Coordination and Innovation in Judiciaries: Correct Law vs. Consistent Law

Mehdi Shadmehr* Sepehr Shahshahani† Charles Cameron‡

Abstract

We identify the coordination consideration among judges who do not have formal authority over each other, and investigate its consequences for their decisions and legal innovations. Coordination concerns arise because judges value the consistent application of law. To mitigate their strategic uncertainty, judges overweight interpretations that are visible throughout the judiciary (e.g., prominent judges' opinions) because their visibility facilitates coordination. This creates a tradeoff between the consistent and correct application of law—the two desiderata of judicial decisionmaking. In particular, anticipating overreactions to their opinions, some prominent judges refrain from expressing their informed opinions. Paradoxically, the propensity to refrain is strongest in prominent judges who care most about the correct application of law. From their perspective, excessive concern for uniformity in the judiciary overrides the informational value of expressing informed opinions. We explore the implications for issuing narrow or broad opinions, the stickiness of precedent, and the practice of stare decisis. We provide concrete examples from contract, property, tort, and constitutional law that support our theoretical mechanisms.

Keywords: non-hierarchical judiciary, informal authority, legal innovation, stare decisis, coordination, beauty contests

*Department of Public Policy, University of North Carolina at Chapel Hill. E-mail: mshadmeh@gmail.com

†School of Law, Fordham University. E-mail: sshahshahani@fordham.edu

‡Department of Politics and Princeton School of Public and International Affairs, Princeton University. E-mail: ccameron@princeton.edu

We wish to thank Pam Bookman and participants at the 2019 meeting of the American Law and Economics Association for their helpful suggestions.

As judges create and apply law, two considerations weigh heavily: (1) correctness: the law should yield good results, in the sense of fairness, efficiency, and conformity to legal norms, and (2) consistency: the law should be applied uniformly, so throughout the judiciary judges dispose of similar cases similarly.[1] Oftentimes the correctness and consistency criteria go hand in hand, for example when all agree on what constitutes the best doctrine and all know that others will apply it too. But sometimes the correctness of prevailing doctrine becomes uncertain, for instance in the face of rapid social change. Then, some judges may begin to see other doctrinal possibilities as superior. In turn, uncertainty about correctness generates strategic uncertainty about consistency because judges cannot be sure which doctrine other judges will find correct. Judges then face difficult decisions. In this paper, we explore the consequences of these twin uncertainties for the decisions of judges as well as the spread of legal innovations among judges who do not have formal authority over each other. Because many American judges stand in such non-hierarchical relation to each other (e.g., in different state jurisdictions), the analysis has wide applicability in the United States; it also applies to other countries that combine common law jurisprudence with federalism (e.g., Canada and Australia).

How do common law judges decide when faced with uncertainty about the correct law? According to Judge Benjamin Cardozo, a judge "must get his knowledge just as the legislator gets it, from experience and study and reflection; in brief, from life itself" (1921, 113). That is, judges use their experience and legal skills to apply laws correctly. But because their knowledge and experiences differ, so do their assessments of the correct law. These differences might undermine the consistent application of law. Thus, when rendering decisions, judges may modify their own assessments to make their judgments closer to what they expect other judges will do. They want to choose the correct law, but they also want to choose the same law.

Because judges care about legal consistency, their decisions involve a strategic element of estimating what others will do. We argue that a key consequence of this strategic environment is that judges overreact to public information about the correct interpretation of the law—e.g., the opinion of a judge who is highly visible in the judiciary (like Benjamin Cardozo of the New York Court of Appeals). Judges overreact in the sense that they put more weight on public information than they would if they only cared about deciding what the correct law is,

[1]Judges care about both considerations, and each consideration has had its advocates in different times. For example, a concern for correctness is reflected in Justice Holmes's famous declaration that the law should reflect "the felt necessities of the times" (Homles 1881, 1), and consistency is reflected in Justice Brandeis's statement that "in most matters it is more important that the applicable rule of law be settled than that it be settled right" (*Burnet v. Coronado Oil & Gas Co.*, 285 U.S. 393, 406 (1932) (Brandeis, J., dissenting)).

and not also about getting their decisions close to each other to increase consistency in the judiciary. They put more weight on public information because it allows them to coordinate better. The reason is that public information is common knowledge: Everyone observes it, and everyone knows that everyone observes it, and so on. Thus, to bring their decisions closer together, judges downplay their private information, which others cannot see and hence cannot use to predict a judge's decision, and emphasize their public information instead. In a sense, because visible opinions about the law function as a focal point, judges put more weight on them than is merited by the information they communicate about the correct law, thereby partially sacrificing correctness to consistency.

The overreaction of judges to public information has important consequences for the spread of legal innovations. Some judges, by virtue of their positions and skillful reasoning, are more visible than others. They are leaders in one field of the law or another, not because they have formal authority to overrule or reverse other judges, but because their opinions are visible and persuasive. As a *Washington Post* editorial said of Learned Hand, the preeminent judge of the Second Circuit Court of Appeals during the early and mid twentieth century, "He has won recognition as a judges' judge. His opinions command respect wherever our law extends, not because of his standing in the judicial hierarchy, but because of the clarity of thought and the cogency of reasoning that shape them" (quoted in Gunther 1994, 574). Leader judges' opinions, books, and law review articles become public information among judges facing similar cases. Judges see these leaders' views of the correct law and, critically, know that other judges see those views as well. That is, the views of leader judges become common knowledge among judges who face similar decisions. But as we discussed above, because judges care about legal consistency, they overweight all public information, including leader judges' opinions. As a result, leader judges who care more about the correct application of the law may counterintuitively refrain from expressing opinions which contain valuable information for finding the correct law. They would rather wait to make their views more accurate so that the gains from additional information outweigh the overreaction to public information.

An example of such restraint can be seen in a bankruptcy decision by Judge Henry Friendly of the Second Circuit, who was considered "the greatest judge of his era" (Dorsen 2012). Friendly acknowledged that in making a narrow decision the court was "not giving bankruptcy judges the guidance which they doubtlessly desire and it is our duty to provide if we properly can." "But," he continued, "it is better to fail in this respect than to attempt to give guidance

without having seen the variety of factual situations, having heard from the adversarial presentations, and having the benefit of the scholarly community which time will undoubtedly afford." *In re B.D. Int'l Disc. Corp.*, 701 F.2d 1071, 1077 (2d Cir. 1983). Friendly was building on other judges' similar reluctance to issue noisy opinions, quoting Justice Harlan that such opinions "suffer the danger of pitfalls that usually go with judging in a vacuum ... they are apt in their application to carry unintended consequences which once accomplished are not always easy to repair." *Id.* (quoting *Sanders v. United States*, 373 U.S. 1, 32 (1963) (Harlan, J., dissenting)). We identify one such "unintended consequence": the overreaction of other judges, who are not bound by the decision, to the public information produced by the decision, stemming from the desire for legal consistency.

The analysis sheds light on the rate and spread of legal innovation in the face of societal and technological changes that demand a corresponding change in the law. It shows, counterintuitively, that judges who care more about adapting the law to changed circumstances may be more hesitant to contribute to that process. The analysis also pinpoints an important strategic rationale for writing "narrow" opinions—opinions that skirt an important legal issue—a topic much discussed in the literature.[2]

Our analysis also reveals a qualitatively new channel for the practice of stare decisis—following precedent—to take hold in courts. When there is common knowledge about a set of precedents, any departure from them can increase inconsistency across judges, because judges remain uncertain about each other's doctrinal assessments, and hence cannot coordinate as effectively. This implies that precedents have inertia: even when changes in the world require changes in the law, and even when knowledge of better law is present in the judiciary as a whole, the aggregate behavior of the judiciary is biased toward precedents. Each judge puts extra weight on precedent because common knowledge about the precedent allows judges to coordinate more effectively.

The bias toward precedent, in turn, highlights the critical role of prominent judges whose doctrinal signals are public information in the judiciary. By virtue of being common knowledge,

[2]Our quotations from Judge Friendly and Justice Harlan show how important is the question of whether to craft a narrow or broad opinion and, for courts that have a discretionary docket (like the federal and most state supreme courts), the question of whether to take a case at all. As Justice Brandeis remarked to future Justice Frankfurter, "the most important thing we do is not doing" (quoted in Urofsky 1985, 313). See also Bickel (1962, 111-198) (discussing the judicial arts of avoiding certain issues or dealing with them narrowly, which he famously called "the passive virtues"). Moreover, our rationale is very different from the delegation mechanism in the literature (e.g., Staton and Vanberg 2008).

the prominent judges' doctrinal changes can allow the judiciary to coordinate on a new doctrine rather than continue with one that many have come to believe is inferior. Thus, the overweighting of new doctrinal information provided by prominent judges can counter the overweighting of old doctrinal information in precedents. However, as noted above, this very overreaction to common knowledge can make prominent judges hesitant to communicate their information when they believe it is insufficiently precise.[3] This mechanism contrasts with explanations of stare decisis grounded in judicial hierarchy (e.g., Bueno de Mesquita and Stephenson 2002). In our setting there is no judicial hierarchy, and no auditing or enforcement by higher courts. Rather, the interest in consistency acts as a disciplining mechanism that pulls judges toward upholding precedent.

Although prominent judges can help overcome the inertia in the old ways, their communicated information can sometimes be harmful. As Judge Friendly appreciated, a prominent judge's assessment of the best doctrine may well be inaccurate. When the leader nonetheless acts, it can lead follower judges into coordinating on an inferior doctrine. Our model predicts that "mistakes" of this kind will occasionally happen. If so, one may see a subsequent doctrinal reversal given accumulating signals. This phenomenon is hard to explain in models in which stare decisis reflects a power relationship between superiors and inferiors (in the absence of changes in the composition or preferences of the apex court). Moreover, in hierarchy models the "correctness" of law is often nothing more than the preferences of the apex court, whereas in our model correctness has objective content distinct from (though reflected in) judicial preferences, which allows us to give a fuller account of how doctrine may go astray. We provide examples not only of successful coordination around prominent judges' opinions, but also of unsuccessful ones leading to subsequent doctrinal reversal.

The literature on interactions among courts focuses primarily on judicial hierarchy, studying a variety of topics including strategic auditing (Cameron, Segal, and Songer 2000), breaking with precedent (Bueno de Mesquita and Stephenson 2002), whistle-blowing (Beim, Hirsch, and Kastellec 2014), and doctrinal compliance (Westerland et al. 2010; Carrubba and Clark 2012; Shahshahani (forthcoming 2021)).[4] However, the vast majority of judges do not have authority over each other, and in this sense are placed horizontally. Our focus is on the interaction among

[3]Bueno de Mesquita and Shadmehr (2020) apply this idea to "social norms and social change."

[4]More broadly, there is a vast literature on various aspects of the judiciary, including the evolution of precedent (Baker and Mezzetti 2012), herding (Daughtey and Reinganum 1999; Talley 1999), judicial review (Fox and Stephenson 2011), opinion obfuscation (Staton and Vanberg 2008), and decisionmaking in collegial courts (Kornhauser 1992).

these judges. Our key insight is to identify the coordination consideration and its consequences for decisionmaking among the multitude of judges who do not have formal authority over each other. We make use of advances in the literature on coordination with incomplete information to capture the nature and consequences of these interactions. In particular, our formalization adapts the framework of Morris and Shin's (2002) seminal paper, "The Social Value of Public Information." Their framework has been applied in finance and macroeconomics (Angeletos and Lian 2016), and to study leadership in party conferences (Dewan and Myatt 2008) and organizations (Bolton et al. 2013; Landa and Tyson 2017). This paper is the first to adapt their insights to the study of courts. A distinguishing feature of our analysis is that the leader judge generally weighs correctness and consistency differently than regular judges, and hence is not a social planner.

Our approach is related to, but distinct from, a literature that examines the coordinative role of law within societies (Hadfield and Weingast 2012; Acemoglu and Jackson 2017; Mailath et al. 2017). In this approach, law coordinates actions by private agents, thereby facilitating personal and commercial interactions and even decentralized punishment of norm-breakers in stateless societies. These papers focus on private agents and treat the judiciary as a unitary actor. We consider the coordination problem facing multiple judges but we do not address social coordination. Hence, the two distinct approaches, while focusing on coordination among different actors, are complementary.

In the following section, we fix ideas by providing two case studies. We then present a stylized model to formalize the logic and intuitions. Appendixes provide additional examples of leader judges and doctrinal coordination, both successful and unsuccessful, as well as proofs and extensions.

Examples

Before we proceed to the model, we provide two examples to illustrate the horizontal spread of legal innovations and the role of leader judges. The first doctrinal move was widely emulated and has proven durable. The second was initially influential but was subsequently found wanting and reversed.

Example 1: *Javins v. First National Realty.* Can a tenant be evicted for not paying rent even if the nonpayment is due to the landlord's failure to make the rented premises

habitable? Is the lease of a residential unit the conveyance of an interest in land, subject to the peculiar requirements of real property law, or is it a contract that should be construed like any other contract? In the pure property interpretation, the landlord has little obligation beyond physically delivering the unit to the tenant; he need not make the unit habitable, and his failure to provide such essentials as heat and sanitation does not excuse the tenant's failure to pay rent. E.g., *Hoy v. Holt*, 91 Pa. 88 (1879). In a middle-ground interpretation, located somewhere between property and contract law, the landlord does have a duty to make the premises habitable, but the tenant cannot raise the landlord's breach of this duty as a defense in an action for eviction for nonpayment of rent; rather, he must bring a separate suit against the landlord to vindicate his right to a habitable unit. E.g., *Posnanski v. Hood*, 46 Wis. 2d 172 (1970). In a more contract-centered view, not only does the landlord have an obligation to make the premises habitable, but "the tenant's obligation to pay rent is dependent upon the landlord's performance of his obligations," so the court in an action for eviction must consider whether the tenant's nonpayment of rent was justified by the failure of the landlord's duties. E.g., *Javins v. First National Realty Corp.*, 428 F.2d 1071 (D.C. Cir. 1970).

Early American caselaw advocated a property-centered interpretation, under the assumption that what matters is the land itself, not the condition of the dwelling constructed on the land (see generally 1 American Law of Property § 3.78). That assumption might have been reasonable for an agrarian society, but in time it became disconnected from a reality in which apartments were leased primarily as places to live in, not as structures on land to be tilled. In response, courts toward the middle of the Twentieth Century came to place greater obligations on landlords and to conceptualize a residential lease less as a conveyance of an interest in land than an ordinary contract.

The best-known landmark in this conceptual transformation is Judge James Skelly Wright's opinion in *Javins*. The opinion elucidated the reasons for the move from property to contract, explaining how the principles "derived from feudal property law" had become inapposite for "the modern apartment dweller," who seeks in an apartment a "well known package of goods and services—a package which includes not merely walls and ceilings, but also adequate heat, light and ventilation, serviceable plumbing facilities, secure windows and doors, proper sanitation, and proper maintenance" (*Javins*, 428 F.2d at 1074). The lease of an apartment is thus like a contract for a standard manufactured good and "should be interpreted and construed like any other contract" (1075). Just like contracts for goods are subject to implied warranties of

fitness and merchantability, leases should be subject to an implied warranty that the premises are habitable (1075-77). Moreover, promises in contracts are mutually dependent, such that one party's failure to carry out her promise excuses the other party from performance. "Under contract principles," then, "the tenant's obligation to pay rent is dependent upon the landlord's performance of his obligations, including his warranty to maintain the premises in habitable condition" (1082).

Judge Skelly Wright's opinion in *Javins* had no binding authority beyond the D.C. Circuit, but it was immensely influential in the nationwide transformation of landlord-tenant law. It was one of the first decisions to recognize the implied warranty of habitability, and apparently the first to allow it as a defense in a landlord's eviction action (Chused 2004, 193). State courts around the country soon adopted both holdings, often citing *Javins* as persuasive, though not binding, precedent (see, e.g., Chused 2004; Rabin 1984). Today *Javins* is viewed as a pioneering decision, and it is often excerpted in property law casebooks.

Example 2: *United States v. Garcia*. Does law enforcement's long-term monitoring of a car with the aid of a GPS device attached to it constitute a "search" within the meaning of the Fourth Amendment?[5] If the answer is yes, then the police need a warrant or at least some form of individualized suspicion before subjecting someone to GPS surveillance. If no, then the police's use of GPS monitoring is not subject to constitutional privacy requirements. Chilling as the specter of judicially unsupervised GPS surveillance might seem, its unconstitutionality was not obvious when cheap GPS monitoring became technologically feasible in the early 2000s.

The closest Supreme Court precedent on point was a case decided in 1983 in which the police surreptitiously placed a beeper in a car and followed the car with the help of a device that received the beepers' beeps, which got louder as the police neared the car. The Court held that the monitoring did not constitute a search, reasoning that the beeper merely made more efficient the police's in-person visual surveillance of the car on public roads, which precedent had established not to be a search (*United States v. Knotts*, 460 U.S. 276, 281-84 (1983)). Responding to concerns that "twenty-four hour surveillance of any citizen of this country will be possible, without judicial knowledge or supervision," the Court noted that "such dragnet-type law enforcement practices" were not at issue in the case and their constitutionality would be assessed when they actually occurred (*id.* at 283-84). *Knotts* did not provide clear guidance

[5]The Fourth Amendment to the U.S. Constitution guarantees "the right of the people to be secure ... against unreasonable searches and seizures," and has been interpreted to protect citizens' privacy against governmental intrusion.

to courts that had to determine the applicability of the Fourth Amendment to warrantless GPS surveillance two decades later. One option was to follow *Knotts*'s logic that GPS simply allowed police officers to technologically "augment[] the[ir] sensory faculties" (*id.* at 282), and to dismiss concerns about mass surveillance as farfetched. Another option was to hold that the "dragnet-type" practices conjectured in *Knotts* were presently at hand, and to distinguish GPS from old beeper technology on the grounds that it did not merely enhance in-person visual surveillance but obviated the need for in-person, on-the-street involvement.

Most courts selected the first option. Judge Richard Posner, arguably the most distinguished late-Twentieth Century American jurist, seems to have been influential in coordinating this selection. *United States v. Garcia*, 474 F.3d 994 (7th Cir. 2007), the first federal appellate decision to squarely face the issue, held that GPS monitoring is not a search. Writing for a unanimous panel of the Seventh Circuit, Posner analogized GPS technology to surveillance cameras and satellite imaging, which substitute for police activity (following a car on a public street) that would not have been a search if conducted in person (*id.* at 997). He acknowledged the dangers of GPS-enabled mass surveillance, but dismissed them by noting that there was no evidence of a mass surveillance program in the case (*id.* at 998).

Courts in other jurisdictions soon faced the same legal issue. They were not obligated to follow Posner's lead (a horizontal relationship), but they did. The impact of Posner's earlier opinion cannot be precisely estimated; but we know that all of the state appellate courts and all but one of the federal courts to decide the issue reached the same conclusion based on very similar reasons—and they all cited Posner's opinion as persuasive authority (see *United States v. Pineda-Moreno*, 591 F.3d 1212 (9th Cir. 2010); *United States v. Marquez*, 605 F.3d 604 (8th Cir. 2010); *United States v. Sparks*, 750 F. Supp. 2d 384 (D. Mass. 2010); *Foltz v. Commonwealth*, 57 Va. App. 68 (2010), aff'd on other grounds, 58 Va. App. 107 (2011) (en banc); *State v. Sveum*, 2009 WI App 81). Indeed, some of these later opinions mention Posner by name in an apparent effort to bolster their conclusions (*Marquez*, 605 F.3d at 610; *Foltz*, 57 Va. App. at 84). Only one federal appellate court held that GPS monitoring is a search (*United States v. Maynard*, 615 F.3d 544 (D.C. Circuit 2010)).[6]

The issue was finally settled by the Supreme Court in 2012. The Court held, 9-0, that GPS monitoring is indeed a search. *Jones* has subsequently come to be seen as a sensible

[6]Some state supreme courts also invalidated warrantless GPS surveillance, but they did so under state constitutional guarantees rather than the Fourth Amendment. See *State v. Jackson*, 150 Wash. 2d 251 (2003); *People v. Weaver*, 12 N.Y.3d 433 (2009); *Commonwealth v. Connolley*, 454 Mass. 808 (2009).

interpretation of Fourth Amendment guarantees in a new technological context. The validity of its conclusion that GPS monitoring is not exempt from constitutional privacy requirements has been widely accepted.

Examples 3-5. In Appendix I we provide three more examples of judicial coordination around new doctrine announced by leader judges. The examples are from contract, tort, and the contract-tort interface. Two of the innovations have proven successful; one has been severely criticized.

Model and Analysis

There are two judges indexed by $i \in \{1, 2\}$. Judge i must take an action $a_i \in \mathbb{R}$. He aims to make his action close to the state of the world $\theta \in \mathbb{R}$ (reflecting a concern for correct law), and to the other judge's action a_j (reflecting a concern for consistency). In particular, judge i's payoff, u_i, from taking action a_i is:

$$u_i(a_i, a_j, \theta) = -(1-r)\ (a_i - \theta)^2 - r\ (a_i - a_j)^2,\ i \neq j,\ r \in (0, 1). \tag{1}$$

The state of the world θ is unknown to judges and they have a common (improper) prior that it is distributed uniformly on $\mathbb{R}$. Judge i observes a noisy private signal about the state of the world, $x_i = \theta + \epsilon_i$, where $\epsilon_i \sim N(0, \sigma_\epsilon^2)$ and ϵ_is are distributed independently of each other and θ. In addition to their private signals, the judges observe a public signal $y = \theta + \eta$, where $\eta \sim N(0, \sigma_\eta^2)$, and η is distributed independently of θ and ϵ_is. Given their signals, judges simultaneously choose their actions. The parameter r indexes how much the judges value legal consistency as compared to correctness.

A judge i's strategy is a mapping from his information set $I_i = (x_i, y)$ to an action a_i, and the equilibrium concept is Bayesian Nash equilibrium.

Interpretation. Before we proceed to the analysis, we interpret the key aspects of the model. The state of the world θ represents the correct choice of legal doctrine. In the *Javins* case, for example, θ captures the correct legal framework for analyzing residential leases, where lower θ corresponds to more elements of contract law and higher θ corresponds to more elements of real property law. The action a_i is judge i's choice of doctrine. The choice of doctrine also determines the disposition of the case (i.e., who wins the lawsuit). In *Javins*, for example, a high θ would imply that a tenant who has not paid rent due to the landlord's failure to provide

habitable premises can be evicted (landlord wins), whereas a low θ would allow the tenant to raise the breach of the warranty of habitability as a defense (tenant wins). We are concerned with legal innovation—which is a question of legal doctrine, not who won any particular lawsuit—so the judge's action space is modeled as continuous.

The timing of the game reflects two realities. First, judges with overloaded dockets and limited resources rarely have time to delve into decisions that have no binding effect on them (with the exception of decisions by highly prominent judges, which we analyze later), and until recently did not even have ready access to caselaw from outside their jurisdiction (Caldeira 1985). Second, uncertainty about the correct law typically arises in the face of social or technological changes, which occur concurrently in different jurisdictions, so courts are often grappling with the same issues at around the same time. Our simultaneous-move model is equivalent to a model with sequential moves in which later-moving judges observe the private signals of earlier-moving judges with a great deal of noise; the same results therefore go through with sequential moves as long as the noise is sufficiently large.[7]

Analysis. We focus on the natural class of symmetric linear strategies

$$a_i(I_i) = kx_i + (1-k)y, \;\; k \in [0,1], \tag{2}$$

so that higher signals x_i or y, indicating a higher value of θ, both raise the judge's action, albeit possibly with different weights. To characterize the equilibrium in these linear strategies one must show that there exists a weight k that makes the corresponding strategies best responses to each other. Following the steps of Morris and Shin (2002), we show in the Appendix that such an equilibrium exists and it is unique.

Proposition 1 ***(Morris and Shin 2002)*** *There is a unique equilibrium in which a judge i with private signal x_i and a public signal y chooses an action*

$$a_i(I_i) = \frac{ay + b(1-r)x_i}{a + b(1-r)}, \tag{3}$$

where $a = 1/\sigma_\eta^2$ is the precision of the public signal and $b = 1/\sigma_\epsilon^2$ is the precision of private signals.

[7]One could contemplate an extended version of the game in which each judge observes a noisy private signal of the other judge's private signal. Our results are the limit of this extended game where the noise in additional signals become very large, and hence our results go through when the noise in these additional signals is above a threshold. Similarly, the tradeoff between correctness and consistency, and the overreaction to public information, remain present in the alternative, sequential game in which a judge observes noisy signals of preceding judges' actions or signals.

Proposition 1 reveals the judges' coordination incentives. If judges did not have any concern for coordination (i.e., judicial coherence or the rule of law), then $r = 0$, and $a_i(I_i) = E[\theta|x_i, y] = \frac{ay+bx_i}{a+b}$, reflecting the judges' pure consideration for legal correctness—using all their information to estimate the location of θ. But as judges value judicial coherence more (i.e., as r rises), each judge pays less attention to his private assessment of the case, which is unknown to the other judges, and puts more weight on what is common knowledge throughout the judiciary—in order to reach conclusions and take actions that are closer to other judges.

Leader Judges. This framework allows us to study some aspects of judicial leadership and the spread of legal innovations among judges who do not have formal authority over each other. The essential attribute of leader judges is their visibility. This visibility may be the result of any number of factors, such as the judge's previous political, judicial, lawyerly, or academic activities. For example, Judge Skelly Wright, whose work we discussed in the previous section, had been a famous trial judge before becoming an appellate judge. Serving on the Eastern District of Louisiana, Skelly Wright ordered and presided over the desegregation of a number of public institutions during the 1950s. His desegregation decisions met massive resistance by the state legislature and governor, as well as white public opinion. The judge became a pariah in his hometown of New Orleans, with a cross being burned on his home lawn and his family requiring bodyguards for safety (Bernick 1980, Bass 1981). The Southern chairman of the Senate Judiciary Committee called Skelly Wright "a no-good son of a bitch" (Miller 1984, 83), and Southern Senators' opposition was such that he could not be nominated to the Fifth Circuit Court of Appeals. On the other hand, the judge's stance endeared him to Northern liberals and made him a national figure—gaining positive media coverage, an honorary degree from Yale, and letters of support from Supreme Court Justices Black and Douglas—so President Kennedy nominated him for a seat on the D.C. Circuit Court of Appeals. By the time Skelly Wright began serving in that position in 1962, he possessed a leader's profile, and it was natural for other judges to take account of his opinions. A less costly path to judicial leadership is through academic prominence.[8]

For our purposes, what matters is the fact of leadership, not how the judge got there (though we briefly discuss the latter's potential strategic implications in the Conclusion). Thus, a judge whose opinions often make national news is a simple example of a leader judge in our setting,

[8]For example, Judge Richard Posner of the Seventh Circuit Court of Appeals and Judge Guido Calabresi of the Second Circuit were well-known pioneers of the law and economics movement before appointment to the bench (e.g., Posner 1973; Calabresi 1970). Judge Frank Easterbrook had been prominent both as an academic and as an advocate before appointment to the Seventh Circuit. Of course, these judges later cemented their leadership position by distinguished judicial output.

because those opinions likely become common knowledge in the judiciary. The visibility of a leader judge makes his opinions useful as a device for strategic coordination. An ordinary judge may voluntarily adopt the leader judge's legal opinions, knowing that other ordinary judges may do so as well. Thus, a key aspect of the spread of legal innovations is whether and when the leader judge tends to announce his opinions and reasoning.

Analysis with Leader Judges. To study these questions, we extend the model by positing that, in addition to regular judges discussed above, there is a prominent judge (the "leader judge") with three characteristics: (1) he is able to obtain more accurate information about the correct law (θ), at a cost, (2) his obtained information is publicly observed by other judges in the judiciary, and (3) he cares about the judiciary as a whole.[9] The second characteristic captures the visibility attribute of leader judges that we discussed above. Our model assumes that the leader judge's opinion readily becomes common knowledge in the judiciary. One can imagine a more general model in which ordinary judges, at some cost, can observe each other's opinions (signals) with some noise, and they can observe the leader judge's opinion with a smaller noise at a smaller cost. Our setting corresponds to the case in which observing other ordinary judges' signals is too costly and noisy, while observing the leader judge's signal is not noisy and has negligible costs. Thus, our model abstracts from these information acquisition features and from the endogenous emergence of leader judges, leaving them to future work.

In the modified game, first, the leader judge chooses a, the precision of the public signal y. Then, regular judges observe the public signal y and their private signals x_i, and simultaneously decide which action to take. The equilibrium concept is perfect Bayesian equilibrium.

The first step can be modeled in two ways—first, where the leader judge starts with a signal with a given precision and must decide whether to marginally improve it; second, where the leader judge decides the level of precision. (In both cases, acquiring precision may be costly.) Proposition 2 below is phrased in terms of marginal improvements to precision, but we also say more about levels of precision that may or may not be chosen in equilibrium.

The substantive interpretation of a is how thoroughly the leader judge goes into the legal issue occupying the judiciary. At the extreme, $a = 0$ can be interpreted as the leader judge not taking the case or writing an opinion that completely evades the issue by deciding the case on

[9]That is, even when the leader judge decides a case, the value of making the correct judgment in that one case is negligible to him relative to the effect of the information on the actions of many judges in the judiciary. To simplify the exposition, our model considers only two regular judges in the judiciary, but the qualitative results extend routinely to $N > 2$ judges with $u_i = -(1-r)\ (a_i - \theta)^2 - r\ \frac{\sum_{j \neq i}(a_i - a_j)^2}{N-1}$.

other grounds. More generally, higher a connotes an opinion that delves deeply into the judge's thinking about how the issue must be resolved, whereas lower a connotes a "narrow" opinion that barely engages with the issue and provides little guidance for others, as in our example with Judge Friendly.

Let R be the leader judge's weight on consistency, and $C(a)$ be the costs of obtaining the (public) signal y with precision a, with $C(0) = 0$, and $C'(a)$, $C''(a) > 0$. The leader judge's payoff U_L, from obtaining the (public) signal y with precision a is $U_L(a_1, a_2, \theta, a) = u_L(a_1, a_2, \theta) - C(a)$, with

$$u_L(a_1, a_2, \theta) = -(1 - R)\ [(a_1 - \theta)^2 + (a_2 - \theta)^2] - R\ [(a_1 - a_2)^2 + (a_2 - a_1)^2], \tag{4}$$

where we recognize that the equilibrium actions of other judges, a_1 and a_2, are influenced by the precision of the leader judge's public signal.

To understand the key strategic forces, we first consider the extreme cases of $R = 0$ and $R = 1$. When $R = 1$, the leader judge only cares about consistent application of the law throughout the judiciary: he simply wants all the judges to make the same decision ($a_1 = a_2$), whatever that decision may be. But ordinary judges also care about the correct (interpretation of the) law—they want their decisions to be close to θ. Critically, it is the judges' private information that causes their actions to diverge. If the judges did not have any private information, a public signal y would lead them to perfectly coordinate on action $a_i = E[\theta|y]$. Of course, the judges do have private information, but the more accurate is the public signal, the less weight the judges will put on their private information, leading to closer actions and more consistency. Thus, absent any direct cost of obtaining a precise public signal, a marginal increase in the precision of the public signal is always desirable for a leader judge who only cares about consistency.

Next, consider the case of $R = 0$, in which the leader judge only cares about correct law. From his perspective, judges should use all their information to take actions as close as possible to the expectation of θ. That is, he would have a judge i choose $a_i = E[\theta|y, x_i] = \frac{ay+bx_i}{a+b}$. But as Proposition 1 shows, the equilibrium action of judge i is $a_i(y, x_i) = \frac{ay+b(1-r)x_i}{a+b(1-r)}$: Because judges also care about coordinating, they underweight their private information, putting too much emphasis (from the perspective of a pure concern for correct law) on public information. Thus, from the leader judge's perspective, more precise public information is a double-edged sword: It is beneficial because more information allows judges to have a more accurate estimate of the law, but it can also be harmful because judges overreact to public information. As a result, from the leader judge's perspective, a somewhat informative public signal can be less desirable

than a completely uninformative public signal (or no public signal) with $a = 0$, because in the latter case the judges would not underweigh their private signals and would choose $a_i = x_i$.

To extend these intuitions and make them more precise, we must investigate the effect of higher public signal precision on the leader judge's expected payoff by tracing its influence on the equilibrium behavior of other judges. One can calculate the leader judge's marginal benefit of raising the public signal's precision in the absence of precision-improvement costs (see the proof of Proposition 2 in the Appendix):

$$\frac{\partial E[u_L]}{\partial a} = \frac{2(1-R)b}{[a+(1-r)b]^3} \left(\frac{a}{b} - f(r,R)\right), \tag{5}$$

where $f(R,r) \equiv -\frac{1-r}{1-R}\,[1-2r+(3-2r)R]$.

When the information advantage of the leader judge relative to other judges is high enough (i.e., when a/b is larger than $f(r,R)$), the leader judge always benefits from revealing marginally more precise information to other judges (absent information acquisition costs). How high is high enough? That depends on the weights, R and r, that the leader judge and other judges put on legal consistency and correctness, as these weights determine the threshold $f(r,R)$. When $f(r,R) < 0$, even when the leader judge's information is very imprecise ($a \approx 0$), he still benefits from revealing more precise information. However, when $f(r,R) > 0$ this is not true anymore, and $f(r,R) > 0$ if and only if $R < R^*(r) \equiv \frac{2r-1}{3-2r}$. Proposition 2 summarizes these results.

Proposition 2 *Even absent information acquisition costs, revealing marginally more precise information harms the leader judge if and only if the relative precision of his information is sufficiently low. Formally,*

$$\frac{\partial E[u_L]}{\partial a} < 0 \quad \Leftrightarrow \quad \frac{a}{b} < f(R,r).$$

A necessary condition is for the leader judge to care sufficiently more about correctness than about consistency. Formally, there is an increasing function $R^*(r)$ such that, for a given r, $f(r,R) > 0$ if and only if $R < R^*(r)$.

Proposition 2 reveals that even absent any direct costs, the leader judge would not disseminate marginally more precise (public) information whenever (1) he cares about the judiciary's best interpretation of the law sufficiently more than he cares about its consistency ($R < R^*(r)$),[10] *and* (2) he believes that his information will not be sufficiently precise relative

[10]That $R^*(r)$ is increasing reflects that as judges care more about consistency (as r increases), they overweight the public signal more, raising the costs (due to overreaction) of revealing public information.

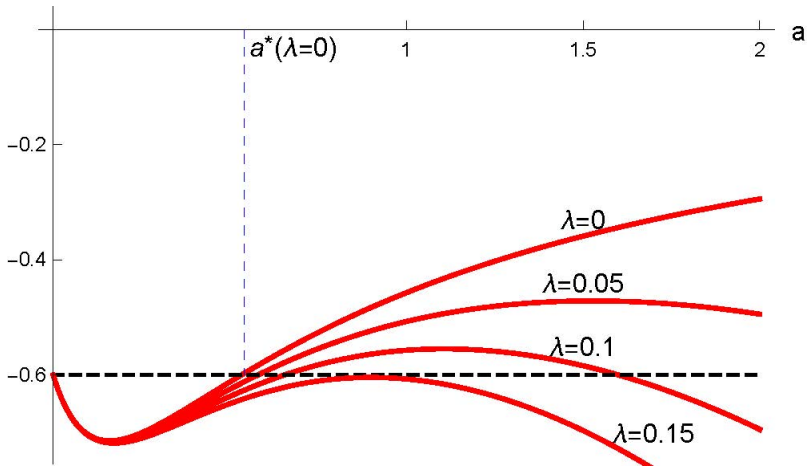

Figure 1: The leader judge's expected payoff as a function of the precision, a, of his public signal. Different curves illustrate different costs of obtaining more precise signals: $C(a) = \lambda a^2$. Even absent any direct costs of acquiring more precise information, there is a threshold $a^*(\lambda = 0)$ such that the leader judge prefers not to disseminate an informative public signal $(a = 0)$ rather than to disseminate a moderately informative public signal $a \in (0, a^*(\lambda = 0))$. Equation (6) specifies this threshold. Moreover, when $\lambda > \lambda^* \approx 0.15$, the cost of acquiring more precise information is sufficiently high that the leader judge does not acquire any informative public signal to disseminate.

to the information of other judges $(a/b < f(R, r))$. Then, the leader judge would not like the judges to observe an informative public signal unless it is sufficiently informative. As Figure 1 illustrates, the leader judge prefers a completely uninformative public signal (with $a = 0$) to any moderately informative signal with $a \in (0, a^*)$. The threshold a^* is the unique non-zero solution to $E[u_L(a)] = E[u_L(0)]$, which exists whenever $f(r, R) > 0$:

$$a^* = \frac{1 - R}{1 + R} \frac{f(r, R)}{1 - r}\, b. \tag{6}$$

Moreover, suppose $C(a) = \lambda c(a)$ with $\lambda > 0$, and $c(0) = c'(0) = 0$ and $c'(a),\ c''(a) > 0$ for $a > 0$. Parameter λ captures the "primeness" or "readiness" of the informational environment, with a high λ indicating that good information is costly to produce and a low λ indicating that it is cheap. For example, a case presenting an issue never thought of before might have a high λ whereas a case presenting an issue on which there has been much caselaw and scholarly commentary might have a low λ. (Alternatively, λ can be interpreted as capturing the competence of the leader judge, with lower λ indicating higher competence.) Focusing on the case with $f(r, R) > 0$, there exists a λ^* such that the leader judge chooses a positive a if and only if $\lambda < \lambda^*$. Figure 1 illustrates.

In sum, two broad conclusions emerge. When the leader judge cares about the correct ap-

plication of law sufficiently more than the consistent application of law (i.e., $R < R^*(r)$ so that $f(r, R) > 0$), a leader judge does not reveal information (send a public signal) unless (1) the signal is sufficiently precise (i.e., we will not observe $a < a^*$); *and* (2) the informational environment is sufficiently prime (i.e., $\lambda < \lambda^*$). This is not to say that less precise public signals are not valuable to the judiciary (in the first case), or that information revelation in less prime environments (with $\lambda > \lambda^*$) cannot be valuable to the judiciary (in the second case); rather, *from the leader judge's perspective*, the overreaction of the judiciary to his opinion outweighs the value of making informative statements. In failing to speak, leader judges deprive the judiciary of their guidance. But from the perspective of leader judges like Friendly, who put significant weight on the correctness of laws, "It is better to fail in this respect than to attempt to give guidance without having seen the variety of factual situations, having heard from the adversarial presentations, and having the benefit of the scholarly community which time will undoubtedly afford."

Heterogeneous values. The foregoing analysis assumes common values—meaning that although judges may differ in their private information about correct law, they all have the same underlying understanding of what "correct law" means. They might see the target differently, but they are all trying to find the same target (θ). In some contexts, for example in areas of law with high "ideological" content, it might be appropriate to assume more fundamental disagreement over values (i.e., what correct law means). Appendix III extends the model to such private value settings, showing that the main results carry over.

Interestingly, with private values, it is more difficult for our reticence effect to kick in. That is, it takes an even greater concern for correctness (as compared to consistency) for the leader judge not to speak (formally, $R_p^*(r) < R_c^*(r)$, where p and c denote private and common values respectively). This is because ordinary judges do not put as much weight on the public signal in the private value setting as in the common value setting—which, in turn, is due to the fact that in the former setting the public signal is valuable to ordinary judges only for coordination and not also for estimating correct law.

Conclusion

Formal analysis of courts has focused primarily on judges' interactions along the vertical dimension of the judicial hierarchy, where some judges have legal authority over others (Kastellec 2017). This paper explores the interactions between the vast majority of judges, who have

no authority relationship with one another, such as judges in different jurisdictions. Despite the absence of formal authority, such judges interact because they face a fundamental tension present in all common law judicial systems: the desire to employ law well-suited to "the felt necessities of the times" on the one hand and the desire to maintain consistency in adjudication on the other. We studied the strategic considerations that arise from these concerns and investigated their consequences for the horizontal spread of legal innovations. In particular, we showed that a desire for coordination leads judges to overweight publicly visible legal interpretations, including those enunciated by judges in a position of (informal) leadership. Knowing this, a leader judge who is concerned about the correct interpretation of law might refrain from communication, even when he has valuable information to communicate, because regular judges would place more weight on his opinion than its informational content merits. The greater the regular judges' concern for consistency, the greater the risk of overreliance, and therefore the larger the pool of leader judges who refrain from innovation. And the greater the leader judge's concern for correctness (vis-á-vis consistency), the greater his aversion to coordination on an incorrect doctrine, and therefore the greater the area of the parameter space for which he will not innovate. Our analysis sheds light on how the twin concerns for correctness and consistency influence the initiation and spread of legal innovation. We have pinpointed an important strategic reason for writing narrow opinions. And we have supplied a new informational-coordinative explanation for the stickiness of precedent and the practice of stare decisis.

Several directions for future research stand out. One could combine the interactions of judges along both horizontal and vertical dimensions, for example judges who interact both with peers in different jurisdictions and with superiors in the judicial hierarchy. Another direction is to incorporate judicial interactions on a network. Network structure may facilitate or hinder the spread of information depending on the location of judges relative to each other. Caldeira (1985) and Bird and Smythe (2012) offer suggestive empirical findings; adapting a computational approach to study networks (Siegel 2011, 2013) can facilitate progress in this direction. A third direction would endogenize the emergence of leader judges. Our model is concerned with what leader judges do once they have become secure in their leadership perch, so the only considerations that guide their decision whether to innovate are the innovation's impacts on the correctness and consistency of law. But a leader judge in the making might also care about writing opinions that are apt to promote him to leadership. This may counteract the reticence-inducing effect of caring greatly about correctness if the judge wants to increase

his visibility, or may amplify the effect if the judge is keen to develop a reputation for speaking only when he knows the answer. The effect of such career concerns on the spread of legal innovations would be interesting to explore, and would speak to the legal literature on great judges (e.g., Posner 1990; Gunther 1994; Landes, Lessig, and Solimine 1998; Ursin 2009; Dorsen 2012). These directions, which can be explored theoretically and empirically, are left to the future.

Appendix I: More Examples

Example 3: *MacPherson v. Buick.* Does a lawsuit arising from injuries sustained by defectively manufactured products belong in (or, as the lawyers says, "sound in") contract or tort? If the action sounds in contract, then liability can attach only to a person with whom the injured party has a contract (the "privity of contract" requirement); but if it sounds in tort, then no privity of contract is required, and the manufacturer of a defective product may be liable to anyone who uses the product in a reasonable way. Thus, when deciding cases regarding injuries from defective products, judges must choose a location on a spectrum with contract requirements toward one end and tort requirements toward the other. The traditional view had been closer to the former. The contract-tort distinction assumed special importance in the early twentieth century with the rise of mass-produced goods.

In a landmark opinion, Judge Benjamin Cardozo of the New York Court of Appeals rejected the old common law notion that the duty of guarding against the dangers attendant to manufactured products "grows out of contract and nothing else," and "put the source of the obligation" in tort law (*MacPherson v. Buick Motor Co.*, 217 N.Y. 382, 390 (1916)). The manufacturer of a negligently made product is thus liable to one who is injured by it, "irrespective of contract" (*id.* at 389-90). In the context of the case, this new interpretation allowed a person who was injured while driving a Buick car to sue the Buick Motor Company itself, rather than being limited to suing the dealer from whom he had bought the car (and with whom he was in privity of contract).

The New York Court of Appeals was the leading state court during Cardozo's tenure, with several prominent judges on its bench (Posner 1990). Moreover, Cardozo was involved in founding the American Law Institute, and his innovations appeared in the Institute's restatements of tort law. A judge who was interested in advances in tort law would thus naturally turn to the New York Court of Appeals for guidance. Indeed, Graham's (2015) study shows that Cardozo's opinion in *MacPherson* had a significant impact on tort law as it was adopted by judges throughout the country. Today there is no doubt that a person injured by a defective product may recover from the manufacturer regardless of whether they have a contract.

Example 4: *Escola v. Coca Cola.* When should a manufacturer be held liable for an injury caused by its products? What should be the test of liability and who should carry the burden of proof? Toward one end of the spectrum, negligence is required for fault, and the burden of proving negligence is on the injured party. Toward the other end of the spectrum,

the manufacturer is strictly liable: negligence is irrelevant, and the manufacturer must pay for injuries from the reasonable use of its products. In a concurring opinion in a 1944 case, *Escola v. Coca Cola Bottling Co.*, Roger Traynor of the California Supreme Court argued that "a manufacturer of goods [should] be responsible for their quality regardless of negligence" (24 Cal. 2d 453, 463). Traynor justified his stance on the grounds that the manufacturer is in the best position to know about the dangers inherent in its products and how to reduce such dangers. "As handicrafts have been replaced by mass production," wrote Traynor, an ordinary customer has lost the ability to "investigate for himself the soundness of a product" (467). "Manufacturing processes, frequently valuable secrets, are ordinarily either inaccessible to or beyond the ken of the general public" (467). Therefore, fixing the responsibility for injury solely on the manufacturer "will most effectively reduce the hazards to life and health inherent in defective products that reach the market" (462). The majority had relied on a doctrine in tort law (*res ipsa loquitur*) that placed a *presumption* of liability on the manufacturer. But Traynor argued that such doctrines were insufficient because the manufacturer could rebut the presumption by introducing evidence that it took proper care in manufacturing the product, and "An injured person ... is not ordinarily in a position to refute such evidence or identify the cause of the defect, for he can hardly be familiar with the manufacturing process as the manufacturer himself is" (463).

Traynor made his arguments for strict liability again in 1949, 1958, and 1960. By the 1960s, California had become the leading state supreme court (Dear and Jessen 2007; Friedman et al. 1981), and Traynor was a prominent judge, soon to become chief justice of the California Supreme Court. Finally, in the landmark case *Greenman v. Yuba Power Products*, 59 Cal. 2d 57 (1963), the court as a whole adopted strict liability. Soon thereafter, the American Law Institute's *Restatement (Second) of Torts* (1965) adopted Traynor's arguments in *Escola*. Graham's (2015) data show that after Traynor's innovation was adopted in California, it spread rapidly to other states. According to Bird and Smythe (2012, 568), it has now been adopted in almost all states.

Example 5: *ProCD v. Zeidenberg* and *Hill v. Gateway.* Are consumers bound by "shrinkwrap" terms—that is, terms enclosed in standard-form documents, often exceedingly long, that the consumer does not see until after paying for the product? The question is of immense importance today, when many contracts do not follow the bargained-out offer-and-acceptance sequence imagined as the foundation of contract law but are instead standardized forms proffered by one party on a take-or-leave basis. The answer could be a categorical yes

or no, or it could be a more nuanced doctrine specifying what kinds of terms are enforceable under what circumstances and given what kinds of notice.

Judge Frank Easterbrook, one of the most prominent judges in recent decades (see footnote 10), answered yes in a pair of cases. In *ProCD, Inc. v. Zeidenberg*, 86 F.3d 1447 (7th Cir. 1996), the seller sued the buyer for violating the terms enclosed in the box containing the product. The district court held that the shrinkwrap agreement was not enforceable because its terms became known to the customer only *after* paying for the product and opening the box, and one cannot be bound by terms that are unknown at the time of purchase. Easterbrook reversed, holding that, given the small-print notice on the outside of the box that the transaction was subject to a license, the terms of the license became part of the purchase agreement. If the customer did not like the terms, he could return the product. "Notice on the outside, terms on the inside, and a right to return the software for a refund if the terms are unacceptable (a right that the license expressly extends), may be a means of doing business valuable to buyers and sellers alike" (*id.* at 1451). Easterbrook enumerated different purchases—airline tickets, insurance policies, concert tickets—that in his view followed this "valuable" model, noting that a requirement to recite all terms before purchase would crowd out more pertinent information and waste everybody's time (*id.*). In the end, the terms are as much part of the "product" as the contents of the package, and consumers are protected against bad products through "[c]ompetition among vendors, not judicial revision of a package's contents" (*id.* at 1453). *Hill v. Gateway 2000, Inc.*, 105 F.3d 1147 (7th Cir. 1997), involved similar facts, though this time there was no pre-purchase notice of additional terms. Easterbrook again overruled the district court to enforce the shrinkwrap terms, relying on *ProCD*.

Being a federal court's interpretation of Wisconsin contract law, *ProCD* is not only not binding on other states; it is not even a definitive pronouncement on Wisconsin law.[11] Nevertheless, the decision has had great impact, with Issacharoff and Marotta-Wurgler (forthcoming 2020-2021) calling it "without doubt the most influential opinion in this area of law." By the same authors' count, *ProCD* and *Gateway* are the most and second-most influential decisions on shrinkwrap contracts, as measured by both total and annual out-of-state citations.

This popularity, however, has not been accompanied by critical acclaim. Though the opinions are not without defenders (see Posner 2010), academic commentary has been widely critical

[11]Under the doctrine of *Erie Railroad Co. v. Tompkins*, 304 U.S. 64 (1938), state courts are the ultimate interpreters of state law, and the role of federal courts sitting in diversity is to predict the interpretation that the highest state court would choose.

(e.g., Pitet 1997; Bern 2004; White 2004). For one thing, commentators have found Easterbrook's analysis of the Uniform Commercial Code[12] to be sloppy. UCC § 2-207 deals with contract formation and contractual terms when the sale of goods is accompanied by terms asserted by one side but not clearly accepted by the other side. This provision, sometimes dubbed a "battle of the forms" provision, was summarily dismissed by Easterbrook as "irrelevant" because "our case has only one form" (*ProCD*, 86 F.3d at 1452; *accord Gateway*, 105 F.3d at 1150). But, as commentators have pointed out, § 2-207 is not limited to situations involving more than one form. And, although Easterbrook is correct that "a vendor, as master of the offer, may invite acceptance by conduct"—namely, by paying, reading the terms at home, and returning the goods if the terms are not acceptable—there was no such invitation in either case. According to the principles dictating when a contract is formed (UCC § 2-206), the most natural conclusion in the circumstances of both cases is that the contract was formed when the money changed hands. The contrary conclusion is especially startling in *Gateway*, where there was not even pre-purchase notice of the existence of additional terms (let alone their content), so Easterbrook's model of "notice on the outside, terms on the inside" does not apply.

In short, Easterbrook used a contrived reading of the UCC to pretend that the parties had, through their conduct, agreed on the shrinkwrap terms. It is important to note that, contrary to Easterbrook's suggestion, the absence of such agreement would not render the contract "term-free" (105 F.3d at 1148). Recourse would have to be had to contract doctrines that specify what terms govern in the absence of clear mutual assent, and under such doctrines the case might still have come out in the seller's favor (especially in *ProCD*). So, by insisting on the fiction that the parties had agreed to the shrinkwrap terms, Easterbrook gave sellers (who can, after all, easily provide clearer notice) lots of help they did not really need—and opened the door to the enforcement of terms more objectionable than those at issue in *ProCD* and *Gateway*.[13] Although Easterbrook's economic intuitions about the benefits of not requiring a full pre-purchase recitation of terms seem mostly sound, achieving those benefits did not necessitate the draconian rule that a seller can, with minimal or no notice, dictate its terms by making their rejection impossible except on the inconvenience of effectuating a prompt post-purchase return. Easterbrook's conjecture that consumers are protected by vendors competing over shrinkwrap

[12] A joint project of the National Conference of Commissioners on Uniform State Laws and the American Law Institute, the UCC is a model code drafted to harmonize and rationalize the law of sales and other commercial transactions. It was first published in 1952 and has now been adopted by all states except Louisiana.

[13] Unconscionable terms are not enforceable, but the doctrine of unconscionability outlaws very little.

terms seems questionable given the terms' low salience and high cost of deciphering, the well-known fact that nobody reads them (e.g., Bakos et al. 2014), and the systematic differences in purchase terms of identical goods offered in-store and online (Marks 2017).[14] Today *Gateway* is often taught the first-year contracts class as a flawed opinion.

[14]Indeed, the conjecture proves too much: The same logic would dictate, for example, that there should be no implied warranty of merchantability in the sale of goods, nor an implied warranty of habitability in residential leases.

Appendix II: Proofs

Proof of Proposition 1: From equation (1), player i's expected payoff given his information $I_i = (x_i, y)$ is:

$$E[u_i(a_1, a_2, \theta)|I_i] = -(1-r)\ (a_i^2 - 2a_i\ E[\theta|I_i] + E[\theta^2|I_i]) - r\ (a_i^2 - 2a_i\ E[a_j|I_i] + E[a_j^2|I_i]).$$

Player i chooses his action a_i to maximize this expected payoff. The first order condition is:

$$\frac{\partial E[u_i(a_1, a_2, \theta)|I_i]}{\partial a_i} = -(1-r)\ 2(a_i - E[\theta|I_i]) - r\ 2(a_i - E[a_j|I_i]) = 0,$$

so player i's best response is:

$$a_i(I_i) = (1-r)\ E[\theta|I_i] + r\ E[a_j|I_i]. \tag{7}$$

Recall that $a = \frac{1}{\sigma_\eta^2}$ is the precision of public signal and $b = \frac{1}{\sigma_\epsilon^2}$ is the precision of private signals. Then, from the properties of Normal distribution,

$$E[\theta|I_i] = E[x_j|x_i, y] = \frac{ay + bx_i}{a+b},\ \text{for } j \neq i. \tag{8}$$

Moreover, given the linear strategy of judge j, $a_j(I_j) = kx_j + (1-k)y$,

$$E[a_j|I_i] = E[kx_j + (1-k)y|x_i, y] = k\ E[x_j|x_i, y] + (1-k)\ y = k\ \frac{ay + bx_i}{a+b} + (1-k)y. \tag{9}$$

Substituting from (8) and (9) into (7) and collecting the coefficients of x_i and y yields:

$$a_i(I_i) = \frac{b(rk + 1 - r)}{a+b}\ x_i + \left(1 - \frac{b(rk+1-r)}{a+b}\right)\ y. \tag{10}$$

By comparing the coefficients of x_i and y from equations (2) and (10), we can solve for k:

$$k = \frac{b(rk+1-r)}{a+b},\ \text{and hence } k = \frac{b(1-r)}{a + b(1-r)}.$$

Substituting this equilibrium value of k into $a_i(I_i)$ in equation (2) and simplifying yields: $a_i(I_i) = \frac{ay + b(1-r)x_i}{a+b(1-r)}$. □

Proof of Proposition 2: To ease calculations, it helps to write u_L in terms of θ and the noise in public and private signals. From equation (3), recall that, given a public signal y and his private signal x_i, judge i's equilibrium strategy is $a_i(I_i) = \frac{ay + b(1-r)x_i}{a+b(1-r)}$. Substituting $y = \theta + \eta$

and $x_i = \theta + \epsilon_i$ into this equilibrium strategy yields $a_i(y, x_i) = \theta + \frac{a\eta + b(1-r)\epsilon_i}{a+b(1-r)}$. Substituting from this into (4) yields:

$$\begin{aligned} u_L(a_1, a_2, \theta) &= -(1-R)\left[\left(\frac{a\eta + b(1-r)\epsilon_1}{a+b(1-r)}\right)^2 + \left(\frac{a\eta + b(1-r)\epsilon_2}{a+b(1-r)}\right)^2\right] \\ &\quad -2R\left(\frac{a\eta + b(1-r)\epsilon_1}{a+b(1-r)} - \frac{a\eta + b(1-r)\epsilon_2}{a+b(1-r)}\right)^2. \end{aligned}$$

Because η and ϵ_is are independent, we have $E[\eta\epsilon_i] = E[\epsilon_1\epsilon_2] = 0$, and hence:

$$\begin{aligned} \frac{E[u_L]}{2} &= -(1-R)\,\frac{a^2\,E[\eta^2] + b^2(1-r)^2\,E[\epsilon_i^2]}{[a+b(1-r)]^2} - R\,\frac{2b^2(1-r)^2\,E[\epsilon_i^2]}{[a+b(1-r)]^2} \\ &= -(1-R)\,\frac{a+b(1-r)^2}{[a+b(1-r)]^2} - R\,\frac{2b(1-r)^2}{[a+b(1-r)]^2}, \end{aligned} \tag{11}$$

where the second equality follows by recognizing that $E[\eta^2] = \frac{1}{a}$ and $E[\epsilon_i^2] = \frac{1}{b}$. Equation (11) allows us to study the effect of increases in the precision of the leader judge's public signal on the judiciary as a whole and hence on the leader judge's expected payoff. Differentiating $E[u_L]$ with respect to a and simplifying the result yields:

$$\frac{1}{2}\frac{\partial E[u_L]}{\partial a} = \frac{1-R}{[a+(1-r)b]^3}\,[a - f(r,R)\,b],$$

where $f(R,r) \equiv -\frac{(1-r)\,[1-2r+(3-2r)R]}{1-R}$. Thus, if $[1-2r+(3-2r)R] \geq 0$, then $\frac{dE[u_L]}{da} > 0$. In this case, absent the costs of acquiring more precise information, the leader judge would always like to obtain more precise information. In contrast, if $[1-2r+(3-2r)R] < 0$, then $f(r,R) > 0$, and the leader judge would like more precise public information if and only if $a > f(r,R)\,b$. Finally, observe that $[1-2r+(3-2r)R] < 0$ if and only if $R < R^*(r) \equiv \frac{2r-1}{3-2r}$. □

Appendix III: Heterogeneous Judges

In the text, we considered a common value setting, so that all judges want to target the same state θ. Here, we show that results and tradeoffs qualitatively extend to private value settings, in which judges differ on the correct law. Consider the setting analyzed in the text, but alter judge i's payoff to:

$$u_i(a_i, a_j, x_i) = -(1-r)\ (a_i - x_i)^2 - r\ (a_i - a_j)^2,\ i \neq j,\ r \in (0,1). \tag{12}$$

Recall that $x_i = \theta + \epsilon_i$, where $\epsilon_i \sim N(0, \sigma_\epsilon^2)$ and the state of the world θ is unknown to judges and they have a common (improper) prior that it is distributed uniformly on $\mathbb{R}$. Thus, judges know their own view of correct law, but their views differ (e.g., they value equity and efficiency differently): to judge i, the correct law is x_i. That is, judges have different, private, but correlated preferences for correct law. In some contexts, one might call x_i a judge's "ideology." Although judges have private values, they still use the public signal y to learn about other judges' private value and hence behavior. Mirroring the analysis of Proposition 1, we have:

Proposition 3 *There is a unique equilibrium in which a judge i chooses an action*

$$a_i^{private} = \frac{ray + (1-r)(a+b)x_i}{a + b(1-r)}. \tag{13}$$

When the public signal is uninformative ($a \to 0$) or when judges do not care about coordination ($r \to 0$), each judge chooses what he considers the correct law, i.e., x_i. In the other extreme, when the public signal is very accurate ($a \to \infty$) or a judge's preference for correct law x_i is uninformative about the other's ($b \to 0$), each judge chooses $ry + (1-r)x_i$. In equilibrium, a judge will never put a weight more than r on the public signal, because the public signal matters only for coordination and the judges' weight on coordination is r. This contrasts with the common value setting, in which a judge is also uncertain about correct law, and the public signal is also valuable in estimating the correct law. Then, if either $b \to 0$ or $a \to \infty$, judges choose y.

Next consider the setting with a leader judge. The leader judge's preferences remain the same as before: he wants judges to target θ and be close to each other. That is, the leader judge believes that the correct law is the expected value of what a typical judge would consider the correct law: $E[x_i] = \theta$.[15] Mirroring the analysis of Proposition 2, we have:

$$\frac{E[u_L(a)]}{2} = -(1-R)\ \frac{r^2 a + (1-r)^2(a+b)^2/b}{(a+b(1-r))^2} - R\ \frac{2(1-r)^2(a+b)^2/b}{(a+b(1-r))^2}.$$

[15] With a continuum of judges over a unit interval, θ is the average of the judges' preferences: $\theta = \int_i x_i di$.

Differentiating with respect to a yields:

$$\frac{1}{2r}\frac{\partial E[u_L(a)]}{\partial a} = \frac{a(2-3r+2r^2+(2-5r+2r^2)R) - b(1-r)(-2+3r-(2-r)R)}{(a+(1-r)b)^3}.$$

Because the coefficient of a is always strictly positive for $r < 1$, we have

$$\frac{\partial E[u_L(a)]}{\partial a} < 0 \quad \Leftrightarrow \quad \frac{a}{b} < \frac{(1-r)(-2+3r-(2-r)R)}{2-3r+2r^2+(2-5r+2r^2)R}.$$

Thus, we have a counterpart to Proposition 2:

Proposition 4 *Even absent information acquisition costs, revealing marginally more precise information harms the leader judge if and only if the relative precision of his information is sufficiently low. Formally,*

$$\frac{\partial E[u_L]}{\partial a} < 0 \quad \Leftrightarrow \quad \frac{a}{b} < f_p(R,r),$$

where $f_p(R,r) = \frac{(1-r)(-2+3r-(2-r)R)}{2-3r+2r^2+(2-5r+2r^2)R}$. *A necessary condition is for the leader judge to care sufficiently more about correctness than about consistency. Formally, there is an increasing function* $R_p^*(r) = \frac{3r-2}{2-r}$ *such that, for a given* r, $f_p(r,R) > 0$ *if and only if* $R < R_p^*(r)$.

From the proof of Proposition 2, $R^*(r) = \frac{2r-1}{3-2r}$. Thus,

Corollary 1 $R_p^*(r) < R^*(r)$.

That is, it is more difficult for our reticence effect to arise in the private value setting. This reflects, in part, that judges do not react to public signals as much in the private value setting as in the common value setting—and sometimes under-react. In particular, from Propositions 1 and 3, we have:

$$a_i^{private} = \frac{ray + (1-r)(a+b)x_i}{a+b(1-r)} \quad , \quad a_i^{common} = \frac{ay + (1-r)bx_i}{a+b(1-r)}.$$

Now, consider a leader judge who only cares about the correct law, so that $R = 0$. From his perspective, the appropriate weight on the public signal is $a/(a+b)$. In the common value setting, judges always put a higher weight on the public signal: $\frac{a}{a+b(1-r)} > \frac{a}{a+b}$ for $r > 0$. In contrast, in the private value setting, judges put a higher weight on the public signal than the leader judge desires whenever $\frac{ra}{a+b(1-r)} > \frac{a}{a+b}$, i.e., $\frac{a}{b} < \frac{2r-1}{1-r}$. Indeed, in Proposition 4, $f_p(R=0,r) < \frac{2r-1}{1-r}$, so we arrive at the expected conclusion that for a marginal increase in public signal precision to harm the leader judge, a necessary condition is that judges overreact to the public signal.

References

Acemoglu, Daron, and Matthew O. Jackson. 2017. "Social Norms and the Enforcement of Laws." *Journal of the European Economic Association* 12: 245-95.

Angeletos, George-Marios, and Chen Lian. 2016. "Incomplete Information in Macroeconomics: Accommodating Frictions in Coordination." *Handbook of Macroeconomics, Vol. 2*: 1065–1240.

Baker, Scott, and Claudio Mezzetti. 2012. "A Theory of Rational Jurisprudence." *Journal of Political Economy* 120 (3): 513-51.

Bakos, Yannis, Florencia Marotta-Wurgler, and David Trossen. 2014. "Does Anyone Read the Fine Print? Consumer Attention to Standard-Form Contracts." *Journal of Legal Studies* 43: 1-35.

Bass, Jack. 1981. *Unlikely Heroes*. New York, NY: Simon & Schuster.

Beim, Deborah, Alexander Hirsch, and Jonathan Kastellec. 2014. "Whistleblowing and Compliance in the Judicial Hierarchy." *American Journal of Political Science* 58 (4): 904-18.

Bern, Roger. 2004. "'Terms Later' Contracting: Bad Economics, Bad Morals, and a Bad Idea for a Uniform Law, Judge Easterbrook Notwithstanding." *Journal of Law and Policy* 12: 641-796

Bernick, Michael. 1980. "The Unusual Odyssey of J. Skelly Wright." *Hastings Constitutional Law Quarterly* 7: 971-999.

Bickel, Alexander. 1962. *The Least Dangerous Branch: The Supreme Court at the Bar of Politics*. New York, NY: Bobbs-Merrill.

Bird, Robert C., and Donald J. Smythe. 2012. "Social Network Analysis and the Diffusion of the Strict Liability Rule for Manufacturing Defects, 1963-87." *Law and Social Inquiry* 37 (3): 565-94.

Bolton, Patrick, Markus K. Brunnermeier, and Laura Veldkamp. 2013. "Leadership, Coordination and Corporate Culture." *Review of Economic Studies* 80: 512-37.

Bueno de Mesquita, Ethan, and Mehdi Shadmehr. 2020. "Social Norms and Social Change." Mimeo.

Bueno de Mesquita, Ethan, and Matthew Stephenson. 2002. "Informative Precedent and Intrajudicial Communication." *American Political Science Review* 96: 755-66.

Calabresi, Guido. 1970. *The Costs of Accidents: A Legal and Economic Analysis*. New Haven,

CT: Yale University Press.

Caldeira, Gregory A. 1985. "The Transmission of Legal Precedent: A Study of State Supreme Courts." *American Political Science Review* 79 (1): 178-94.

Cameron, Charles, Jeffrey Segal, and Donald Songer. 2000. "Strategic Auditing in a Political Hierarchy: An Informational Model of the Supreme Court's Certiorari Decisions." *American Political Science Review* 94 (1): 101-16.

Cardozo, Benjamin. 1921. *The Nature of the Judicial Process.* New Haven, CT: Yale University Press.

Carrubba, Clifford, and Tom Clark. 2012. "Rule Creation in a Political Hierarchy." *American Political Science Review* 106 (3): 622-43.

Chused, Richard. 2004. "Saunders (a.k.a. Javins) v. First National Realty Corporation." *Georgetown Law Journal on Poverty Law & Policy* 11:191-247.

Daughety, Andrew F. and Jennifer F. Reinganum. 1999. "Stampede to Judgment: Persuasive Influence and Herding Behavior by Courts." *American Law and Economics Review* 1:158-189.

Dear, Jake, and Edward W. Jessen. 2007. "'Followed Rates' and Leading State Cases, 1940-2005." *UC Davis Law Review* 41:683-711.

Dewan, Torun, and David P. Myatt. 2008. "The Qualities of Leadership: Direction, Communication, and Obfuscation." *American Political Science Review* 102: 351-68.

Dorsen, David M. 2012. *Henry Friendly: Greatest Judge of His Era.* Cambridge, MA: Belknap Press of Harvard University Press.

Fox, Justin, and Matthew Stephenson. 2011. "Judicial Review as a Response to Political Posturing." *American Political Science Review* 105: 397-414.

Friedman, Lawrence, Robert Kagin, Bliss Cartwright, and Stanton Wheeler. 1981. "State Supreme Courts: A Century of Style and Citation." *Stanford Law Review* 33 (5): 773-818.

Graham, Kyle. 2015. "The Diffusion of Doctrinal Innovations in Tort Law." *Marquette Law Review* 99 (1): 74-177.

Gunther, Gerald. 1994. *Learned Hand: The Man and the Judge.* New York, NY: Alfred A. Knopf.

Hadfield, Gillian and Barry Weingast. 2012. "What Is Law? A Coordination Model of the Characteristics of Legal Order." *Journal of Legal Analysis* 4(2): 471–514.

Holmes, Oliver Wendell, Jr. 1881. *The Common Law.* Boston, MA: Little, Brown & Co.; Repr. The Lawbook Exchange, Ltd., 2009.

Kastellec, Jonathan. 2017. "The Judicial Hierarchy: A Review Essay." *Oxford Research Encyclopedia of Politics.*

Issacharoff, Samuel and Florencia Marotta-Wurgler. Forthcoming 2020-2021. "The Hollowed Out Common Law." *UCLA Law Review.* `https://ssrn.com/abstract=3261372`.

Kornhauser, Lewis A. 1992. "Modeling Collegial Courts II: Legal Doctrine." *Journal of Law, Economics, and Organization* 8: 441-70.

Landa, Dimitri, and Scott Tyson. 2017. "Coercive Leadership." *American Journal of Political Science* 61: 559-74.

Landes, William, Lawrence Lessig, and Michael Solimine. 1998. "Judicial Influence: A Citation Analysis of Federal Courts of Appeals Judges." *Journal of Legal Studies* 27 (2): 271-332.

Mailath, George, Stephen Morris, and Andrew Postlewaite. 2017. "Laws and Authority." *Research in Economics* 71: 32-42.

Marks, Colin. 2017. "Online and 'As Is.' " *Pepperdine Law Review* 45: 1-53.

Miller, Arthur Selwyn. 1984. *A "Capacity for Outrage": The Judicial Odyssey of J. Skelly Wright.* Westport, CT: Greenwood Press.

Morris, Stephen, and Hyun Song Shin. 2002. "The Social Value of Public Information." *American Economic Review* 92: 1521-34.

Pitet, Christopher. 1997. "The Problem with Money Now, Terms Later: *ProCd, Inc. v. Zeidenberg* and the Enforceability of Shrinkwrap Software Licenses." *Loyola of Los Angeles Law Review* 31: 325-352.

Posner, Eric. 2010. "*ProCD v. Zeidenberg* and Cognitive Overload in Contractual Bargaining." *University of Chicago Law Review* 77: 1181-1194.

Posner, Richard. 1973. *Economic Analysis of Law.* Boston, MA: Little, Brown and Company.

Posner, Richard. 1990. *Cardozo: A Study in Reputation.* Chicago, IL: University of Chicago Press.

Rabin, Edward. 1984. "The Revolution in Residential Landlord-Tenant Law: Causes and Consequences." *Cornell Law Review* 69: 517-84.

Shahshahani, Sepehr. Forthcoming 2021. "The Fact-Law Distinction: Strategic Factfinding

and Lawmaking in the Judicial Hierarchy." *Journal of Law, Economics, & Organization* 37. https://ssrn.com/abstract=3189833.

Siegel, David. 2011. "When Does Repression Work? Collective Action in Social Networks." *Journal of Politics* 73 (4): 993-1010.

Siegel, David. 2013. "Social networks and the mass media." *American Political Science Review* 107 (4): 786-805.

Staton, Jeffrey, and Georg Vanberg. 2008. "The Value of Vagueness: Delegation, Defiance, and Judicial Opinions." *American Journal of Political Science* 52: 504-19.

Talley, Eric. 1999. "Precedential Cascades: An Appraisal." *Southern California Law Review* 73: 87-137.

Urofsky, Melvin. 1985. "The Brandeis-Frankfurter Conversations." *Supreme Court Review* 1985:299-339.

Ursin, Edwin. 2009. "How Great Judges Think: Judges Richard Posner, Henry Friendly, and Roger Traynor on Judicial Lawmaking." *Buffalo Law Review* 57: 1267-1360.

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

White, James. 2004. "Contracting under Amended 2-207." *Wisconsin Law Review* 2004: 723-751.