American Economic Journal: Microeconomics 2017, 9(2): 188–225
https://doi.org/10.1257/mic.20140272

Disclosure and Legal Advice†

By Yeon-Koo Che and Sergei Severinov*

This paper examines how the advice that lawyers provide to their clients affects the disclosure of evidence and the outcome of adjudication, and how the adjudicator should allocate the burden of proof in light of these effects. Despite lawyers' expertise in assessing the evidence, their advice is found to have no effect on adjudication if the lawyers follow the strategies of disclosing all favorable evidence. A lawyer's advice can *influence the outcome in his client's favor, either if (s)he can credibly advise his client to suppress some favorable evidence or if legal advice is costly. The effect is socially undesirable in the former case, but it is desirable in the latter case. Our results provide a general perspective for understanding the role of private information and expert advice in disclosure.* (*JEL* C72, D71, D72, D82)

Lawyers play a prominent role in the modern day adjudication process. One notable aspect of their role involves advising clients on disclosing information to the court. Lawyers can advise their clients which evidence is unfavorable and should be withheld and which evidence is favorable and thus should be disclosed. Although lawyers often have a disclosure duty before the tribunal (particularly in civil cases), the rules of confidentiality and attorney-client privilege enable them to suppress evidence during discovery and trial, particularly when the opposing party and the tribunal are unaware of the existence of the evidence.[1] The goal of this paper is to understand whether and to what extent the lawyers can affect the outcome of a trial by influencing the amount and the nature of information reaching the court.

To this end, we study adjudication of a dispute by a judge between two parties, say defendant and a plaintiff, who may obtain legal advice. Formally, the dispute is modeled as an evidence disclosure game. The judge decides whether to "convict" or "acquit" the defendant based on all information available to her. Part of the judgment-relevant information is the *evidence* the parties themselves may (or may

*Che: Department of Economics, Columbia University, 420 West 118th Street, 1029 IAB, New York, NY 10027 (e-mail: yc2271@columbia.edu); Severinov: Vancouver School of Economics, University of British Columbia, 6000 Iona Drive, Vancouver, British Columbia V6T1L4 (e-mail: sseverinov@gmail.com). The authors thank Andrew Daughety, Bob Hall, Ken Judd, Navin Kartik, Bart Lipman, Mitchell Polinsky, and Kathy Spier, seminar participants at the University of Arizona, Stanford Law, and Economics Seminar and numerous conferences for helpful comments. We also thank an anonymous referee for helpful suggestions. Sergei Severinov acknowledges financial support from Social Sciences and Humanities Research Council of Canada through SSHRC grant # 410-2009-1070. Yeon-Koo Che was supported by the Global Research Network program through the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF project number 2016S1A2A2912564). The standard disclaimer applies.

†Go to https://doi.org/10.1257/mic.20140272 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

[1] The attorney-client privilege protects privileged information in testimony at trial. Federal Rules of Civil Procedure (Rules 26(b)(1) and 26(b)(3)) limit discovery of privileged information and trial preparation materials.

not) possess. The evidence is observed privately by the possessing party but can be disclosed verifiably to the court. The main strategic decision for a party is whether to disclose evidence truthfully or to withhold it. The judge's ruling depends also on another piece of information, *legal strength* of the case, which reflects the legal rules and standards of interpreting that evidence and other public evidence surrounding that case. The legal strength of the case is observed only by the judge and by the lawyer a party may retain. Thus, a lawyer can assess precisely whether a party's evidence is favorable or unfavorable and how strong his case would be without its disclosure. A party advised by a lawyer can thus make a more informed decision about disclosure.

If the evidence is disclosed by either party, then the judge rules on the basis of the evidence and the legal strength. If the evidence is not disclosed, the judge rules solely on the basis of the legal strength and her belief about the evidence given the disclosure strategies the parties are perceived to employ. The advice of laywers influences the latter, and may potentially affect the judge's ruling in case neither party discloses evidence. Our goal is to understand the equilibrium outcome in terms of parties' disclosure behavior and the judge's ruling, with and without lawyer advising.

The resulting model introduces rich strategic interactions in a disclosure game. First, the lack of common knowledge about the existence of evidence makes the judge's inference nontrivial, since nondisclosure need not imply a party's concealment of unfavorable information. Hence, an equilibrium typically would not involve full disclosure, much in contrast to the unraveling that is typical in verifiable disclosure games (see Milgrom 1981 and Grossman 1981). Second, a judge's inference is influenced by his perception of each party's disclosure strategy, and the latter depends on whether the party has obtained legal advice or not. In this sense, lawyer advising adds a new dimension both to the parties' strategic disclosure and the quality of the judge's inference and her ruling.

Finally, the act of seeking lawyer advice itself, especially when it is costly, may signal whether a party possesses relevant evidence, and if so, what that evidence may be. This signal generally influences the judge's inference and her ruling. Our model accommodates these strategic interactions.

Our model produces several surprising results about the role of advising on disclosure.[2] First, we find lawyer advising to be irrelevant—both privately and socially—under a baseline scenario where the advice is costless and a lawyer employs a disclosure strategy that satisfies a certain credibility requirement, that is, to disclose information if and only if it is favorable to her client. Such a disclosure strategy is weakly dominant for a lawyer-advised party. Therefore, there is a strong rationale for selecting this strategy and focusing on the resulting equilibrium outcome in case parties obtain legal advice. Such an ability to fine-tune disclosure based on the legal strength of the case is lacking for a party without legal advice.

[2] To understand these results and the various effects discussed below, it is important to focus on the parties' disclosure behavior and the judge's ruling in case neither party discloses the evidence. These two aspects completely characterize the outcome, since disclosure by at least one party gives the judge complete information about the case and eliminates all ambiguity about her ruling.

Hence, one may expect that an unadvised party will suffer a strictly worse ruling at least in some instances than if he can play the weakly dominant lawyer-advised disclosure strategy. Our surprising finding is that the equilibrium outcome in the lawyer-advised game is identical in every state of the world (i.e., for every realization of evidence, legal standard, and information possessed by the parties) to the unique equilibrium outcome of the disclosure game played by unrepresented parties.

A rough intuition, made precise in the text, is that in equilibrium the judge follows the same ruling strategy in case no evidence is disclosed, regardless of whether parties obtain legal advice or not. This strategy is characterized by a threshold, with the judge ruling for one party if the legal strength is above this threshold and for the other if the legal strength is below that threshold. Faced with this strategy of the judge, an unadvised party is able to neutralize any disadvantage relative to a lawyer-advised party, in the sense that he makes only irrelevant mistakes in disclosure: he withholds favorable evidence (unwittingly) only when the judge would rule favorably without disclosure and discloses unfavorable evidence (also unwittingly) only when withholding it would have led to the same unfavorable ruling by the judge.

We next extend our analysis to identify two scenarios in which legal advice regarding disclosure *does* affect the outcome of a trial. First, a lawyer's advice matters if he can credibly follow a strategy of suppressing some favorable evidence, when it does not harm the client. Such a strategy is weakly dominated by the strategy of disclosing all favorable evidence. However, by following this dominated strategy a lawyer can skew the inference by the judge and thereby her ruling in his favor in the event that neither party discloses evidence. Compared with the case in which all favorable evidence is disclosed, the judge's inference becomes more favorable toward the party who withholds some favorable evidence, since nondisclosure in the latter case is attributed less to the selective withholding of unfavorable evidence by this party.

Playing a weakly dominated strategy requires a certain degree of credibility, since a lawyer may "tremble at the last moment" and switch to a dominant strategy of disclosing all favorable evidence. If the judge suspects this, her inference and ruling will not be swayed and will remain the same as in the equilibrium where the dominant disclosure strategies are used. This highlights the importance of a lawyer's reputation, because arguably only lawyers who have strong reputations as "no-nonsense" litigators—that do not present redundant evidence—can credibly follow a strategy that involves sometimes suppressing favorable but redundant evidence. This role of lawyers generates a private incentive for hiring them.[3]

Lawyer advising can also affect the adjudication outcome when hiring a lawyer is costly. In this case, obtaining costly legal advice serves as a "signal" that can affect the judge's inference and change her ruling in case no evidence is disclosed. In particular, we demonstrate that the parties who do not possess any evidence (good or bad) incur this cost to credibly signal their lack of evidence and avoid a prejudicial inference by the adjudicator following their nondisclosure of evidence. Thus, hiring a lawyer buys "the right to be silent without prejudice" for those who have no evidence

[3] Welfare implications of the lawyers' disclosure strategies are studied in Che and Severinov (2014), where we show that the inefficiencies resulting therefrom can be remedied by placing restrictions on the judges' inferences and the allocations of the burden of proof.

to disclose. The parties with unfavorable evidence also hire lawyers and often end up withholding evidence. However, the parties with moderately unfavorable evidence—those who would seek legal advice had it been free—do not hire a lawyer and disclose their evidence. Overall, the cost of legal advice increases the disclosure of private information, which in turn improves the quality of adjudication—in fact, more so as the cost increases. Specifically, as the cost of legal advice increases, a party will opt for disclosure (instead of hiring a lawyer) for a greater range of evidence. This result hinges on the judge making a negative inference about the party who neither discloses nor retains a lawyer. Such a negative inference appears plausible, as the judge may regard an unrepresented party as less reliable, and therefore attribute such a party's nondisclosure to her withholding unfavorable evidence.

Our analysis has several notable implications. First, our model provides a useful framework for analyzing the advisory role of lawyers in dispute resolution. Admittedly, legal representation in the real world includes several aspects not captured in our simple model. Yet, the advisory role of lawyers in disclosure is an important one, and our model identifies ways in which this role may (or may not) affect the outcome of adjudication. In this sense, our model can serve as a useful benchmark for studying different aspects of legal representation.

Furthermore, as we argue in the next section, our analysis applies beyond the comparison of self-representation and legal representation regimes, and more broadly to regimes differing in terms of the quality of legal advice. Therefore, our results can also be used to inform the choice of the quality of lawyers and legal advice.

Our modeling framework and the results are useful for understanding the role of advising more broadly, in settings other than dispute resolution. In particular, the insights we develop hold equally well when there is only one party. This is notable because in a number of situations decisions that have significant consequences for a party must be made based on the information provided by that party. Promotion and grant allocation, college admission, and job application are relevant examples. The applicants in these and similar contexts often seek advice from mentors, counselors or consultants regarding strategies of information revelation, and evaluators make decisions based on filtered information, often aware of the advising that may have led to the filtering. Our results offer insights regarding the effect of such advised disclosure.

The issue of legal advice has received relatively little formal treatment in the literature. Legal scholars have recognized the factors favoring and disfavoring the lawyer-aided adversarial system but disagree on the relative importance of those factors. Proponents argue that vigorous adversarial competition among lawyers leads the court to focus on relevant evidence, thus making judicial fact-finding efficient (Luban 1983, Bundy and Elhauge 1991, 1993). Critics point out that lawyers can mislead as much as inform the court (Frank 1973). In particular, Kaplow and Shavell (1989) point out that while the lawyers' ability to suppress evidence based on legal expertise undoubtedly benefits their clients, its social implications are ambiguous. Although the current paper is similar in spirit to the last study, there are important distinctions. First, these authors do not perform a full-fledged equilibrium analysis, focusing instead on the effect of legal advice when possible outcomes are exogenously fixed. Second, they treat the adjudicators' inferences as exogenous, while we allow the inferences to depend on the players' strategies. In another related contribution, Iossa and Jullien (2012) study the market for lawyers focusing on the match between the nature of the legal dispute and the quality of the lawyers hired by the litigants, as well as on the effect of the lawyer's reputation on the adjudicators.

On the theoretical side, our paper contributes to the literature on verifiable disclosure games. The literature originated from the seminal contributions by Grossman (1981), Milgrom (1981), and Milgrom and Roberts (1986), and was further developed by Lipman and Seppi (1995) and Seidmann and Winter (1997). The key result of this literature is the so-called "unraveling," namely that conflicting interests can lead to full revelation of the parties' private, but non-falsifiable, information. The common knowledge of an agent's possession of information is crucial for this result. We relax this common knowledge assumption, as in Verrecchia (1983) and Shin (1994, 1998). As mentioned above, the relaxation of common knowledge makes the judge's inference nontrivial. Lewis and Poitevin (1997) study disclosure of verifiable information in regulatory proceedings. More recently, Kartik, Suen, and Xu (2013) analyze disclosure of verifiable information in the context of a persuasion game. Bull and Watson (2004, 2007) and Deneckere and Severinov (2008) study disclosure of hard evidence in the mechanism design setting.

Our paper is also related to the literature on the adversarial system of disclosure. Dewatripont and Tirole (1999) study the desirability of the adversarial system in a broad organization design context. Shin (1998) compares adversarial and inquisitorial litigation systems. Froeb and Kobayashi (1996, 2001) assess the implications of endogenous evidence production. See also Daughety and Reinganum (2000). Sobel (1985); Hay and Spier (1997); Sanchirico (1998, 2000, 2001); and Bernardo, Talley, and Welch (2000) explore the allocation of the burden of proof and evidence production from the standpoint of litigation costs and deterrence. Sanchirico and Triantis (2008) study costly evidence fabrication and efficiency of adjudication in the context of contractual disputes.[4] Bull (2008) studies costly evidence production and disclosure in a complete information setting when hard evidence can be obtained at a moderate cost. Bull (2009) compares inquisitorial and adversarial litigation systems when a litigant may suppress evidence by incurring a cost. Sanchirico (2004) discusses evidence tempering and the laws against it. Che and Kartik (2009) examine incentives to acquire verifiable, but suppressible, information in a sender-receiver relationship. Shchepetova (2013) is a comparative study of costly evidence production in inquisitorial and adversarial systems.

Seidmann (2005), Mialon (2005), and Leshem (2010) investigate the effect of the defendant's right to silence, with and without adverse inference by the adjudicator, on the adjudication outcomes and welfare.[5]

In the follow-up paper, Turkay (2013) studies how the severity of legal punishment affects evidence disclosure behavior. Hadfield and Leshem (2012) provide a comprehensive review of the law and economics literature on attorney-client relationship and, in particular, the role of the confidentiality rules. None of these papers deal with the role of lawyers in disclosure—the focus of this paper.

[4] Levy (2005) studies the effect of career concerns on judges' decision making.

[5] Also related is the literature on cheap talk, which includes Crawford and Sobel (1982) and Krishna and Morgan (2001) among others.

I. Model

Two parties, 1 and 2, are in a dispute, which is adjudicated by an adjudicator in a tribunal. It is convenient to interpret parties 1 and 2 as a defendant and a plaintiff in a litigation. However, our model applies equally well to a number of different settings. The adjudicator in our model can be either a judge or a jury or a combined entity, whom we shall call simply "the judge" throughout.[6] Lawyers provide legal advice, if hired by the parties.

There are two pieces of judgment-relevant information that pertain to the case. First, there is *evidence* $s \in [0,1] =: S$, which may be observed only by the parties to the dispute. As will become clear, this variable is defined from party 2 (plaintiff)'s perspective, so the higher s is, the more favorable the evidence is for party 2. The evidence is observed with probability p_{00} by neither party, with probability p_{11} by both parties, and with probability p_{10} (respectively, p_{01}) by party 1 only (respectively, party 2 only).[7] Obviously, $\sum_{i,j=0,1} p_{ij} = 1$, and we assume that $p_{ij} > 0$ for all $i, j = 0, 1$. We allow for possible correlation in the parties' abilities to observe evidence. The evidence is "hard" in the sense that, while it can be concealed, it cannot be fabricated or manipulated. For instance, the evidence can take the form of an unforgeable document or a non-perjuring witness. Equivalently, the evidence may be soft but perjury laws prevent the possessor of the evidence from falsifying it. It is well known that the non-falsifiability of information leads to full revelation of information (Grossman 1981, Milgrom and Roberts 1986). Unraveling of this kind will not occur in our setting, however, since the possession of evidence is no longer common knowledge.

There is another piece of judgment-relevant information, *legal strength* $\theta \in [0,1] =: \Theta$ *of the case*, which is observed only by the lawyers and the judge. Similarly to the evidence, the higher θ is, the stronger party 2's case is from the legal standpoint. The legal strength θ of the case represents the judge's interpretation of the laws and legal standards, either universally or in application to the current case. Additionally, θ could also reflect the court's interpretation of external circumstances surrounding the case, such as basic uncontested facts, etc. Thus, when s is disclosed, the judge's ruling depends on both s and θ, and when s is not disclosed, the ruling depends only on θ.[8]

The disputing parties have limited knowledge of the law and incomplete understanding of the legal process, so they can learn the legal strength θ only by hiring lawyers. Lawyers understand the body of the law, as well as the judge's interpretation of the law and her possible biases. For instance, the lawyer and the judge may be able to assess more accurately how strong or weak the mitigating circumstances are. Ultimately, the lawyers' ability—and the litigants' *in*ability—to observe θ introduces a potentially useful role for the lawyers.

[6] For jury interpretation to apply, the jury must be given instructions regarding the content and application of the law by the judge.

[7] "Observing" s means either possessing that evidence or having proof of its existence.

[8] Posner (1999) discusses a class of "bare bones cases" in which very little evidence is presented by the parties, and the adjudicator has to rule on the basis of the law and a few uncontested facts. Such "bare bones" cases fit the description of situations where s is not disclosed.

We assume that (s, θ) is drawn from $S \times \Theta$ according to an absolutely continuous cdf, $F(s, \theta)$ with a positive density $f(s, \theta)$ in the interior of $S \times \Theta$. From an ex ante perspective, θ is random because the law and legal standards as well as the circumstances perceived by the court may vary across cases.

Concerning the probability distribution of (s, θ), two cases appear to be natural and intuitive: (i) s and θ are independently distributed since s embodies case evidence, and θ reflects the law; (ii) s and θ are positively correlated, or affiliated, since they pertain to the same case. This is particularly salient if we allow θ to also include legal interpretation of uncontested facts or circumstances. Fortunately, all the results of this paper hold in both of these cases. The statistical property which includes the case of independently distributed (s, θ) as well as positively correlated/affiliated ones is (weak) Monotone Likelihood Ratio Property (MLRP):

ASSUMPTION 1 (MLRP): *For all* $s' \geq s$ *and* $\theta' \geq \theta, f(s', \theta')/f(s, \theta') \geq f(s', \theta)/f(s, \theta)$.

To understand the value of legal advice, we will compare two regimes. In the first regime, the parties are not represented by lawyers and do not receive any legal advice. In the second regime, both parties are represented by lawyers, at no cost to them. Self-representation serves as a benchmark necessary for our analysis, but it is not without practical relevance. Although few parties represent themselves in civil or criminal trials in state or federal courts in the United States, many litigants do so in municipal courts and administrative trial procedures. In small claims courts—which comprise a significant share of trials in the United States—legal representation is expressly forbidden in most states (California, New York, Arizona, and others).[9] Macfarlane (2013) documents that self-represented litigants are increasingly common across various civil courts in Canada.

Further, our comparison should not be interpreted narrowly as pertaining only to these two regimes. Rather, it applies more broadly to the role of quality differences in legal advice. For this, both regimes should be interpreted as involving lawyer advising but of varying quality. Specifically, self-representation in our model corresponds to retaining a low quality, non-specialist lawyer, who provides incomplete advice, so that residual uncertainty about θ remains. Legal representation then corresponds to retaining a high-quality lawyer who provides a litigant with full information about θ.

The time line of the events in both regimes is as follows. At date 0, (s, θ) is realized. At date 1, parties 1 and 2 observe the evidence s with probabilities $p_{10} + p_{11}$ and $p_{01} + p_{11}$, respectively, while the judge and the lawyers learn θ. At date 2 a trial is held, at which parties 1 and 2 simultaneously and independently decide whether to disclose the evidence s to a judge, provided that the respective party has observed it. In the lawyer advising regime, this decision is made with the help of a lawyer providing legal advice. At the end of the trial or after it, the judge rules either for party 1 or for party 2.

[9] See Spurrier (1980) for details. The problem of withholding evidence is particularly relevant in small claims courts, since the discovery process is very limited and the trials focus on a few key elements of evidence.

Evidence Disclosure Behavior.—If a party does not hire a lawyer, then his decision to disclose s is based solely on s. In contrast, if a party hires a lawyer, he can make the disclosure decision based on the lawyer's advice, i.e., his knowledge of θ.

A lawyer prefers his client to prevail in court. There are no agency issues in the attorney-client relationship, so a client will communicate s to his lawyer truthfully, and the lawyer will explain the legal issues, i.e., communicate θ to the client truthfully. Therefore, a lawyer-advised party can simply be viewed as informed of both s (if he observes s) and θ.

Formally, party i's disclosure strategy is a function ρ_i mapping $S \times \Theta$ to $[0,1]$, with $\rho_i(s,\theta)$ representing the probability that party $i \in \{1,2\}$ discloses s for given θ. If a party is unrepresented, he does not observe θ, so $\rho_i(\,\cdot\,,\theta)$ must satisfy the condition $\rho_i(\,\cdot\,,\theta) = \rho_i(\,\cdot\,,\theta')$ for any θ, θ'.

Judge's Adjudication Behavior.—In the last stage of the game, the judge makes a binary decision, ruling either for party 1 or party 2. For instance, in a criminal trial the judge convicts or acquits the defendant. A binary decision is common, and is more general than may appear at first glance. For instance, there may be no ambiguity about the size of damages, leaving the liability as the only object of dispute.[10]

The judge's ruling depends on (s,θ) if s has been disclosed, and on the legal strength θ alone if s has not been disclosed. The judge's decision given (s,θ) is described by a function $g(s,\theta)$, interpreted as her assessment of party 1's (defendant's) culpability, or his "culpability index." If $g(s,\theta) > 0$, then the judge finds party 1 culpable and rules for party 2. If $g(s,\theta) < 0$, the judge finds party 1 innocent and rules for him. The judge is indifferent if $g(s,\theta) = 0$, but since the distribution $F(s,\theta)$ is absolutely continuous, how a tie is broken has no real consequence.

We assume that the function $g(\,\cdot\,,\cdot\,)$ is common knowledge between all players, including the lawyers and parties 1 and 2,[11] and that $g(s,\theta)$ is increasing and continuous in both arguments. So, lower s and θ are more favorable for party 1, and vice versa. In a tort setting, s may indicate the degree of defendant's responsibility in causing harm to the plaintiff, the extent of his negligence or of the plaintiff's contributory negligence.[12] Meanwhile, θ may correspond to the evidence standard for establishing causation and the liability rule adopted by the court, for instance, its willingness to apply strict liability versus negligence or to consider the defense based on the plaintiff's contributory negligence. It is not uncommon that the same evidence s could lead to the defendant being found liable in one case but not in the other depending on the particular evidence standard and the liability rule adopted by the court.

[10]The binary feature can also be justified in an idealistic Beckerian world in which any defendant found liable is subject to a sanction equal to his maximum wealth limit.

[11]In practice, the parties have some information about the relative strength of their case based on its circumstances even before they consult lawyers and/or initiate a lawsuit. Such prior knowledge is reflected in our model by the structure of the culpability index (i.e., the function $g(\,\cdot\,,\cdot\,)$) as well as the distribution of θ and s. It is also plausible that the parties may have some private information, beyond evidence, prior to filing a lawsuit. We thank the referee for this suggestion, which we leave for future research.

[12]For example, s could represent a piece of correspondence or memorandum written by an employee of a pharmaceutical company discussing potential side-effects of its drug, a subject matter of the litigation.

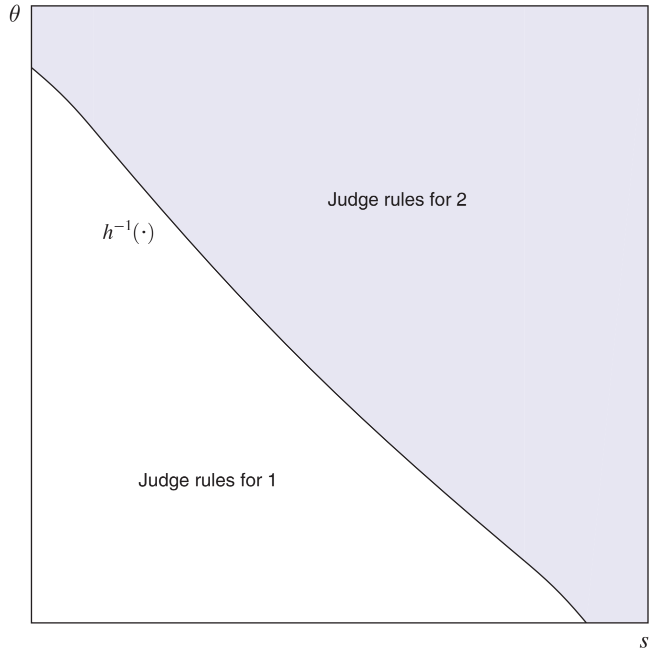

Figure 1. Judge's Ruling in the Event of Disclosure

To make the judge's decision problem nontrivial, we assume that $\int g(s,1)f(s\,|\,1)\;ds > 0$ and $\int g(s,0)f(s\,|\,0)\;ds < 0$, where $f(s\,|\,\theta)$ denotes the density of s conditional on θ. This implies that legal standards and commonly known facts have enough inherent variability that the judge's unconditional belief about the culpability swings from one side to the other as θ changes from the most favorable for party 1 ($\theta = 0$) to the least favorable ($\theta = 1$).[13] Since $g(s,\theta)$ is monotonically increasing in both arguments, there exists a strictly decreasing continuous function $s = h(\theta)$ such that $g(h(\theta),\theta) = 0$ for all $\theta \in [\underline{\theta},\overline{\theta}]$, where $\overline{\theta} := \max\{\theta\,|\,\exists s' \in S$ such that $g(s',\theta) = 0\}$ and $\underline{\theta} := \min\{\theta\,|\,\exists s'' \in S$ such that $g(s'',\theta) = 0\}$. The graph of this function partitions the (s,θ) space into two regions where the judge rules for party 1 and party 2, respectively, when she observes both s and θ, as depicted in Figure 1.[14]

The adjudication criterion $g(s,\theta)$ can be rationalized as a certain societal objective followed by the judge. Suppose the society would like to minimize the cost associated with a wrong decision, i.e., "convicting the innocent or exonerating the guilty." Let c_1 and c_2 be the cost of ruling mistakenly for party 1, the defendant, ("exonerating the guilty") and for party 2 ("convicting the innocent"), respectively, and let $\pi(s,\theta)$ be the probability that party 1 is guilty for given (s,θ). Then, if the judge convicts party 1 with probability z, the expected cost of a mistake is

$$(1 - \pi(s,\theta))\,c_2 z + \pi(s,\theta)\,c_1(1 - z).$$

[13] This assumption merely rules out trivial cases.
[14] The two regions have nonempty interiors under the above assumption.

To minimize this cost, the judge should choose $z = 1$ if $\pi(s,\theta) - \frac{c_2}{c_1 + c_2} > 0$ and should choose $z = 0$ otherwise. Our model accommodates this behavior if we let $g(s,\theta) := \pi(s,\theta) - \frac{c_2}{c_1 + c_2}$.[15] We assume throughout that the judge follows the criterion g whenever the evidence s is disclosed by either party.

If no party discloses s, then the judge still makes a decision given her belief about g, based on the legal strength θ of the case she observes and on the inference about the parties' disclosure decisions. The adjudicator's decision rule in the event of nondisclosure, henceforth referred to as *default ruling strategy*, is described by the function $\delta : \Theta \mapsto [0,1]$, where $\delta(\theta)$ denotes the probability with which she rules for party 2 if she observes signal θ and no evidence is disclosed.

The judge's default ruling strategy depends on her *posterior assessment* of party 1's culpability given nondisclosure:[16]

$$(1) \qquad E[g|\rho_1(\cdot), \rho_2(\cdot), \theta] := p_{00} E_{\emptyset}[g|\theta] + p_{10} E_1[g|\theta] + p_{01} E_2[g|\theta] + p_{11} E_{12}[g|\theta],$$

where $E_{\emptyset}[g|\theta] := \int_0^1 g(s,\theta) f(s|\theta)\, ds$, $E_i[g|\theta] := \int_0^1 g(s,\theta)(1 - \rho_i(s,\theta)) f(s|\theta)\, ds$, $i = 1, 2$, and $E_{12}[g|\theta] := \int_0^1 g(s,\theta)(1 - \rho_1(s,\theta))(1 - \rho_2(s,\theta)) f(s|\theta)\, ds$. Plainly, the judge's posterior is a weighted average of the expected culpability criterion based on four alternative scenarios of evidence observability. The first term $E_{\emptyset}[g|\theta]$ is party 1's expected culpability given that no party has observed the evidence s. The term $E_i[g|\theta]$ is the expectation of g given that only party $i \in \{1,2\}$ has observed s but has not disclosed it. The last term $E_{12}[g|\theta]$ reflects the case in which both parties have observed s but neither has disclosed it.

For later, it is useful to note that when the parties follow *cutoff strategies*: party 1 discloses evidence if and only if $s < \hat{s}_1$ and party 2 discloses if and only if $s > \hat{s}_2$. The judge's posterior assessment given such strategies reduces to (with a slight abuse of notation):

$$(2) \qquad E[g|\hat{s}_1, \hat{s}_2, \theta] := p_{00} \int_0^1 g(s,\theta) f(s|\theta)\, ds + p_{10} \int_{\hat{s}_1}^1 g(s,\theta) f(s|\theta)\, ds + p_{01} \int_0^{\hat{s}_2} g(s,\theta) f(s|\theta)\, ds + p_{11} \int_{\hat{s}_1 \le s \le \hat{s}_2} g(s,\theta) f(s|\theta)\, ds.$$

[15] Different standards of proof and evidence adopted by the courts are consistent with this model. Indeed, let $\alpha := \frac{c_2}{c_1 + c_2}$. If $\alpha = 0.51$, then the judge can be said to follow the rule of preponderance of evidence. An α that lies somewhere in the interval $(0.6, 0.7)$ may be seen as corresponding to the standard of "clear and convincing evidence." According to Posner (1999), judges associate threshold probability levels between 0.75 and 0.9 with the standard of "proof beyond a reasonable doubt." See also Thompson (1989), Tribe (1971), and Wells (1992) on this.

[16] Note that this notion captures unconditional expectation of g. A (more standard) conditional expectation of g equals $E[g|\rho_1(\cdot), \rho_2(\cdot), \theta]$ divided by the probability of nondisclosure:

$$p_{00} + p_{10} \int_0^1 (1 - \rho_1(s,\theta)) f(s|\theta)\, ds + p_{01} \int_0^1 (1 - \rho_2(s,\theta)) f(s|\theta)\, ds + p_{11} \int_0^1 (1 - \rho_1(s,\theta))(1 - \rho_2(s,\theta)) f(s|\theta)\, ds.$$

As will be seen, our analysis depends only on the sign of the expectation, so the difference—whether the expectation is conditional or unconditional—is immaterial. Here we focus on unconditional expectation for simplicity. The current formulation is based on the Bayesian updating. Che and Severinov (2014) shows that our framework can be extended to accommodate non-Bayesian updating by the judge.

Equilibrium Concept and Outcome.—In each regime, we focus on perfect Bayesian equilibria in the parties' disclosure strategies and the judge's default ruling strategy, summarized by a triple, (ρ_1, ρ_2, δ). We assume that, whenever the evidence is disclosed by either party, the judge follows the criterion g. In case s is disclosed, the criterion $g(s, \theta)$ can be viewed as an immutable legal rule, and any deviation from it would constitute an "error" of law. There is more ambiguity and greater scope for the judge's discretion when crucial evidence is not disclosed. Thus, we will need to focus on and examine a nontrivial inference problem facing the judge in the event of nondisclosure.

Our ultimate interest is in the equilibrium outcome of the trial given the information available to the parties. Formally, an *adjudication outcome* is a function, $\phi : X_1 \times X_2 \times S \times \Theta \mapsto [0,1]$, that maps the state of the world (x_1, x_2, s, θ) into the probability that the judge rules for party 2, where $x_i \in \{0,1\},\ i = 1,2$, with $x_i = 1$ if party i observes s and $x_i = 0$ if party i does not observe s. In particular, an equilibrium (ρ_1, ρ_2, δ) induces the following outcome function:

$$\begin{aligned}\phi(x_1, x_2, s, \theta) = {} & \delta(\theta)\,(1 - x_1\,\rho_1(s,\theta))\,(1 - x_2\,\rho_2(s,\theta)) \\ & + \mathbf{1}_{\{g(s,\theta) \geq 0\}}\big[1 - (1 - x_1\,\rho_1(s,\theta))(1 - x_2\,\rho_2(s,\theta))\big],\end{aligned}$$

where $\mathbf{1}_{\{A\}}$ has value 1 in the event A and zero otherwise. We are interested in comparing the adjudication outcomes induced by equilibria under different legal regimes.

II. Irrelevance of Lawyer Advising

In this section, we characterize equilibrium outcomes under legal regimes that differ in the availability of (costless) legal advice. We then compare them.

A. *No Advising*

In this regime, neither party has a lawyer. So, each party must decide whether to disclose the evidence s without being certain about the legal strength θ, and thus without knowing whether this disclosure will lead to a favorable or an unfavorable ruling by the judge.

We shall establish that there exists a unique perfect Bayesian equilibrium. In this equilibrium, both parties and the judge adopt cutoff strategies. In particular, there is a common equilibrium threshold $\hat{s}$ such that party 1 discloses s if and only if $s < \hat{s}$, and party 2 discloses s if and only if $s > \hat{s}$. Absent disclosure, the judge rules for party 1 if $\theta < \hat{\theta}$ and for party 2 if $\theta > \hat{\theta}$, where $\hat{\theta}$ is the judge's equilibrium threshold. The judge uses a cutoff strategy because her posterior $E[g \mid \hat{s}, \hat{s}, \theta]$ is monotonically increasing in θ, which itself follows from two effects. First, a higher θ is a stronger evidence of 1's culpability, holding s fixed. Second, there is also an inference effect. Weak monotone likelihood ratio property (MLRP) implies that a higher value of s is equally or more likely under a higher θ than under a lower θ. So nondisclosure is weakly more likely to be a result of party 1's concealment of unfavorable s (rather than his not observing s), given the parties' cutoff strategies.

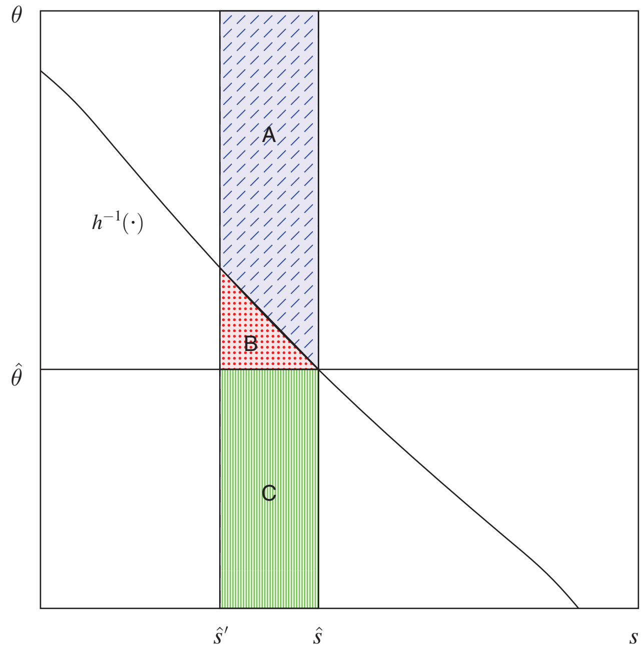

FIGURE 2. OPTIMAL DISCLOSURE STRATEGY

Obviously, this inference effect adds to the judges's suspicion of 1's culpability. Figure 2 graphs the two thresholds $\hat{s}$ and $\hat{\theta}$.

Crucially, the two thresholds $\hat{s}$ and $\hat{\theta}$ cross each other on the curve h^{-1}, i.e., $g(\hat{s}, \hat{\theta}) = 0$, or $\hat{s} = h(\hat{\theta})$. The intuition for this is as follows. Suppose that, facing the judge's threshold $\hat{\theta}$, party 1 deviates by withholding some $s \leq \hat{s}$, say an interval $[\hat{s}', \hat{s})$. Let us show that such a deviation is not profitable. Without loss, assume that party 2 does not disclose, or else party 1's disclosure would not matter. If θ is either below $\hat{\theta}$ (region C in Figure 2) or above $h^{-1}(s)$ (region A in Figure 2), then this deviation makes no difference, for the judge's ruling will be the same whether party 1 discloses s or not. But if θ happens to be between $h^{-1}(s)$ and $\hat{\theta}$ (region B), then withholding s will result in a ruling against party 1 whereas disclosing it would result in a ruling in his favor. So withholding any $s < h(\hat{\theta})$ is never profitable. A similar argument shows that disclosing $s > h(\hat{\theta})$ is suboptimal for party 1.

This argument shows why party 1 and party 2 will adopt cutoff strategies with threshold $\hat{s} = h(\hat{\theta})$: party 1 discloses the evidence if and only if $s < \hat{s}$, and party 2 discloses if and only if $s > \hat{s}$. Substituting this into (2), the judge's equilibrium posterior becomes $E[g \,|\, h(\hat{\theta}),\, h(\hat{\theta}),\, \hat{\theta}]$. Hence, her cutoff threshold is given by:

$$\hat{\theta}^* := \inf\left\{\hat{\theta} \in \Theta \,|\, E\left[g \,|\, h(\hat{\theta}), h(\hat{\theta}), \hat{\theta}\right] > 0\right\}, \tag{3}$$

where $\hat{\theta}^* := 1$ if the set on the right-hand side is empty.

It is instructive to study how the judge's threshold varies as her posterior assessment changes, particularly because of the changes in the probabilities, p_{10}, p_{01}, and p_{11}, of different evidence observability scenarios. First, note that a change in p_{11} has no effect on the judge's posterior assessment and hence on the threshold $\hat{\theta}^*$ because $\hat{s}_1 = \hat{s}_2 = \hat{s}$ and therefore the value of the last integral in (2) is zero. Next, if p_{10} falls relative to p_{01}, the value of (2) decreases and hence $\hat{\theta}$ increases. It is intuitive to describe this change as a shift of the burden of proof from party 1 to party 2, as the judge becomes more suspicious of party 2's strategic withholding. This, in turn, causes a decrease in the parties' common disclosure threshold, $h(\hat{\theta}^*)$. In other words, the party with an increased burden of proof discloses more evidence, and the party with a decreased burden discloses less evidence. The next proposition establishes the threshold nature of the equilibrium and its properties formally. Its proof is provided in the Appendix.

PROPOSITION 1: *If neither party is advised by a lawyer, there exists a unique perfect Bayesian equilibrium. In this equilibrium, party* 1 *discloses s if and only if* $s < h(\hat{\theta}^*)$, *and party* 2 *discloses s if and only if* $s > h(\hat{\theta}^*)$. *Following nondisclosure, the judge rules for party* 1 *if* $\theta < \hat{\theta}^*$ *and for party* 2 *if* $\theta > \hat{\theta}^*$. *The threshold* $\hat{\theta}^*$ *decreases in* p_{10} *and increases in* p_{01}, *while threshold* $\hat{s}$ *increases in* p_{10} *and decreases in* p_{01}; *namely, party* 1 *discloses more and party* 2 *discloses less as the burden of proof shifts to party* 1.

B. *Two-Sided Advising*

In this regime, both parties receive lawyer advice and learn θ. Hence, unlike in the no-advising case, the parties make their disclosure decisions based on both s and θ. Recall that the judge's ruling in case of disclosure follows the criterion $g(s, \theta)$. So, party 1 has a weakly dominant strategy of disclosing s if and only if $s < h(\theta)$, or $g(s, \theta) < 0$. Disclosing $s < h(\theta)$ leads to a sure win for party 1, whereas withholding it may entail an unfavorable ruling. Likewise, withholding $s > h(\theta)$ is a dominant strategy for party 1 because the judge may rule for party 1 without disclosure but will rule against him for sure if s is disclosed. By the same logic, party 2's weakly dominant strategy is to disclose s if and only if $s > h(\theta)$, or $g(s, \theta) > 0$.

Dominant strategies have an intuitive appeal in our model, particularly with lawyer advising. In most legal systems, a defense lawyer in a criminal trial has a positive duty to explore all avenues of defense, and withholding exculpatory evidence would contravene this obligation. Getting a client's consent for withholding favorable evidence may also be problematic. Furthermore, the judge could simply refuse to believe that a lawyer is not following a dominant strategy, in which case a deviation from the dominant strategy would not have any benefit. Finally, if there is even a small uncertainty about the judge's default ruling, then disclosing all favorable evidence and withholding all unfavorable evidence is the unique optimal strategy for either party. For these reasons, we focus on the dominant disclosure strategies here. Later, we will consider what happens when these arguments do not apply and examine equilibria supported by weakly dominated disclosure strategies.

Note that the disclosure behavior of a represented party using a dominant strategy is different from the equilibrium behavior of an unrepresented party, since the disclosure threshold of a represented party varies with θ. In the no-advising case, the parties employ a common disclosure threshold $\hat{s} = h(\hat{\theta}^*)$ that does not vary with θ. Consequently, the judge's posterior is $E[g \,|\, h(\theta), h(\theta), \theta]$ in the two-sided advising case, and is $E[g|\hat{s}, \hat{s}, \theta]$ with $\hat{s} \equiv h(\hat{\theta}^*)$ in the no-advising case. These posteriors differ in their magnitudes for almost all θ, but, remarkably, they have the same sign for all θ, so they cross zero at the same threshold $\hat{\theta}^*$. In particular,

$$E[g \,|\, h(\theta),\, h(\theta),\, \theta] \gtreqless 0 \quad \text{if} \quad \theta \gtreqless \hat{\theta}^*.$$

This leads the judge to adopt the same default ruling under both regimes. Intuitively, the reason is that whenever the realized value of θ is equal to the judge's threshold $\hat{\theta}$, the disclosure behavior of the parties is identical in both regimes. Party 1 discloses $s < h(\hat{\theta})$, whether she is represented or not. Similarly, party 2 discloses $s > h(\hat{\theta})$ in both cases. Hence, we arrive at the following result.

PROPOSITION 2: *If both parties are advised by lawyers, there exists a unique equilibrium in weakly dominant disclosure strategies. In this equilibrium, party* 1 *discloses s if and only if* $g(s,\theta) < 0$, *and party* 2 *discloses s if and only if* $g(s,\theta) > 0$. *Absent disclosure, the judge rules for party* 1 *if* $\theta < \hat{\theta}^*$ *and for party* 2 *if* $\theta > \hat{\theta}^*$.

The fact that the threshold $\hat{\theta}^*$ is the same in the no-advising and the two-sided advising cases is crucial for the irrelevance result presented below.

C. *One-Sided Advising*

The results in subsections IIA and IIB generalize to the regime in which only one side (say, without loss of generality, party 1) hires a lawyer. Focusing as before on weakly dominant strategies, party 1 will disclose s if and only if $s < h(\theta)$, just as in subsection IIB. As established in subsection IIA, party 2's unique optimal strategy is to disclose s if and only if $s > h(\hat{\theta})$ where $\hat{\theta}$ denotes the threshold used by the judge in her default ruling strategy when s is not disclosed.

So, when the judge observes θ but not s, her posterior becomes

$$E[g \,|\, h(\theta),\, h(\hat{\theta}),\, \theta].$$

Since this posterior is monotonically increasing in θ and changes sign from negative to positive at $\hat{\theta}^*$, the following result is immediate.

PROPOSITION 3: *If only one party hires a lawyer, there exists a unique equilibrium in weakly dominant disclosure strategies. In this equilibrium, if party* 1 (*party* 2) *obtains legal advice, she discloses s if and only if* $g(s,\theta) < 0$ $(g(s,\theta) > 0)$. *If party* 1 (*party* 2) *does not obtain legal advice, she reveals s if and only if* $s < h(\hat{\theta}^*)$ $(s > h(\hat{\theta}^*))$. *The judge uses threshold* $\hat{\theta}^*$—*defined in* (3)—*in her cutoff default ruling strategy.*

A change in p_{10} or in p_{01} affects the judge's default ruling strategy and the strategy of an unrepresented party in the same way as in Proposition 1*, without affecting the strategy of a represented party.*

The key result of subsequent interest is that the judge follows the same default ruling strategy as in the other regimes. As explained above, this is due to the fact that the parties' different disclosure strategies do not affect the sign of the judge's equilibrium posterior assessment, although they do affect its magnitude.[17]

D. *Irrelevance of Lawyer Advising*

A striking feature emerging from our analysis is that the judge's equilibrium default ruling strategy is the same in all three regimes. At the same time, Propositions 1, 2, and 3 establish that the parties follow different disclosure strategies across the regimes. However, we will show below that the differences in the parties' disclosure behavior do not produce any real differences in the outcome of the trial.

In fact, this irrelevance is a consequence of a more general property of our disclosure/adjudication game, which is described in the following Lemma. Let $\phi^{\hat{\theta}}(x_1, x_2, s, \theta)$ denote the probability with which the judge rules for party 2 in the state (x_1, x_2, s, θ) when she follows a default ruling strategy with threshold $\hat{\theta}$.

LEMMA 1 (Decision Equivalence): *Suppose that the judge adopts a cutoff strategy with threshold $\hat{\theta} \in \Theta$ in her default ruling. Regardless of the legal regime, i.e., whether either party obtains legal advice or not, mutual best responses by the two parties in disclosure lead to the same outcome characterized by the following outcome function $\phi^{\hat{\theta}}(x_1, x_2, s, \theta)$:*

$$(4) \qquad \phi^{\hat{\theta}}(x_1, x_2, s, \theta) = \begin{cases} \mathbf{1}_{\{g(s,\theta) \geq 0\}} & \text{if } x_1 = x_2 = 1 \\ \mathbf{1}_{\{g(s,\theta) \geq 0 \text{ and } \theta \geq \hat{\theta}\}} & \text{if } x_1 = 1, x_2 = 0 \\ \mathbf{1}_{\{g(s,\theta) \geq 0 \text{ or } \theta \geq \hat{\theta}\}} & \text{if } x_1 = 0, x_2 = 1 \\ \mathbf{1}_{\{\theta \geq \hat{\theta}\}} & \text{if } x_1 = x_2 = 0 \end{cases}.$$

The Decision Equivalence Lemma establishes that the judge's cutoff strategy uniquely determines the equilibrium adjudication outcome, regardless of the parties' use of legal advice. Figure 3 illustrates the outcomes in alternative evidence observability cases when the judge follows a default ruling strategy with threshold $\hat{\theta}$. The shaded area depicts the set of states (s, θ) in which the judge rules for party 2, regardless of the legal regime. Consider the case $(x_1, x_2) = (1, 0)$ in panel B. Here party 1 observes a signal but party 2 does not. Since the default ruling of the judge is to rule for party 2 for any $\theta > \hat{\theta}$ without any disclosure, the best party 1 can hope to achieve is to win the case whenever (s, θ) falls in the unshaded region. With legal

[17] In fact, this result holds even more generally. Suppose each party randomizes in hiring a lawyer, the judge has some arbitrary—possibly inaccurate—belief about the representation choices. The main logic determining the equilibrium behavior remains unchanged, leading to the same characterization of the judge's default ruling strategy.

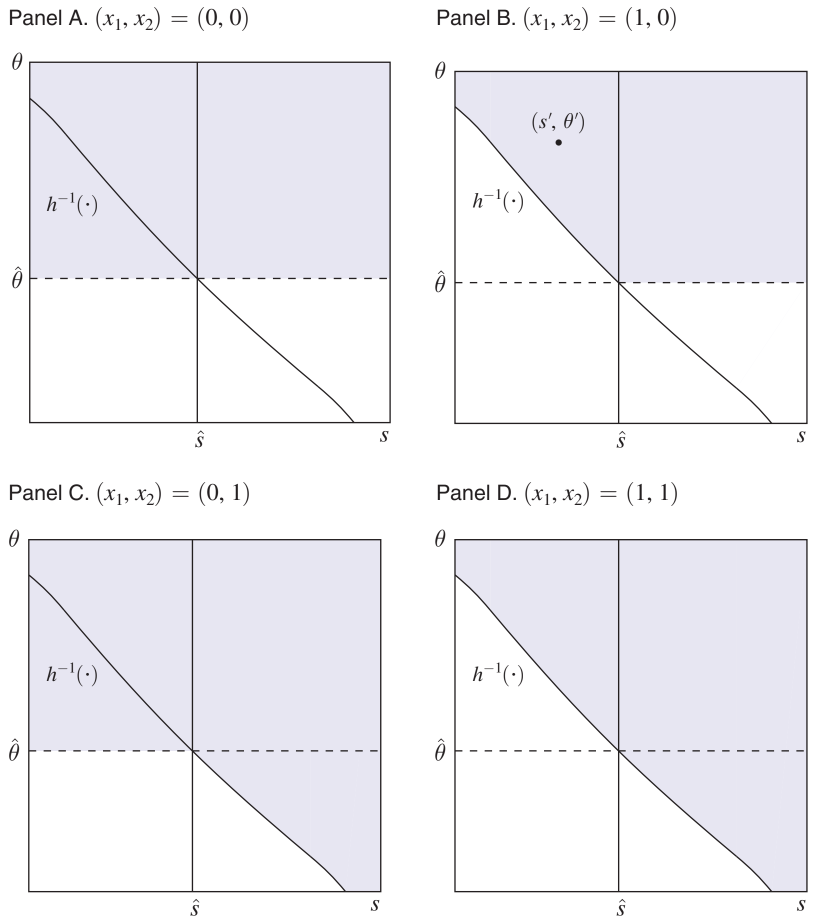

Figure 3. Decision Equivalence

Note: Ruling for party 2 is in the shaded area.

advice, he can attain this outcome by disclosing only in states (s, θ) that lie below $h^{-1}(\cdot)$. Less trivially, the party can achieve the same outcome even without access to legal advice. By disclosing s if and only if $s < \hat{s}$, he can still ensure winning in all "unshaded" states. Although not fully aware of how his s would be seen by the judge, the party nevertheless manages to disclose his evidence when he must (the unshaded area above $\hat{\theta}$) and withhold it when he must (the shaded area below $\hat{\theta}$). For instance, in state (s', θ'), the signal is unfavorable to party 1, but without knowing this, he discloses. Legal advice would have led him to suppress such evidence. Remarkably, however, the outcome is the same since the judge rules against him without any disclosure. In sum, Decision Equivalence Lemma suggests that a party does not gain from legal advice at all if the judge follows the same threshold default ruling strategy. The only condition for this result is that the judge's default ruling strategy is rationally anticipated and the culpability of party 1 increases in (s, θ).[18]

[18] Also notice that the result does not require a Bayesian judge; all it requires is that the judge applies the same threshold across different legal regimes. See Che and Severinov (2014).

Importantly, Propositions 1–3 imply that the judge applies the same threshold in equilibrium across legal regimes, Hence, combining these Propositions with Lemma 1, we obtain our key result:

PROPOSITION 4 (Irrelevance of Legal Advice): *Suppose that represented parties employ weakly dominant strategies in disclosure. Then in every advising regime there is a unique equilibrium characterized by the outcome function* $\phi^{\hat{\theta}^*}(\cdot)$ *described by* (4).

It is worth pointing out that our irrelevance result, established here for the case of a Bayesian judge, holds more broadly, even if the judge is non-Bayesian. In fact, it is not hard to show that the irrelevance result of Proposition 4 as well as Propositions 1–3 remain valid when, instead of the Bayesian inference rule given by (1), the judge follows one of the large class of inference rule obtained from (1) by replacing "correct Bayesian" weights $(p_{00}, p_{10}, p_{01}, p_{11})$ with some arbitrary nonnegative weights $(w_{00}, w_{10}, w_{01}, w_{11})$. Such non-Bayesian updating will lead to different threshold value $\hat{\theta}$, but the threshold will be the same in all legal regimes, and thus their equivalence, due to Lemma 1 (see Che and Severinov 2014).

Our analysis focuses on one particular aspect of legal representation—the role of lawyers as gatekeepers of information reaching the court. In practice, lawyers perform a number of valuable tasks that are not captured by our model, so our result should not be interpreted as suggesting that legal representation is never useful. Nevertheless, the robustness of our irrelevance result is surprising. To the extent that information disclosure matters in adjudication, our irrelevance result clarifies and qualifies the sense in which lawyers influence this process. Also, by scrutinizing the assumptions behind the irrelevance result we can start to identify the circumstances under which *legal advice matters*. The next section considers one such case.

III. Relevance of Lawyer Advising: Withholding Favorable Evidence

Thus far, we have focused on one class of equilibria in which a lawyer discloses all ex post-favorable evidence. Although such a strategy is weakly dominant and therefore can be seen as focal, this section shows that an alternative strategy of suppressing favorable evidence may influence the judge's posterior belief in a way that may favor the party.

To illustrate this point, suppose both parties have hired lawyers. Suppose further that party 1 follows a strategy of never disclosing any evidence while party 2 discloses favorable evidence. This will induce the judge to treat party 1 more favorably in case of nondisclosure, since nondisclosure by party 1 is neutral about the content of the evidence, while party 2's nondisclosure may have arisen from his concealment of evidence favorable to party 1. Specifically, the judge's posterior in case of nondisclosure becomes $E[g \mid 0, h(\theta), \theta]$, which is less than $E[g \mid h(\theta), h(\theta), \theta]$, and is therefore more favorable to party 1. Likewise, if party 1 adopts the dominant strategy but party 2 adopts the strategy of never disclosing, then the judge forms a posterior $E[g \mid h(\theta), 1, \theta]$, which is more favorable for party 2 than if both adopt their dominant strategies. Of course, a party's strategy of suppressing potentially favorable evidence makes sense only if the judge rules favorably for this party despite the suppression. Since the judge will not rule favorably if θ is sufficiently unfavorable, the strategy will work only for a range of θs. Particularly, let

$$\hat{\theta}_+ := \inf\{\theta \mid E[g \mid 0, h(\theta), \theta] > 0\};$$

$$\hat{\theta}_- := \inf\{\theta \mid E[g \mid h(\theta), 1, \theta] > 0\}.$$

Observe that $\hat{\theta}_- \leq \hat{\theta}^* \leq \hat{\theta}_+$, and that $\hat{\theta}_- < \hat{\theta}_+$.[19] If $\theta > \hat{\theta}_+$, then $E[g \mid 0, h(\theta), \theta] > 0$, so the judge will rule for party 2 unless party 1 presents evidence $s < h(\theta)$. Hence, it will never be optimal for party 1 to suppress evidence $s < h(\theta)$ for such θ. But when $\theta < \hat{\theta}_+$, party 1 may be able to sway the judge's posterior by suppressing all evidence, both favorable and unfavorable, since the judge will rule for party 1 even without disclosure. By the same logic, party 2 can credibly sway the judge's posterior by suppressing all evidence (given that party 1 follows his dominant strategy), if $\theta > \hat{\theta}_-$.

These arguments suggest that the parties can manipulate the judge's posterior within the interval $[\hat{\theta}_-, \hat{\theta}_+]$. Indeed, any $\hat{\theta} \in [\hat{\theta}_-, \hat{\theta}_+]$ can be supported as an equilibrium threshold for the judge's default ruling when both parties are advised by lawyers.

The benefit of withholding seemingly favorable evidence is most evident when only one party has access to legal advice, the case illustrated in (ii) of Proposition 5 below and in Figure 4. Suppose initially neither party retains a lawyer. Then, by Proposition 1, the judge employs a threshold of $\hat{\theta}^*$ in her default ruling. Now, suppose only party 1 hires a lawyer, who advises the party to withhold any s when $\theta < \hat{\theta}_+$. The resulting disclosure strategy is depicted as the darker shaded (triangular-shaped) area in Figure 4. Under the new threshold $\hat{\theta}_+$, party 2 adjusts his disclosure strategy to the lighter shaded area to the right of $\hat{s}_+$. Then the judge's belief upon nondisclosure under any θ below $\hat{\theta}_+$ becomes favorable to party 1, because the only party who strategically withholds evidence in this case is party 2.

More generally, the judge can be induced to employ any threshold in $[\hat{\theta}^*, \hat{\theta}_+]$. Importantly, any equilibrium threshold in this situation is more favorable for party 1 than $\hat{\theta}^*$, the threshold used when neither party hires a lawyer. Specifically, we have:

PROPOSITION 5 (Relevance of Legal Advice with Weakly Dominated Strategies):

(*i*) *Suppose that both parties retain lawyers. Then, for any* $\hat{\theta} \in [\hat{\theta}_-, \hat{\theta}_+]$, *there exists a perfect Bayesian equilibrium in which, absent disclosure, the judge rules for party* 1 *if* $\theta < \hat{\theta}$ *and for party* 2 *if* $\theta > \hat{\theta}$; *party* 1 *discloses evidence* s *if and only if* $s < h(\theta)$ *and* $\theta > \hat{\theta}$, *and party* 2 *discloses* s *if and only if* $s > h(\theta)$ *and* $\theta < \hat{\theta}$.

[19] Using equation (2), it is easy to see that $\hat{\theta}_- \leq \hat{\theta}^* \leq \hat{\theta}_+$ because $E[g \mid 0, h(\theta), \theta] < E[g \mid h(\theta), h(\theta), \theta] < E[g \mid h(\theta), 1, \theta]$. Further, $\hat{\theta}^+ > \hat{\theta}^-$ if $\hat{\theta}^+ > 0$ and $\hat{\theta}^- < 1$. The former holds if $Eg(s, 0) < 0$ and the latter holds if $Eg(s, 1) > 0$, both of which are satisfied by assumption.

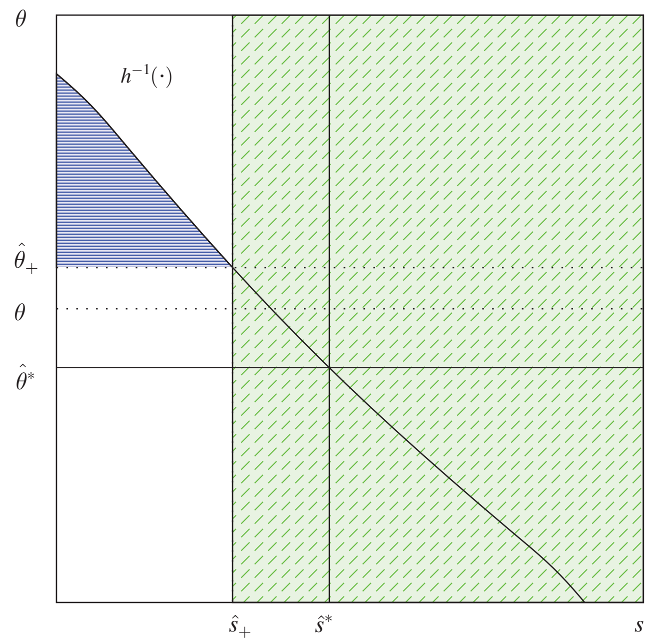

Figure 4. Withholding Evidence to Influence Judge's Belief

Conversely, any equilibrium cutoff default ruling strategy has a threshold in $[\hat{\theta}_-, \hat{\theta}_+]$.

(ii) *Suppose only party* 1 *retains a lawyer. Then, for any* $\hat{\theta} \in [\hat{\theta}^*, \hat{\theta}_+]$, *there exists a perfect Bayesian equilibrium in which, absent disclosure, the judge rules for party* 1 *if* $\theta < \hat{\theta}$ *and for party* 2 *if* $\theta > \hat{\theta}$; *party* 1 *discloses evidence* s *if and only if* $s < h(\theta)$ *and* $\theta > \hat{\theta}$, *and party* 2 *discloses* s *if and only if* $s > h(\hat{\theta})$.

Conversely, any cutoff default ruling strategy by the judge has a threshold in $[\hat{\theta}^*, \hat{\theta}_+]$. *An analogous characterization holds if only party* 2 *retains a lawyer.*

Proposition 5 shows that a range of different posteriors by the judge—and thus a range of different default rulings—is sustainable when the parties adopt legal strategies of withholding seemingly favorable evidence along with unfavorable evidence. Since such a strategy requires conditioning the disclosure decision on θ, it cannot be played without the legal expertise of a lawyer. In this sense, we have identified a source of private value of legal advice—namely, the ability to advise a party to withhold seemingly favorable evidence.

Clearly, the relevance of legal advice rests on the credibility of playing the weakly dominated strategy of withholding favorable evidence. One way in which a lawyer can achieve such credibility is by building a reputation through repeated trials. Casual observation suggests that lawyers are concerned about their reputations and undertake specific steps to enhance and maintain them. For example, some criminal defense lawyers are known to call very few witnesses. Sometimes, a lawyer may rest the case without presenting any evidence or calling any witnesses at all if he or she believes that the case has not been proven by the prosecutors. Our results indicate that such behavior helps to build a lawyer's reputation, which could skew the court's ruling in favor of that lawyer and her clients. From this perspective, the above result can be interpreted in terms of the relative reputation of the lawyers representing the two sides.

A lawyer with a good reputation can make one better off, while a lawyer with a bad reputation can make one worse off, relative to the no-advising case. Interestingly, good reputation in our context means being known for presenting limited evidence, while bad reputation is being known for presenting too much evidence, sometimes unnecessarily. While a party benefits from gaining legal advice unilaterally, it is impossible for both parties to be strictly better off from legal advice, relative to the case with no legal advice for either side. This fact implies that the game of hiring lawyers has the structure of a prisoner's dilemma, which may explain why both parties would hire lawyers in equilibrium.

How does the strategic withholding of favorable evidence affect social welfare? Since the effect of strategic withholding is to manipulate the equilibrium threshold $\hat{\theta}$ the judge employs,[20] we can ask this question via the following general mechanism framework. Specifically, suppose the judge "chooses" a threshold $\hat{\theta}$, and the parties best respond to the chosen threshold under any arbitrary legal regime. By Lemma 1, the parties' best responses lead to a unique state-contingent outcome, regardless of the legal regime. Hence, the judge's problem is succinctly described as follows:

$$(WP) \qquad \max_{\hat{\theta}\in\Theta} \sum_{i,j=0,1} p_{ij}\, E\Big[\phi^{\hat{\theta}}(i,j,s,\theta)g(s,\theta)\Big],$$

where $\phi^{\hat{\theta}}$ is the outcome induced by the cutoff rule with threshold $\hat{\theta}$ (see Lemma 1). This objective function puts a positive value on the ruling for party 2 if and only if party 1 is truly culpable, i.e., $g(s,\theta) > 0$ (see the discussion of the Judge's Adjudication Behavior in Section I). The analysis of (*WP*) yields the following result.

PROPOSITION 6: *The cutoff $\hat{\theta}^*$ defined in* (3) *is socially optimal in the sense that it solves* (*WP*). *In other words, any other threshold $\hat{\theta}$ induced via lawyer manipulation strictly harms the welfare.*

The intuition behind the result can be seen in Figure 4. Given threshold $\hat{\theta}_+$, the welfare loss can be represented by two areas: the area below the line $\theta = \hat{\theta}_+$ and to the right of the curve $g(s,\theta) = 0$ is where the loss arises from the withholding by party 1 of unfavorable evidence in state $(x_1, x_2) = (1, 0)$, and the darker shaded area is where the loss arises from the withholding by party 2 of unfavorable evidence in state $(x_1, x_2) = (0, 1)$. Raising the threshold, and thus increasing the

[20] Specifically, the interval $[\hat{\theta}_-, \hat{\theta}_+]$ can be interpreted as scope for manipulation by the lawyers.

burden of proof against party 2, increases the former loss and reduces the latter loss. The equilibrium threshold $\hat{\theta}^*$ without lawyer manipulation balances the two losses at the margin optimally; that is, the judge is indifferent between the relative losses at the threshold.[21]

IV. Can Money Buy Justice? A Signaling Role of Costly Legal Advice

So far, we have assumed that legal advice is available for free. This assumption helps to isolate the effect of lawyers' expertise and provides a foundation for normative analysis. In reality, lawyer advice is often costly, more so if a lawyer has a good reputation or a high profile. It is thus natural to extend our model to introduce positive lawyer cost. The most important finding of this section is that lawyer advice, when available at a cost, changes the parties' disclosure behavior in a way that affects both private and social welfare. It also enables us to provide some insights regarding the possibility that prominent lawyers can skew justice in favor of their clients, which is a commonly expressed concern.

To extend our model in this direction, we need to add more structure. In particular, the cost of lawyer advising must be measured in units comparable to the value of winning. To this end, we assume that the cost of legal advice is $w > 0$, and that a party derives a value $v > 0$ when (s)he wins the dispute and zero when (s)he loses, all measured in monetary units. We then consider a game in which each party with a signal s chooses whether to hire a lawyer at the cost w, followed by the disclosure game studied earlier. As already mentioned, the model is more aptly interpreted as pertaining to the quality of legal advice. Namely, a litigant can either hire a standard lawyer, at a cost normalized to zero, but such advice is incomplete and leaves residual uncertainty about θ; or a litigant can hire an expensive expert lawyer whose advice costs $w > 0$ and provides perfect information about θ. Representation choices of the parties are observable.

Further, we simplify our model in several ways. First, we consider a symmetric environment in which each party observes evidence with probability $p \in (0,1)$, independent of the other party (i.e., $p_{10} = p_{01} = p(1-p), p_{11} = p^2$, $p_{00} = (1-p)^2$). Next, we assume s and θ to be independently distributed according to cdf $F(s,\theta) = K(s)L(\theta)$ with $K(\cdot)$ and $L(\cdot)$ having densities k and l, respectively. Next, we assume that the nature of the dispute is symmetric between the two parties.

ASSUMPTION 2 (Symmetry): *For all* $(s,\theta) \in [0,1]^2$, $k(s) = k(1-s)$, $l(\theta) = l(1-\theta)$, *and* $g(s,\theta) = -g(1-s, 1-\theta)$.

Among other things, this assumption means that the strength of the case for party 1 in state (s,θ) is the same as the strength of the case for party 2 in state $(1-s, 1-\theta)$. The assumption implies that $g\left(\frac{1}{2},\frac{1}{2}\right) = 0$, or $h^{-1}\left(\frac{1}{2}\right) = \frac{1}{2}$.

[21] By contrast, under the manipulated strategies, the judge is never made indifferent at the threshold; for instance, when the threshold $\hat{\theta} > \hat{\theta}^*$ is induced by party 1's lawyer's strategy, the judge's posterior never crosses zero; the judge finds party 1 strictly culpable for any $\theta \geq \hat{\theta}$ and strictly non-culpable for any $\theta < \hat{\theta}$.

With costly legal advice, there are two types of equilibria: "no signaling" equilibrium in which legal advice plays no signaling role, and a "signaling" one in which lawyer advising becomes a signal. We begin with the no signaling equilibrium.

PROPOSITION 7 (No Signaling Equilibrium): *There exists an equilibrium in which no party retains a lawyer, and the parties follow the same strategies as in the equilibrium described in Proposition* 1.

The proof of this proposition is straightforward. Suppose that the judge follows a default ruling strategy with threshold $\hat{\theta}^*$ defined in (3), regardless of whether any party has retained a lawyer. Then, it is a best response for parties not to hire lawyers and to follow the disclosure strategies with threshold $h(\hat{\theta}^*)$. Such strategies in turn rationalize the judge's beliefs and her default ruling behavior. If a party, say party 1, deviates and hires a lawyer, then disclosing s if and only if $g(s,\theta) < 0$ is his dominant strategy. Thus, the judge's ruling is sequentially rational after this deviation also. Hence, the no signaling equilibrium can be sustained.

There exists another, more interesting, equilibrium in which legal representation plays a nontrivial signaling role, In this equilibrium, a lawyer is hired for two reasons. First, a party without evidence hires a lawyer to add "credibility" to his nondisclosure. Second, a party with sufficiently weak (likely unfavorable) evidence retains a lawyer to make sure that they disclose only favorable evidence and to imitate those without any evidence. Specifically, for any cost $w > 0$, and the associated threshold $\hat{s}(w) \in \left[\frac{1}{2}, 1\right]$ (defined below), consider the following *lawyer signaling strategies*:

- Party 1 retains a lawyer either if he observes no evidence or if he observes $s > \hat{s}(w)$. In the latter case, he discloses s if and only if $g(s,\theta) < 0$ (the region below C in Figure 5). Party 2 employs a symmetric strategy, retaining a lawyer in the event of no evidence or evidence $s < 1 - \hat{s}(w)$.
- If party 1 observes $s \in [0, \hat{s}(w)]$, he does not hire a lawyer and discloses s (the area to the left of $\hat{s}(w)$ in Figure 5). Symmetrically, party 2 does not hire a lawyer and discloses s if he observes $s > 1 - \hat{s}(w)$.
- If s is disclosed, then the judge rules for party 1 if and only if $g(s,\theta) < 0$. If s is not disclosed, and either both sides have retained lawyers or no side has retained a lawyer, then the judge rules for 1 if and only if $\theta < \frac{1}{2}$. If s is not disclosed and only party 1 (party 2) has retained a lawyer, then the judge rules for party 1 if and only if $g(0,\theta) < 0$ ($g(1,\theta) < 0$).

Note that the judge's beliefs off the equilibrium path are prejudiced against a party without a lawyer. Particularly, *if there is no disclosure and only party* 1 (*party* 2) *has retained a lawyer*, *the judge believes that* $s = 0$ ($s = 1$), *i.e.*, *the evidence is the worst for party* 2 (*party* 1).

To complete the description, we need to characterize the threshold signal $\hat{s}(w)$. Specifically, party 1 with signal $s = \hat{s}(w)$ must be indifferent between hiring a lawyer *and* withholding s and not hiring a lawyer *and* disclosing s. This requires

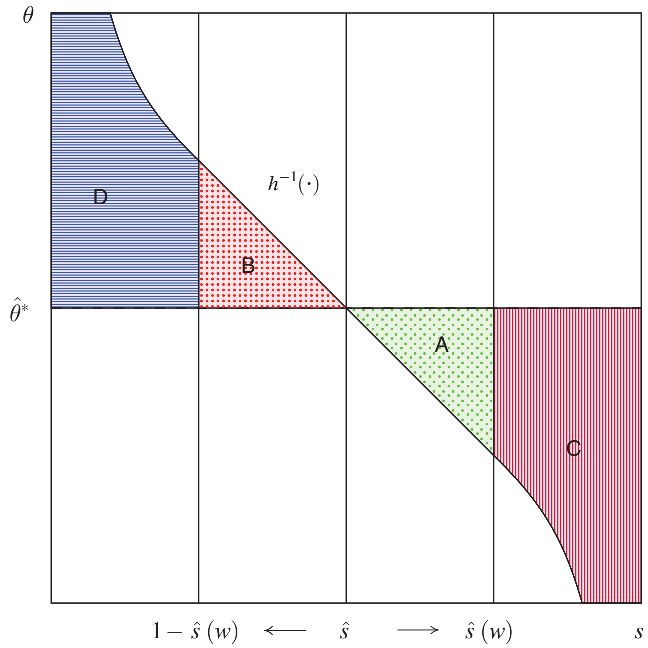

Figure 5. Disclosure in the Lawyer Signaling Equilibrium

$$v(1-p)\left[L\left(\frac{1}{2}\right) - L(h^{-1}(\hat{s}(w)))\right] = w. \tag{5}$$

First, the right-hand side of (5) is the cost w of hiring the lawyer. Its left-hand side is the value of hiring the lawyer for type $s = \hat{s}(w)$: the value v of winning multiplied by the increase in the winning probability due to the lawyer. The latter comprises the probability $1 - p$ that party 2 does not observe the signal. (If party 2 observes $\hat{s}(w)$, then he will disclose it because $\hat{s}(w) \geq \frac{1}{2}$ and hence $\hat{s}(w) \in [1 - \hat{s}(w), 1]$, so party 1's decision would not make any difference.) Assuming party 2 does not observe $\hat{s}(w)$, if party 1 does not hire a lawyer and discloses $\hat{s}(w)$, he wins if $g(\hat{s}(w), \theta) < 0 \Leftrightarrow \theta < h^{-1}(\hat{s}(w))$, whereas hiring a lawyer allows him to win if $\theta < \frac{1}{2}$. Hence, hiring a lawyer increases party 1's probability of winning by $(1-p)\left[L\left(\frac{1}{2}\right) - L(h^{-1}(\hat{s}(w)))\right]$, explaining the left-hand side of (5).

Since $\hat{s}(w) \leq 1$, the cost of lawyer advising must not exceed $\overline{w} := v(1-p)\left[L\left(\frac{1}{2}\right) - L(h^{-1}(1))\right]$ for a lawyer to ever be hired. For any $w \in [0, \overline{w}]$, there is a unique solution $\hat{s}(w)$ to (5): its left-hand side is nondecreasing in $\hat{s}(w)$, is zero at $s = \frac{1}{2}$, and equals $\overline{w}$ at $s = 1$. Hence, $\hat{s}(w)$ lies in $\left[\frac{1}{2}, 1\right]$ for any $w \in [0, \overline{w}]$. Further, $\hat{s}(w)$ is increasing in w and equals 1 when $w = \overline{w}$.

We are now in a position to state the main result of this section.

PROPOSITION 8 (Lawyer Signaling Equilibrium): *If* $w \in [0, \overline{w}]$, *then the lawyer signaling strategies constitute a perfect Bayesian equilibrium.*

In the signaling equilibrium, almost all types who hire a lawyer strictly prefer to do so. So, legal advice *does* have private value.[22] Specifically, by hiring a lawyer, a party without evidence buys "the right to remain silent without prejudice," since a non-disclosing party without a lawyer suffers from a very negative inference.[23]

The concern that those who can afford prominent lawyers can buy justice is real in this equilibrium. To see it, consider party 1 with evidence $\hat{s}(w) - \epsilon$ and the same party with evidence $\hat{s}(w) + \epsilon$, for small ϵ. The latter type incurs the cost of hiring a lawyer (whereas the former does not) and consequently wins with a higher probability than the former, despite having a weaker case. In this sense, expensive lawyers can buy justice. But the marginal type does not benefit from representation because he has to pay legal fees for an increased chance of winning.

The signaling equilibrium also affects social welfare. To the extent that the lawyers' fees are neutral transfers from the litigants to the lawyers, net social welfare can be measured by the adjudicator's objective. By this criterion, the effect of costly lawyer signaling is quite intuitive. A higher cost of nondisclosure (in terms of negative inference) means that the parties are compelled to disclose more. More disclosure leads to more accurate adjudication. So lawyer signaling—hence costly legal advice—improves welfare. More precisely, since in the lawyer signaling equilibrium party 1 (resp, party 2) does not hire a lawyer and discloses s if $s \leq \hat{s}(w)$ (resp, $s \geq 1 - \hat{s}(w)$) where $\hat{s}(w) > \frac{1}{2}$, this equilibrium entails extra disclosure by the amount represented in Figure 5 with a rightward shift of party 1's threshold and a leftward shift of party 2's threshold.

Compared with the no signaling equilibrium (which is equivalent to the equilibrium without representation), the lawyer signaling equilibrium entails different outcomes in two cases: party 1 alone observes evidence and the state is in area A (so party 1 does not disclose); or party 2 alone observes evidence and the state is in area B (so party 2 does not disclose). In both cases, wrong adjudication decisions are made in the no signaling equilibrium but correct adjudication is made in the lawyer signaling equilibrium. Still, the lawyer signaling equilibrium involves incorrect decisions if party 1 (resp, 2) alone observes evidence and the state is in area C (resp, D). But as w increases toward $\overline{w}$, the regions of erroneous ruling shrink, and the social welfare increases. When $w \geq \overline{w}$, there is full disclosure, so adjudication decisions are always correct. The following result shows that costly legal advice can be socially valuable.

[22] The signaling equilibrium is robust to the possible hiring of a lawyer secretly. A secret representation has no benefit when a party ends up not disclosing evidence, for he will be subject to a prejudicial inference (and while paying the lawyer cost of w). The same is true when the party ends up disclosing. Hence hiring a lawyer secretly is no better than hiring a lawyer openly.

[23] The off-equilibrium beliefs in Propositions 7 and 8 survive the tests based on the intuitive criterion of Cho and Kreps (1987) and D1 criterion of Banks and Sobel (1987). The key observation is that on the equilibrium path the expected payoff to a party is greater when her evidence is more favorable to her. Therefore, whenever the judge's response to an off-equilibrium action makes a deviation profitable for a party with some evidence, this response gives the same payoff to the party with the most unfavorable evidence and hence this deviation is also profitable for the latter party. So the beliefs that attribute a unilateral deviation to a litigant having the most unfavorable evidence survive the tests of D1 and intuitive criterion.

COROLLARY 1: *The lawyer signaling equilibrium yields higher social welfare than the no signaling equilibrium. The social welfare is increasing in the lawyer fee* w *on* $(0, \overline{w}]$ *and attains the first-best at* $w = \overline{w}$, *in which case all evidence is disclosed.*

The social value of lawyers in our model comes from greater disclosure. Legal fees serve as the cost of signaling lack of evidence without prejudice. So, increased legal fees put more pressure on the parties to disclose their evidence instead of buying off the right to be silent.[24]

The results of this section also inform the debate regarding regulating adjudicators' inferences. In particular, they imply that negative inferences from nondisclosure can improve the disclosure incentives which leads to a better outcome. At the minimum, this suggests that a careful examination of the rules governing adjudicators' inferences is warranted. This is notable since adjudicators are often prohibited from drawing negative inferences against parties refusing to disclose their information, lest such inference distort the judgment. However, little is known about how such inferences affect the parties' disclosure incentives.

Thus far, we have characterized two equilibria. A natural question is whether there are any other equilibria. We show that, given plausible restrictions, no other equilibrium exists.

PROPOSITION 9: *Suppose that parties do not randomize in their representation choices,*[25] *the represented parties always follow the dominant strategy of disclosing all favorable evidence and withholding all unfavorable evidence, and the judge follows a threshold strategy in her default ruling. Then there are no other equilibria besides the lawyer signaling equilibrium of Proposition* 8 *and the no signaling equilibrium of Proposition* 7.

V. Conclusions

In this paper, we have studied the effect of lawyer advising on disclosure and adjudication. Our analysis was concerned with the role of lawyers as gatekeepers of information reaching the court. We have shown that lawyer advising does not affect the adjudication outcome if legal advice is costless and the lawyers cannot credibly suppress favorable evidence. At the same time, lawyer advising can affect the adjudication outcome if lawyers can leverage their legal expertise to engage in more sophisticated strategic behavior. First, they can do so by credibly withholding some favorable evidence and thereby affecting the judge's inference in their client's favor when evidence is not disclosed. Second, lawyer advising can play a signaling

[24] Note that in our lawyer signaling equilibrium, a party without a lawyer always discloses evidence. A represented party discloses only some of the time. This feature seems to accord well with anecdotal evidence. Specifically, it is well-known that lawyers often advise their clients not to communicate or divulge information to police and prosecutors in criminal cases, and/or not to testify during trials.

[25] There is only one other possible equilibrium scenario when such randomization is allowed: an uninformed party randomizes between hiring and not hiring a lawyer. Although we have not been able to completely rule out this type of equilibrium, it is easy to show that the qualitative nature of the outcome in such an equilibrium would be the same as in the lawyer signaling equilibrium of Proposition 8.

role. In our signaling equilibrium, hiring a lawyer allows a litigating party to avoid a negative inference from nondisclosure, so legal representation "buys the right to be silent without prejudice," which ultimately affects the litigation outcome.

Besides illuminating the lawyers' role in disclosure, this paper yields useful insights regarding several related issues. First, our findings help to understand the implications of quality differences in legal advice. There is a concern that high-profile lawyers may influence the outcome to the point of jeopardizing fair adjudication. Our model can be used to understand the effect of quality differences in lawyering, once we interpret self-representation as representation by an inexperienced lawyer. Second, our study also helps to understand the role of the adjudicators' interpretation of evidence and the inferences they may draw from nondisclosure of evidence. These aspects of legal proceedings often affect a court's ruling but remain poorly understood.[26]

More broadly, our results shed light on the role of advising in settings other than legal disputes. Agents and divisions often compete for resources within organizations. Resource allocation decisions—which could take such forms as merit assignment, promotion of the employees, and budget allocation between divisions—in turn depend on the information provided by those closely affected by the decisions. Advising agents regarding disclosure of information can affect both the quality of information transmission and the resource allocation decision itself. Our results offer basic insights regarding the role of advising in such circumstances.

For simplicity, we have assumed that the adjudicators update evidence (or lack thereof) in the Bayesian fashion. While this is a natural modeling approach, in practice judges and juries make non-Bayesian inferences often constrained by laws and regulations designed to avoid certain prejudices. Che and Severinov (2014) show that our main results are robust to such non-Bayesian updating. Nevertheless, the procedural rules on the adjudicators' inferences raise interesting and important questions that require deeper understanding. We leave this topic for future research.

Appendix: Proofs

We first establish several lemmas that will be used in the proofs.

LEMMA A1: *For any* $\theta' > \theta$,

$$\int_0^{\hat{s}} g(s,\theta) f(s\,|\,\theta')\; ds \;\geq\; \min\left\{0, \int_0^{\hat{s}} g(s,\theta) f(s\,|\,\theta)\; ds\right\}.$$

PROOF:

Both $\frac{F(s\,|\,\theta)}{F(\hat{s}|\theta)}$ and $\frac{F(s|\theta')}{F(\hat{s}|\theta')}$ are cdfs on $[0,\hat{s}]$. By MLRP, $\frac{F(s|\theta')}{F(\hat{s}|\theta')}$ first-order stochastically dominates $\frac{F(s\,|\,\theta)}{F(\hat{s}|\theta)}$. Therefore, since $g(s,\theta)$ is increasing in s, we have:

[26] Che and Severinov (2014) explore the implications of our model to study the regulations and the restrictions placed on them.

$\int_0^{\hat{s}} g(s,\theta)\frac{f(s|\theta')}{F(\hat{s}|\theta')}\,ds \geq \int_0^{\hat{s}} g(s,\theta)\frac{f(s\,|\,\theta)}{F(\hat{s}|\theta)}\,ds$, which can be rewritten as: $\int_0^{\hat{s}} g(s,\theta)f(s|\theta')\,ds \geq \frac{F(\hat{s}|\theta')}{F(\hat{s}|\theta)}\int_0^{\hat{s}} g(s,\theta)f(s\,|\,\theta)\,ds$. The latter inequality implies that, if $\int_0^{\hat{s}} g(s,\theta)f(s\,|\,\theta)\,ds \geq 0$, then $\int_0^{\hat{s}} g(s,\theta)f(s|\theta')\,ds \geq 0$. Also, note that $\frac{F(\hat{s}|\theta')}{F(\hat{s}|\theta)} \leq 1$. Hence, if $\int_0^{\hat{s}} g(s,\theta)f(s\,|\,\theta)\,ds \leq 0$, then $\int_0^{\hat{s}} g(s,\theta)f(s|\theta')\,ds \geq \int_0^{\hat{s}} g(s,\theta)f(s\,|\,\theta)\,ds$. ■

LEMMA A2: *Fix any* $\hat{s} \in (0,1)$. *If* $E[g|\hat{s},\hat{s},\theta] \geq 0$, *then* $E[g|\hat{s},\hat{s},\theta'] > 0$ *for any* $\theta' > \theta$.

PROOF:

Suppose that $E[g|\hat{s},\hat{s},\theta] \geq 0$. Recall that

$$E[g|\hat{s},\hat{s},\theta] := p_{00}\int_0^1 g(s,\theta)f(s\,|\,\theta)\,ds + p_{10}\int_{\hat{s}}^1 g(s,\theta)f(s\,|\,\theta)\,ds + p_{01}\int_0^{\hat{s}} g(s,\theta)f(s\,|\,\theta)\,ds. \tag{6}$$

The result follows from several observations. By MLRP and monotonicity of $g(\,\cdot\,,\cdot\,)$, for $\theta' > \theta$,

$$\int_0^1 g(s,\theta')f(s\,|\,\theta')\,ds > \int_0^1 g(s,\theta)f(s\,|\,\theta)\,ds \tag{7}$$

and

$$\frac{\int_{\hat{s}}^1 g(s,\theta')f(s\,|\,\theta')\,ds}{1-F(\hat{s}\,|\,\theta')} > \frac{\int_{\hat{s}}^1 g(s,\theta)f(s\,|\,\theta)\,ds}{1-F(\hat{s}\,|\,\theta)}. \tag{8}$$

Note that $\int_{\hat{s}}^1 g(s,\theta)f(s\,|\,\theta)\,ds \geq 0$, for otherwise (6) would imply that $E[g|\hat{s},\hat{s},\theta] < 0$, in contradiction to our original assumption. Further, by MLRP, $1-F(\hat{s}\,|\,\theta') \geq 1-F(\hat{s}\,|\,\theta)$. It thus follows from (8) that

$$\int_{\hat{s}}^1 g(s,\theta')f(s\,|\,\theta')\,ds \geq \int_{\hat{s}}^1 g(s,\theta)f(s\,|\,\theta)\,ds. \tag{9}$$

By (7), the first term $\int_0^1 g(s,\theta)f(s\,|\,\theta)\,ds$ in (6) is strictly increasing in θ, and by (9), its second term $\int_{\hat{s}}^1 g(s,\theta)f(s\,|\,\theta)\,ds$ is nondecreasing in θ.

Now consider the third term. Suppose first that $\int_0^{\hat{s}} g(s,\theta)f(s\,|\,\theta)\,ds \geq 0$. Then by Lemma A1, $\int_0^{\hat{s}} g(s,\theta')f(s|\theta')\,ds \geq 0$. This implies that both $\int_0^1 g(s,\theta')f(s|\theta')\,ds > 0$ and $\int_{\hat{s}}^1 g(s,\theta')f(s|\theta')\,ds > 0$. Hence, $E[g|\hat{s},\hat{s},\theta'] > 0$. Suppose next $\int_0^{\hat{s}} g(s,\theta)f(s\,|\,\theta)\,ds < 0$. Then, by Lemma A1 $\int_0^{\hat{s}} g(s,\theta')f(s|\theta')\,ds \geq \int_0^{\hat{s}} g(s,\theta)f(s\,|\,\theta)\,ds$. Combining this fact with (7) and (9), we again conclude that $E[g|\hat{s},\hat{s},\hat{\theta}'] > 0$. ■

LEMMA A3: *If* $E[g\,|\,h(\theta), h(\theta), \theta] \geq 0$, *then* $E[g\,|\,h(\theta'), h(\theta'), \theta'] > 0$ *for any* $\theta' > \theta$.

PROOF:
Recall that

$$(10) \quad E[g\,|\,h(\theta), h(\theta), \theta] = p_{00}\int_0^1 g(s,\theta)f(s\,|\,\theta)\;ds + p_{10}\int_{h(\theta)}^1 g(s,\theta)f(s\,|\,\theta)\;ds + p_{01}\int_0^{h(\theta)} g(s,\theta)f(s\,|\,\theta)\;ds.$$

The first term of (10) is increasing in θ by (7). Now consider the second term. We have:

$$(11) \quad \int_{h(\theta')}^1 g(s,\theta')f(s\,|\,\theta')\;ds = \int_{h(\theta)}^1 g(s,\theta')f(s\,|\,\theta')\;ds + \int_{h(\theta')}^{h(\theta)} g(s,\theta')f(s|\,\theta')\;ds > \int_{h(\theta)}^1 g(s,\theta)f(s\,|\,\theta)\;ds,$$

where the inequality follows from (9) and the fact that $h(\theta') < h(\theta)$ and $g(s,\theta') > 0$ if $s > h(\theta')$.

Now consider the third term. Since $\int_0^{h(\theta)} g(s,\theta)f(s\,|\,\theta)\;ds < 0$, and $g(s,\theta)$ is strictly increasing in θ, Lemma A1 implies that

$$(12) \quad \int_0^{h(\theta)} g(s,\theta')f(s|\,\theta')\;ds > \int_0^{h(\theta)} g(s,\theta)f(s\,|\,\theta)\;ds.$$

Differentiating $\int_0^{h(\theta)} g(s,\theta)f(s\,|\,\theta)$ with respect to θ, we have

$$h'(\theta)g(h(\theta),\theta)f(h(\theta)\,|\,\theta) + \frac{d}{d\tilde{\theta}}\left[\int_0^{h(\theta)} g(s,\tilde{\theta})f(s|\tilde{\theta})\;ds\right]_{\tilde{\theta}=\theta} \geq 0,$$

since the first term vanishes and the second term is nonnegative by (12).

Combining the observations, we conclude that (10) is strictly increasing in θ. ■

PROOF OF PROPOSITION 1:
The proof consists of several steps.

Step 1: In any equilibrium, parties 1 and 2 use cutoff strategies with the same threshold, i.e., there exists $\hat{s}$ such that party 1 discloses evidence s if (only if) $s < (\leq)\hat{s}$, and party 2 discloses evidence s if (only if) $s > (\geq)\hat{s}$.

PROOF:
Fix any equilibrium and suppose that party 1 has observed evidence s. Party 1's disclosure decision affects the outcome of the trial only if party 2 does not disclose the evidence. Let $P_0(s)$ denote the probability that the judge rules for party 1 in that

equilibrium if s is not disclosed. This probability depends on s, because the judge's decision depends only on the value of θ, and s and θ are (weakly) affiliated. On the other hand, if party 1 discloses s, then the judge will rule for party 1 if $g(s,\theta) < 0$, or $\theta < h^{-1}(s)$. Thus, party 1 discloses s in that equilibrium if (only if)

$$P_0(s) < (\leq) \Pr\{\theta < h^{-1}(s) \,|\, s\}. \tag{13}$$

Similarly, party 2 discloses s if (only if)

$$P_0(s) > (\geq) \Pr\{\theta < h^{-1}(s) \,|\, s\}. \tag{14}$$

Thus, the disclosure incentives of the two parties are precisely the opposite.

To establish that parties 1 and 2 use cutoff strategies in any equilibrium, we show that, for any $s' > s$, $P_0(s') > \Pr\{\theta < h^{-1}(s') \,|\, s'\}$ if $P_0(s) = \Pr\{\theta < h^{-1}(s) \,|\, s\}$. Suppose the judge follows a default ruling strategy, $\delta(\theta)$, i.e., she rules for party 2 with probability $\delta(\theta)$ given θ and nondisclosure. Let $f(\theta \,|\, s)$ denote (with slight abuse of notation) the conditional density of θ conditional on s. Then, we have:

$$\begin{aligned}
P_0(s') &\equiv \int_0^1 (1-\delta(\theta)) f(\theta \,|\, s')\, d\theta \\
&= \int_0^{h^{-1}(s)} (1-\delta(\theta)) f(\theta \,|\, s')\, d\theta + \int_{h^{-1}(s)}^1 (1-\delta(\theta)) \frac{f(\theta | s')}{f(\theta \,|\, s)} f(\theta \,|\, s)\, d\theta \\
&\geq \int_0^{h^{-1}(s)} (1-\delta(\theta)) f(\theta \,|\, s')\, d\theta + \int_{h^{-1}(s)}^1 (1-\delta(\theta)) \frac{f(h^{-1}(s) \,| s')}{f(h^{-1}(s) \,|\, s)} f(\theta \,|\, s)\, d\theta \\
&= \int_0^{h^{-1}(s)} (1-\delta(\theta)) f(\theta \,|\, s')\, d\theta + \int_0^{h^{-1}(s)} \delta(\theta) \frac{f(h^{-1}(s) \,| s')}{f(h^{-1}(s) \,|\, s)} f(\theta \,|\, s)\, d\theta \\
&\geq \int_0^{h^{-1}(s)} (1-\delta(\theta)) f(\theta \,|\, s')\, d\theta + \int_0^{h^{-1}(s)} \delta(\theta) \frac{f(\theta | s')}{f(\theta \,|\, s)} f(\theta \,|\, s)\, d\theta \\
&= \int_0^{h^{-1}(s)} f(\theta \,|\, s')\, d\theta = \Pr\{\theta < h^{-1}(s) \,|\, s'\} > \Pr\{\theta < h^{-1}(s') \,|\, s'\}.
\end{aligned}$$

The first and the last two equalities in this sequence hold by definition. The two non-strict inequalities hold by MLRP. The equality between them holds because

$$\begin{aligned}
P_0(s) = \Pr\{\theta < h^{-1}(s) \,|\, s\} &\Leftrightarrow \int_0^1 (1-\delta(\theta)) f(\theta \,|\, s)\, d\theta = \int_0^{h^{-1}(s)} f(\theta \,|\, s)\, d\theta \\
&\Leftrightarrow \int_{h^{-1}(s)}^1 (1-\delta(\theta)) f(\theta \,|\, s)\, d\theta = \int_0^{h^{-1}(s)} \delta(\theta) f(\theta \,|\, s)\, d\theta.
\end{aligned}$$

The lone strict inequality holds because $h^{-1}(\,\cdot\,)$ is strictly decreasing, and s and θ are affiliated.

A symmetric argument establishes that, for all $s'' < s$, $P_0(s'') < \Pr\{\theta < h^{-1}(s'') \mid s''\}$ if $P_0(s) = \Pr\{\theta < h^{-1}(s) \mid s\}$.

In combination, these results imply the existence of a common threshold $\hat{s} \in [0,1]$ such that party 1 discloses (withholds) s if $s < \hat{s}$ ($s > \hat{s}$) and party 2 discloses (withholds) s if $s > \hat{s}$ ($s < \hat{s}$).[27] ∎

Step 2: In any equilibrium, the judge follows a cutoff strategy in her default ruling; i.e., there exists $\hat{\theta}$ such that $\delta(\theta) = 0$ if $\theta < \hat{\theta}$ and $\delta(\theta) = 1$ if $\theta > \hat{\theta}$.

PROOF:

By Step 1, the parties follow cutoff disclosure strategies with some common threshold $\hat{s}$. Hence, the judge's posterior on party 1's culpability when she observes θ is given by $E[g|\hat{s}, \hat{s}, \theta]$. Then, by Lemma A2, there exists $\hat{\theta} \in \Theta$ such that $\delta(\theta) = 0$ if $\theta < \hat{\theta}$ and $\delta(\theta) = 1$ if $\theta > \hat{\theta}$. ∎

Step 3: If $\hat{s}$ is the parties' common threshold and $\hat{\theta}$ is the judge's threshold, then $\hat{s} = h(\hat{\theta})$.

PROOF:

Since the parties' strategies must constitute best responses to the judge's default ruling strategy with threshold $\hat{\theta}$, we must have

$$P_0(s) = \Pr\{\theta < \hat{\theta} \mid s\}.$$

Hence, the optimality of the cutoff strategies with threshold $\hat{s}$, together with (13) and (14), implies that $\Pr\{\theta < \hat{\theta} \mid s\} < (\leq) \Pr\{\theta < h^{-1}(s) | s\}$ if (only if) $s < (\leq) \hat{s}$. Similarly, $\Pr\{\theta < \hat{\theta} \mid s\} > (\geq) \Pr\{\theta < h^{-1}(s) | s\}$ if (only if) $s > (\geq) \hat{s}$. Therefore, $\hat{s} = h(\hat{\theta})$. ∎

Step 4: It is an equilibrium for the judge to follow a cutoff strategy with threshold $\hat{\theta}^*$ and for the parties to follow cutoff strategies with a common threshold $h(\hat{\theta}^*)$.

PROOF:

Recall from (3) that $\hat{\theta}^* := \inf\{\theta \in \Theta \mid E[g \mid h(\theta), h(\theta), \theta] > 0\}$. It then follows from Lemma A2 that

$$E[g \mid h(\hat{\theta}^*), h(\hat{\theta}^*), \theta] \gtreqless 0 \quad \text{if} \quad \theta \gtreqless \hat{\theta}^*.$$

So, the judge's cutoff strategy with threshold $\hat{\theta}^*$ is optimal when the parties adopt cutoff strategies with common threshold $h(\hat{\theta}^*)$. Likewise, Steps 1 and 3 show that the parties' cutoff strategies with common threshold $h(\hat{\theta}^*)$ are best responses to the

[27] If some party, say party 1, has a strict incentive for disclosing all s, then the statement remains valid with $\hat{s} = 1$.

judge's cutoff strategy with threshold $\hat{\theta}^*$. Hence, this strategy profile constitutes a perfect Bayesian equilibrium. ∎

Step 5: The equilibrium described in Step 4 is unique.

PROOF:

The uniqueness follows from the uniqueness of the judge's threshold, which in turn follows from Lemma A3.

The comparative statics result (the last statement) follows from the inspection of (10). Suppose p_{10} rises and (p_{00}, p_{01}, p_{11}) falls in an arbitrary way to satisfy $p_{10} + p_{00} + p_{01} + p_{11} = 1$. Then, the indicator function $\mathbf{1}_{\{E[g|h(\theta), h(\theta), \theta] \geq 0\}}$, which takes value 1 if $E[g\,|\,h(\theta), h(\theta), \theta] \geq 0$ and 0 otherwise, can only increase. In particular, $E[g\,|\,h(\theta), h(\theta), \theta]$ becomes strictly positive at θ for which $E[g\,|\,h(\theta), h(\theta), \theta] = 0$ before the change. This proves that $\hat{\theta}^*$ must decrease. A symmetric argument implies that $\hat{\theta}^*$ increases as p_{01} increases. ∎

PROOF OF PROPOSITION 2:

The weak dominance of the parties' disclosure strategies is already established in the text. Given the disclosure strategies, when the judge observes θ, her posterior of party 1's culpability is given by $E[g\,|\,h(\theta), h(\theta), \theta]$. Recall that

$$\hat{\theta}^* := \inf\{\theta \in \Theta\,|\,E[g\,|\,h(\theta), h(\theta), \theta] > 0\}.$$

Lemma A3 then implies that

$$E[g\,|\,h(\theta), h(\theta), \theta] \gtreqqless 0 \quad \text{if} \quad \theta \gtreqqless \hat{\theta}^*,$$

proving that the judge's cutoff default ruling strategy with threshold $\hat{\theta}^*$ is optimal. The uniqueness of the equilibrium follows from the uniqueness of the equilibrium threshold, which in turn follows from Lemma A3 and the definition of $\hat{\theta}^*$. ∎

PROOF OF PROPOSITION 3:

Suppose without loss of generality that party 1 has hired a lawyer but party 2 has not. (The opposite case is completely symmetric.) Then, party 1 has a dominant strategy of disclosing (withholding) s if $s > h(\theta)$ ($s < h(\theta)$). Just as in Proposition 1, party 2 will adopt a cutoff strategy with some threshold $\hat{s} \in S$.

Consider next the judge's default ruling strategy. Given θ and nondisclosure of s, the judge's posterior becomes $E[g\,|\,h(\theta), \hat{s}, \theta]$. Lemmas A1 and A2 imply that this posterior is ordinally monotonic: i.e., $E[g\,|\,h(\theta), \hat{s}, \theta] \geq 0$ implies $E[g\,|\,h(\theta'), \hat{s}, \theta'] > 0$ for $\theta' > \theta$. Hence, the judge adopts a cutoff strategy with some threshold $\hat{\theta}$. Then, the same argument as in Proposition 1 can be used to establish that $\hat{s} = h(\hat{\theta})$. It then follows that $\hat{\theta} = \hat{\theta}^*$. Further, the equilibrium threshold $\hat{\theta}^*$ is unique by Lemma A2. The comparative statics of $\hat{\theta}^*$ with respect to p_{10} and p_{01} follows immediately from Proposition 1. ∎

PROOF OF PROPOSITION 4:

Suppose that the judge follows a cutoff strategy with a threshold $\hat{\theta} \in \Theta$. We show that any combination of the parties' best response disclosure strategies leads to the same outcome, regardless of whether either party has obtained legal advice. To begin, given the threshold $\hat{\theta}$, let $\mathcal{S}_1^{\hat{\theta}}$ be a set of party 1's disclosure strategies such that: (i) if $\rho_1(s,\theta) \in \mathcal{S}_1^{\hat{\theta}}$, then $\rho_1(s,\theta) = 1$ for almost every (s,θ) with $\theta > \hat{\theta}$ and $g(s,\theta) < 0$; (ii) $\rho_1(s,\theta) = 0$ for almost every (s,θ) with $\theta < \hat{\theta}$ and $g(s,\theta) > 0$. Similarly, let $\mathcal{S}_2^{\hat{\theta}}$ be the set of disclosure strategies for party 2 such that, if $\rho_2(s,\theta) \in \mathcal{S}_2^{\hat{\theta}}$, then $\rho_2(s,\theta) = 0$ for almost every (s,θ) with $\theta > \hat{\theta}$ and $g(s,\theta) < 0$, and $\rho_2(s,\theta) = 1$ for almost every (s,θ) with $\theta < \hat{\theta}$ and $g(s,\theta) > 0$. In words, a party $i = 1,2$ following a strategy in $\mathcal{S}_i^{\hat{\theta}}$ will always present evidence that would overturn an unfavorable default ruling and will never present evidence that would overturn a favorable ruling. Such a strategy is optimal for each party, regardless of the opponent's disclosure strategy. If the opponent discloses, then a party's strategy has no effect, whereas if the opponent does not disclose, then no other strategy can make the party strictly better off. Therefore, if the judge follows a default ruling strategy with a cutoff threshold $\hat{\theta}$, then any pair of parties' disclosure strategies $(\rho_1,\rho_2) \in \mathcal{S}_1^{\hat{\theta}} \times \mathcal{S}_2^{\hat{\theta}}$, induces the outcome in (4).

If a party $i = 1,2$ has obtained legal advice, clearly all strategies in $\mathcal{S}_i^{\hat{\theta}}$ are feasible. Importantly, party i without legal advice also has access to a strategy in $\mathcal{S}_i^{\hat{\theta}}$. This can be seen by the fact that a simple cutoff strategy $\hat{\rho}_1(s,\theta) := \mathbf{1}_{\{s<h(\hat{\theta})\}}$ does not depend on θ (so it is a feasible strategy for party 1 without lawyer advice), yet it belongs to $\mathcal{S}_1^{\hat{\theta}}$. Likewise, $\hat{\rho}_2(s,\theta) := \mathbf{1}_{\{s>h(\hat{\theta})\}}$ is feasible for party 2 when he has no legal advice, but it belongs to $\mathcal{S}_2^{\hat{\theta}}$.

Finally, to complete the proof, fix any legal regime, and suppose $\rho_i(s,\theta)$ is party i's best response to some strategy of player j and the judge's threshold strategy $\hat{\theta}$. Then, we must have $\rho_i \in \mathcal{S}_i^{\hat{\theta}}$. Otherwise, one can show that the strategy is strictly worse for party i than the simple cutoff strategy $\hat{\rho}_i(s,\theta)$, which is available for that party in every legal regime. The argument for this result is essentially the same as the one provided prior to Proposition 1. ∎

PROOF OF PROPOSITION 5:

Before proceeding, it is useful to establish some preliminary results. The arguments employed in Lemmas A2 and A3 can be combined to show that $\forall \theta' > \theta$:

$$E[g\,|\,0,h(\theta),\theta] \geq 0 \Rightarrow E[g\,|\,0,h(\theta'),\theta'] > 0;$$

$$E[g\,|\,h(\theta),1,\theta] \geq 0 \Rightarrow E[g\,|\,h(\theta'),1,\theta'] > 0.$$

From these, it follows that $E[g\,|\,0,h(\theta),\theta] > 0$ if and only if $\theta > \hat{\theta}_+$, and $E[g\,|\,h(\theta),1,\theta] > 0$ if and only if $\theta > \hat{\theta}_-$.

We first prove (i). Fix any $\hat{\theta} \in [\hat{\theta}_-,\hat{\theta}_+]$. We shall prove that there exists an equilibrium in which the judge adopts a cutoff strategy with threshold $\hat{\theta}$. In this equilibrium, party 1 discloses s if and only if $s < h(\theta)$ and $\theta > \hat{\theta}$, whereas party 2 discloses s if and only if $s > h(\theta)$ and $\theta < \hat{\theta}$. Given these disclosure strategies, the judge's posterior becomes $E[g \,|\, 0, h(\theta), \theta] < 0$ if $\theta < \hat{\theta}$ and $E[g \,|\, h(\theta), 1, \theta] > 0$ if $\theta > \hat{\theta}$. Hence, it is optimal for the judge to rule for party 1 if and only if $\theta < \hat{\theta}$. Given the default ruling by the judge, party i's $(i = 1, 2)$ disclosure strategy is in $\mathcal{S}_i^{\hat{\theta}}$ and hence constitutes a best response. The first statement is thus proven.

Next, consider the converse. To prove that any equilibrium threshold $\hat{\theta}$ must be in $[\hat{\theta}_-, \hat{\theta}_+]$, suppose to the contrary that there exists an equilibrium strategy combination $(\hat{\theta}, \rho_1, \rho_2)$ such that, $\hat{\theta} \notin [\hat{\theta}_-, \hat{\theta}_+]$. Without loss, let $\hat{\theta} > \hat{\theta}_+$. Then, since the judge rules for party 1 for any $\tilde{\theta} \in (\hat{\theta}_+, \hat{\theta})$, we must have $E[g|\, \rho_1, \rho_2, \tilde{\theta}] \leq 0$. But, comparing $E[g|\, \rho_1, \rho_2, \tilde{\theta}]$ and $E[g \,|\, 0, h(\tilde{\theta}), \tilde{\theta}]$ term by term,[28] we conclude that $E[g|\, \rho_1, \rho_2, \tilde{\theta}] > 0$ if $E[g \,|\, 0, h(\tilde{\theta}), \tilde{\theta}] > 0$. Since $\tilde{\theta} > \hat{\theta}_+$, $E[g \,|\, 0, h(\tilde{\theta}), \tilde{\theta}] > 0$, so $E[g|\, \rho_1, \rho_2, \tilde{\theta}] > 0$, which is a contradiction. Hence, we conclude that $\hat{\theta} \leq \hat{\theta}_+$. A symmetric argument proves that $\hat{\theta} \geq \hat{\theta}_-$.

Turning now to part (ii), let us without loss of generality focus on the case in which only party 1 obtains legal advice. Choose any $\hat{\theta} \in [\hat{\theta}^*, \hat{\theta}_+]$. Consider the disclosure strategies whereby party 1 discloses s if and only if $s < h(\theta)$ and $\theta > \hat{\theta}$, and party 2 (who does not have legal advice) discloses s if and only if $s > h(\hat{\theta})$. This pair of strategies constitute best responses given the judge's threshold $\hat{\theta}$. Under these disclosure strategies, the judge's posterior is $E[g \,|\, 0, h(\hat{\theta}), \theta]$ if $\theta < \hat{\theta}$ and $E[g \,|\, h(\theta), h(\hat{\theta}), \theta]$ if $\theta > \hat{\theta}$. But $E[g \,|\, 0, h(\hat{\theta}), \theta] < 0$ and $E[g \,|\, h(\theta), h(\hat{\theta}), \theta] > 0$ because $\hat{\theta} \in [\hat{\theta}^*, \hat{\theta}_+]$. Hence, the judge's cutoff strategy is optimal. The proof of the converse is analogous to that for part (i) and is therefore omitted. ∎

PROOF OF PROPOSITION 6:

The objective function (WP) can be rewritten as follows:

$$
\begin{aligned}
& p_{11} \int_{s>h(\theta)} g(s,\theta) f(s \,|\, \theta) l(\theta) \; ds d\theta + p_{10} \int_{\theta > \hat{\theta}} \int_{s > h(\theta)} g(s,\theta) f(s \,|\, \theta) l(\theta) \; ds d\theta \\
& \quad + \; p_{01} \left\{ \int_{\theta > \hat{\theta}} \int_0^1 g(s,\theta) f(s \,|\, \theta) l(\theta) \; ds d\theta + \int_{\theta \leq \hat{\theta}} \int_{s > h(\theta)} g(s,\theta) f(s \,|\, \theta) l(\theta) \; ds d\theta \right\} \\
& \quad + p_{00} \int_{\theta > \hat{\theta}} \int_0^1 g(s,\theta) f(s \,|\, \theta) l(\theta) \; ds d\theta,
\end{aligned}
$$

where $l(\,\cdot\,)$ is the marginal density of θ.

[28] First, the two expressions differ only in two terms. (The last term vanishes since in both cases at least one party discloses—a consequence of Lemma 1.) The comparison in the two terms is as follows. First, $\int_0^1 (1 - \rho_2(s, \tilde{\theta}))\, g(s, \tilde{\theta}) f(s|\tilde{\theta}) \; ds \geq \int_0^{h(\theta)} g(s, \tilde{\theta}) f(s|\tilde{\theta}) \; ds$. Next, since $\rho_1 \in \mathcal{S}_2^{\hat{\theta}}$, by Lemma 1, $\rho_1(s, \tilde{\theta}) = 0$ if $g(s, \tilde{\theta}) > 0$ and $\tilde{\theta} < \hat{\theta}$. Hence, $\int_0^1 (1 - \rho_1(s, \tilde{\theta}))\, g(s, \tilde{\theta}) f(s|\tilde{\theta}) \; ds \geq \int_0^1 g(s, \tilde{\theta}) f(s|\tilde{\theta}) \; ds$.

Differentiating this with respect to $\hat{\theta}$ yields

$$-p_{00}\int_0^1 g(s,\hat{\theta})f(s\,|\,\theta)\;dsl(\hat{\theta}) - p_{10}\int_{h(\hat{\theta})}^1 g(s,\hat{\theta})f(s|\hat{\theta})\;dsl(\hat{\theta}) - p_{01}\int_0^{h(\hat{\theta})} g(s,\hat{\theta})f(s|\hat{\theta})\;dsl(\hat{\theta})$$
$$\gtreqless 0 \quad \text{if} \quad \hat{\theta} \lesseqgtr \hat{\theta}^*.$$

So the objective function (WP) attains its maximum at $\hat{\theta}^*$. ∎

PROOF OF PROPOSITION 8:

Given the strategies adopted by parties 1 and 2, the judge's ruling strategy is rational under Bayesian beliefs. In particular, the symmetry between the two parties implies that the judge's threshold of one-half is optimal when both parties are represented. We next show that the disputing parties' strategies are sequentially rational. Given the symmetry, it suffices to check only party 1's incentives for a deviation. The proof consists of several steps.

Step 1: If party 1 observes $s \in [0,1]$ and does not retain a lawyer and the judge and party 2 follow the candidate equilibrium strategies, then it is optimal for party 1 to disclose s.

PROOF:

If party 1 discloses s, he wins if and only if $\theta < h^{-1}(s)$. So, his expected payoff is equal to

$$vL(h^{-1}(s))\,. \tag{15}$$

If party 1 does not disclose, the expected outcome depends on the value of s. Suppose, first, that $s < 1 - \hat{s}(w)$. With probability $1 - p$, party 2 does not observe s. He then hires a lawyer and makes no disclosure. With probability p, party 2 observes s. He then hires a lawyer and discloses s if $\theta > h^{-1}(s)$. In either case, whenever party 2 never discloses, the off-the-path belief is $s = 1$. Hence, the judge rules for party 1 if and only if $\theta < h^{-1}(1)$. So, party 1's payoff from nondisclosure is

$$vL(h^{-1}(1))\,. \tag{16}$$

Next, suppose that $s > 1 - \hat{s}(w)$. Then, if party 2 observes s, he does not retain a lawyer and discloses s. If party 2 does not observe s, then he retains a lawyer and makes no disclosure. Hence, given the judge's strategy, party 1's payoff from nondisclosure is

$$vpL(h^{-1}(s)) + v(1-p)L(h^{-1}(1)). \tag{17}$$

Since $L(h^{-1}(s))$ is nonincreasing in s, each of (16) and (17) is less than (15). Therefore, if party 1 observes $s \in [0,1]$ and does not retain a lawyer, it is optimal for him to disclose. ∎

Step 2: If party 1 observes $s \in [0,1]$ and the judge and party 2 follow the candidate equilibrium strategies, then it is optimal for party 1 to retain a lawyer if and only if $s \geq \hat{s}(w)$.

PROOF:

If party 1 retains a lawyer, then it is optimal for him to disclose s if and only if $\theta \leq h^{-1}(s)$. Let us compute party 1's expected payoff associated with this strategy. If party 2 also observes s (which happens with probability p), then he discloses either if $s > 1 - \hat{s}(w)$ (without hiring a lawyer) or if $s \leq 1 - \hat{s}(w)$ and $\theta > h^{-1}(s)$ (after hiring a lawyer). Either way, party 1 wins if and only if $\theta \leq h^{-1}(s)$.

If party 2 does not observe s (which happens with probability $1 - p$), he retains a lawyer and does not disclose. Given the judge's default ruling strategy when both sides are represented, party 1 wins if and only if $\theta < \max\left\{h^{-1}(s), \frac{1}{2}\right\}$. Thus, party 1's expected payoff when he hires a lawyer after observing s is

$$vpL(h^{-1}(s)) + v(1-p)L\left(\max\left\{h^{-1}(s), \frac{1}{2}\right\}\right) - w. \tag{18}$$

Suppose next that party 1 does not retain a lawyer after observing s. By Step 1, he will then always disclose s and receive the payoff given by (15). By (5), (18) is greater than (15) if and only if $s \geq \hat{s}(w)$. So the strategy of hiring a lawyer if and only if $s \geq \hat{s}(w)$ is, indeed, optimal. ∎

Step 3: If party 1 does not observe s and the judge and party 2 follow the candidate equilibrium strategies, then it is optimal for party 1 to retain a lawyer.

PROOF:

Suppose, that party 1 retains a lawyer. Given the other players' strategies, party 1 wins if $\theta \leq \frac{1}{2}$ and party 2 does not disclose, and if $\theta \leq h^{-1}(s)$ and party 2 discloses s. Hence, party 1's expected payoff is equal to

$$vp\left[K(1-\hat{s}(w))L\left(\frac{1}{2}\right) + \int_{1-\hat{s}(w)}^{1} L(h^{-1}(s))k(s)\,ds\right] + v(1-p)L\left(\frac{1}{2}\right) - w. \tag{19}$$

Meanwhile, if party 1 does not retain a lawyer, then his payoff becomes

$$\begin{aligned} &vp\left[K(1-\hat{s}(w))L\left(h^{-1}(1)\right) + \int_{1-\hat{s}(w)}^{1} L\left(h^{-1}(s)\right)k(s)\,ds\right] \\ &\quad + v(1-p)L\left(h^{-1}(1)\right). \end{aligned} \tag{20}$$

Subtracting (20) from (19) gives

$$\begin{aligned} v[pK(1-\hat{s}(w)) + 1 - p]\left(L\left(\frac{1}{2}\right) - L\left(h^{-1}(1)\right)\right) - w &> v(1-p)\left(L\left(\frac{1}{2}\right) - L\left(h^{-1}(1)\right)\right) - w \\ &\geq v(1-p)\left(L\left(\frac{1}{2}\right) - L\left(h^{-1}(s(w))\right)\right) - w = 0. \end{aligned}$$

Hence, it is optimal for party 1 to retain a lawyer in this case. ∎

Steps 1–3 establish that party 1's candidate equilibrium strategy constitutes his best response to party 2's and the judge's strategies. By symmetry, the same is true for party 2. The optimality of the judge's strategy was established earlier. ∎

PROOF OF PROPOSITION 9:

We first show that there is no equilibrium in which each party hires a lawyer with probability 1. The proof is by contradiction. Suppose to the contrary that such an equilibrium exists. Let $\hat{\theta}$ be the judge's threshold characterizing his default strategy in case of nondisclosure. If $h^{-1}(0) > \hat{\theta}$, then party 1 strictly prefers to remain unrepresented and disclose s when s is sufficiently close to zero. On the other hand, if $h^{-1}(0) \leq \hat{\theta}$ then $h^{-1}(1) \leq \hat{\theta}$ because $h^{-1}(\cdot)$ is decreasing. In this case, party 2 strictly prefers to remain unrepresented and disclose s when s is sufficiently close to 1.

We next rule out the existence of equilibria in which some party, say party 1, remains unrepresented with probability 1 when he does not learn s, while some informed types of this party hire a lawyer with positive probability. The proof is also by contradiction. So, suppose such an equilibrium exists. Since party 1's actions do not affect the outcome when party 2 discloses, let us focus on the case in which party 2 does not disclose s. On the equilibrium path, party 2 does not disclose with positive probability, because he does not learn s with a positive probability.

Nondisclosure by both parties when party 1 is represented and party 2 has not deviated from the candidate equilibrium leads the judge to conclude that party 1 possesses evidence s such that $g(s, \theta) > 0$. This is because party 1 hires a lawyer only when he is informed and, by assumption, represented parties disclose all favorable evidence. So, the judge will always rule against represented party 1 if there is no disclosure, and party 2's nondisclosure occurs on the equilibrium path. Therefore, in the candidate equilibrium, represented party 1 wins the dispute if and only if $g(s, \theta) < 0$. Yet, he could have done strictly better by not hiring a lawyer and disclosing with probability 1, generating a contradiction. ∎

REFERENCES

Banks, Jeffrey S., and Joel Sobel. 1987. "Equilibrium Selection in Signaling Games." *Econometrica* 55 (3): 647–61.

Bernardo, A. E., E. Talley, and I. Welch. 2000. "A theory of legal presumptions." *Journal of Law, Economics, and Organization* 16 (1): 1–49.

Bull, Jesse. 2008. "Costly Evidence Production and the Limits of Verifiability." *B.E. Journal of Theoretical Economics* 8 (1): Article 18.

Bull, Jesse. 2009. "Costly Evidence and Systems of Fact-Finding." *Bulletin of Economic Research* 61 (2): 103–25.

Bull, Jesse, and Joel Watson. 2004. "Evidence disclosure and verifiability." *Journal of Economic Theory* 118 (1): 1–31.

Bull, Jesse, and Joel Watson. 2007. "Hard evidence and mechanism design." *Games and Economic Behavior* 58 (1): 75–93.

Bundy, Stephen McG., and Einer Richard Elhauge. 1991. "Do Lawyers Improve the Adversary System? A General Theory of Litigation Advice and Its Regulation." *California Law Review* 79 (2): 313–420.

Bundy, Stephen McG., and Einer Elhauge. 1993. "Knowledge about Legal Sanctions." *Michigan Law Review* 92: 261–335.
Che, Yeon-Koo, and Navin Kartik. 2009. "Opinions as Incentives." *Journal of Political Economy* 117 (5): 815–60.
Che, Yeon-Koo, and Sergei Severinov. 2014. "Disclosure and Legal Advice." http://www.columbia.edu/~yc2271/files/papers/Lawyer_advice_12-02-2014.pdf.
Cho, In-Koo, and David M. Kreps. 1987. "Signaling Games and Stable Equilibria." *Quarterly Journal of Economics* 102 (2): 179–221.
Crawford, Vincent P., and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica* 50 (6): 1431–51.
Daughety, Andrew F., and Jennifer F. Reinganum. 2000. "Appealing Judgments." *Rand Journal of Economics* 31 (3): 502–25.
Deneckere, Raymond, and Sergei Severinov. 2008. "Mechanism design with partial state verifiability." *Games and Economic Behavior* 64 (2): 487–513.
Dewatripont, Mathias, and Jean Tirole. 1999. "Advocates." *Journal of Political Economy* 107 (1): 1–39.
Frank, Jerome. 1973. *Courts on Trial: Myth and Reality in American Justice*. Princeton: Princeton University Press.
Froeb, Luke M., and Bruce H. Kobayashi. 1996. "Naive, Biased, yet Bayesian: Can Juries Interpret Selectively Produced Evidence?" *Journal of Law, Economics, and Organization* 12 (1): 257–76.
Froeb, Luke M., and Bruce H. Kobayashi. 2001. "Evidence production in adversarial vs. inquisitorial regimes." *Economics Letters* 70 (2): 267–72.
Grossman, Sanford J. 1981. "The Informational Role of Warranties and Private Disclosure about Product Quality." *Journal of Law and Economics* 24 (3): 461–83.
Hadfield, Gillian K., and Shmuel Leshem. 2012. "Attorney-Client Confidentiality." In *Procedural Law and Economics*, edited by Chris Sanchirico, 52–66. Cheltenham, UK: Edward Elgar Publishing.
Hay, Bruce L., and Kathryn E. Spier. 1997. "Burdens of Proof in Civil Litigation: An Economic Perspective." *Journal of Legal Studies* 26 (2): 413–31.
Iossa, Elisabetta, and Bruno Jullien. 2012. "The market for lawyers and quality layers in legal services." *Rand Journal of Economics* 43 (4): 677–704.
Kaplow, Louis, and Stephen Shavell. 1989. "Legal Advice about Information to Present in Litigation: Its Effects and Social Desirability." *Harvard Law Review* 102 (3): 565–615.
Kartik, Navin, Wing Suen, and Frances Xu. 2013. "Multi-Sender Disclosure of Verifiable Information." http://web.law.columbia.edu/sites/default/files/microsites/contract-economic-organization/files/S14WK/kartik_-_multi-sender_disclosure.pdf.
Krishna, Vijay, and John Morgan. 2001. "A Model Of Expertise." *Quarterly Journal of Economics* 116 (2): 747–75.
Leshem, Shmuel. 2010. "The benefits of a right to silence for the innocent." *Rand Journal of Economics* 41 (2): 398–416.
Levy, Gilat. 2005. "Careerist Judges." *Rand Journal of Economics* 36 (2): 275–97.
Lewis, Tracy, and Michael Poitevin. 1997. "Disclosure of Information in Regulatory Proceedings." *Journal of Law, Economics, and Organization* 13 (1): 50–73.
Lipman, Barton L., and Duane J. Seppi. 1995. "Robust Inference in Communication Games with Partial Provability." *Journal of Economic Theory* 66 (2): 370–405.
Luban, David. 1983. "The Adversary System Excuse." In *The Good Lawyer: Lawyers' Roles and Lawyers' Ethics*, 83–122. Totowa, NJ: Rowman and Allanheld.
Macfarlane, Julie. 2013. *The National Self-Represented Litigants Project: Identifying and Meeting the Needs of Self-Represented Litigants, Final Report*. Treasurer's Advisory Group on Access to Justice. Toronto, May.
Mialon, Hugo M. 2005. "An Economic Theory of the Fifth Amendment." *Rand Journal of Economics* 36 (4): 833–48.
Milgrom, Paul R. 1981. "Good News and Bad News: Representation Theorems and Applications." *Bell Journal of Economics* 12 (2): 380–91.
Milgrom, Paul, and John Roberts. 1986. "Relying on the Information of Interested Parties." *Rand Journal of Economics* 17 (1): 18–32.
Posner, Richard A. 1999. "An Economic Approach to the Law of Evidence." *Stanford Law Review* 51: 1477–1546.
Sanchirico, Chris William. 1998. "The burden of proof in civil litigation: A simple model of mechanism design." *International Review of Law and Economics* 17 (3): 431–47.

Sanchirico, C. W. 2000. "Games, information, and evidence production: With application to English legal history." *American Law and Economics Review* 2 (2): 342–80.

Sanchirico, Chris William. 2001. "Relying on the Information of Interested—and Potentially Dishonest—Parties." *American Law and Economics Review* 3 (2): 320–57.

Sanchirico, Chris William. 2004. "Evidence Tampering." *Duke Law Journal* 53: 1215–1336.

Sanchirico, Chris William, and George Triantis. 2008. "Evidentiary Arbitrage: The Fabrication of Evidence and the Verifiability of Contract Performance." *Journal of Law, Economics, and Organization* 24 (1): 72–94.

Seidmann, Daniel J. 2005. "The Effects of a Right to Silence." *Review of Economic Studies* 72 (2): 593–614.

Seidmann, Daniel J., and Eyal Winter. 1997. "Strategic Information Transmission with Verifiable Messages." *Econometrica* 65 (1): 163–69.

Shchepetova, Anastasiya. 2013. "Inquisitorial vs. Adversarial system and the Right to be Silent." https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=IIOC2013&paper_id=602.

Shin, Hyun Song. 1994. "The Burden of Proof in a Game of Persuasion." *Journal of Economic Theory* 64 (1): 253–64.

Shin, Hyun Song. 1998. "Adversarial and Inquisitorial Procedures in Arbitration." *Rand Journal of Economics* 29 (2): 378–405.

Sobel, Joel. 1985. "Disclosure of evidence and resolution of disputes: Who should bear the burden of proof?" In *Game-theoretic models of bargaining*, edited by Alvin E. Roth, 341–62. Cambridge: Cambridge University Press.

Spurrier, R. L. 1980. *Inexpensive Justice: Self Representation in the Small Claims Court*. Port Washington, NY: Kennikat Press.

Thompson, William C. 1989. "Are Juries Competent to Evaluate Statistical Evidence?" *Law and Contemporary Problems* 52 (4): 9–41.

Tribe, Laurence H. 1971. "Trial by Mathematics: Precision and Ritual in the Legal Process." *Harvard Law Review* 84 (6): 1329–93.

Turkay, Evsen. 2011. "Evidence Disclosure and Severity of Punishments." https://mpra.ub.uni-muenchen.de/31821/.

Verrecchia, Robert E. 1983. "Discretionary disclosure." *Journal of Accounting and Economics* 5: 179–94.

Wells, Gary L. 1992. "Naked statistical evidence of liability: Is subjective probability enough?" *Journal of Personality and Social Psychology* 62 (5): 739–52.