief that $x \neq \omega$ and $\underline{\alpha}$ is defined in the proof of Lemma 7.

[38] For example, in the Ninth Circuit between 1995 and 2013, only about 52% of the nearly 52,000 civil appeals filed were reviewed on the merits.

[39] This occurs under a knife-edge condition $\alpha_L^* = \underline{\alpha}$ that can assumed away without changing any of the qualitative results.

Next, I consider the appellant's decision making. Let the appellate court's equilibrium review probability be r_L^*. The litigant optimally exerts high effort if the expected benefit of doing so is greater than the cost.

Lemma 9. Let λ_L be L's belief that $\omega = L$. In equilibrium, if $\alpha_0 < \underline{\alpha}$ and $\lambda_L > \frac{1}{2}$, then L's equilibrium appeal is $a_L^* = 0$. Otherwise, L's equilibrium appeal is a function of his cost:

$$a_L(c_L) = \begin{cases} 1 & \text{if } \bar{c}_L \geq c_L \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

where c_L is defined in the proof.

By Lemma 7, the appellate court only reverses when it believes the judgment x^* is incorrect and α_L^* is sufficiently low. Whenever the litigant learns that $\omega = L$, he is always weakly better off revealing this information, as he will secure reversal for sure. Otherwise, the litigant is better off sending an uninformative signal.

Lemma 10. L's equilibrium message is

$$m_L(s_L) = \begin{cases} \omega & \text{if } s_L = \omega = L \\ \phi & \text{otherwise} \end{cases}$$

Next, I analyze the trial process. The trial judge anticipates the behavior of both the appellant and the appellate court in deciding how to rule in each of its cases. Then, she rules in favor of the plaintiff if her expected utility from doing so is higher than her expected utility from ruling in favor of the defendant. Let μ_L be the trial judge's belief that $\omega \neq L$. Given the appellate court's reversal strategy in Lemma 7 and the appellant's appeal strategy in Lemma 9, there are two cases to consider: (1) $\alpha^0 < \underline{\alpha}$ and (2) $\alpha^0 \geq \underline{\alpha}$. In the first case, the appellate court always reverses if $x = D$ and $s_J = \phi$, regardless of what the appellant does. Then, the trial judge is strictly better off preventing an appeal by ruling in a manner consistent with her belief. That is, $x^* = D$ if and only if $s_J = D$.

Next, consider the second case. Now, the appellate court reverses a decision if and only if the litigant provides hard information that the judgment is incorrect. That is, if $m_L = \omega = L$. As a result of this, the trial judge has more flexibility in making her judgment. Let χ^* be the trial judge's belief that her judgement x is incorrect. Then, her judgment is derived from the following maximization problem:

$$\max_{x \in \{P,D\}} (1-\chi^*) - \chi^*(\kappa-1) r_L^*(\alpha_L^*)\alpha_L^* \tag{14}$$

Label the trial judge's net expected benefit from ruling in favor of the plaintiff instead of the defendant whenever she is uninformed ($s_J = \phi$) as Δ. Substituting in equilibrium values into (14) and rearranging, I express Δ as a function of the model's parameters.

$$\Delta(\pi, \kappa, c_A, c_P, c_D) = 2\pi - 1 + \frac{\kappa-1}{c_A}\left[\pi^2(\alpha_P^*)^2 - (1-\pi)^2(\alpha_D^*)^2\right]$$

Using the function $\Delta(\cdot)$, we can characterize the trial judge's optimal judgment. I assume that, when indifferent, she rules in favor of the plaintiff.[40]

Lemma 11. Given the optimal strategies of A and L, J's equilibrium judgment strategy is:

$$x(\pi, \kappa, c_A, c_P, c_D) = \begin{cases} D & \text{if } (\alpha_0 \geq \underline{\alpha} \text{ and } \Delta(\pi, \kappa, c_A, c_P, c_D) < 0) \text{ or } s_J = D \\ P & \text{if } (\alpha_0 \geq \underline{\alpha} \text{ and } \Delta(\pi, \kappa, c_A, c_P, c_D) \geq 0) \text{ or } s_J = P \\ P & \text{otherwise} \end{cases}$$

Now that I have characterized the optimal behavior of each player, I show that an equilibrium exists.

Proposition 9. There exists a perfect Bayesian equilibrium where:

- J's judgment is given by $x(\cdot)$.
- A uses review strategy $r_L(\cdot)$ and reversal strategy $R(\cdot)$.
- L exerts effort according to $a_L(\cdot)$ and sends a message $m_L(\cdot)$.

[40] Again, this assumption is innocuous, as it rules out equilibria driven by knife-edge conditions and does not substantively alter the results.

- Beliefs are updated by Bayes rule on the equilibrium path, and all information sets are reached in equilibrium.

This equilibrium can be summarized briefly. First, note that there are two regions of the parameter space that generate different outcomes. First, if $\alpha_0 < \underline{\alpha}$, then the equilibrium features the trial judge ruling consistent with her information, litigants mounting weak appeals, and the appellate court reversing only when a losing defendant provides hard information that the trial judge made an error. Second, if $\alpha_0 \geq \underline{\alpha}$, then the resource asymmetry between the litigants (in addition to information about the merits) influences the trial judge's decision making. She rules in favor of relatively more powerful litigants more often, even if her belief is that the merits support the weaker litigant. Moreover, the appellate court still reverses only when the losing litigant provides hard information that an error was made. In the next section, I show how this equilibrium outcome can be inefficient.

3 Inefficient Appeals

The equilibrium characterized in Proposition 9 has some interesting implications, which I explore in this section. The main goal of this model is to illustrate how endogenous appeals can generate inefficient judgments. Moreover, these inefficient judgments are *not* driven by the trial judge's bias, but rather the structure of the appeals process. It is important to underscore that in this model, endogenous appeals are not generated by litigants' private information as they are in other models of the litigation (see, for example, Cameron and Kornhauser 2006; Cooter and Ulen 2012). Instead, the endogenous appeals studied in this model involve costly effort by the appellant to discover that the judge had made an error in their initial decision. As a result, this model is much closer to models of information acquisition, such as Dewatripont and Tirole (1999), where litigants' contribution comes from the effort the exert rather than the signals they send.

To study the potential for inefficient appeals, I first establish the necessary conditions for such inefficient appeals. As is apparent from the expression for $\Delta(\cdot)$ above, it is feasible to have equilibria where $x^*(s_J = \phi) = P$ or where $x^*(s_J = \phi) = D$. That is, depending on the parameters of the model, it is possible for the trial judge to rule in favor of either the plaintiff or the defendant whenever uninformed about the state of the world. This is an important observation because whenever the trial judge is uninformed, her prior belief $\pi > \frac{1}{2}$ indicates that she should rule in favor of the plaintiff if she is maximizing the probability of a correct decision. The fact that she sometimes has an incentive to rule for the defendant in this scenario implies that she sometimes rules *against* her belief about ω. I now characterize the conditions under which the trial judge rules against her information. First, low resource litigants have to be sufficiently bad at mounting appeals *or* cases have to be very certain. I refer to this as this condition C1.

> **Lemma 12.** [condition C1] If $\alpha_0 < \underline{\alpha}$ or if $\pi > \overline{\pi}$ (where $\overline{\pi}$ is defined in the proof), the trial judge always makes a judgment consistent with her information.

Now, suppose that C1 holds and that in equilibrium, $x^* = D$. By Lemma 11, this only occurs if (1) the trial judge learned that the state of the world indeed favored the defendant or (2) if the plaintiff will mount a weaker appeal than the defendant. I refer to (2) as condition C2.

> **Lemma 13.** [condition C2] If $x^* = D$ and $s_J \neq D$, then $(\alpha_P^*, \alpha_D^*) = (\alpha_0, \alpha_1)$.

The preceding lemmas demonstrate that there are circumstances where the trial judge ignores her information about the case when she makes her judgment. In particular, such an equilibrium exists whenever cases are sufficiently uncertain or the litigants are relatively bad at mounting effective appeals (C1), and whenever the plaintiff is relatively weaker than the defendant (C2).

> **Corollary 2.** If $\pi \in (\frac{1}{2}, \overline{\pi})$, and if $c_D < \bar{c}_D$ and $c_P > \bar{c}_P$, then there is an

equilibrium where $x(s_J = \phi) = D$.

For clarity, I label this equilibrium strategy profile $\mathcal{S}_D^*$ to indicate that its outcome is tilted in favor of the defendant. What are the implications of such an equilibrium? First, note that the courts' payoffs in that equilibrium, which I denote with a * are as follows.

$$u_A^* = \delta + (1-\delta)[(1-\pi) + \pi\alpha_0 r_P^*] - \frac{c_A}{2}(r_P^*)^2$$
$$u_J^* = \delta + (1-\delta)[(1-\pi) - \pi(\kappa-1)\alpha_0 r_P^*]$$

As a counter-factual exercise, suppose instead that the trial judge were to rule in favor of the plaintiff in the absence of definitive information, and the defendant does not mount a strong appeal. Label this strategy profile $\mathcal{S}'$. If the trial judge is sufficiently reversal averse, then the equilibrium $\mathcal{S}_D^*$ does not maximize the utility of the courts.

Proposition 10. The equilibrium $\mathcal{S}_D^*$ is inefficient.

There is an important caveat to Proposition 10. In the proof, I show that $\mathcal{S}_D^*$ is inefficient by demonstrating that an alternative strategy profile, $\mathcal{S}'$ yields a higher total payoff. That strategy profile, however, is not a *Pareto improvement* over $\mathcal{S}_D^*$, as clearly it makes one or more players worse off (otherwise there would be a profitable deviation from $\mathcal{S}_D^*$ and it would not be an equilibrium). For example, note that the defendant is worse off under $\mathcal{S}'$ than under $\mathcal{S}_D^*$ because he is more likely to lose at trial and is barred from making a strong appeal (his best response). However, because $\mathcal{S}_D^*$ is inefficient, all the players could be strictly better off under $\mathcal{S}'$ if the those players could be compensated for their reduced payoff.

How could we achieve more efficient equilibria? There could be at least two ways to do this. First, notice that the existence of $\mathcal{S}_D^*$ requires that the defendant be significantly more powerful (*i.e.*, have lower costs) than the plaintiff. Therefore, a corollary of Proposition 10 is that inefficient equilibria emerge when litigants are not evenly matched (and the direction of the asymmetry benefits the litigant with the "weaker" case as represented by the prior π). Institutional arrangements that reduce severe imbalances in litigant resources would eliminate inefficient equilibria of the form $\mathcal{S}_D^*$.

3.1 Banning Endogenous Appeals

It may be infeasible to significantly reduce the asymmetry between litigants. A second way to compensate for deviations from the inefficient equilibrium $\mathcal{S}_D^*$ could be to prevent endogenous appeals all together. To explore this possibility, I first establish a benchmark result. In the absence of endogenous appeals, the trial judge always rules in a manner consistent with her information, and the appellate court never reviews.

> **Proposition 11.** (No Appeals) Suppose that litigants are unable to mount appeals. Then, a perfect Bayesian equilibrium of the game:
>
> - J rules in favor of the plaintiff if and only if her belief is that ω is more likely to be P than D.
> - A never reviews cases.
> - Beliefs are pinned down by Bayes' rule and all information sets are reached in equilibrium.

I label this equilibrium $\mathcal{S}^B$. In this equilibrium, the trial judge always rules for the plaintiff when she is not successful at discovering ω. That is, in this equilibrium, $x^*(s_J = \phi) = P$ and the trial judge's judgment always matches her belief about the state of the world. The courts' payoffs in this benchmark equilibrium, which I denote with a superscript B, are as follows.

$$u_A^B = u_J^B = \delta + (1 - \delta)\pi$$

Given this simple game form, the equilibrium is efficient. Moreover, the proportion of cases resolved correctly is $\delta + (1 - \delta)\pi$. How does this compare to the players' payoffs in the full version of the model with endogenous appeals? The players are better off when the litigants are barred from making endogenous appeals than they are in the inefficient equilibrium $\mathcal{S}_D^*$.

> **Lemma 14.** Suppose $\kappa > \widehat{\kappa}$ or $c_A > \widehat{c}_A$ where $\widehat{\kappa}$ and $\widehat{c}_A$ are defined in the proof. Then, the players are collectively better off banning endogenous appeals, relative to the equilibrium $\mathcal{S}_D^*$.

As long as the trial judge is sufficiently reversal averse (or the appellate court's review costs are sufficiently high), Lemma 14 indicates that banning endogenous appeals would leave the players better off. This result provides a counterargument to more optimistic theories of endogenous appeals and underscores that certain, plausible institutional contexts may yield inefficient outcomes when advocates are able to acquire information to appeal decisions.

3.2 Social Optimality

The previous result provides a rationale for banning endogenous appeals, but that rationale depends on a particular notion of efficiency. In particular, when trial judge's reversal costs are sufficiently high, $\mathcal{S}^B$ is a more efficient equilibrium precisely because of the cost borne by the trial judge. However, it may be more desirable to focus on alternative notions of social optimality that prioritize how well the legal system functions. I examine this issue using two separate notions: (1) *Is the appellate court better off banning endogenous appeals than in the inefficient equilibrium?* and (2) *Are cases more likely to be resolved correctly if endogenous appeals are banned, as compared to the inefficient equilibrium?*

First, I analyze whether the appellate court only is better off under the inefficient equilibrium $\mathcal{S}_D^*$ or when endogenous appeals have been banned, $\mathcal{S}^B$. The justification for this notion of social optimality is that the appellate court's preferences encompass both a benefit from resolving cases correctly, but also a cost of reviewing the trial judge. This latter cost can be seen as wasteful given that the trial court has no *ex ante* reason to subvert outcomes, but is induced to do so *solely* because of endogenous appeals. Thus, the audit of the trial judge becomes a necessity, when it would not be in the absence of appeals. As long as c_A is sufficiently high, we see that the appellate court is worse off under $\mathcal{S}_D^*$ than banning endogenous appeals.

> **Lemma 15.** If the appellate court's review is sufficiently costly (high c_A), then the appellate court is better off banning endogenous review than in the inefficient equilibrium $\mathcal{S}_D^*$.

Next, I examine whether cases more likely to be resolved correctly with endogenous appeals banned, $\mathcal{S}^B$ than under the inefficient equilibrium $\mathcal{S}_D^*$. In the inefficient equilibrium $\mathcal{S}_D^*$, the probability that a case is resolved correctly is as follows.

$$\delta + (1-\delta)(1-\pi) + (1-\delta)\pi\alpha_0 r_P^*$$

Then, society is better off with banning endogenous appeals if:

$$\delta + (1-\delta)\pi > \delta + (1-\delta)(1-\pi) + (1-\delta)\pi\alpha_0 r_P^*$$

Substituting in equilibrium values and rearranging, this simplifies to:

$$2\pi - 1 > \frac{\pi^2\alpha_0^2}{c_A}$$

Given that $c_A > 0$, this condition only holds if c_A is sufficiently high.

> **Lemma 16.** If c_A is sufficiently high, then more cases are resolved correctly when endogenous appeals are banned than in the inefficient equilibrium.

Lemmas 15 and 16 each demonstrate that when the appellate court's review cost is sufficiently high, that outcomes can be better when endogenous appeals are banned. These results reinforce the argument that endogenous appeals can sometimes be welfare reducing. Moreover, to the extent that we see appellate courts constrained in their ability to review cases and the conditions for the inefficient equilibrium $\mathcal{S}_D^*$ are present, then cases may be resolved better overall *without* endogenous appeals.

4 Discussion

I have shown that there can be reductions in social welfare due to endogenous appeals. However, the *benefits* of endogenous appeals have been noted by others (*e.g.*, Dewatripont and Tirole 1999; Cameron and Kornhauser 2006; Shavell 2006; Cooter and Ulen 2012). In particular, endogenous appeals may (1) discipline a biased trial judge who may otherwise make incorrect decisions; or (2) provide important information about litigants' private information about their guilt or liability. Due to significant differences between these models and the model presented here, a direct comparison is beyond the scope of this paper.[41] However, to the extent that (1) trial judges dislike reversals, and (2) litigants engage in post-judgment information acquisition, the dynamics of this model will be present.

The rules of the U.S. federal court system suggest that the model is empirically relevant. The Federal Rules of Civil Procedure and the Federal Rules of Appellate Procedure structure the appeal process in such a way that both enables—and requires—losing parties to exert effort to learn whether or not the district judge made a (legal) error. However, the fact that the U.S. appellate system resembles the the model laid out in this paper does not automatically mean that inefficiency will occur. As Lemmas 12 and 13 make clear, additional requirements are necessary to ensure the existence of an inefficient equilibrium. In particular, if cases are sufficiently uncertain or litigants are sufficiently bad at mounting strong appeals, *and* there is a resource asymmetry between the plaintiff and defendant, then the model demonstrates how trial judges orient their decision making away from their own information about the merits of the case and toward powerful litigants.

[41] For example, in Cameron and Kornhauser (2006), there is no information acquisition as in this paper.

The basic argument underscores a strategic tradeoff facing trial judges who dislike being reversed, but still do prefer to "get it right." If there is a sufficiently large asymmetry in the capability of the two litigants, the trial judge is tempted to rule in favor of the powerful litigant in order to head off a possibly forceful appeal and potentially embarrassing reversal. Perversely, that judge would face no such temptation if (1) litigants were more evenly matched, or (2) litigants were banned from making appeals. Thus, if appellate courts are sufficiently burdened by their case loads, then the model suggests that more cases could be resolved correctly simply by disallowing endogenous appeals.

References

Alliance for Justice. 2014. *Broadening the Bench: Judicial Nominations and Professional Diversity.* Accessed August 15, 2015. `http://www.afj.org/broadening-the-bench-judicial-nominations-and-professional-diversity`.

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

Ashcroft v. Iqbal. 2009. 556 U.S. 662. LexisNexis Academic (November 25, 2016).

Ashworth, Scott, and Kenneth W. Shotts. 2011. "Challengers, Democratic Contestation, and Electoral Accountability." `http://ssrn.com/abstract=1901510`.

Bazemore v. Friday. 1986. 478 U.S. 385. LexisNexis Academic (August 31, 2015).

Beim, Deborah. 2016. "Learning in the Judicial Hierarchy." *Journal of Politics.* `Forthcoming`.

Bell Atlantic Corp. v. Twombly. 2007. 550 U.S. 544. LexisNexis Academic (November 25, 2016).

Blank v. Sullivan & Cromwell. 1975. 418 F. Supp 1. LexisNexis Academic (August 22, 2015).

Boyd, Christina L. 2015. "Representation on the Courts? The Effect of Trial Judges' Sex and Race."

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. 2005. "Decision Rules in a Judicial Hierarchy." *Journal of Institutional and Theoretical Economics* 161 (2): 264–292.

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,* 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.

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. 2012a. "A Theory of Opinion Writing in a Political Hierarchy." *Journal of Politics* 74 (2): 584–603.

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

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

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

Chin, Denny. 2012. "Summary Judgment in Employment Discrimination Cases: A Judge's Perspective." *New York Law School Law Review* 57:671–682.

Choi, Stephen J., Mitu Gulati, and Eric A. Posner. 2011. "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 Jonathan P. Kastellec. 2013. "The Supreme Court and Percolation in the Lower Courts: An Optimal Stopping Model." *Journal of Politics* 75 (1): 150–168.

Clermont, Kevin M., and Stewart J. Schwab. 2004. "How Employment Discrimination Plaintiffs Fare in Federal Court." *Journal of Empirical Legal Studies* 1 (2): 429–458.

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

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.

Dolan, Maura. 2013. *Prop. 8: 'I have done my part,' retired Judge Vaughn Walker says.* San Francisco. `http://articles.latimes.com/2013/mar/26/local/la-me-ln-prop-8-judge-vaughn-walker-20130326`.

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

Emons, Winand, and Claude Fluet. 2007. “Accuracy versus Falsification Costs: The Optimal Amount of Evidence under Different Procedures.” *Journal of Law, Economics, and Organization* 25 (1): 134–156.

Epp, Charles R. 1998. *The Rights Revolution: Lawyers, Activists, and Supreme Courts in Comparative Perspective.* Chicago: University of Chicago Press.

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.

Farhang, Sean. 2010. *The Litigation State: Public Regulation and Private Lawsuits in the U.S.* Princeton, NJ: Princeton University Press.

Farhang, Sean, and Gregory J. Wawro. 2004. “Institutional Dynamics on the U.S. Court of Appeals: Minority Representation Under Panel Decision Making.” *Journal of Law, Economics, and Organization* 20 (2): 299–330.

Fischman, Joshua B., and Max M. Schanzenbach. 2011. “Do Standards of Review Matter? The Case of Federal Criminal Sentencing.” *Journal of Legal Studies* 40 (2): 405–437.

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. 2013a. *Learning While Governing: Expertise and Accountability in the Executive Branch.* Chicago: University of Chicago Press.

Gailmard, Sean, and John W. Patty. 2013b. “Participation, Process, & Policy: The Informational Value of Politicized Judicial Review.”

Hazelwood School District v. United States. 1977. 433 U.S. 299. LexisNexis Academic (August 31, 2015).

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. 2015. "The Downsides of Dispassionate Judges: A Theory of Costly Law Enforcement in District Courts."

Ifill, Sherrilyn A. 2010. "Challenging Judge Walker." *The Root* (). http://www.theroot.com/articles/politics/2010/08/critics%7B%5C_%7Dsay%7B%5C_%7Dprop%7B%5C_%7D8%7B%5C_%7Djudge%7B%5C_%7Dis%7B%5C_%7Dbiased.html.

In re: Seroquel Products Liability Litigation. 2007. No. 6:06-md-01769-ACC-DAB. (M.D. Fla., August 21, 2007). http://www.flmd.uscourts.gov/MDL/documents/393.pdf.

Kagan, Robert A. 1991. "Adversarial Legalism and American Government." *Journal of Policy Analysis and Management* 10 (3): 369–406.

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

Kim, Pauline T. 2007. "Lower Court Discretion." *New York University Law Review* 82 (2): 383–442.

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.

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

Kornhauser, Lewis A. 1995. “Adjudication by a Resource-Constrained Team: Hierarchy and Precedent in a Judicial System.” *Southern California Law Review* 68:1605–1629.

Kornhauser, Lewis A., and Lawrence G. Sager. 1986. “Unpacking the Court.” *The Yale Law Journal* 96 (1): 82–117.

Landa, Dimitri, and Jeffrey R. Lax. 2009. “Legal Doctrine on Collegial Courts.” *Journal of Politics* 71 (3): 946–963.

Lax, Jeffrey R. 2007. “Constructing Legal Rules on Appellate Courts.” *American Political Science Review* 101 (3): 591–604.

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

Lax, Jeffrey R., and Charles M. Cameron. 2007. “Bargaining and Opinion Assignment on the US Supreme Court.” *Journal of Law, Economics, and Organization* 23 (2): 276–302.

Li, Victor. 2014. “Looking Back on Zubulake, 10 Years Later.” *ABA Journal* ().

Lithwick, Dahlia. 2010. “A Brilliant Ruling.” *Slate* (). `http://www.slate.com/articles/news%7B%5C_%7Dand%7B%5C_%7Dpolitics/jurisprudence/2010/08/a%7B%5C_%7Dbrilliant%7B%5C_%7Druling.html`.

Macfarlane, Katherine A. 2014. “The Danger of Nonrandom Case Assignment: How the Southern District of New York’s ”Related Cases” Rule Shaped Stop-and-Frisk Rulings.” *Michigan Journal of Race and Law* 19 (2): 199–246.

McDonnell Douglas Corp. v. Green. 1973. 414 U.S. 811. LexisNexis Academic (August 31, 2015).

McKoski, Raymond J. 2014. “Disqualifying Judges When Their Impartiality Might Reasonably Be Questioned: Moving Beyond a Failed Standard.” *Arizona Law Review* 56 (2): 411–478.

O’Scannlain, Diarmuid F. 2002. “Takings Clause Jurisprudence: Muddled, Perhaps; Judicial Activism, No.” *The Georgetown Journal of Law and Public Policy* 1:129–132.

Posner, Richard A. 2008. *How Judges Think.* Cambridge, MA: Harvard University Press.

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.

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

Rutherglen, George. 2007. *Employment Discrimination Law: Visions of Equality in Theory and Doctrine.* 2nd. New York: Foundation Press.

Schanzenbach, Max M., and Emerson H. Tiller. 2006. “Strategic Judging Under the U.S. Sentencing Guidelines: Positive Political Theory and Evidence.” *Journal of Law, Economics, and Organization* 23 (1): 24–56.

Segal, Jeffrey A., and Harold J. Spaeth. 1993. *The Supreme Court and the Attitudinal Model.* New York: Cambridge University Press.

Sen, Maya. 2015. "Is Justice Really Blind? Race and Reversal in US Courts." *The Journal of Legal Studies* 44 (S1): S187–S229.

Shavell, Steven. 2006. "The Appeals Process and Adjudicator Incentives." *Journal of Legal Studies* 35 (1): 1–29.

Spitzer, Matthew, and Eric Talley. 2013. "Left, Right, and Center: Strategic Information Acquisition and Diversity in Judicial Panels." *Journal of Law, Economics, and Organization* 29 (3): 638–680.

Stein, Sam. 2010. "Tony Perkins: Prop. 8 Judge Should Have Recused Himself Because Of His Sexuality." *The Huffington Post* (). `http://www.huffingtonpost.com/2010/08/08/tony-perkins-prop-8-judge%7B%5C_%7Dn%7B%5C_%7D674788.html`.

Stephenson, Matthew C. 2006. "A Costly Signaling Theory of Hard Look Theory." Cambridge, MA. `http://lsr.nellco.org/harvard%7B%5C_%7Dolin/539`.

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

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.

Toobin, Jeffrey. 2013. "Rights and Wrongs." *The New Yorker* ().

Totenberg, Nina. 2013. "Judge Who Struck Down Proposition 8 Knew Case Would Go Far." *National Public Radio* (). `http://www.npr.org/2013/06/29/196765535/judge-who-struck-down-proposition-8-knew-case-would-go-far`.

Turner, Ian R. 2015. "Working Smart and Hard? Agency Effort, Judicial Review, and Policy Precision."

United States v. Gagnon. 1985. 470 U.S. 522. LexisNexis Academic (November 26, 2016).

Wards Cove Packing Co. v. Atonio. 1989. 490 U.S. 642. LexisNexis Academic (August 31, 2015).

Weiser, Benjamin, and Joseph Goldstein. 2013. "Federal Panel Softens Tone on Judge It Removed From Stop-and-Frisk Case." *New York Times* ().

Wright, Charles Alan, and Andrew D. Leipold. 2008. *Federal Practice and Procedure.* 4th ed. Eagan, MN: Thomson-West.

Zubulake v. UBS Warburg. 2004. 229 F.R.D. 422. LexisNexis Academic (August 15, 2015).

Appendices

A Proofs

Proof of Lemma 1. The optimal effort follows directly from the first order condition of U_{NR} and by the fact that the second order condition $-c''(e)$ is negative. Because $c'(0) = 0$, $c''(e) > 0$ and $\lim_{e \to 1} c'(e) = \infty$, it follows that e^*_{NR} is strictly between zero and one. Finally, since $c'(\cdot), c''(\cdot) > 0$, as β increases, e^*_{NR} must increase for the equality to hold. $\square$

Proof of Lemma 2. Restrict attention to pure strategies, which may be pooling or separating. First, notice that A's sequentially rational review decision is to affirm if and only if its posterior belief that the judgment is consistent with its ideal point is greater than one half. Moreover, this implies that it is sequentially rational for A to affirm any judgment in favor of the plaintiff since its prior belief is that the plaintiff should prevail (and updating via Bayes's Rule makes A *more* confident that the plaintiff should prevail).

Pooling equilibria. In a pooling equilibrium, regardless of her information D either votes in favor of the defendant or the plaintiff. A's belief about ω is its prior and therefore, A's best response is to affirm pro-plaintiff and reverse pro-defendant judgments on the equilibrium path. In equilibrium, then, D's reversal aversion ensures that she always rules in favor of the plaintiff. Off the equilibrium path, A's reversal of pro-defendant decisions prevents deviations by D since a pro-defendant ruling is never upheld and simply imposes the reversal cost, k. Such off equilibrium behavior can be supported by belief that $\Pr(\omega < \widehat{\omega}_A | x = \text{defendant}) < \frac{1}{2}$.

If the district judge's equilibrium effort is sufficiently low, this subgame features a unique pooling equilibrium where the district judge rules in favor of the plaintiff independent of her information and the appellate court affirms judgments in favor of the plaintiff and reverses judgments in favor of the defendant. It is obvious why there is no pooling equilibrium that involves pro-defendant judgments on the equilibrium path. In such an equilibrium, certain types of D always has an incentive to deviate: if A reverses off the equilibrium path, deviations increase some D-types' payoffs by β, and if A affirms off the equilibrium path, deviations increase all D-types' payoffs by k.

It is less clear why there are no separating equilibria when e^* is sufficiently low. By contradiction, suppose there is some separation in an equilibrium, so that some types rule in favor of the plaintiff and some types rule in favor of the defendant. If A reverses pro-defendant and affirms pro-plaintiff judgments, then separation cannot be sustained: the types who rule in favor of the defendant are better off deviating. If instead A affirms pro-defendant rulings and reverses pro-plaintiff rulings, then again, separation cannot sustained: the types who rule in favor of the plaintiff are better off deviating. Finally, if A always reverses or always affirms, then the reversal cost becomes irrelevant for D's decision making and in equilibrium, D gets its preferred outcome. However, given the conflict of interest between A and D, A may actually have an incentive to deviate from such a strategy.

It is not optimal for the appellate court to let the district court do what it wants whenever the district court's effort is sufficiently low. Condition (2) is therefore a necessary condition for a pooling equilibrium.

Separating equilibria. In a separating equilibrium, some types adjudicate in favor of the plaintiff and others rule in favor of the defendant.

First, we show that in any separating equilibrium, reversals by A do not occur on the equilibrium path. By contradiction, suppose that they did. Then, (1) A either reverses after one decision but not the other or (2) A always reverses. For (2) to be an equilibrium, it must be that A prefers the opposite of what D decides in any equilibrium, and would never reverse if each D-type made the opposite decision. D has a profitable deviation that eliminates its cost k but ensures the same outcome. Finally, note that (1) implies that D-types that get reversed have a profitable deviation to issue the judgment that isn't reversed.

Therefore, any separating equilibrium involves A affirming all decisions on the equilibrium path. Then, D's best response is to rule in favor of the plaintiff if $s_D = \omega > \widehat{\omega}_D$, and the defendant otherwise.

However, if the condition fails, then there is a separating equilibrium where the appellate court defers completely to the district court: it always affirms the judgment of the district judge.[42] If the two courts' ideal points are far away from each other—formally, if $\frac{F_A}{F_D} < \frac{1}{2}$—then condition (2) holds for all $e_R^* \in [0, 1]$ and every equilibrium of the simple review regime entails D pooling on a pro-plaintiff decision. $\square$

Proof of Lemma 3. ($\Rightarrow$) By contradiction, suppose D exerts strictly positive effort and that this implies that either $\frac{F_A}{F_D} < \frac{1}{2}$ or $\beta < \tilde{\beta}$. If $\frac{F_A}{F_D} < \frac{1}{2}$, then by Lemma 2, D rules in favor of the plaintiff independent of its information, and it is optimal to set $e_R^* = 0$. If $\beta < \tilde{\beta}$, then the maximum effort that is optimal for D is $e_R^* < \tilde{e}_R$ and condition (2) holds. Again, by Lemma 2, D rules in favor of the plaintiff independent of its information, and it is optimal to set $e_R^* = 0$. Therefore, $e_R^* > 0 \Rightarrow \frac{F_A}{F_D} > \frac{1}{2}$ and $\beta < \tilde{\beta}$. ($\Leftarrow$) By contradiction, suppose $\frac{F_A}{F_D} > \frac{1}{2}$ and $\beta > \tilde{\beta}$, and that this implies that D's effort is zero. If $e_R^* = 0$, then Lemma 2 implies that D always rules in favor of the plaintiff regardless of her information. Her expected payoff is $\beta(1 - F_D)$. If, however, she exerts strictly positive effort, her best response ignoring A's strategy would be to set e_R^* so that $c'(e_R^*) = \beta(1-F_D)$, where $e_R^* > \tilde{e}_R$ by the fact that $\beta > \tilde{\beta}$. Thus, condition (2) fails and by Lemma 2, A defers to D's judgment and D adjudicates in her preferred manner. D's expected payoff is then $e_R^*\beta + (1 - e_R^*)F_D\beta = \beta(F_D + e_R^*(1 - F_D))$. Since $F_D > \frac{1}{2}$ and $e_R^* \geq 0$, this payoff is strictly higher than the payoff D receives with no effort. Therefore, $e_R^* > 0 \Leftrightarrow \frac{F_A}{F_D} > \frac{1}{2}$ and $\beta < \tilde{\beta}$. $\square$

Proof for Proposition 1. In text. $\square$

[42]Technically, there are two symmetric equilibria: one where A always affirms, as described in the text and one where A always reverses. The strategic calculation of D is the same in both equilibria, as are the equilibrium outcomes. The always-reverse equilibrium is inefficient and substantively uninteresting. If A incurred even a small cost to reverse, the always-reverse equilibrium vanishes.

Definition 4. Let $\underline{F_A}$ be defined as:

$$\underline{F_A} \equiv \frac{1}{4} + \frac{1}{4}\sqrt{-1 + 4F_D - 2F_D^2}$$

And conversely, let $\underline{F_D}$ be defined as:

$$\underline{F_D} \equiv \begin{cases} 1 - 2\sqrt{F_A - 2F_A^2} & \text{if } F_A > \underline{F_A} \\ \frac{1}{2} & \text{otherwise} \end{cases}$$

Then, let $\overline{e}, \underline{e} \in [0,1]$ be defined as follows:

$$\underline{e} \equiv \begin{cases} \dfrac{1 - F_D - \sqrt{1 - 4F_A + 8F_A^2 - 2F_D + F_D^2}}{2F_A} & \text{if } F_D \leq \underline{F_D} \\ 1 & \text{otherwise} \end{cases}$$

$$\overline{e} \equiv \begin{cases} \dfrac{1 - F_D + \sqrt{1 - 4F_A + 8F_A^2 - 2F_D + F_D^2}}{2F_A} & \text{if } F_D \leq \underline{F_D} \\ 1 & \text{otherwise} \end{cases}$$

And let $\overline{\beta}, \underline{\beta} \geq 0$ be defined as follows:

$$(1 - F_D)\underline{\beta} = c'(\underline{e}) \qquad\qquad (1 - F_D)\overline{\beta} = c'(\overline{e})$$

where $\overline{\beta} = \infty$ if $\overline{e} = 1$ and $\underline{\beta} = \infty$ if $\underline{e} = 1$.

Lemma 17. In any equilibrium with endogenous information, the appellate court reverses a pro-defendant judgment if and only if (1) $m = \phi$ and either $F_A < \underline{F_A}$ or $\beta \notin (\underline{\beta}, \overline{\beta})$, or (2) if $m = \omega > \widehat{\omega}_A$. The appellate court reverses a pro-plaintiff judgment if and only if $m = \omega < \widehat{\omega}_A$.

Proof for Lemma 17. Upon receiving a hard information signal from D, is straight forward to observe that it is a best response for A to reverse if and only if $x =$ defendant and $\omega > \widehat{\omega}_A$ or $x =$ plaintiff and $\omega < \widehat{\omega}_A$. It remains to be shown that reversing a pro-defendant judgment is a best response if and only if $m = \phi$ and either $F_A < \underline{F_A}$ or $\beta \notin (\underline{\beta}, \overline{\beta})$.

Let $x =$ defendant.

($\Leftarrow$) First, suppose $F_A < \underline{F_A}$. A's posterior belief that x is decided correctly is

$$\Pr(\omega < \widehat{\omega}_A | x = \text{defendant}, m = \phi) = \frac{F_A(1-e^2)}{1 - e(1 - F_D + eF_A)}$$

Since $F_A < \underline{F_A}$, then $\Pr(\omega < \widehat{\omega}_A | x = \text{defendant}, m = \phi) < \frac{1}{2}$ and A believes that the judgment is less likely to be correct than incorrect. Given this belief, it is a best response to reverse. Second, suppose $F_A \geq \underline{F_A}$ and $\beta \notin (\underline{\beta}, \overline{\beta})$. Then either $\beta < \underline{\beta}$ or $\beta > \overline{\beta}$ and by Definition 4, $e < \underline{e}$ or $e > \overline{e}$. Using the quadratic equation, it follows that in order for condition (3) to hold, $e \in [\underline{e}, \overline{e}]$, which is a contradiction. Since condition (3) fails, A's does not defer to D and reverses a pro-defendant decision.

($\Rightarrow$) Suppose that A reverses a pro-defendant judgment. Then, by sequential rationality, A must believe that $\Pr(\omega < \widehat{\omega}_A | x = \text{defendant}, m = \phi) < \frac{1}{2}$. Simplifying this yields the converse of condition (3) in text:

$$e(1 - F_D - eF_A) < 1 - 2F_A$$

This always holds if $F_A < \underline{F_A}$. Otherwise, the quadratic equation indicates that $e < \underline{e}$ or $e > \overline{e}$, which in turn implies $\beta < \underline{\beta}$ or $\beta > \overline{\beta}$. □

Lemma 18. In any equilibrium with endogenous information, the district judge's judgment is as follows:

- If the appellate court is non-deferential, she rules in favor of the defendant when she has hard information that $\omega < \widehat{\omega}_A$, and rules in favor of the plaintiff otherwise.
- If the appellate court is deferential, she rules in favor of the plaintiff when she learns that $\omega > \widehat{\omega}_D$ (regardless of whether it is hard information) and rules in favor of the defendant otherwise.

Proof for Lemma 18. We show that D has no incentive to deviate. Clearly D has no incentive to deviate by ruling in favor of the plaintiff when the case facts are sufficiently low that both courts agree that the defendant should prevail. Then, consider a potential deviation where D rules in favor of the defendant when the equilibrium strategy indicates she should rule in favor of the plaintiff. Such a deviation leads to a reversal and an outcome in favor of the plaintiff. There is no benefit from the deviation, but D pays the cost of reversal. $\square$

Proof of Lemma 4. First, note that $\hat{\beta}$ is defined as

$$\hat{\beta}(1 - F_D) = c'\left(\frac{1 - F_D}{2F_A}\right)$$

where $\frac{1-F_D}{2F_A}$ is the level of effort where the marginal benefit from deference equals the marginal benefit from non-deference. Therefore, for all $\beta > \hat{\beta}$ it follows that effort under non-deference is larger than effort under deference, given that $c'()$ is strictly convex. Moreover, for all $\beta < \hat{\beta}$ it follows that effort under deference is larger than effort under deference, given that $c'()$ is strictly convex. $\square$

Proof of Proposition 2. First note that there is no $e \in [0, 1]$ such that condition (3) holds if

$$F_A < \frac{1}{4} + \frac{\sqrt{-1 + 4F_D - 2F_D^2}}{4}$$

Therefore, there is no deference if F_A is sufficiently low. Moreover, recall that deference is suboptimal in the case of simple review if

$$F_A < \frac{1}{2}F_D$$

Since

$$\frac{1}{4} + \frac{\sqrt{-1 + 4F_D - 2F_D^2}}{4} > \frac{1}{2}F_D$$

it follows that deference under endogenous information requires that F_A be strictly higher than deference under simple review. This suffices to demonstrate that there is strictly less deference under endogenous information than under simple review. □

Lemma 19. Under review with endogenous information, the district judge exerts strictly positive effort.

Proof for Lemma 19. In text. □

Lemma 20. Strong activists exert weakly greater effort under endogenous information than under simple review. Weak activists exert weakly less effort under endogenous information than under simple review if and only if they are sufficiently activist ($\beta > \tilde{\beta}$).

Proof of Lemma 20. Now, we consider four cases.

(1) Activist judge and $\hat{\beta} < \beta < \tilde{\beta}$. Then she faced non-deference under simple review, and exerted zero effort. By Lemma 19, she exerts positive effort under endogenous information.

(2) Activist judge and and $\beta \geq \hat{\beta} > \tilde{\beta}$. Then the activist judge faced deference under simple review. If she faces deference under endogenous information, then she exerts the same level of effort as under simple review. If she faces non-deference under endogenous information, then by Lemma 4 she exerts strictly greater effort.

(3) Non-activist judge and $\beta < \hat{\beta} \leq \tilde{\beta}$. Then the weak activist judge exerted zero effort under simple review and strictly positive effort under endogenous information (again, by Lemma 19).

(4) Non-activist judge and $\tilde{\beta} < \beta < \hat{\beta}$. Then the weak activist judge exerted strictly positive effort under simple review and strictly positive effort under endogenous information. However, by Lemma 4, she exerts strictly lower effort under endogenous information than under simple review. □

Proof of Proposition 3. Let $F_A < \underline{F_A}$ or $\beta > \overline{\beta}$. Then, by Lemma 17 the appellate court adopts a non-deferential review strategy in an endogenous information equilibrium, and by

Lemmas 4 and 20, equilibrium effort is strictly higher. Because the appellate court gets its preferred disposition more often, it is strictly better off.

Next, let $F_A > \underline{F_A}$ and $\underline{\beta} < \beta < \overline{\beta}$. Then, by Lemma 17 the appellate court adopts a deferential review strategy in equilibrium, and is no better off than under simple review.

Finally, let $F_A > \underline{F_A}$ and $\tilde{\beta} < \beta < \overline{\beta}$. Then by Lemma 20, equilibrium effort is lower, and the appellate court is worse off if and only if

$$\underbrace{F_A + (1 - F_D)e^*_R}_{\text{simple review}} > \underbrace{F_D + (1 - F_D)e^*_{\bar{R}}}_{\text{endogenous information}}$$

This reduces to:

$$e^*_R - e^*_{\bar{R}} > \frac{F_D - F_A}{1 - F_D}$$

□

Proof for Proposition 4. Let $F_A < \underline{F_A}$. Then, under endogenous review, there is no deference, whereas under simple review, A deferred for all $\beta > \tilde{\beta}$. Suppose $\min\{\tilde{\beta}, \hat{\beta}\} < \infty$. Then for all $\beta > \min\{\tilde{\beta}, \hat{\beta}\}$, D faces deference under simple review and non-deference under endogenous information. Accordingly, D gets her preferred outcome less often, and is induced to exert more or less effort than her unconstrained decision problem. She is strictly worse off under endogenous information than simple review, and her best response is $y =$ pre-trial.

Next, let $F_A > \underline{F_A}$. Then, by Lemma 17, the appellate court does not defer to judges in the endogenous information environment if $\beta < \beta < \underline{\beta}$ or if $\beta > \overline{\beta}$. However, by Lemma 2 and the definition of $\tilde{\beta}$, A defers to these kinds of judges under simple review. Again, because D gets her preferred outcome less often, and is induced to exert more or less effort than her unconstrained decision problem, she is strictly worse off under endogenous information than simple review. Her best response is $y =$ pre-trial. □

Lemma 21. Let $\hat{x}$ be J's optimal judgment under the assumption that A never reverses in the absence of information. In the absence of hard information, it is optimal for A to affirm $\hat{x}$ if and only if $\max\{\tilde{\pi}_0, \tilde{\pi}_1\} < \dot{\pi}$, where $\dot{\pi}$ is defined in the proof.

Proof of Lemma 21. ($\Rightarrow$) Suppose A's belief is not degenerate (which occurs after history $(x^*, s_J = m_L = \phi)$), and that it affirms $\hat{x}$ in equilibrium. Affirming is a best response to the following beliefs:

$$\underbrace{\frac{\pi}{\pi + (1-\pi)(1-a_D^*)}}_{\mu_A(x^*=P, s_J=m_L=\phi)} \geq \frac{1}{2} \qquad\qquad \underbrace{\frac{\pi(1-a_P^*)}{\pi(1-a_P^*) + 1 - \pi}}_{\mu_A(x^*=D, s_J=m_L=\phi)} \leq \frac{1}{2}$$

The left condition holds for all $\pi > \frac{1}{2}$. Since a type $t = 0$ judge only rules in favor of the plaintiff for some values of $\pi > \frac{1}{2}$, the left condition always holds for a $t = 0$ judge. However, for sufficiently high β, a $t = 1$ judge may rule $x^* = P$ for some $\pi < \frac{1}{2}$. Therefore, the condition holds for $t = 1$ judge if and only if $\tilde{\pi}_1 > \hat{\pi}$, where $\hat{\pi} \equiv \sqrt{(c_D - 1)c_D} - (c_D - 1) < \frac{1}{2}$ is the value of π where the left condition holds with equality. Substituting $\tilde{\pi}_1$, it is straightforward to verify that the condition always holds.

Define $\dot{\pi} \in (0, 1)$ to be the value of π such that the right condition holds with equality. Then, A's belief is consistent if and only if $\pi \leq c_P - \sqrt{(c_P - 1)c_P} \equiv \dot{\pi}$. Since type-$t$ J's equilibrium decision involves $x^* = D$ for all $\pi < \tilde{\pi}_t$, as long as $\max\{\tilde{\pi}_0, \tilde{\pi}_1\} < \dot{\pi}$, A does not reverse J's decision. Figure 16 illustrates the logic.

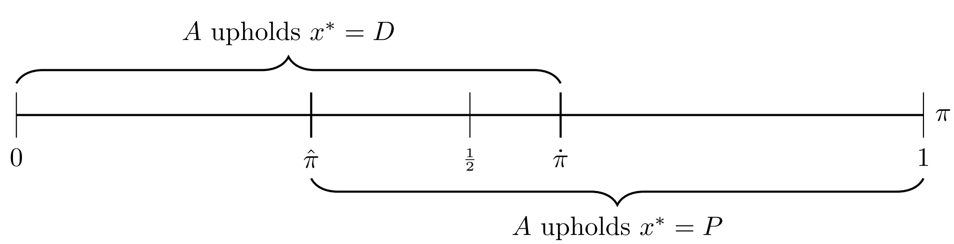


Figure 16: A's optimal affirmances

($\Leftarrow$) Suppose $\max\{\tilde{\pi}_0, \tilde{\pi}_1\} < \dot{\pi}$. Since A upholds a decision $x^* = D$, then for all $\pi < \min\{\tilde{\pi}_0, \tilde{\pi}_1\}$, A affirms $\hat{x} = D$. For all $\pi \in \tilde{\pi}_t$, a type-t's judgment $\hat{x} = P$. Since $\min\{\tilde{\pi}_0, \tilde{\pi}_1\} > \hat{\pi}$ (by logic above), A upholds all $\hat{x} = P$. $\square$

Proof of Proposition 5. Because A's utility is maximized when the final judgment is consistent with ω, it is easy to see that it reverses cases that it believes to be incorrect. Moreover, A's beliefs upon learning ω are straightforward. However, if $s_J = m_L = \phi$, so that A is uninformed, its belief is formed by Bayes' rule.

Since L pays a cost ε_L for revealing information that undermines its case, then $m_L^* = \omega$ if and only if $s_L = \omega = L$. Anticipating A's reversal strategy and L's equilibrium message, J minimizes the chance of reversal in equilibrium by making a judgment according to equation (8), following the logic in the text and Lemma 21. Notice that this judgment is constrained by appellate review when $\tilde{\pi}_t < \dot{\pi}$.

Since L gets a reversal only when $s_J = \phi$ and $m_L^* = \omega = L$, L exerts effort according to equation (5) if $s_J = \phi$, and exerts zero effort otherwise.

J's effort in equation (10) is sequentially rational because it is the effort level that maximizes the *ex ante* expected utility of J given in equation equation (9). $\square$

First, I restate Proposition 6 formally:

> **Proposition 6′.** If $c_P - c_D \leq 1$ and $k > \hat{k}$ (as defined by equation 15 below), then $|\frac{1}{2} - \tilde{\pi}_0| \geq |\frac{1}{2} - \tilde{\pi}_1|$.

Proof of Proposition 6. Suppose that $c_P - c_D \leq 1$ and that $k > \hat{k}$, where $\hat{k}$ is defined below.

To prove the claim, I consider two cases: $\beta \leq k$ and $\beta > k$. First, I prove that this holds for $\beta \leq k$ by directly showing that $\frac{1}{2} \leq \tilde{\pi}_1 < \tilde{\pi}_0$.

- $\tilde{\pi}_0 > \frac{1}{2}$ if and only if:

$$\frac{\sqrt{c_D c_P} - c_P}{c_D - c_P} > \frac{1}{2}$$

This condition holds if $c_D < c_P$, which is true by assumption.

- $\tilde{\pi}_1 > \frac{1}{2}$ if and only if:

$$\frac{-c_P(\beta(c_D-1)+k)+\sqrt{c_P(\beta(c_D-1)+k)c_D(\beta(c_P-1)+k)}}{(c_P-c_D)(\beta-k)} > \frac{1}{2}$$

This condition holds if $c_D < c_P$, which is true by assumption.

- $\tilde{\pi}_0 > \tilde{\pi}_1$ if and only if:

$$\frac{\sqrt{c_D c_P}-c_P}{c_D-c_P} > \frac{-c_P(\beta(c_D-1)+k)+\sqrt{c_P(\beta(c_D-1)+k)c_D(\beta(c_P-1)+k)}}{(c_P-c_D)(\beta-k)}$$

This condition holds if $c_D < c_P$, which is true by assumption.

Now, I prove the claim when $\beta > k$. Using the definitions of $\tilde{\pi}_0$ and $\tilde{\pi}_1$, to prove the claim, I show that the following holds.

$$\left|\frac{1}{2}-\frac{c_P-\sqrt{c_D c_P}}{c_P-c_D}\right| > \left|\frac{1}{2}-\frac{-c_P(\beta(c_D-1)+k)+\sqrt{c_P(\beta(c_D-1)+k)c_D(\beta(c_P-1)+k)}}{(c_P-c_D)(\beta-k)}\right|$$

The argument of the absolute value function on the left hand side is negative, whereas the argument of the absolute value function on the left hand side is positive, so the condition can be rewritten as:

$$\frac{c_P-\sqrt{c_D c_P}}{c_P-c_D}+\frac{-c_P(\beta(c_D-1)+k)+\sqrt{c_P(\beta(c_D-1)+k)c_D(\beta(c_P-1)+k)}}{(c_P-c_D)(\beta-k)} > 1$$

This reduces to:

$$\frac{c_P(\beta-k)-(\beta-k)\sqrt{c_D c_P}-c_P(\beta(c_D-1)+k)+\sqrt{c_P(\beta(c_D-1)+k)c_D(\beta(c_P-1)+k)}}{(c_P-c_D)(\beta-k)} > 1$$

This condition is satisfied if one of the following conditions hold:

$$\left[c_P > 2 \text{ and } c_D \geq \frac{c_P}{c_P - 1}\right] \quad \text{or} \quad \left[\beta < k\left(\frac{c_D + c_P}{c_D + c_P - c_D c_P}\right)\right]$$

Define $\hat{k}$ to be the minimum k that ensures one of these conditions hold. Then:

$$\hat{k} = \begin{cases} 0 & \text{if } c_P > 2 \text{ and } c_D \geq \dfrac{c_P}{c_P - 1} \\ \beta\left(\dfrac{c_D + c_P - c_D c_P}{c_D + c_P}\right) & \text{otherwise} \end{cases} \tag{15}$$

Finally, we check that A's equilibrium reversal strategy does not constrain x^*. By Lemma 21 and the fact that $\tilde{\pi}_0 > \tilde{\pi}_1$, the proposition holds if:

$$\tilde{\pi}_0 \leq \dot{\pi} \Leftrightarrow \frac{c_P \pm \sqrt{c_D c_P}}{c_P - c_D} \leq c_P - \sqrt{(c_P - 1)c_P}$$

This reduces to $c_P - c_D \leq 1$ $(= b)$, thus completing the proof.[43] □

Lemma 22. The equilibrium effort of J has the following characteristics:

- $e^*(\pi)$ is strictly increasing on the interval $[0, \min\{\tilde{\pi}_t, \dot{\pi}\})$ and strictly decreasing on the interval $(\min\{\tilde{\pi}_t, \dot{\pi}\}, 1]$.
- J's equilibrium effort is strictly increasing in the personal benefit, β and the cost of reversal, k; and strictly decreasing in the cost of effort c_J.

Proof of Lemma 22. First notice that Proposition 5 establishes that $x^*_\phi = D$ for all $\pi \leq \min\{\tilde{\pi}_t, \dot{\pi}\}$ and $x^*_\phi = P$ for all $\pi > \min\{\tilde{\pi}_t, \dot{\pi}\}$. Then, first and second derivatives of equation (10) with respect to π are given by:

[43] This is more restrictive than what is necessary. For example, if $k > \beta$, then the necessary condition is $\tilde{\pi}_1 < \dot{\pi}$. However, the condition in the proof ensures existence for all $\beta, k > 0$. If it is violated, we require $[\hat{\beta} > \beta > k$ and $c_D < \frac{c_P^2}{2c_P - 1}]$ or $[\beta > \hat{\beta} > k$ and $c_D \geq \frac{c_P^2}{2c_P - 1}]$, where $\hat{\beta} \equiv \left(\frac{c_D - c_D c_P + c_P^2}{c_D - 2c_D c_P + c_P^2}\right) k > k$

$$\frac{\partial e^*}{\partial \pi} = \begin{cases} \dfrac{c_P\beta t + 2\pi(k-\beta t)}{c_J c_P} \\ -\dfrac{c_D\beta t + 2(1-\pi)(k-\beta t)}{c_J c_D} \end{cases} \quad \text{and} \quad \frac{\partial (e^*)^2}{\partial^2 \pi} = \begin{cases} \dfrac{2(k-\beta t)}{c_J c_P} & \text{if } \pi \le \min\{\tilde{\pi}_t, \dot{\pi}\} \\ \dfrac{2(k-\beta t)}{c_J c_D} & \text{if } \pi > \min\{\tilde{\pi}_t, \dot{\pi}\} \end{cases}$$

If $t = 0$ or if $k \ge \beta$, then it is apparent from inspection of the first derivative that the equilibrium effort is strictly increasing when $\pi \le \min\{\tilde{\pi}_t, \dot{\pi}\}$ and strictly decreasing when $\pi > \min\{\tilde{\pi}_t, \dot{\pi}\}$. If $t = 1$ and $\beta > k$, then to prove the claim, it suffices to show that $c_P\beta + 2\pi(k-\beta) > 0$ for all $\pi \le \min\{\tilde{\pi}_1, \dot{\pi}\}$ and $-c_D\beta - 2(1-\pi)(k-\beta) < 0$ for all $\pi > \min\{\tilde{\pi}_1, \dot{\pi}\}$. Thus, the proof proceeds by showing that

$$\tilde{\pi}_1 < \frac{c_P\beta}{2(\beta-k)} \quad \text{and} \quad \frac{2(\beta-k) - c_D\beta}{2(\beta-k)} < \dot{\pi}$$

Using Lemma 21 and equation (24) in Appendix C, these conditions hold. The final set of comparative statics in the second bullet are easily obtained by inspecting equation (10). $\square$

Proof of Lemma 5. To prove this result, I show that the equilibrium effort of type $t = 1$ is greater than the equilibrium effort of type $t = 0$ for all $\pi \in [0, 1]$. Since $c_P > c_D$, there are three cases to consider: $\pi < \tilde{\pi}_1$, $\pi \in [\tilde{\pi}_1, \tilde{\pi}_0]$, $\pi > \tilde{\pi}_0$. The first and third cases involve both types adjudicating in the same way: in favor of the defendant and the plaintiff respectively. Suppose $\beta > 0$ and Assumption 4 hold. It is straightforward to see from inspection of equation (10) that a $t = 0$ judge exerts strictly less effort than a $t = 1$ judge when they have the same predisposition.

Now consider the second case, $\pi \in [\tilde{\pi}_1, \tilde{\pi}_0]$. By Lemma 22, $e_1^*(\pi)$ is strictly decreasing and $e_0^*(\pi)$ is strictly increasing for all $\pi \in (\tilde{\pi}_1, \tilde{\pi}_0)$. Then to show that $e_1^*(\pi) > e_0^*(\pi)$ in this interval, it suffices to show that $e_1^*(\dot{\pi}) > e_0^*(\dot{\pi})$:

$$\frac{1}{c_J c_D}\left[(1-\dot{\pi})(c_D-(1-\dot{\pi}))\beta t+(1-\dot{\pi})^2 k\right] > \frac{k(1-\dot{\pi})^2}{c_J c_D}$$

The condition holds if $c_D > 1$, thus completing the proof. □

Next, we formally restate Proposition 7.

Proposition 7′. $\xi(x_\phi^*|t=1) > \xi(x_\phi^*|t=0)$.

Proof of Proposition 7. Since $c_P > c_D$, there are three cases to consider: (I) $\pi < \min(\tilde{\pi}_0, \tilde{\pi}_1)$, (II) $\pi > \max(\tilde{\pi}_0, \tilde{\pi}_1)$ and (III) $\pi \in [\tilde{\pi}_1, \tilde{\pi}_0]$. I show that the *ex ante* probability of a correct decision is greater for a $t=1$ than for a $t=0$ judge in each case.

First, define $\xi(x_\phi^*|t)$ by substituting in the equilibrium values:

$$\xi(x_\phi = P|t) = \pi + \frac{(1-\pi)^2}{c_D} + \frac{(1-\pi)^2(c_D-(1-\pi))(\beta t(c_D-(1-\pi))+k(1-\pi))}{c_D^2 c_J} \tag{16}$$

Similarly,

$$\xi(x_\phi = D|t) = 1-\pi+\frac{\pi^2}{c_P}+\frac{\pi^2(c_P-\pi)(\beta t(c_P-\pi)+k\pi)}{c_P^2 c_J} \tag{17}$$

Cases (I) and (II). From equations (16) and (17), it is apparent that $\xi(x_\phi^*|t=1)$ is strictly greater than $\xi(x_\phi^*|t=0)$ since $(1-\pi)^2(c_D-(1-\pi)) > 0$ and $\pi^2(c_P-\pi) > 0$.

Case (III). By contradiction, suppose that $\xi(x_\phi = P|t=1)$ were weakly less than $\xi(x_\phi = D|t=0)$ at some point in the interval $[\tilde{\pi}_1, \tilde{\pi}_0]$. Then, by the definition of $\xi(\cdot)$, this implies the *ex ante* probability that $y^* = \omega$ is lower for the $t=1$ judge than for the $t=0$ judge at such a point. Using the fact that, when uninformed, a $t=0$ judge rules for the defendant and a $t = 1$ judge rules for the plaintiff, we can rewrite the *ex ante* probability as follows:

$$\xi(x^*_\phi|t) = \begin{cases} e^*(x^*_\phi, t) + [1 - e^*(x^*_\phi, t)][1 - \pi(1 - a^*_P)] & \text{if } t = 0 \\ e^*(x^*_\phi, t) + [1 - e^*(x^*_\phi, t)][1 - (1 - \pi)(1 - a^*_D)] & \text{if } t = 1 \end{cases}$$

By Lemma 5, $e^*(x^*_\phi = P, t = 1) > e^*(x^*_\phi = D, t = 0)$. In order for $\xi(x_\phi = P|t = 1) \leq \xi(x_\phi = D|t = 0)$ for some $\pi \in [\tilde{\pi}_1, \tilde{\pi}_0]$ as assumed, it must be that:

$$1 - \pi(1 - a^*_P) > 1 - (1 - \pi)(1 - a^*_D)$$

Substituting equilibrium values and simplifying yields:

$$1 > \frac{(1 - \pi)^2}{c_D} - \frac{\pi^2}{c_P} + 2\pi$$

The right hand side is increasing in π, and thus is at its smallest when $\tilde{\pi}_1$:

$$1 > \frac{\beta^2(c_P - c_D) + \beta k(2c_Dc_P + c_D - 3c_P) - 2k\left(\sqrt{c_Dc_P(\beta(c_D - 1) + k)(\beta(c_P - 1) + k)} - c_Pk\right)}{(\beta - k)^2(c_P - c_D)}$$

The right hand side is strictly greater than one and the condition fails, a contradiction. Therefore, there exists no $\pi \in [\tilde{\pi}_1, \tilde{\pi}_0]$ such that $\xi(x_\phi = P|t = 1) \leq \xi(x_\phi = D|t = 0)$. □

Proof of Lemma 6. First, in all subgames where the uncertainty is resolved for A or L, the equilibrium behavior is as in Proposition 5. Otherwise, there are two kinds of (uninformed) equilibria, one where J rules in favor of P and one where J rules in favor of D.

Candidate equilibrium #1: P-Equilibrium. Suppose A affirms decisions when uninformed (on and off the equilibrium path) and that J does not conceal information in favor of D.

Given $x^* = P$, e^* is the solution to the following maximization problem:

$$\max_{e} \quad e\pi\delta + e\varepsilon_J + (1-e)(1 - a_D^*(1-\pi))\delta - ka_D^*(1-\pi) - \frac{c_J}{2}e^2$$

This yields an optimal effort level:

$$e^* = \frac{\delta\pi + \varepsilon_J - \delta(1 - a_D^*(1-\pi))}{c_J} < 0$$

Since ε_J is small (see section 5), e^* is negative. Moreover since $e \in [0,1]$, we have a corner solution. This allows us to pin down a_D and a_P, which are the same as in Proposition 5.

Next, we check that $x^* = P$ is optimal. Since J exerts no effort, $s_J = \phi$ and thus $x^* = P$ is an equilibrium best response if:

$$U_J(x^* = P) = \delta\Big((1 - a_D) + a_D\pi\Big) - ka_D(1-\pi) > (\delta - k)a_P\pi = U_J(x' = D)$$

which reduces to

$$\delta > \left(\frac{a_D^*(1-\pi) - a_P^*\pi}{1 - a_D^*(1-\pi) - a_P^*\pi}\right)k \tag{18}$$

Given that $\delta > 0$, equation (18) becomes $\delta > \hat{\delta}$, where

$$\hat{\delta} \equiv \max\left\{0, \left(\frac{a_D^*(1-\pi) - a_P^*\pi}{1 - a_D^*(1-\pi) - a_P^*\pi}\right)k\right\}$$

Because of ε_L and ε_J, it is straightforward to see that the equilibrium messages of L and J are

$$m_J^* = \begin{cases} \omega & \text{if } s_J = \omega \\ \phi & \text{otherwise} \end{cases} \qquad\qquad m_L^* = \begin{cases} L & \text{if } s_L = L \\ \phi & \text{otherwise} \end{cases}$$

Next, we verify that A's equilibrium beliefs are consistent. At public history ($x^* = P, m_J^* = m_D^* = \phi$), $r^* = 0$ is a best response if and only if

$$\frac{\pi(1-e^*)}{\pi(1-e^*) + (1-\pi)(1-a_D^*)} \geq \frac{1}{2}$$

Substituting in $a_D^* = \frac{1-\pi}{c_D}$ and $e^* = 0$

$$0 \geq (c_D - 1)(1-2\pi) - \pi^2$$

This reduces to:

$$\pi \geq \hat{\pi} \equiv \sqrt{(c_D - 1)c_D} + 1 - c_D$$

Now, suppose that A reverses off equilibrium path deviations to $x' = D$. Then off the equilibrium path, $a'_P = 0$, and J has no incentive to deviate from the equilibrium for all $\delta > 0$.

Finally, we consider the off equilibrium beliefs of A. We assume that a deviation to $x' = D$ (with no information) leads A to make an inference that J is uninformed since otherwise, J would have revealed her information. (This inference is correct in light of the fact that $e^* = 0$.) If A affirms, its posterior belief off the equilibrium path is thus:

$$\mu^{\text{off}} = \frac{\pi(1-a'_P)}{\pi(1-a'_P) + (1-\pi)}$$

For the belief to be consistent, then:

$$\mu^{\text{off}} = \frac{\pi(1-a'_P)}{\pi(1-a'_P) + (1-\pi)} \leq \frac{1}{2}$$

which reduces to

$$\pi \leq c_P - \sqrt{(c_P - 1)c_P} \equiv \dot{\pi} \in (0.5, 1)$$

Thus, for $\pi \in [\hat{\pi}, \dot{\pi}]$ there exists a P-Equilibrium supported by off equilibrium affirmances and for $\pi \in (\dot{\pi}, 1]$ there exists a P-Equilibrium supported by off equilibrium reversals.

Candidate equilibrium #2: D-Equilibrium. First, suppose that in an equilibrium, A affirms if $x^* = D$ and reverses after a deviation to $x' = P$. As a result, P's optimal effort is given by $a_P^* = \frac{\pi}{c_P}$ as in Proposition 5, and off the equilibrium path $a_D' = 0$.

Next, if J discovers ω, then she is better off setting $x^* = \omega$ and revealing this information. Then, given that J's judgment is $x^* = D$ when uninformed, J's interim utility over effort is:

$$U_J(e) = e\pi(\delta + \varepsilon_J) + e(1 - \pi)\varepsilon_J + (1 - e)a_P^*\pi(\delta - k) - \frac{c_J}{2}e^2$$

The first order condition yields:

$$e^* = \frac{\pi\delta + \varepsilon_J - a_P^*\pi(\delta - k)}{c_J} = \frac{\pi\delta(c_P - \pi) + \varepsilon_J c_P + \pi^2 k}{c_J c_P} \tag{19}$$

Thus, $e^* > 0$. Because of ε_L and ε_J, it is straightforward to see that the equilibrium messages of L and J are

$$m_J^* = \begin{cases} \omega & \text{if } s_J = \omega \\ \phi & \text{otherwise} \end{cases} \qquad\qquad m_P^* = \begin{cases} P & \text{if } s_P = P \\ \phi & \text{otherwise} \end{cases}$$

Finally, we consider A's beliefs. On the equilibrium path, consistency of beliefs requires

$$\Pr(D|h^D) = \frac{(1 - \pi)(1 - e^*)}{(1 - \pi)(1 - e^*) + \pi(1 - e^*)(1 - a_P^*)} \geq \frac{1}{2}$$

which reduces to

$$\pi \leq c_P - \sqrt{(c_P - 1)c_P} \equiv \dot{\pi}$$

Now suppose that A affirms judgments off the equilibrium path: $x' = P$. Then, $a'_D = \frac{1-\pi}{c_D}$ as in Proposition 5 and J has no incentive to deviate if $\delta \leq \hat{\delta}$, as given in equation (18).

Finally, we consider the off equilibrium beliefs of A. By a similar logic as for the P-Equilibrium, we assume that a deviation to $x' = P$ (with no information) leads A to make an inference that J is uninformed. Otherwise, J would have revealed her information. If A reverses, then $a'_D = 0$, and A's posterior belief off the equilibrium path is thus $\mu^{\text{off}} = \pi$. Reversing off the equilibrium path is thus sequentially rational if $\pi \leq \frac{1}{2}$.

However, if $\pi > \frac{1}{2}$, affirming is sequentially rational given $x' = P$ and $a'_D = 0$, but now D has an incentive to exert positive effort, $a'_D > 0$. Then, A's posterior is

$$\mu^{\text{off}} = \frac{\pi}{\pi + (1-\pi)(1-a'_D)}$$

To affirm $x' = P$, the following must hold

$$\mu^{\text{off}} = \frac{\pi}{\pi + (1-\pi)(1-a'_D)} \geq \frac{1}{2}$$

which reduces to

$$\pi \geq \hat{\pi} \equiv \sqrt{(c_D - 1)c_D} + 1 - c_D$$

Since $\hat{\pi} < \frac{1}{2}$, it follows that affirming when $\pi > \frac{1}{2}$ is sequentially rational off the equilibrium path. Finally, for $\pi > 0.5$, we check that $x = D$ is optimal even with off equilibrium affirmances:

$$U_J(x^* = D) = (\delta - k)a_P\pi \geq \delta\Big((1-a_D) + a_D\pi\Big) - ka_D(1-\pi) = U_J(x' = P)$$

This is the reverse of the condition for the P-Equilibrium. Therefore, for $\pi \in [0, \frac{1}{2}]$ there exists a D-Equilibrium supported by off equilibrium reversals, and for $\pi \in [\frac{1}{2}, \dot{\pi}]$ and $\delta \leq \hat{\delta}$, there exists a D-Equilibrium supported by off equilibrium affirmances. Note that, for $\pi \in [\tilde{\pi}_d, \dot{\pi}]$, $\hat{\delta} = 0$, so the latter condition can be rewritten: for $\pi \in [\frac{1}{2}, \tilde{\pi}_d]$ and $\delta \leq \hat{\delta}$, there exists a D-Equilibrium supported by off equilibrium affirmances. □

First, we restate the proposition formally.

> **Proposition 8′.** Let $\pi < \frac{1}{2}$. Then, there exists an equilibrium where $\xi(x_\phi|\delta > 0) > \xi(x_\phi|\delta = 0)$. Moreover, suppose that $\delta < \hat{\delta}$ or that $\pi < \hat{\pi}$ (or both). Then, $\xi(x_\phi|\delta > 0) > \xi(x_\phi|\delta = 0)$ in all equilibria.

Proof of Proposition 8. Existence of a D-Equilibrium is shown in Lemma 6. Moreover, such an equilibrium exists for all $\pi < \bar{\pi}$, where

$$\bar{\pi} = \begin{cases} \dfrac{1}{2} & \text{if } \delta > \hat{\delta} \\ \pi_d & \text{if } \delta \leq \hat{\delta} \end{cases}$$

Notice that for all $\pi \in [0, \bar{\pi}]$, the dispassionate judge rules in favor of the defendant. The equilibrium effort of a dispassionate judge is given by equation (10):

$$e_d^* = \frac{\pi^2 k}{c_J c_P}$$

The equilibrium effort of a biased judge is given by equation (19):

$$e_b^* = \frac{\pi^2 k}{c_J c_P} + \frac{\pi\delta(c_P - \pi) + \varepsilon_J c_P}{c_J c_P} > e_d^*$$

In a D-Equilibrium, the quality of decision making is given by:

$$\xi(x^*_\phi|t) = \begin{cases} e^*_d(D) + [1 - e^*_d(D)][1 - \pi(1 - a^*_P)] & \text{if judge is dispassionate} \\ e^*_b(D) + [1 - e^*_b(D)][1 - \pi(1 - a^*_P)] & \text{if judge is biased} \end{cases}$$

Since $e^*_b > e^*_d$ the latter probability is larger than the former and a biased judge produces higher quality outcomes than a dispassionate judge.

Finally, from Lemma 6, it is apparent that if $\delta < \hat{\delta}$ or $\pi < \hat{\pi}$, then the only equilibria are D-Equilibria. □

Proof of Lemma 7. Let A's belief that x is wrong after observing s_J is w'_L. Then, after observing m_L, her posterior belief about whether $x \neq \omega$ is:

$$\frac{(1 - \alpha^*_L)w'_L}{(1 - \alpha^*_L)w'_L + (1 - w'_L)}$$

Given that she receives a benefit of 1 if the final outcome is ω, this posterior is also her expected benefit from reversing. Therefore, she reverses if:

$$\frac{(1 - \alpha^*_L)w'_L}{(1 - \alpha^*_L)w'_L + (1 - w'_L)} > \frac{1}{2}$$

This reduces to

$$\alpha^*_L < \frac{2w'_L - 1}{w'_L} =\equiv \underline{\alpha}$$

If the condition holds with equality, A is indifferent between reversing and affirming. Moreover, since $\alpha^*_L \in (0, 1)$, A never reverses if $w'_L \leq \frac{1}{2}$.

Finally, note that $R = 0$ if A does not review. □

Proof of Lemma 8. Given Lemma 7, A's maximization problem is

$$\max_{r_L \in [0,1]} \left[-\frac{c^A}{2} r_L^2 + \begin{cases} w_L^* r_L & \text{if } w_L^* > \dfrac{1}{2} \text{ and } \alpha < \underline{\alpha} \\ \alpha_L^* w_L^* r_L & \text{otherwise} \end{cases} \right]$$

The first order condition (crefeq:review in text) yields an optimum that is a maximum because the second order condition is negative. □

Proof of Lemma 9. Assume $\lambda_L^* \leq \frac{1}{2}$ or $\alpha_0 > \underline{\alpha}$. Suppose A's probability of reviewing a case lost by L is $r_L^*(\alpha_L^*)$. Then, if L exerts effort, he pays cost c_L and learns the state of the world with probability α_1. If he learns the state of the world, he learns that $x \neq \omega$ with probability λ_L. By Lemma 7, then in equilibrium, L's expected benefit from effort is $b\lambda_L r_L^*(\alpha_1)\alpha_1$. Conversely, his expected benefit from no effort is $b\lambda_L r_L^*(\alpha_0)\alpha_0$. L is weakly better off making effort if:

$$b\lambda_L r_L^*(\alpha_1)\alpha_1 - c_L \geq b\lambda_L r_L^*(\alpha_0)\alpha_0$$

Substituting equilibrium values and rearranging:

$$\bar{c}_L \equiv \frac{b\lambda_L^2(\alpha_1^2 - \alpha_0^2)}{c_A} \geq c_L$$

Next, assume $\lambda_L^* > \frac{1}{2}$ and $\alpha_0 < \underline{\alpha}$. By Lemma 7 and Lemma 8, A always reverses and L's effort is not consequential. Therefore, L is strictly better off not making effort, $a_L^* = 0$. □

Proof of Lemma 10. Due to his advocacy payoffs described in Assumption 8, it is straight forward to see that L is better off sending $m_L = \omega$ if and only if $\omega = L$. □

Proof of Lemma 11. We show that J has no incentive to deviate. We consider three cases: (1) $\alpha_0 < \underline{\alpha}$, (2a) $\alpha_0 \geq \underline{\alpha}$ and $s_J = \omega$ and (2b) $\alpha_0 \geq \underline{\alpha}$ and $s_J = \phi$.

Case 1. If $\alpha_0 < \underline{\alpha}$, then it is optimal to set $x^* = P$, since a deviation to $x' = D$ induces A to reverse whenever it reviews.

Case 2a. If $\alpha_0 \geq \underline{\alpha}$ and $s_J = \omega$, then J knows the state of the world, ω. Given the review and reversal strategy of A, she is strictly better off setting $x = \omega$ than $x \neq \omega$. By contradiction, suppose this is not true. Then, she is weakly better off setting $x \neq \omega$. Then, by (12), the appellate court reviews the case with probability $r_L^* > 0$ and by Lemma 7 reverses the decision. J suffers a reversal cost κ and is strictly worse off than if she had ruled $x = \omega$. If A does not review, then the outcome of the case does not match the state of the world, and J is strictly worse off than if she had ruled $x = \omega$. Finally, J cannot be indifferent setting $x = \omega$ because she is strictly worse off both with and without review. Since these review and non-review are exhaustive, then J is never indifferent between setting $x \neq \omega$ and $x = \omega$.

Case 2b. If $\alpha_0 \geq \underline{\alpha}$ and $s_J = \phi$, then J only has a (prior) belief about the state of the world, ω. Let $\Delta(\cdot)$ be the expected net benefit of $x = P$ over $x = D$. (See text.) If J rules for P when indifferent, it is straight forward to see that $\Delta(\cdot) \geq 0$ implies $x = P$, $\Delta(\cdot) < 0$ implies $x = D$. □

Proof of Proposition 9. Lemmas 7 to 11 establish the equilibrium strategies. What remains to be shown is that the beliefs are consistent. Note that all information sets are on the equilibrium path due to the fact that $\pi, \delta, \alpha_0, \alpha_1, r_L^* \in (0,1)$. Therefore, we need not apply refinements to off equilibrium beliefs. Moreover, the beliefs on the equilibrium path are derived by Bayes' rule. First, J's signal is public, so the players' posterior beliefs are either degenerate or equal to the prior belief after s_J is chosen. Finally, A's belief at her reversal information set after reviewing is either degenerate or the prior since $m_L \in \{\omega, \phi\}$. Therefore, we have established the existence of a perfect Bayesian equilibrium. □

Proof of Lemma 12. First, define $\overline{\pi}$ to be the level of π such that $\Delta(\cdot) = 0$:

$$\overline{\pi} \equiv \frac{\alpha_1^2(\kappa-1)+c_A}{(\kappa-1)(\alpha_1^2-\alpha_0^2)} - \sqrt{\frac{(\alpha_0^2(\kappa-1)+c_A)(\alpha_1^2(\kappa-1)+c_A)}{(\kappa-1)^2(\alpha_0^2-\alpha_1^2)^2}}$$

Then, the result is simply a corollary of Lemma 11. $\square$

Proof of Lemma 13. Suppose $x^* = D$. Then by Lemma 11, $\Delta(\cdot) \leq 0$ or $s_J = D$. From the definition of $\Delta(\cdot)$, it is clear that $\Delta(\cdot) \leq 0 \Leftrightarrow \alpha_P^* < \alpha_D^*$. Since $\alpha_L^* \in \{\alpha_0, \alpha_1\}$, this implies $(\alpha_P^*, \alpha_D^*) = (\alpha_0, \alpha_1)$. $\square$

Proof of Proposition 10. To prove this statement, it is sufficient to show that there exists a strategy profile where the sum of payoffs is greater than under the equilibrium outcome. Consider the strategy profile $\mathcal{S}'$, which features the following deviations from $\mathcal{S}_D^*$: $x'(s_J = \phi, \Delta(\cdot) < 0) = P$ and $a_D' = 0$. We show that the sum of payoffs under $\mathcal{S}'$ is greater than under $\mathcal{S}_D^*$. By contradiction, suppose not, then:

$$u_A^* + u_J^* + b \geq u_A' + u_J' + b \tag{20}$$

The first two terms on the left and right of the inequality are the payoffs for A and J, respectively, and the remaining terms characterize the sum of payoffs to the litigants. Note that b is obtained by one or the other litigant in every equilibrium (depending on the outcome), and no litigant effort costs are paid in the strategy profiles $\mathcal{S}_D^*$ and $\mathcal{S}'$. However, if $c_A > \alpha_0\pi$, then equation (20) is false and a contradiction is obtained.[44] $\square$

Proof of Proposition 11. This is formally equivalent to the equilibrium in Proposition 9 when $\alpha_0 < \underline{\alpha}$. $\square$

Proof of Lemma 14. It suffices to show that the sum of payoffs under $\mathcal{S}_D^*$ is strictly lower

[44] Verified via `Mathematica`. See Figure 17 in Appendix.

than under $\mathcal{S}^B$. By contradiction, suppose not. Then, the following holds:

$$(1-\delta)[2(1-\pi)+\pi\alpha_0 r_P^*(2-\kappa)] - \frac{c_A}{2}(r_P^*)^2 > 2(1-\delta)\pi$$

Substituting in r_P^* and reducing:

$$4c_A(1-\delta)(1-2\pi)+\alpha_0^2\pi^2[4(1-\delta)-2\kappa(1-\delta)-1] > 0 \tag{21}$$

Define $\widehat{\kappa}$ to be the level of κ such that Equation (21) holds with equality. There are two cases, (I) $\kappa \geq \widehat{\kappa}$ and (II) $\kappa < \widehat{\kappa}$. For case (I), the condition does not hold since the LHS is weakly negative and a contradiction is obtained. For case (II), there exists a value of α_0 such that the condition holds with equality. Label this value $\widehat{\alpha}$. Then, for all $\alpha_0 < \widehat{\alpha}$, the condition does not hold, and a contradiction is obtained for case (II). □

Proof of Lemma 15. To see this, it must be that:

$$\delta + (1-\delta)\pi > \delta + (1-\delta)[(1-\pi)+\pi\alpha_0 r_P^*] - \frac{c_A}{2}(r_P^*)^2$$

This reduces to

$$c_A > \frac{\alpha_0^2\pi^2}{4\pi-2}$$

$$or$$

$$\delta > \frac{2c_A+\alpha_0^2\pi^2-4c_A\pi}{2c_A+2\alpha_0^2\pi^2-4c_A\pi}$$

□

Proof of Lemma 16. In text. □

B District Judge's Interior Effort

For J to exert interior effort in equilibrium, the following must hold:

$$c_J > \begin{cases} \dfrac{1}{c_P}\left[\pi(c_P - \pi)\beta t + \pi^2 k\right] & \text{if } x_\phi = D \\ \dfrac{1}{c_D}\left[(1-\pi)(c_D - (1-\pi))\beta t + (1-\pi)^2 k\right] & \text{if } x_\phi = P \end{cases} \tag{22}$$

We define a type-dependent function $\tilde{c}_J(\pi, t)$, which returns a threshold value of c_J such that equation (22) holds with equality:

$$\tilde{c}_J(\pi, t) = \begin{cases} \dfrac{\pi(c_P - \pi)\beta t + \pi^2 k}{c_P} & \text{if } \pi \leq \tilde{\pi}_t \\ \dfrac{(1-\pi)(c_D - (1-\pi))\beta t + (1-\pi)^2 k}{c_D} & \text{if } \pi \geq \tilde{\pi}_t \end{cases}$$

There are three properties of this function worth noticing. First, the inequalities in the conditionals are weak. This is because the two components of the piecewise function are equal when evaluated at $\tilde{\pi}_t$. Second, $\tilde{c}_J(\pi, t)$ is strictly increasing when $\pi < \tilde{\pi}_t$ and strictly decreasing when $\pi > \tilde{\pi}_t$. Thus, $\tilde{c}_J(\pi, t)$ achieves its global maximum at $\tilde{\pi}_t$. And third, $\tilde{c}_J(\pi, t = 1)$ is strictly greater than $\tilde{c}_J(\pi, t = 0)$ for all $\pi \in (0, 1)$. Using these facts, in order for J's effort to be interior in in any equilibrium, $c_J > \bar{c}_J$ (Assumption 4), where:

$$\bar{c}_J \equiv \tilde{c}_J(\pi = \tilde{\pi}_1, t = 1) = \frac{c_P \beta \tilde{\pi}_1 - (\beta - k)\tilde{\pi}_1^2}{c_P} \tag{23}$$

and $\tilde{\pi}_1$ is given by equation (26). For the case of a biased judge, replace β with δ in Equation (23).

C Characterization of $\tilde{\pi}_t$

We explicitly characterize the values as the roots of the quadratic equation $\chi(\pi, t, \cdot) = 0$:

$$(2\pi - 1)\beta t + (\beta t - k)\left(\frac{(1-\pi)^2 c_P - \pi^2 c_D}{c_D c_P}\right) = 0$$

Using the quadratic formula, the solution to this quadratic equation is the element in Π that is in the unit interval, where Π is defined as:

$$\Pi = \left\{\frac{c_P(\beta t(c_D - 1) + k) \pm \sqrt{c_D c_P(\beta t(c_D - 1) + k)(\beta t(c_P - 1) + k)}}{(\beta t - k)(c_D - c_P)}\right\} \tag{24}$$

Given that $c_P > c_D$, then equation (24) reduces to the following:

$$\tilde{\pi}_0 = \frac{c_P - \sqrt{c_D c_P}}{c_P - c_D} \tag{25}$$

if $t = 0$, and

$$\tilde{\pi}_1 = \frac{c_P(\beta(c_D - 1) + k) - \sqrt{c_D c_P(\beta(c_D - 1) + k)(\beta(c_P - 1) + k)}}{(k - \beta)(c_P - c_D)} \tag{26}$$

if $t = 1$.

D An Alternative Definition of Impartiality

In the main text of the paper, "impartiality" is defined relative to the equilibrium judgment of a district judge. In particular, an impartial judge was one who made a decision (when uninformed) about the case that was consistent with the prior probability that the plaintiff experienced illegal discrimination.

An alternative way to understand impartiality is that an impartial judge is one whose decision making across a large number of cases would match the probability that the plaintiff experienced illegal discrimination. That is, an impartial judge's *ex ante* probability of ruling in favor of the plaintiff matches the prior belief.

> **Definition 5** (Normative Benchmark: Impartiality (2))**.** A district judge's adjudication is "impartial" if and only if the *ex ante* probability of ruling in favor of the plaintiff matches the prior probability that the plaintiff experienced illegal discrimination.

Denote the *ex ante* probability that a type t judge rules in favor of the plaintiff on a case with a prior of π by $\gamma_t(\pi)$. Then using the optimal judgment given by equation (8):

$$\gamma_t(\pi) \equiv \Pr(x^* = P|\pi, t) = \begin{cases} e_t^* \pi & \text{if } \pi \leq \min\{\tilde{\pi}_t, \dot{\pi}_t\} \\ 1 - (1-\pi)e_t^* & \text{if } \pi > \tilde{\pi}_t \end{cases}$$

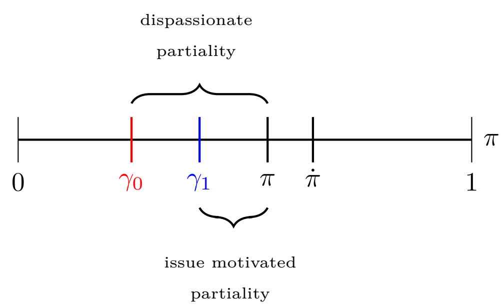
where $\dot{\pi}$ is defined in Lemma 21. For a given π, we know that $e_0^*(\pi) < e_1^*(\pi)$ by Lemma 5. Since $c_P > c_D$, it follows that $\tilde{\pi}_1 < \tilde{\pi}_0$.

Suppose $\tilde{\pi}_1 < \tilde{\pi}_0 < \dot{\pi}$. Then, if $\pi \notin (\tilde{\pi}_1, \tilde{\pi}_0)$, it is easy to see that $\gamma_0(\pi > \tilde{\pi}_0) > \gamma_1(\pi > \tilde{\pi}_1) > \pi$ and $\gamma_0(\pi \leq \tilde{\pi}_0) < \gamma_1(\pi \leq \tilde{\pi}_1) < \pi$. If, however, $\pi \in (\tilde{\pi}_1, \tilde{\pi}_0)$, then the issue motivated judge is more impartial if:

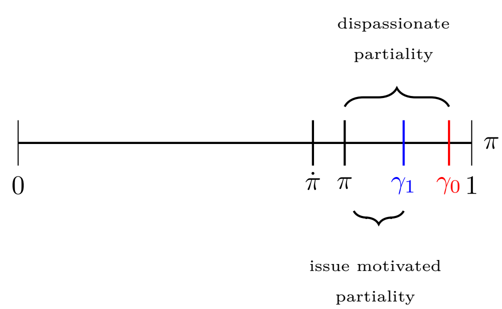
$$\underbrace{\pi - e_0^* \pi}_{|\gamma_0 - \pi|} > \underbrace{1 - \pi - (1-\pi)e_1^*}_{|\gamma_1 - \pi|}$$

This condition holds as long as $\pi > \frac{(1-e_1^*)}{(1-e_1^*)+(1-e_0^*)}$. This holds for all $\pi \geq \tilde{\pi}_1$.

The following figures illustrates this result. In each figure, I label the amount of partiality (*i.e.*, how far she or he is from being impartial) of each type of judge. First, suppose that $\pi < \dot{\pi}$:



Next, suppose that $\pi > \dot{\pi}$:



As is apparent from these diagrams, the dispassionate type judge is less impartial according to Definition 5.

E Supplemental Code

Figure 17: Supplemental Code for Proposition 10

```
(* S' YIELDS HIGHER PAYOFF THAN SD *)
Simplify[Reduce[
  (1 - d) * ((1 - p) + p * a0 * rP0) - (cA / 2) * rP0 ^ 2 + (1 - d) * ((1 - p) - p * (k - 1) * a0 * rP0) >
    (1 - d) * (p + (1 - p) * a0 * rD0) - (cA / 2) * rD0 ^ 2 + (1 - d) * (p - (1 - p) * (k - 1) * a0 * rD0)
   && 0 < d < 1
   && k ≥ 1
   && 0.5 < p < 1 / (2 - a0)
   (*&&(cA≥a0*p&&rD0==((1-p)*a0)/cA&&rP0==(p*a0)/cA)*)
   (*&&(0<cA<a0*p&&rD0==1&&rP0==1)*)
   && (a0 * p ≥ cA ≥ a0 * (1 - p) && rD0 == ((1 - p) * a0) / cA && rP0 == 1)
   && 0 < a0 < a1 < 1]]
```

Reduce::ratnz: Reduce was unable to solve the system with inexact coefficients
The answer was obtained by solving a corresponding exact system and numericizing the result ≫

```
False
```