



Columbia University

*Department of Economics
Discussion Paper Series*

On the Value of Legal Representation

*Yeon-Koo Che
Sergei Severinov*

Discussion Paper No.: 0506-06

*Department of Economics
Columbia University
New York, NY 10027*

October 2005

On the Value of Legal Representation*

Yeon-Koo Che[†] Sergei Severinov[‡]

October 5, 2005

ABSTRACT: In this paper we investigate 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 the effect. Despite lawyers' expertise in assessing the evidence, their advice is found to have no effect on adjudication in a broad set of circumstances, if legal advice is costless and the lawyers follow undominated strategies in disclosure. A lawyer's advice *can* influence the outcome to his client's favor, either if he can credibly advise his client to suppress some favorable evidence or if there is a cost associated with legal advice. The effect is socially undesirable in the former case, but it is desirable in the latter case although the benefit rests on its purely dissipative role as a "money burning" device rather than on his expertise.

KEYWORDS: Legal advice, disclosure of evidence, burden of proof allocation, regulating adjudicators' inferences.

*The authors thank Bob Hall, Ken Judd, Bart Lipman, Mitch Polinsky, and Kathy Spier, and the participants at the Decentralization Conference, seminar participants at the University of Arizona, Stanford Law and Economics Seminar, and Hoover Brown Bag Lunch, for helpful comments. This paper was completed while the second author was a National Fellow at the Hoover Institution, Stanford University, and he thanks the Hoover Institution for hospitality. The first author acknowledges financial support from National Science Fellowship (SES 0319061) and Shoemaker Fellowship.

[†]Department of Economics, Columbia University and University of Wisconsin-Madison, email: yc2271@columbia.edu

[‡]Fuqua School of Business, Duke University, email: sseverin@duke.edu

1 Introduction

Lawyers play a highly prominent and visible role in the modern day adjudication process. An important aspect of their role involves advising their clients over disclosing information to the court. It is well documented that litigating parties withhold from the court information which they perceive as unfavorable (See Rhode (1985)). Lawyers can advise their clients to withhold unfavorable information and disclose favorable one. Although lawyers often (particularly, in civil cases) have a disclosure duty before the tribunal, 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 Further, as Kaplow and Shavell (1989) noted, “the combination of carefully crafted responses, limited testimony and the adversary’s inability to conceive of (or expend the resources to ask) every possible question may well result in a significant gap between the information learned by the adversary’s lawyer and that possessed by the client.” To the extent that lawyers can affect the amount and nature of information reaching the court, they could affect the outcome of a trial. Our objective is to understand this effect.

To this end, we develop a model of dispute adjudication in which two disputing parties, say defendant and a plaintiff, are tried by an adjudicator, called “judge.” In the trial, each party presents his privately-held evidence, and the judge rules either to “convict” or “acquit” the defendant based on the disclosed evidence (information). The party’s evidence is “hard” in that he cannot manipulate it, but he can withhold it. Hence, the main decision facing each party is whether to disclose his evidence or not. This decision is not trivial, however, since the judge’s ruling depends not just on the evidence itself but also on the legal rules and standards that are applied to interpret that evidence and the other public evidence surrounding that case. Lacking the knowledge of the latter, the party cannot assess how favorable his evidence will be if he presents it to the court. By contrast, lawyers’ special expertise lies in their knowledge of the law. They can assess the clients’ evidence in light of the law and legal standards and determine whether this evidence is favorable or unfavorable for that client. They also have better understanding of the strength of the case without disclosure. A lawyer-represented party can thus learn how favorable his evidence will be and can make a more informed decision about disclosure. We study the effect of lawyer representation by comparing the adjudication process with and without lawyer representation.

At first glance, lawyers’ expertise appears to leave no doubt about the value of their advice, at least to their clients. Armed with the knowledge of the law, a lawyer should be able to improve

¹The 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.

her client’s disclosure decision. Surprisingly, however, this benefit does not always materialize. We find legal advice to be irrelevant — both privately and socially — when it can be obtained for free and the parties employ undominated strategies in disclosure (i.e., of disclosing evidence if and only if it is favorable). In that case, legal advice indeed affects one’s disclosure behavior, but it does *not* affect the outcome of adjudication. This irrelevance is surprisingly robust, as it holds regardless of whether one or both parties obtain legal advice and whether the adjudicator makes a Bayesian inference based on the parties’ disclosure strategies or she simply follows an ad-hoc rule satisfying some reasonable properties.

Lawyer representation can make a difference when the two conditions are relaxed. First, a lawyer representation matters if a lawyer can credibly follow a (weakly dominated) strategy of suppressing some favorable evidence. Such a strategy can skew the inference by the court, and thus the adjudication outcome, in favor of his client. Such a role of lawyers generates a private incentive for hiring lawyers, but the aggregate welfare of the parties falls if both parties hire lawyers, which suggests a “prisoners’ dilemma” type explanation for the prevalence of lawyer representation. Moreover, this role of lawyers distorts parties’ disclosure in a socially undesirable way. We show that this harm can be remedied if the adjudicator commits *ex ante* to the way in which she assigns the burden of the proof and thus the way in which she draws an inference about the guilt. This last result provides some rationale for restrictions the law places on how judges and juries should interpret evidence or lack thereof.

Legal representation can also affect the adjudication outcome when it comes at a cost. The cost of representation provides a means with which parties without additional information — either incriminating or exculpatory — to credibly signal the lack of evidence and avoid a prejudicial inference by the adjudicator. Although a party with extremely unfavorable evidence mimics the same behavior (i.e., hires a lawyer and withholds evidence), other parties who possess less unfavorable evidence — but still unfavorable enough that they would have withheld it had legal advice been free — choose not to hire a lawyer and are compelled to disclose their evidence. In essence, hiring a lawyer buys one “the right to be silent without prejudice.” 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. Although legal advice has both private and social value in this case, this value does not arise from the lawyer’s knowledge of the law but rather from its role as a “money burning” device.

Our paper also yields useful insights regarding various the restrictions on inferences that they can draw from nondisclosure of evidence. To our knowledge, the current paper is the first to offer a normative analysis of this issue. First, we show that no such restrictions are warranted with or without legal representation, when the adjudicator is Bayesian and the lawyers disclose

all favorable evidence. In this case, the equilibrium outcome is socially optimal. However, this conclusion no longer holds when the lawyers can credibly use strategies of withholding some favorable evidence. In this case, the extent of potential harm to social welfare can be reduced by committing the adjudicator to a rule which allocates all burden of proof to one party.

The issue of legal advice has received very little formal treatment in the literature. Legal scholars have recognized the factors favoring and disfavoring lawyer-aided adversarial system but disagree on the relative importance of those factors. Proponents argue that a vigorous adversarial competition among lawyers leads the court to focus on relevant evidence, thus making judicial fact-finding efficient (Luban, 1983; Bundy and Elhauge, 1992 and 1993). Critics point out that lawyers can mislead as much as inform the court (Frank, 1973). In particular, Kaplow and Shavell (1989) point out, via illustrative examples, that while the lawyers' ability to suppress evidence based on legal expertise undoubtedly benefits their clients, its social implications are ambiguous, thus casting doubt on its social benefit. Although the current paper is similar in spirit to the last study, there are several important distinctions. First, these studies do not perform a full-fledged equilibrium analysis of the disclosure game, focusing rather on the effect of legal advice when possible outcomes are exogenously fixed. Further, the issue of adjudicator's inferences is not treated in the literature. In contrast, we explicitly analyze the interplay of incentives of different actors in our model. Many interesting issues of policy relevance can be addressed only with an approach such as ours that treats the adjudicator's inference as a policy choice variable. Inasmuch as we show the social optimality of following the Bayesian approach in adjudication, our paper contributes to the active debate on the use of Bayesian methods in law. Finally, the previous studies restrict their attention to the social value of legal advice, largely taking private incentives for granted. However, as we demonstrate, the issue of private value of legal advice is not trivial and warrants a careful examination.

This paper is also closely related to the economics literature on disclosure of non-falsifiable information. Grossman (1981), Milgrom (1981), Milgrom and Roberts (1986) and Lipman and Seppi (1995) find that conflicting interests can lead to full revelation of commonly shared information by the agents. Shin (1994, 1998) studies information revelation when the possession of information is not common knowledge, which makes the inference (or "burden of proof") non-trivial. In particular, our basic model is related to Shin (1998), who compares adversarial and inquisitorial litigation systems. Froeb and Kobayashi (1996, 2001) and Daughety and Reinganum (2000b) assess the implications of endogenous evidence production arising from the adversarial system. Sobel (1985), Hay and Spier (1997) and Sanchirico (1997, 1998, 2001) explore allocation of the burden of proof and evidence production from the standpoint of litigation costs and/or

deterrence.² None of these papers deal with the role of lawyers in disclosure — the focus of the current paper.

2 Model

Two parties, 1 and 2, are in dispute regarding some matter or event, which is adjudicated by a judge/jury in a tribunal. It is convenient to interpret parties 1 and 2 as a defendant and a plaintiff in a litigation, so we will invoke this interpretation throughout the paper. However, our model is fairly general and can apply 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. 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 a piece evidence $s \in [0, 1] =: S$ which may only be observed by the parties to the dispute. The evidence is observed with probability p_{00} by neither party, with probability p_{11} by both parties; and with probability p_{10} [resp. p_{01}] by party 1 only [resp. party 2 only].³ Obviously, $p_{ij} \geq 0$ for all $i, j = 0, 1$ and $\sum_{i,j=0,1} p_{ij} = 1$. Note that we allow for possible correlation in the parties’ abilities to observe the evidence. No others observe this evidence directly. 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.

The other piece of judgment-relevant information, $\theta \in [0, 1] =: \Theta$ can only be observed by the lawyers and the judge. The variable θ represents the existing laws and legal standards, as well as how the judge applies them and how she interprets the external circumstances and all commonly known evidence surrounding the case, such as basic uncontested facts, police reports, the testimony by neighbors, etc. When s is not disclosed, the adjudicator’s ruling depends only on θ .⁴ We assume that lawyers’ expertise is needed to assess θ , which reflects the special knowledge that the officers of the law (the lawyers and the judge) possess, and the notion that the parties to the dispute typically have limited knowledge of the law and incomplete understanding of the

²Dewatripont and Tirole (1999) study the desirability of adversarial system in a broad organization design context. Levy (2005) studies the effect of career concerns on judges’ decision making. Also related, albeit with less relevance in the litigation setting, is the literature on cheap talk (or communication of falsifiable information), which includes Crawford and Sobel (1982) and Krishna and Morgan (2001).

³“Observing” s means either possessing that evidence or having proof of its existence (perhaps in the opponent’s possession).

⁴Posner (1999) describes 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.

legal process.

Thus, potential value of the lawyers derives from their expertise and abilities to understand the body of the law in the jurisdiction where they practice, as well as the judge’s interpretation of the law and possible biases. For instance, the lawyer and the judge may be able to assess more accurately how strong or weak the mitigating circumstances are for a litigant. The lawyers’ ability — and the litigants’ inability — to observe θ creates *potential* scope for legal advice to be valuable.

We assume that (s, θ) is drawn from $S \times \Theta$ according to an absolutely continuous cdf, $F(s, \theta)$ which has a positive density $f(s, \theta)$ in the interior of $S \times \Theta$. From the ex-ante perspective, θ is random because it describes the realized state of the law and legal standards -one out of many possibilities, as well as a particular realization of commonly known evidence. Since s and θ reflect the nature of underlying case, they may be correlated. We assume that s and θ satisfy the (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 U.S., many litigants do so in municipal courts and administrative trial procedures. Also, in small claims courts — which comprise a significant share of trials in the U.S. — legal representation is expressly forbidden in most states (California, New York, Arizona, and others).⁵ Further, our comparison should not be narrowly interpreted as pertaining only to the two regimes; but it rather applies to any increase in the quality of lawyer representation. For instance, one could view the two regimes as involving lawyer representation but differing only in the quality of representation.

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 (trial), party 1 and party 2 simultaneously and independently decide whether to disclose the evidence s to a judge, provided that the respective party has observed it. In the representation regime, this decision is taken with the help of a lawyer providing legal advice. At date 3, the judge rules either for party 1 or for party 2. The decision whether to hire a lawyer will at first be left outside the model. Later, we introduce the

⁵See Spurrier (1980) for detail. The problem of withholding evidence is particularly relevant in this case, since the discovery process is very limited and the trials focus on a few key elements of evidence.

possibility of the parties deciding on the basis of their private evidence whether to obtain legal representation at a cost.

To complete the description of the model, we need to consider the evidence disclosure behavior by parties 1 and 2, as well as the judge's adjudication behavior.

• **Evidence disclosure behavior:**

If a party is not represented by a lawyer, then his decision whether to disclose s is based solely on s . In contrast, if a party is represented by a lawyer, the lawyer can provide advice based on his knowledge of θ . In particular, the lawyer can advise the client whether disclosing s is beneficial to him for given θ .

We assume that a lawyer would prefer his client to prevail in court, and there are no agency issues in the attorney-client relationship. Given the congruence of interests between the lawyer and the client, the client will have full incentives to communicate s to his lawyer, and likewise the lawyers will have the incentives to explain the legal issues, i.e. to communicate θ to the client truthfully. Therefore, the represented party can simply be viewed as informed of both s (if he observes s) and θ .

Thus, the difference between representation and no representation in our model boils down to the information available to the party when he makes the disclosure decision. In the regime of no representation, the disclosure decision is made solely on the basis of s itself, while in the regime of representation it is based on both s and θ . Formally, party i 's disclosure strategy is a function ρ_i that maps $S \times \Theta$ to $[0, 1]$, with $\rho_i(s, \theta)$ representing the probability that party $i = 1, 2$'s discloses s , in state θ . If the party is not represented, he does not observe θ , so $\rho_i(\cdot, \theta)$ is constant in 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.⁶

The judge's ruling depends on (s, θ) if s has been disclosed, and on θ only if s has not been disclosed. If θ is disclosed, the judge's decision depends on her criterion of party 1 (defendant)'s culpability, $g(s, \theta)$. Specifically, if $g(s, \theta) > 0$, then the judge finds party 1 culpable and thus rules for party 2, and 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 in this case has no real consequence.

⁶Our binary model is more general than appears at first glance. For instance, there may be no ambiguity about the size of damages in case the plaintiff prevails, leaving the liability as the only source of contention. The binary feature can also be justified in an idealistic Beckerian world in which any defendant found liable is subject to a sanction equaling his maximum wealth limit.

We assume that the function $g(s, \theta)$ is increasing and continuous in both arguments. Thus, lower s and θ [resp. higher s and θ] are more beneficial for party 1 [resp. party 2]. In the litigation example, a higher value of s means that the defendant (party 1) is more likely to have caused a certain event, while a higher value of θ indicates that the law and legal standards are more unfavorable toward the defendant. To make the judge's decision problem nontrivial, we assume that $\int g(s, 1)f(s|1)ds > 0$ and $\int g(s, 0)f(s|0)ds < 0$, which means that publicly available information and legal standards 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 to party 1 (i.e., $\theta = 0$) to the most unfavorable (i.e., $\theta = 1$).⁷ 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 $\theta \in [\underline{\theta}, \bar{\theta}]$, where $\bar{\theta} := \max\{\theta | \exists s' \in S \text{ s.t. } g(s', \theta) = 0\}$ and $\underline{\theta} := \min\{\theta | \exists s'' \in S \text{ s.t. } g(s'', \theta) = 0\}$. This function partitions the evidence/legal environment 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 1.⁸

The adjudication criterion $g(s, \theta)$ can be derived from an optimization problem. For example, the judge may wish to minimize the cost of making wrong decisions, i.e. “convicting the innocent and exonerating the guilty,” with appropriate costs assigned to each type of mistake. Let c_1 (c_2) be the judge's personal or perceived societal cost of ruling mistakenly for party 1 (2), i.e. “exonerating the guilty” (“convicting the innocent”), and let $\pi(s, \theta)$ be the probability that for given (s, θ) the party 1, the defendant, is guilty. Then, if the judge convicts party 1 (the defendant) with probability z , the expected cost of a mistake is

$$(1 - \pi(s, \theta))c_2z + \pi(s, \theta)c_1(1 - z).$$

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 $g(s, \theta) := \pi(s, \theta) - \frac{c_2}{c_1+c_2}$.⁹

We assume that the judge's adjudication criterion $g(\cdot, \cdot)$ is common knowledge to all players, including the lawyers and parties 1 and 2. This assumption is not very demanding with respect to the unrepresented parties' rationality because, importantly, the unrepresented parties are uncertain about the actual adjudicator's criterion $g(\cdot, \theta)$. It also does not rule out the possibility that the judge may be biased, as long as the bias is common knowledge. For example, the judge may

⁷This assumption is mainly to simplify exposition. Its only use is to allow for nontrivial analysis in Section 5.

⁸The two regions have nonempty interiors given the above assumption.

⁹Different 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. The interval of (0.6, 0.7) corresponds to the standard of “clear and convincing evidence.” According to Posner (1999) judges associate probability level between 0.75 and 0.9 with the standard of “proof beyond a reasonable doubt.”

adopt a standard of proof, as represented by $\frac{c_2}{c_1+c_2}$, which may differ from the one she is supposed to apply.¹⁰

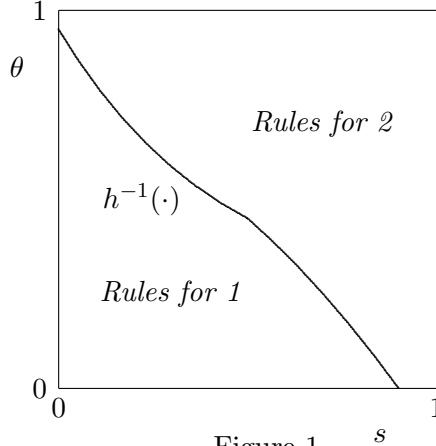


Figure 1

If no party discloses s , then the adjudicator rules on the basis of θ and, possibly, her inference about the parties' disclosure decisions. The adjudicator's decision rule in this case, 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 (s)he rules for party 2 if she observes signal θ and no evidence is disclosed. The judge's strategy depends on her *posterior assessment of party 1's culpability*, or simply her *posterior*, given by

$$\mathbb{E}[g|\rho_1(\cdot), \rho_2(\cdot), \theta; a, b_1, b_2, c] := a\mathbb{E}_0[g|\theta] + b_1\mathbb{E}_1[g|\theta] + b_2\mathbb{E}_2[g|\theta] + c\mathbb{E}_{12}[g|\theta], \quad (1)$$

which is a weighted average of expected culpability criterion based on alternative evidence scenarios, with nonnegative constants a , b_1 , b_2 and c used as weights. The first term, $\mathbb{E}_0[g|\theta] := \int_0^1 g(s, \theta) f(s|\theta) ds$, is party 1's expected culpability given the presumption that no party has observed the evidence s ; $\mathbb{E}_i[g|\theta] := \int_0^1 g(s, \theta) (1 - \rho_i(s, \theta)) f(s|\theta) ds$ is the (normalized) expectation of g given the presumption that only party $i = 1, 2$ has observed s but has not disclosed is consistently with his disclosure strategy ρ_i ;¹¹ and the last expectation term, $\mathbb{E}_{12}[g|\theta] := \int_0^1 g(s, \theta) (1 - \rho_1(s, \theta)) (1 - \rho_2(s, \theta)) f(s|\theta) ds$, is based on the presumption that both have observed s but neither has disclosed it. The judge's default ruling is based on (1) so that she rules for party

¹⁰The bias can be of a more general form. Letting $\Delta(s, \theta)$ denote the degree of bias in favor of party 2, the judge may follow the criterion $g(s, \theta) + \Delta(s, \theta)$ where $g(s, \theta)$ denotes an unbiased criterion. The subsequent analysis goes through as long as the new criterion $g(s, \theta) + \Delta(s, \theta)$ preserves the monotonicity property.

¹¹This expectation as well as subsequent ones are normalized in the sense that the expression is not divided by the probability of nondisclosure, $\int_0^1 (1 - \rho_i(s, \theta)) f(s|\theta) ds$. Normalization has no consequence since its effect is captured by the coefficients a, b_1, b_2, c .

2 (i.e., $\delta(\theta) = 1$) if $\mathbb{E}[g|\rho_1(\cdot), \rho_2(\cdot), \theta] > 0$ and for party 1 if $\mathbb{E}[g|\rho_1(\cdot), \rho_2(\cdot), \theta] < 0$. The dependence of the posterior on (a, b_1, b_2, c) will be suppressed when it does not generate any ambiguity.

The coefficients (a, b_1, b_2, c) reflect how the adjudicator weighs alternative evidence scenarios in her inference formation and in her burden of proof assignment. Since only the sign of the posterior matters for the judge's decision, we normalize by setting $a = 1$, and focus on the values of (b_1, b_2, c) .¹²

Depending on the values of these variables, our adjudication criterion in (1) accommodates a variety of different decision procedures and burden-of-proof allocation rules. For example, if $b_1 = b_2 = c = 0$, then the judge bases her decision only on the prior expectation of g . In this case, the judge is *completely non-Bayesian* in the sense that she does not account for the possibility that one of the parties may be withholding evidence. Next, if $b_1 > 0$ and $b_2 = c = 0$, then the judge believes that party 1 alone, if any, could have the evidence, so she never attributes nondisclosure of evidence to party 2's withholding it. That is, the burden of proof is put on party 1, and absence of evidence generally has negative implications for this party. Likewise, if $b_2 > 0$ and $b_1 = c = 0$, then the burden of proof is assigned to party 2. If both b_1 and b_2 are strictly positive, then the judge assigns some weight to either party withholding the evidence, so the burden of proof is split between the two parties. One important case arises when the judge is *fully Bayesian*, i.e. $(b_1, b_2, c) = (\frac{p_{10}}{p_{00}}, \frac{p_{01}}{p_{00}}, \frac{p_{11}}{p_{00}})$. In this case, the judge's posterior assessment assigns accurate probability weights to alternative scenarios of evidence withholding.

The coefficient b_i can be interpreted as a measure of the extent to which the adjudicator is predisposed against party i in interpreting nondisclosure. In particular, we will say that an inference rule (b'_1, b'_2) is *less prejudicial than* an inference rule (b_1, b_2) if $(b'_1, b'_2) < (b_1, b_2)$.

There is an active debate in the legal literature regarding the appropriate allocation of the burden of proof, as well as the applicability of Bayesian approach and the use of statistical evidence and arguments in judicial decision-making. It is widely acknowledged that adjudicators are prone to biases and errors in computing the true statistical odds of events (see Tribe (1971)) and are often reluctant to convict on the basis of simple statistical likelihood.¹³ Therefore, it is important

¹²The only loss is when $a = 0$. This case is arbitrarily closely approximated by $a \approx 0, a > 0$. Further, c will be seen to play a limited role.

¹³One of the most well-known examples extensively discussed in the literature to highlight the problems with this kind of cases and the use of statistical evidence is the so-called Blue Bus/Grey Bus case. In this case, a plaintiff has been negligently hit by a bus in the location where Blue Bus Company operates a greater number of buses than Grey Bus Company. A direct application of 'more likely than not' criterion should lead the court to convict Blue Bus company on the basis of the 'bare bones' statistical evidence that blue buses are more numerous and, therefore, are more likely to have hit the plaintiff. Yet, experimental results (see Wells (1992)) show that judges and members of the jury are very unlikely to make such conviction when only such evidence is presented. Several legal scholars (e.g. Posner (1999) and Thompson (1989)) explain the reluctance to convict by the fact that it is

to allow for non-Bayesian — as well as Bayesian — burden-of-proof allocations. Furthermore, legal rules and procedures often restrict the admissibility of certain types of evidence and limit the inferences which a judge or a jury are allowed to make from certain evidence or lack thereof, because of the concerns about their prejudicial effect. Our model allows for any biases and errors in the allocation of the burden of proof and in the application of Bayesian approach, as well as for any restrictions that may be imposed by the law of evidence on the adjudicators' inference formation. We only insist that the adjudicator apply the same criterion regardless of whether a party is represented or not and that she correctly foresees the strategy of party i if that party bears some burden of proof (i.e. if $b_i > 0$).

For a later purpose, it is useful to consider a *posterior assessment* arising when the parties follow *cutoff* strategies. Suppose party 1 employs a strategy of disclosing her/his evidence if and only if $s < \hat{s}_1$ and party 2 discloses if and only if $s > \hat{s}_2$. We denote the judge's posterior under such cutoff strategies (with slight abuse of notation) by

$$\begin{aligned} \mathbb{E}[g|\hat{s}_1, \hat{s}_2, \theta; b_1, b_2, c] := & \int_0^1 g(s, \theta) f(s|\theta) ds + b_1 \int_{\hat{s}_1}^1 g(s, \theta) f(s|\theta) ds + b_2 \int_0^{\hat{s}_2} g(s, \theta) f(s|\theta) ds \\ & + c \int_{\{\hat{s}_1 \leq s \leq \hat{s}_2\}} g(s, \theta) f(s|\theta) ds. \end{aligned} \quad (2)$$

• **Equilibrium concept and outcome:**

In each regime, we focus on Bayesian Nash equilibria in the parties' disclosure strategies and the judge's ruling strategy. Since the adjudication decision is trivial in the case of disclosure, the strategy profile is summarized by a triple, (ρ_1, ρ_2, δ) . The object of ultimate interest though is the equilibrium outcome of adjudication, i.e., the decision made as a function of 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 not. We are interested in comparing the adjudication outcomes induced by equilibria under different legal regimes distinguished by the availability of legal advice.¹⁴

quite implausible that the statistic is the only evidence available to the plaintiff. That is, absence of other evidence should lead the adjudicator to infer that the plaintiff is concealing some evidence indicating that the bus actually belonged to the other bus company. The latter point of view is consistent with the judges and juries following an adjudication criterion such as (2) with $b_i > 0$.

¹⁴In particular, an equilibrium (ρ_1, ρ_2, δ) induces an outcome function via

$$\phi(x_1, x_2, s, \theta) = \delta(\theta)(1 - x_1\rho_1(s, \theta))(1 - x_2\rho_2(s, \theta)) + 1_{\{g(s, \theta) \geq 0\}} [1 - (1 - x_1\rho_1(s, \theta))(1 - x_2\rho_2(s, \theta))].$$

3 Irrelevance of Legal Advice

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

3.1 No Representation

In this regime, neither party 1 nor party 2 has a lawyer. Thus, they must decide whether to disclose the evidence s without being certain about the value of 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, the parties adopt cutoff strategies with a common threshold \hat{s} : Party 1 discloses s if and only if $s < \hat{s}$, and party 2 discloses s if and only if $s > \hat{s}$. This is intuitive since a lower [resp. higher] s is more likely to lead to a favorable ruling for party 1 [resp. party 2] when disclosed, given monotonicity of $g(\cdot, \cdot)$ and MLRP between (s, θ) .

The judge's equilibrium default ruling strategy is also characterized by a threshold, $\hat{\theta}$. That is, when s is not disclosed, the judge rules for party 1 if $\theta < \hat{\theta}$ and for party 2 if $\theta > \hat{\theta}$. This is also intuitive since $g(s, \theta)$ is more likely to be negative [resp. positive] if θ is low [resp. high].

To see this more clearly, consider the judge's posterior belief in case of nondisclosure. If the parties follow the cutoff strategies with a common threshold, say $\hat{s} \in [0, 1]$, then this posterior is given by $\mathbb{E}[g|\hat{s}, \hat{s}, \theta, b_1, b_2, c]$. The monotonicity of $\mathbb{E}[g|\hat{s}, \hat{s}, \theta, b_1, b_2, c]$ in θ is explained by two factors. First, a higher θ is by itself a stronger evidence of 1's culpability holding s fixed, since $g(s, \theta)$ is increasing in θ . This effect is reinforced by the parties' disclosure strategies and the MLRP between θ and s . That is, the higher θ is the higher s is likely to be, which further increases the likelihood that party 1 is culpable. This monotonicity of posterior with respect to θ explains the judge's cutoff strategy.

Given the judge's cutoff strategy with threshold $\hat{\theta}$, we can pin down each party's cutoff threshold to be $\hat{s} = h(\hat{\theta})$. To see this, suppose party 1 has observed $s < h(\hat{\theta})$. In this case, it is optimal for party 1 to disclose s , regardless of party 2's strategy. Since party 1's disclosure decision matters only if party 2 never discloses s , assume that party 2 does not disclose. Then there are two possibilities. Suppose first the judge observes $\theta < \hat{\theta}$. Then, withholding $s < h(\hat{\theta}) < h(\theta)$ will result in a favorable ruling for party 1. But disclosing the evidence will also secure favorable ruling, since $g(s, \theta) < g(h(\theta), \theta) = 0$, so party 1 will do no worse by disclosing. Suppose next the judge observes $\theta > \hat{\theta}$. In this case, withholding s will surely result in an unfavorable ruling, whereas disclosure can reverse that ruling with positive probability (i.e., whenever $s < h(\theta)$). Hence, disclosing evidence is optimal if $s < h(\hat{\theta})$. A similar argument shows that it is strictly

optimal for party 1 to withhold any $s > h(\hat{\theta})$.

This argument proves that party 1, and by symmetry party 2, will adopt a cutoff strategy with threshold $\hat{s} = h(\hat{\theta})$. Substituting this into (2), the judge's equilibrium posterior becomes $\mathbb{E}[g|h(\hat{\theta}), h(\hat{\theta}), \hat{\theta}, b_1, b_2, c]$. Hence, her cutoff threshold is given by:

$$\hat{\theta}^*(b_1, b_2) := \inf\{\hat{\theta} \in \Theta \mid \mathbb{E}[g|h(\hat{\theta}), h(\hat{\theta}), \hat{\theta}; b_1, b_2, c] > 0\}, \quad (3)$$

where $\hat{\theta}^*(b_1, b_2) := 1$ if the set in the RHS is empty. Note that the threshold does not depend on c since the last expectation term, $\mathbb{E}_{12}[\cdot]$, in (1) vanishes. More importantly, the threshold $\hat{\theta}^*(b_1, b_2)$ increases in b_2 and decreases in b_1 . Intuitively, the judge's default ruling becomes more favorable towards party 2 as the burden of proof shifts from party 2 to party 1. We can now state the following result.

PROPOSITION 1 *If no party is represented by a lawyer, there exists a unique Bayesian Nash equilibrium in which party 1 discloses s if and only if $s < h(\hat{\theta}^*(b_1, b_2))$, party 2 discloses s if and only if $s > h(\hat{\theta}^*(b_1, b_2))$, and the judge rules for party 1 if $\theta < \hat{\theta}^*(b_1, b_2)$ and for party 2 if $\theta > \hat{\theta}^*(b_1, b_2)$ following nondisclosure. As the burden of proof shifts from party 2 to party 1, the latter discloses more and the former discloses less, i.e. $\hat{\theta}^*(b_1, b_2)$ decreases, and hence $h(\hat{\theta}^*(b_1, b_2))$ increases, in $(b_1, -b_2)$.*

The equilibrium is characterized by two thresholds, $\hat{\theta}^*(b_1, b_2)$ and $\hat{s} = h(\hat{\theta}^*(b_1, b_2))$, as depicted in Figure 2. The fact that these thresholds satisfy $g(\hat{s}, \hat{\theta}^*(b_1, b_2)) = 0$ will be important for our results. It is also instructive to see how the equilibrium changes with the shift in the burden of proof from one party to the other. As the burden of proof shifts from party 2 to party 1, the judge's threshold decreases. This, in turn, causes an increase in the parties' common disclosure threshold, $h(\hat{\theta}^*(b_1, b_2))$. Hence, the party with an increased burden of proof is caused to disclose a wider range of evidence and the party with a decreased burden discloses less evidence.

3.2 Full Representation

In this regime, both parties are represented by lawyers and learn θ through legal advice. Hence, unlike the no representation case, the parties make their disclosure decisions based on both s and θ . In this case, party 1 has a unique weakly dominant strategy of disclosing s if and only if $s < h(\theta)$ or, equivalently, $g(s, \theta) < 0$. Similarly, party 2's unique weakly dominant strategy is to disclose s if and only if $s > h(\theta)$ or, equivalently, $g(s, \theta) > 0$. This is so because the adjudicator's ruling in case of full disclosure is completely determined by the value of her criterion $g(s, \theta)$. Hence the default ruling is the only real adjudicator's decision. Then, disclosing any $s < h(\theta)$ is a dominant strategy for party 1 because withholding may entail an unfavorable default ruling by the judge.

Likewise, withholding $s > h(\theta)$ is a dominant strategy for party 1 because the adjudicator may rule for party 1 without disclosure, but will rule against him for sure if s is disclosed.

Dominant strategies have intuitive appeal in our model, particularly with representation. Since a guilty verdict could be very damaging for a client, it would be hard for a lawyer to withhold favorable evidence at the trial, and it is even harder to imagine a client accepting such a legal strategy. In fact, in the U.S. a lawyer is legally obliged to explore all avenues of defense and present all exculpatory evidence. Furthermore, the judge could simply refuse to believe that a lawyer may be following a different strategy. Also, if there is even a smallest 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. Any of these three possibilities rule out any other equilibria under representation. For this reason, we focus on the dominant disclosure strategies here. Later, we shall consider what happens when these arguments do not apply and examine equilibria supported by "dominated" disclosure strategies.

Note that the dominant disclosure strategies do not depend on the judge's inference rule, particularly on how she allocates the burden of proof between the two parties. In this sense, the presence of attorneys decouples the parties' disclosure from the judge's inference rule. In particular, it causes their disclosure behavior to change compared to the case of no representation. Remarkably, this change in disclosure behavior does *not* affect the judge's default ruling. Indeed, given the dominant disclosure strategies by the parties, the judge's posterior is given by $\mathbb{E}[g|h(\theta), h(\theta), \theta]$. This posterior is different in magnitude from that held by the judge in the case of no representation, but the two posteriors always have the same sign, since

$$\mathbb{E}[g|h(\theta), h(\theta), \theta] \gtrless 0 \text{ if } \theta \gtrless \hat{\theta}^*(b_1, b_2).$$

Hence, we have:

PROPOSITION 2 *If both parties are represented, there exists a unique dominant strategy equilibrium in which the judge uses the cutoff strategy with threshold $\hat{\theta}^*(b_1, b_2)$ (described in (3)) in her default ruling: in the absence of disclosure she rules for party 1 (resp. party 2) if $\theta < \hat{\theta}^*(b_1, b_2)$ (resp. $\theta > \hat{\theta}^*(b_1, b_2)$). Party 1 reveals (resp. withholds) s with probability 1 if $g(s, \theta) < 0$ ($g(s, \theta) > 0$). Party 2 reveals (resp. withholds) s with probability 1 if $g(s, \theta) > 0$ (resp. $g(s, \theta) < 0$).*

The fact that the threshold $\hat{\theta}^*(b_1, b_2)$ used by the judge in her default ruling is the same in the no representation and full representation cases is surprising, and it is worth exploring the logic behind it. In the no representation case, the parties employ a (common) threshold $\hat{s} = h(\hat{\theta}^*(b_1, b_2))$ that does not vary with θ , whereas in the full representation case, their threshold $h(\theta)$ varies with θ and thus in general differs from $h(\hat{\theta}^*(b_1, b_2))$. Nonetheless, these two thresholds coincide with

each other when $\theta = \hat{\theta}^*(b_1, b_2)$. Since the sign of the judge's posterior changes from negative to positive as θ rises past $\hat{\theta}^*(b_1, b_2)$, the posterior has the same sign under both regimes, even though their exact magnitudes will typically be different. Hence, the judge's default ruling is the same under both regimes.¹⁵

3.3 Partial Representation

The results of the previous subsections generalize to the regime in which only one side hires a lawyer. Suppose without loss of generality that party 1 hires a lawyer and party 2 does not. Let $\hat{\theta}$ denote the threshold which the judge uses in her default ruling strategy when s is not disclosed. Focusing as before on undominated strategies, party 1 will disclose s if and only if $s < h(\theta)$, just as in Subsection 3.2. As established in Subsection 3.1, party 2's unique optimal strategy is to disclose s if and only if $s > h(\hat{\theta})$. So, when the judge observes θ but not s , her posterior becomes

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

Since this posterior is monotonic in θ and changes the sign from negative to positive at $\hat{\theta}^*(b_1, b_2)$, the following result is immediate.¹⁶

PROPOSITION 3 *If only one party hires a lawyer, there exists a unique equilibrium in undominated strategies. In this equilibrium the judge uses a cutoff strategy with threshold $\hat{\theta}^*(b_1, b_2)$ (defined in (3)) in her default ruling. If party 1 [resp. party 2] obtains legal advice, he discloses s iff $g(s, \theta) < 0$ [resp. $g(s, \theta) > 0$]. If party 1 [resp. 2] does not obtain legal advice, he uses cutoff strategy with threshold $h(\hat{\theta}^*(b_1, b_2))$: party 1 [resp. party 2] reveals s if $s < h(\hat{\theta}^*(b_1, b_2))$ [resp. $s > h(\hat{\theta}^*(b_1, b_2))$]. A shift in the burden of proof affects only the party without a lawyer, in a way described in Proposition 1.*

3.4 Irrelevance of Representation

A striking feature of all three regimes is that the judge's equilibrium default ruling strategy is the same across all three regimes. The judge adopts a cutoff strategy with the same threshold $\hat{\theta}^*(b_1, b_2)$, regardless of whether the parties obtain legal advice. This does not mean that the parties disclose the same evidence. Propositions 1, 2 and 3 clearly show that the set of s 's

¹⁵More precisely, if the judge's equilibrium threshold is interior, her cutoff threshold must satisfy the same condition $\mathbb{E}[g|h(\hat{\theta}^*(b_1, b_2)), h(\hat{\theta}^*(b_1, b_2)), \hat{\theta}^*(b_1, b_2)] = 0$ in both cases. The condition for the corner solution — i.e. judge's equilibrium threshold being either 0 or 1 — is also the same in the two regimes.

¹⁶The result holds more generally in the following sense. Suppose each party randomizes on hiring a lawyer, and the judge has some arbitrary beliefs about the parties' decisions to hire lawyers. The behavior described in Proposition 3 continues to be an equilibrium in this environment.

revealed to the judge differs across the regimes. Nevertheless, we will show that the difference in the parties' disclosure strategies does not amount to any real difference in the outcome of the trial.

This irrelevance is, in fact, a result of a more general property of equilibrium behavior in our disclosure/adjudication game. This property is described in the following lemma.

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. of whether either party obtains legal advice or not, any combination of the best response disclosure strategies by the two parties lead to the same outcome characterized by the following outcome function:*

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

where $x_i = 1$ ($x_i = 0$) if party i observes (does not observe) the evidence and $\phi^{\hat{\theta}}(x_1, x_2, s, \theta)$ is the probability of ruling for party 2 in the state of the world (x_1, x_2, s, θ) given the adjudicator's default ruling strategy.

The Decision Equivalence Lemma shows that the judge's cutoff strategy uniquely determines the equilibrium adjudication outcome, regardless of the parties' use of legal advice.¹⁷ Some insight into this result can be gained from comparing the cases of no representation and full representation. Consider Figure 2 and suppose that the judge's threshold is given by $\hat{\theta}$.

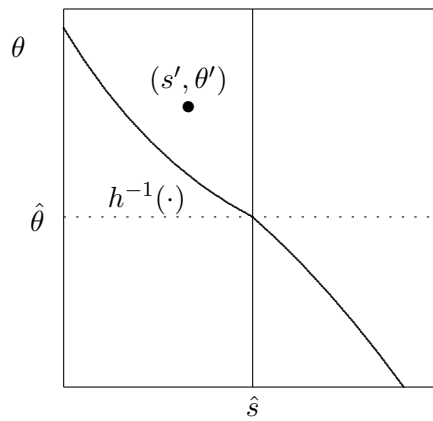


Figure 2

¹⁷This Lemma plays a role akin to that played by the Revenue Equivalence theorem in the auction literature (see Riley and Samuelson (1980) and Myerson (1981)).

Then, by Propositions 1 and 2, the common threshold adopted by the parties under self-representation is given by $\hat{s} = h(\hat{\theta})$, while in the dominant strategy equilibrium with representation party 1 (party 2) discloses when $s < h(\theta)$ ($s > h(\theta)$). Suppose that (s', θ') occurs and party 1 observes s' . With legal advice, party 1 would not disclose s' , but the judge will still rule for party 2. Under no representation, party 1 will disclose s' being unaware of θ' , and the judge will rule for party 2. Hence, despite different disclosure behavior, there is no difference in the adjudication outcome: the judge's ruling is unfavorable to party 1 in either case. Obviously, this outcome does not depend on whether party 2 makes disclosure or not. Combining Propositions 1-3 with Lemma 1, we obtain our key result:

COROLLARY 1 (Irrelevance of legal advice) *Suppose that the judge applies the same inference rule, (b_1, b_2, c) , regardless of the representation regime (i.e. of whether both or any party are represented or not), and the parties employ undominated strategies in disclosure. Then there is a unique equilibrium outcome which does not depend on the representation regime and is characterized by the outcome function $\phi^{\hat{\theta}^*(b_1, b_2)}(\cdot)$.*

Remarkably, the irrelevance result does not depend on the judge's inference rule: it holds if the judge is fully Bayesian (i.e., $(b_1, b_2) = (\frac{p_{10}}{p_{00}}, \frac{p_{01}}{p_{00}})$), or if she follows any other inference rule (b_1, b_2, c) , as long as the same inference rule is applied regardless of the parties' use of legal advice. A change in (b_1, b_2) would typically cause the threshold $\hat{\theta}^*(b_1, b_2)$ to shift, but would not alter the fact that the representation regime has no effect on the adjudication outcome.

The robustness of the irrelevance result is surprising and may appear contradictory to the lawyers' prominent roles in high-profile trials. One should not take the irrelevance result as meaning that legal representation is never useful, for lawyers perform a number of valuable tasks that are not captured by our model and that are likely to be important in complex trials. Our analysis focuses on one particular aspect of legal representation — the role of lawyers as gate-keepers of information reaching the court. To the extent that this role is crucial from the information elicitation perspective, however, our irrelevance result clarifies and qualifies the sense in which lawyers can make a difference. In fact, extensions of our model will explore the situations in which lawyers *do* affect the adjudication outcomes. We turn to these extensions next.

4 Withholding Favorable Evidence and Lawyer Reputation

Thus far, we have focused on equilibria in undominated strategies. Lawyers adopting such strategies never advise their clients to suppress ex post favorable evidence. Although such a strategy is compelling for reasons discussed above, deviating from that strategy — i.e., suppressing favorable evidence — may influence the judge's posterior and thus her ruling favorably for his client. In

this section, we explore this possibility by relaxing the restriction to undominated strategies. In so doing, we continue to focus on equilibria in which the judge follows a cutoff strategy, which is reasonable and intuitive.¹⁸

To see how the adoption of (weakly) dominated strategy influences the judge's inference, suppose that both parties have retained lawyers, and that party 2 (or his lawyer) follows the dominant strategy of disclosing all favorable evidence, but the lawyer for party 1 advises him to withhold s regardless of its value. Given these strategies, the judge's posterior becomes $\mathbb{E}[g|0, h(\theta), \theta]$ which is less than $\mathbb{E}[g|h(\theta), h(\theta), \theta]$, and is therefore more favorable to party 1 than if both adopt their dominant strategies. Intuitively, party 1's withholding of favorable evidence causes the judge to interpret nondisclosure more favorably for that party. Likewise, if party 1 adopts the dominant strategy but party 2 adopts the strategy of never disclosing his private evidence, then the judge forms a posterior $\mathbb{E}[g|h(\theta), 1, \theta] > \mathbb{E}[g|h(\theta), h(\theta), \theta]$, which is more favorable for party 2 than if both adopt their dominant strategies. Define

$$\hat{\theta}_+(b_1, b_2, c) := \inf\{\theta \mid \mathbb{E}[g|0, h(\theta), \theta; b_1, b_2, c] > 0\};$$

$$\hat{\theta}_-(b_1, b_2, c) := \inf\{\theta \mid \mathbb{E}[g|h(\theta), 1, \theta; b_1, b_2, c] > 0\}.$$

Using the same argument as in Lemma A2 in the Appendix, we can show that $\mathbb{E}[g|0, h(\theta), \theta] > 0$ if and only if $\theta > \hat{\theta}_+(b_1, b_2, c)$, and $\mathbb{E}[g|h(\theta), 1, \theta] > 0$ if and only if $\theta > \hat{\theta}_-(b_1, b_2, c)$, and also that $\hat{\theta}_-(b_1, b_2, c) \leq \hat{\theta}^*(b_1, b_2) \leq \hat{\theta}_+(b_1, b_2, c)$. If $(b_1, b_2, c) > (0, 0, 0)$, then at least one inequality is strict, so the region $[\hat{\theta}_-(b_1, b_2, c), \hat{\theta}_+(b_1, b_2, c)]$ is non-degenerate.¹⁹

If $\theta < \hat{\theta}_-(b_1, b_2, c)$, the judge will surely rule for party 1 because the judge's posterior is negative even if she holds the most favorable beliefs for party 2 that party 2 does not disclose any evidence and party 1 discloses all favorable evidence. Similarly, if $\theta > \hat{\theta}_+(b_1, b_2, c)$, then the judge will rule for party 2 in the absence of disclosure because her posterior is positive even if she holds the most favorable belief for party 1. But if $\theta \in [\hat{\theta}_-(b_1, b_2, c), \hat{\theta}_+(b_1, b_2, c)]$, then the judge's default ruling is affected by the parties' disclosure strategies, because for such θ ,

$$\mathbb{E}[g|0, h(\theta), \theta] < 0 < \mathbb{E}[g|h(\theta), 1, \theta],$$

So in this case, the judge's inference and hence her ruling can go either way depending on the equilibrium disclosure strategies.

¹⁸An earlier version of this paper presents other equilibria. Since they do not add any new insight, we omit them here.

¹⁹Note that $\hat{\theta}_-(b_1, b_2, c) \leq \hat{\theta}^*(b_1, b_2) \leq \hat{\theta}_+(b_1, b_2, c)$ because $\mathbb{E}[g|0, h(\theta), \theta] < \mathbb{E}[g|h(\theta), h(\theta), \theta] < \mathbb{E}[g|h(\theta), 1, \theta]$. Further, $\hat{\theta}_+(b_1, b_2, c) > \hat{\theta}^*(b_1, b_2, c)$ if $\hat{\theta}^+(b_1, b_2, c) > 0$ and $\hat{\theta}^-(b_1, b_2, c) < 1$. Given $(b_1, b_2, c) > (0, 0, 0)$, the former holds if $\mathbb{E}[g(s, 0)] < 0$ and the latter holds if $\mathbb{E}[g(s, 1)] > 0$, both of which hold by assumption.

It turns out that any $\hat{\theta} \in [\hat{\theta}_-(b_1, b_2, c), \hat{\theta}_+(b_1, b_2, c)]$ can be sustained as a threshold for the judge's default ruling. Suppose that represented party 1 discloses s if and only if $s < h(\theta)$ and $\theta > \hat{\theta}$, and represented party 2 discloses s if and only if $s > h(\theta)$ and $\theta < \hat{\theta}$. Then, the judge's posterior in the case of no disclosure becomes $\mathbb{E}[g|0, h(\theta), \theta] < 0$ if $\theta < \hat{\theta}$ and $\mathbb{E}[g|h(\theta), 1, \theta] > 0$ if $\theta > \hat{\theta}$. Hence, under these disclosure strategies, it is optimal for the judge to rule for party 1 if and only if $\theta < \hat{\theta}$. Given this default ruling behavior, the parties' disclosure strategies are in turn optimal. We thus report the following observations, for the case of full and partial representations.

PROPOSITION 4 (Relevance of legal advice with dominated strategies) *(i) Suppose that both parties retain lawyers. Then, for any $\hat{\theta} \in [\hat{\theta}_-(b_1, b_2, c), \hat{\theta}_+(b_1, b_2, c)]$, there exists a perfect Bayesian Nash equilibrium such that in the absence of disclosure the judge rules for party 1 [resp. party 2] if $\theta < \hat{\theta}$ [resp. $\theta > \hat{\theta}$]. Conversely, any equilibrium cutoff default ruling strategy has a threshold in $[\hat{\theta}_-(b_1, b_2, c), \hat{\theta}_+(b_1, b_2, c)]$.*

(ii) Suppose only party 1 [resp. party 2] retains a lawyer. Then, for any $\hat{\theta} \in [\hat{\theta}^(b_1, b_2), \hat{\theta}_+(b_1, b_2, c)]$ [resp. $\hat{\theta} \in [\hat{\theta}_-(b_1, b_2, c), \hat{\theta}^*(b_1, b_2)]$], there exists a perfect Bayesian Nash equilibrium such that in the absence of disclosure the judge rules for party 1 [resp. party 2] if $\theta < \hat{\theta}$ [resp. $\theta > \hat{\theta}$]. Conversely, any cutoff default ruling strategy by the judge has a threshold in $[\hat{\theta}^*(b_1, b_2), \hat{\theta}_+(b_1, b_2, c)]$ [resp. $[\hat{\theta}_-(b_1, b_2, c), \hat{\theta}^*(b_1, b_2)]$].*

Proposition 4 shows that a range of different nondisclosure rulings reflecting different posteriors by the judge, is sustainable by legal strategies of withholding favorable evidence. Such strategies, if credible, can manipulate the judge's inference. For instance, a default ruling with a threshold $\hat{\theta} > \hat{\theta}^*(b_1, b_2)$ would be more favorable to party 1 than that sustainable under dominant strategies. Since these strategies require conditioning disclosure of evidence on the realized level of θ , they cannot be played without the legal expertise of a lawyer. In this sense, we have identified a source of private value of legal advice — the ability to advise a party to withhold favorable evidence.

This benefit is most evident when only one party has access to legal advice, the case illustrated in (ii) of Proposition 4. Suppose initially neither party retains a lawyer. Then, by Proposition 1, the judge employs a threshold of $\hat{\theta}^*(b_1, b_2)$ in her default ruling. Suppose now only party 1 hires a lawyer. Provided that the lawyer can credibly advise party 1 to adopt a dominated strategy, the judge can be induced to employ a more favorable threshold in $[\hat{\theta}^*(b_1, b_2), \hat{\theta}_+(b_1, b_2, c)]$. More importantly, any equilibrium threshold in this situation is (weakly) more favorable for party 1 than $\hat{\theta}^*(b_1, b_2)$, the threshold used when neither party hires a lawyer. In this sense, legal advice becomes relevant once (weakly) dominated strategies are available. Nevertheless, if both parties hire lawyers, the parties cannot be both better off than if neither party hired a lawyer. The last 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.

Clearly, the relevance of legal advice rests on the credibility of weakly dominated strategy of withholding favorable evidence. One way in which a lawyer can achieve such credibility could be by building a reputation through repeated trials. Casual observation suggests that lawyers are concerned about their reputations and undertake specific steps to develop/enhance them. For example, some criminal defense lawyers are known to call very few witnesses. Sometimes, a lawyer would rest the case without calling any witnesses at all if (s)he considers that the case has not been proven by the prosecutors. Our results indicate that if such behavior helps to build a lawyer's reputation that could skew the court's ruling in favor of that lawyer and his 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 representation 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.

The interval, $[\hat{\theta}_-(b_1, b_2, c), \hat{\theta}_+(b_1, b_2, c)]$, can be interpreted as a range of lawyer manipulation of the outcomes. It is worth noting that this range depends on the inference rule (b_1, b_2, c) . In particular, this range vanishes as $(b_1, b_2, c) \rightarrow (0, 0, 0)$; i.e., as the judge's inference rule becomes completely non-Bayesian/non-prejudicial. This feature points to a possible rationale for regulating the judge's inference rule — a topic addressed more fully in the next section.

A complete analysis of the reputation-building requires a repeated game framework that goes beyond the scope of this paper. But it is worth pointing out a number of factors that can facilitate the establishment of a reputation. First, it is easier to build a reputation when a lawyer practices in the same court and in front of the same judge. Second, the strategy of withholding favorable evidence is optimal when there exists no doubt that the client's case is very strong purely on legal grounds and the commonly held evidence, yet the criterion for the latter depends on the equilibrium that is being played. So, if it is less than common knowledge that the court's default ruling would be favorable for the client, exculpatory evidence will never be withheld, and only the undominated equilibrium will survive. This concern appears to be particularly relevant in the jury trials. Legal observers indicate that jury trials are subject to considerable uncertainties and their outcome is often hard to predict. In such cases, a lawyer will always disclose all favorable evidence²⁰, and so it will be difficult to build necessary reputation. Finally, reputational concerns also imply that a party hiring a lawyer will never do so secretly, since the value of representation comes from a favorable selection of equilibrium.

²⁰In fact, it is a duty of a lawyer to explore all plausible lines of defense.

5 Welfare implications of legal advice and restrictions on judges' inferences

In this section, we consider the welfare implications of our results. Recall that the judge's inference rule, (b_1, b_2, c) , influences the parties' disclosure behavior, and ultimately determines the equilibrium outcome. It is thus important to understand how the inference rule affects the welfare and how it should be optimally chosen.

Of particular interest is whether Bayesian inference rule $(b_1, b_2, c) = (\frac{p_{10}}{p_{00}}, \frac{p_{01}}{p_{00}}, \frac{p_{11}}{p_{00}})$ is optimal. This issue has special policy relevance for the law of evidence which imposes various restrictions on inferences that adjudicators are allowed to make from evidence or from lack thereof. For instance, the burden of proof is sometimes assigned to one of the parties (e.g., the plaintiff); in criminal cases a negative inference cannot be made when the defendant refuses to testify; a certain inference is to be made even with insufficient evidence (i.e. prima facie rules); and self-interest alone does not constitute a reason for discounting evidence (Daughety and Reinganum, 2000a, 2000b; Posner, 1999). If Bayesian inference rule is optimal, then such restrictions would not be warranted for rational adjudicators who strive to adjudicate in a Bayesian fashion. If not, some restrictions may be warranted and it is important to examine what they are. Despite much debate on this issue, it has not been addressed analytically before.

Our model allows us to address this issue in a simple fashion. It is natural to take the judge's objective as our welfare criterion. When maximizing this criterion, we can treat the judge's threshold as the only choice variable because, by the equivalence result of Lemma 1, it completely pins down the outcome. So, we can formulate our welfare inquiry as the following mechanism design problem:

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

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 in Section 2). The analysis of [WP] yields the following result.

PROPOSITION 5 *A cutoff ruling strategy with a Bayesian threshold, $\hat{\theta}^*(\frac{p_{10}}{p_{00}}, \frac{p_{01}}{p_{00}})$, is socially optimal, i.e. solves [WP].*

According to Proposition 5, the equilibrium outcome is socially optimal if the parties are self-represented or are represented by lawyers disclosing all favorable evidence, and the judge uses the Bayesian inference rule. In this sense, Proposition 5 argues against regulating adjudicators' inferences.

This conclusion does not hold, however, if lawyers can credibly use the strategies of not disclosing all favorable evidence. Proposition 4 shows that, in this case, any threshold $\hat{\theta} \in [\hat{\theta}_-(\frac{p_{10}}{p_{00}}, \frac{p_{01}}{p_{00}}, \frac{p_{11}}{p_{00}}), \hat{\theta}_+(\frac{p_{10}}{p_{00}}, \frac{p_{01}}{p_{00}}, \frac{p_{11}}{p_{00}})]$ can arise in equilibrium even under Bayesian inference rule. Thus, if a lawyer, due to his reputation, credibly uses a disclosure strategy which shifts the threshold away from $\hat{\theta}^*(\frac{p_{10}}{p_{00}}, \frac{p_{01}}{p_{00}})$, then legal representation is privately valuable but moves the outcome away from social optimum. So legal advice is socially harmful when it matters.

In light of this result, the following proposition provides a possible justification for regulating the adjudicators' inferences.

PROPOSITION 6 *(i) If $(p_{10}, p_{01}) \gg (0, 0)$, then there exists $(b_1, b_2) \ll (\frac{p_{10}}{p_{00}}, \frac{p_{01}}{p_{00}})$ with either $b_1 = 0$ or $b_2 = 0$ such that $\hat{\theta}^*(b_1, b_2) = \hat{\theta}^*(\frac{p_{10}}{p_{00}}, \frac{p_{01}}{p_{00}})$. (ii) There exists (b_1, b_2) with either $b_1 = 0$ or $b_2 = 0$, such that*

$$\hat{\theta}^*(\frac{p_{10}}{p_{00}}, \frac{p_{01}}{p_{00}}) \in [\hat{\theta}_-(b_1, b_2, 0), \hat{\theta}_+(b_1, b_2, 0)] \subset [\theta_-(\frac{p_{10}}{p_{00}}, \frac{p_{01}}{p_{00}}, \frac{p_{11}}{p_{00}}), \theta_+(\frac{p_{10}}{p_{00}}, \frac{p_{01}}{p_{00}}, \frac{p_{11}}{p_{00}})]. \quad (5)$$

The first part of Proposition 6 shows that the outcome attained under the Bayesian inference rule can be replicated by an inference rule that assigns no burden of proof to one of the parties, provided that the parties follows undominated strategies. The evidence law often assigns the entire burden of proof to one of the parties, e.g. the plaintiff. Interestingly, a one-sided allocation does not necessarily mean that the adjudication outcome would be different from the one attained under the Bayesian rule. By part (ii) of the Proposition, such a burden of proof allocation has the additional advantage of being less susceptible to manipulation by the lawyers, as the range of equilibrium outcomes around the optimal outcome is narrower than that under the Bayesian one. These two findings imply that committing adjudicators to an inference rule different from the Bayesian could be socially desirable. Thus, our analysis offers some rationale for regulating adjudicators' inferences.

6 Costly Legal Advice as a Signal

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. Moreover, it applies to a large number of criminal cases in which the state pays for the defense. However, to broaden the appeal of our analysis, it is natural to extend our model to the case of costly legal representation. The goal of such extension would be to understand how costly legal advice changes the nature of the disclosure game. In particular, we will explore who is more likely to pay for such advice and how it affects adjudication outcomes, as well as social and private welfare.

To extend our model in this direction, we need to add more structure to it. In particular, we need to quantify the cost of legal representation and the value of winning in comparable units. To this end, we assume that the cost of legal advice is $w > 0$, and that a party derives the value of $v > 0$ when (s)he wins the dispute and zero when (s)he loses, all measured in monetary units.

For ease of exposition, we simplify our model in several ways. First, we assume that $b_1 = b_2 = b$. Likewise, s and θ are assumed 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 says that the strength of the case for party 1 at state (s, θ) is the same as the strength of the case for party 2 at state $(1 - s, 1 - \theta)$. The assumption implies that $g(\frac{1}{2}, \frac{1}{2}) = 0$, or $h^{-1}(\frac{1}{2}) = \frac{1}{2}$.

With costly legal advice, there are two types of equilibria: an “uninformative” equilibrium in which legal advice has no signalling role, and an “informative” one in which legal representation becomes a signal. We begin with the uninformative equilibrium.

PROPOSITION 7 (**Uninformative 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 $\hat{\theta}^*$ characterized by (3), regardless of whether any party has retained a lawyer. Then, it is a best response for unrepresented parties to follow disclosure strategies with threshold $h(\hat{\theta}^*)$ which, in turn, rationalizes the judge’s beliefs and her default ruling behavior under self-representation. Also, if party 1 (party 2) hires a lawyer, the judge believes that this party will follow a dominant strategy of disclosing θ if and only if $g(s, \theta) < 0$ ($g(s, \theta) > 0$). Then, the Decision Equivalence Lemma 1 implies that no party will have an incentive to obtain costly legal advice.

There exists another, more interesting, equilibrium in which legal representation plays a non-trivial signaling role. In this equilibrium, a lawyer can be hired for two reasons. First, a party without evidence hires a lawyer to add “credibility” to his lack of evidence. Second, a party with sufficiently weak (likely to be unfavorable) evidence does so to imitate those who do not have any evidence. Specifically, for any cost $w > 0$, there exists a threshold $\hat{s}(w) \in [\frac{1}{2}, 1]$ (defined below), such that the following *lawyer signaling strategies* constitute an equilibrium:

- Party 1 [resp. party 2] retains a lawyer either if he observes no evidence or if he observes $s > \hat{s}(w)$ [resp. $s < 1 - \hat{s}(w)$]. In the latter case, he discloses s if and only if $g(s, \theta) < 0$ [resp. $g(s, \theta) > 0$].
- If party 1 [resp. party 2] observes $s \in [0, \hat{s}(w)]$ [resp. $s \in [1 - \hat{s}(w), 1]$], he does not hire a lawyer and discloses s .
- 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 [resp. party 2] has retained a lawyer, then the judge rules for party 1 if and only if $g(0, \theta) < 0$ [resp. $g(1, \theta) < 0$].

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 $\hat{s}(w)$. It must satisfy

$$v(1-p)[L(\frac{1}{2}) - L(h^{-1}(\hat{s}(w)))] = w, \quad (6)$$

To understand (6), note that party 1 with signal $s = \hat{s}(w)$ must be indifferent between hiring a lawyer *and* not hiring a lawyer and disclosing $\hat{s}(w)$. The first strategy has cost w but raises party 1's chance of winning if party 2 has not observed $\hat{s}(w)$. (If party 2 observes $\hat{s}(w)$, then he will disclose it because $\hat{s}(w) \in [1 - \hat{s}(w), 1]$, so party 1's decision would not make any difference.) So, suppose that party 2 does not observe $\hat{s}(w)$. Then, 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 the strategy of hiring a lawyer allows him to win if $\theta < \frac{1}{2}$. Hence, after hiring a lawyer party 1's probability of winning increases by $(1-p)[L(\frac{1}{2}) - L(h^{-1}(\hat{s}(w)))]$. Therefore, for party 1 with evidence $\hat{s}(w)$ to be indifferent, the expected gain from winning more often on the LHS of (6) must equal the cost of hiring a lawyer, w , on the RHS of (6).

Since $\hat{s}(w) \leq 1$, the cost of legal representation must not exceed $\bar{w} := v(1-p)[L(\frac{1}{2}) - L(h^{-1}(1))]$ for a lawyer to be ever hired. For any $w \in [0, \bar{w}]$, there is a unique solution $\hat{s}(w)$ to (6): its LHS is nondecreasing in $\hat{s}(w)$ and is equal to zero at $s = \frac{1}{2}$ and equals \bar{w} at $s = 1$. Hence, $\hat{s}(w)$ exists and lies in $[\frac{1}{2}, 1]$ for any $w \in [0, \bar{w}]$. Further, $\hat{s}(w)$ is increasing in w and equals 1 when $w = \bar{w}$.

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

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

In this equilibrium, a lawyer “buys” a party “the right to be silent without prejudice,” in the sense that a party hiring a lawyer does not suffer from an extremely negative inference if he does not disclose. Those who choose to hire a lawyer and “buy this right” either have no evidence to reveal or have sufficiently unfavorable evidence (i.e., party 1 with $s > \hat{s}(w)$ or party 2 with $s < 1 - \hat{s}(w)$). Since these types strictly prefer to hire a lawyer, legal advice *does* have private value.

Recent cases have raised the concern that those who can afford prominent lawyers can buy justice. To see this in our model, 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.²¹ In this sense, expensive lawyers can buy justice. But the represented party himself does not benefit 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 parties 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 non-disclosure (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, 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 $\hat{s}(w) > \frac{1}{2}$. The disclosure of evidence $s \in [\frac{1}{2}, \hat{s}(w)]$ (resp. $s \in [1 - \hat{s}(w), \frac{1}{2}]$) by party 1 (resp. party 2) represents the aforementioned extra disclosure, as this party would consider such s too unfavorable to disclose in the no-representation equilibrium (which gives rise to the representation-neutral outcome). The increased disclosure is depicted in Figure 3 as a rightward shift of party 1’s threshold and a leftward shift of party 2’s threshold.

²¹A simple revealed preference argument shows this. By continuity, when ϵ is small, the probabilities that party 1 wins *after disclosing* $s(w) - \epsilon$ and $s(w) + \epsilon$, respectively, are close to each other. However, party 1 with signal $s(w) + \epsilon$ prefers to incur a positive cost of hiring a lawyer which therefore must be associated with less disclosure but a discrete increase in the probability of winning.

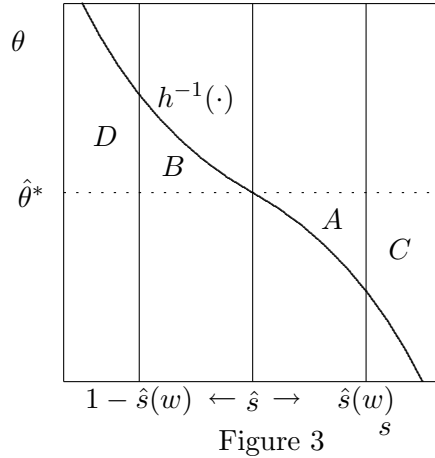


Figure 3

As a result, compared with the representation-neutral outcome, the lawyer signaling equilibrium entails different outcomes in two cases: (i) party 1 alone observes evidence and the state is in area A ; (ii) party 2 alone observes evidence and the state is in area B . In both cases, wrong adjudication decisions are made in the representation-neutral outcome but these mistakes are corrected in the lawyer signaling equilibrium. For instance, suppose that only party 1 observes evidence and the state is in area A . Then, the judge would rule incorrectly for 1 in the representation-neutral outcome because s would not be disclosed, but she will rule correctly for 2 in the lawyer signaling equilibrium because party 1 will disclose s . Still, the lawyer signaling equilibrium involves incorrect decisions if party 1 (2) alone observes evidence and the state is in area C (D). But as w increases toward \bar{w} , the regions of erroneous ruling shrink and the social welfare increases. In the limit where $w = \bar{w}$, there is full disclosure, and so adjudication decisions are always correct.

COROLLARY 2 The lawyer signaling equilibrium yields higher social welfare than the representation-neutral outcome. The social welfare is increasing in the lawyer fee w on $(0, \bar{w}]$ and attains the first-best at $w = \bar{w}$, in which case all private evidence is disclosed.

This result shows that legal advice can be socially valuable. Yet, it is important to recognize that this “social value” is not generated by the lawyers’ expertise. The role of lawyers is purely dissipative. In this sense, our result casts doubt on the social value of lawyers’ expertise in assessing private evidence.

The lawyer signaling equilibrium also has an implication for regulating adjudicators’ inferences. Adjudicators are often prohibited from drawing negative inferences against parties refusing to disclose their information, lest such inference may distort the judgment. Little is known, however, about how such inferences would affect parties’ disclosure incentives. Our analysis shows that

negative inferences on nondisclosure can improve parties' disclosure incentives and lead to a better outcome. In our lawyer signaling equilibrium, an unrepresented party faces a very high burden of proof when not disclosing his evidence and is therefore compelled to disclose more evidence. A recent case illustrates the relevance of this point. In May 2005, financier Ronald Perelman won his fraud case against Morgan Stanley largely due to the judge's decision to shift the burden of proof from Perelman, the plaintiff, to Morgan Stanley, the defendant. The judge took this decision after Morgan Stanley was found to have failed to produce a large number of internal emails relevant to the case.²² At the minimum, our result suggests that a careful examination of the rules governing adjudicators' inferences is warranted.

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 the representation decision,²³ 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 uninformative equilibrium of Proposition 7.

7 Conclusion

In this paper, we have studied the effect of legal representation on the adjudication outcomes. Our analysis was concerned with the role of lawyers as gate-keepers of information reaching the court. We have compared outcomes under legal representation and self-representation, and have shown that legal representation does not affect the adjudication outcome under a fairly broad set of circumstances. This result casts doubt on the significance of the lawyers' expertise for optimal evidence disclosure decisions. At the same time, we have shown that lawyers can affect the adjudication outcome if they can leverage their legal expertise to engage in more sophisticated strategic behavior. First, they can do so by credibly withholding some favorable information and thereby affecting the judge's inference in their clients favor when evidence is not disclosed. Second, legal representation can play a signaling role. In our signalling equilibrium, hiring a lawyer allows

²²See "Perelman Beats Morgan Stanley," *Wall Street Journal*, May 16 2005.

²³If we allow such randomization, then there is only one other possible equilibrium scenario: an uninformed party randomizes between hiring and not hiring a lawyer. Although we have not been able to completely rule out this type of equilibria, it is easy to show that the qualitative nature of the outcome in such an equilibrium would be the same as in the lawyer signalling equilibrium of Proposition 8: an informed party with a 'good' signal (small s for party 1 and high s for party 2) does not hire a lawyer and discloses, while an informed party with a 'bad' signal will hire a lawyer and typically would not disclose.

a litigating party to avoid a negative inference from non-disclosure, so legal representation “buys a right to be silent without being prejudiced” which ultimately affects the litigation outcome.

Besides 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. Recent legal cases have generated 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 various rules regulating the adjudicators interpretation of evidence and restrictions on inferences they can or cannot draw from nondisclosure of evidence.

There are several avenues of extending our model. First, it would be interesting to consider how legal representation affects the outcome when the parties have an opportunity to settle at a pre-trial stage. Second, it would be natural to extend our model to costly evidence production case and explore the effect of legal representation. We intend to address these issues in future research.

8 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 cdf's 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: $\int_0^{\hat{s}} g(s, \theta) \frac{f(s|\theta')}{F(\hat{s}|\theta')} ds \geq \int_0^{\hat{s}} g(s, \theta) \frac{f(s|\theta)}{F(\hat{s}|\theta)} ds$, 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 $\int_0^{\hat{s}} g(s, \theta) f(s|\theta') ds \geq \int_0^{\hat{s}} g(s, \theta) f(s|\theta) ds$. \blacksquare

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

PROOF: Recall that

$$\mathbb{E}[g|\hat{s}, \hat{s}, \theta] := \int_0^1 g(s, \theta) f(s|\theta) ds + b_1 \int_{\hat{s}}^1 g(s, \theta) f(s|\theta) ds + b_2 \int_0^{\hat{s}} g(s, \theta) f(s|\theta) ds. \quad (7)$$

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

$$\int_0^1 g(s, \theta') f(s|\theta') ds > \int_0^1 g(s, \theta) f(s|\theta) ds \quad (8)$$

and

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

To get the second inequality, rewrite the set restriction as $1_{\{s > \hat{s}\}}$, an increasing function of s . Therefore, the first term $\int_0^1 g(s, \theta) f(s|\theta) ds$ is strictly increasing in θ , and the second term $b_1 \int_{\hat{s}}^1 g(s, \theta) f(s|\theta) ds$ 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, together with (8) and (9), implies that $\mathbb{E}[g|\hat{s}, \hat{s}, \theta'] > 0$.

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 (8) and (9), we again conclude that $\mathbb{E}[g|\hat{s}, \hat{s}, \theta'] > 0$. \blacksquare

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

PROOF: Recall that

$$\mathbb{E}[g|h(\theta), h(\theta), \theta] = \int_0^1 g(s, \theta) f(s|\theta) ds + b_1 \int_{h(\theta)}^1 g(s, \theta) f(s|\theta) ds + b_2 \int_0^{h(\theta)} g(s, \theta) f(s|\theta) ds. \quad (10)$$

The first term of (10) is increasing in θ by (8). Now consider the second term. We have:

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

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

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

Differentiating $\int_0^{h(\theta)} g(s, \theta) f(s|\theta) ds$ 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 discloses s in that equilibrium if (only if)

$$P_0(s) < (\leq) \Pr\{\theta < h^{-1}(s) \mid s\}. \quad (13)$$

Similarly, party 2 discloses s if (only if)

$$P_0(s) > (\geq) \Pr\{\theta < h^{-1}(s) \mid s\}. \quad (14)$$

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 $s' > s$, $P_0(s') > \Pr\{\theta < h^{-1}(s') \mid s'\}$ if $P_0(s) = \Pr\{\theta < h^{-1}(s) \mid s\}$. 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. Then, we have:

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

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

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

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 $\hat{s} \in [0, 1]$ s.t. party 1 discloses (withholds) s if $s < \hat{s}$ ($s > \hat{s}$) and party 2 discloses (withholds) s if $s > \hat{s}$ ($s < \hat{s}$).²⁴ ■

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 $\mathbb{E}[g|\hat{s}, \hat{s}, \theta]$. Then, by Lemma A2, there exists $\hat{\theta} \in \Theta$ such that $\delta(\theta) = 0$ if $\theta < \hat{\theta}$ and $\delta(\theta) = 1$ if $\theta > \hat{\theta}$. ■

²⁴If some party, say party 1, has a strict incentive for disclosing all s , then the statement remains valid with $\hat{s} = 1$.

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 $\Pr\{\theta < \hat{\theta} \mid s\} < (\leq) \Pr\{\theta < h^{-1}(s) \mid s\}$ if (only if) $s < (\leq) \hat{s}$. Similarly, $\Pr\{\theta < \hat{\theta} \mid s\} > (\geq) \Pr\{\theta < h^{-1}(s) \mid s\}$ 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 $h(\hat{\theta}^*)$.*²⁵

PROOF: Recall from (3) that

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

It then follows from Lemma A2 that

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

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 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. ■

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 $\mathbb{E}[g|h(\theta), h(\theta), \theta]$. Recall that

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

Lemma A3 then implies that

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

²⁵For brevity, we omit the dependence of $\hat{\theta}^*$ on b_1 , b_2 and c .

proving that the judge's cutoff default ruling strategy with threshold $\hat{\theta}^*$ is optimal. The uniqueness of 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 $\mathbb{E}[g|h(\theta), \hat{s}, \theta]$. Lemmas A1 and A2 imply that this posterior is ordinally monotonic: i.e., $\mathbb{E}[g|h(\theta), \hat{s}, \theta] \geq 0$ implies $\mathbb{E}[g|h(\theta'), \hat{s}, \theta'] > 0$ 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. ■

PROOF OF LEMMA 1: Fix a judge's default ruling strategy to be of the cutoff form with a threshold $\hat{\theta} \in \Theta$. We show that any best responses by the two parties in their disclosure strategies to the judge's strategy lead to the same outcome, regardless of whether each party has obtained legal advice or not. To begin, given a threshold $\hat{\theta}$, consider a set of disclosure strategies of party 1, $\mathcal{S}_1^{\hat{\theta}}$, such that $\rho_1(s, \theta) \in \mathcal{S}_1^{\hat{\theta}}$ if

$$\begin{aligned} \theta > \hat{\theta} \text{ and } g(s, \theta) < 0 &\implies \rho_1(s, \theta) = 1 \\ \theta < \hat{\theta} \text{ and } g(s, \theta) > 0 &\implies \rho_1(s, \theta) = 0. \end{aligned}$$

Similarly, let $\mathcal{S}_2^{\hat{\theta}}$ be the set of disclosure strategies of party 2 such that $\rho_2(s, \theta) \in \mathcal{S}_2^{\hat{\theta}}$ if

$$\begin{aligned} \theta > \hat{\theta} \text{ and } g(s, \theta) < 0 &\implies \rho_2(s, \theta) = 0 \\ \theta < \hat{\theta} \text{ and } g(s, \theta) > 0 &\implies \rho_2(s, \theta) = 1. \end{aligned}$$

A party $i = 1, 2$ following a strategy in $\mathcal{S}_i^{\hat{\theta}}$ will always present an evidence that would overturn an unfavorable default ruling, and would never present an evidence that will overturn a favorable ruling. Such a strategy is optimal for each party, regardless of the opponent's disclosure strategy: If the latter discloses then a party's strategy has no effect, whereas if he does not disclose, then no other strategy can make the party strictly better off. (The argument for this is sketched in the paragraphs preceding Proposition 1.) Hence, if such strategy is feasible, then it will be used by each party. If a party $i = 1, 2$ has obtained a legal advice, clearly all strategies in $\mathcal{S}_i^{\hat{\theta}}$ are feasible. Interestingly, a strategy in $\mathcal{S}_i^{\hat{\theta}}$ is feasible to party i even when he has no legal advice. To see this, consider party 1's strategy $\rho_1(s, \theta) = 1$ if $s < \hat{s}$ and $\rho_1(s, \theta) = 0$ if $s > \hat{s}$, where

$\hat{s} := \inf\{s|g(s, \hat{t}) > 0\}$. Clearly, this strategy is feasible and belongs to $\mathcal{S}_1^{\hat{\theta}}$. (A symmetric cutoff strategy works for party 2.)

Finally, any disclosure strategies $(\rho_1, \rho_2) \in \mathcal{S}_1^{\hat{\theta}} \times \mathcal{S}_2^{\hat{\theta}}$, given the judge's cutoff default ruling strategy with threshold $\hat{\theta}$, induce the outcome in (4), which depends only on the threshold value $\hat{\theta}$. ■

PROOF OF PROPOSITION 4: Before proceeding, it is useful to establish some preliminary results. The arguments employed in Lemmas A2 and A3 can be combined to show that $\forall \theta' > \theta$:

$$\begin{aligned}\mathbb{E}[g|0, h(\theta), \theta] \geq 0 &\implies \mathbb{E}[g|0, h(\theta'), \theta'] > 0 \\ \mathbb{E}[g|h(\theta), 1, \theta] \geq 0 &\implies \mathbb{E}[g|h(\theta'), 1, \theta'] > 0.\end{aligned}$$

From these, it follows that $\mathbb{E}[g|0, h(\theta), \theta] > 0$ if and only if $\theta > \hat{\theta}_+(b_1, b_2, c)$, and $\mathbb{E}[g|h(\theta), 1, \theta] > 0$ if and only if $\theta > \hat{\theta}_-(b_1, b_2, c)$.

We first prove (i). Fix any $\hat{\theta} \in [\hat{\theta}_-(b_1, b_2, c), \hat{\theta}_+(b_1, b_2, c)]$. 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 $\mathbb{E}[g|0, h(\theta), \theta] < 0$ if $\theta < \hat{\theta}$ and $\mathbb{E}[g|h(\theta), 1, \theta] > 0$ 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 $[\hat{\theta}_-(b_1, b_2, c), \hat{\theta}_+(b_1, b_2, c)]$, suppose otherwise i.e. there exists an equilibrium strategy combination $(\hat{\theta}, \rho_1, \rho_2)$ s.t. $\hat{\theta} \notin [\hat{\theta}_-(b_1, b_2, c), \hat{\theta}_+(b_1, b_2, c)]$. At first, consider $\hat{\theta} > \hat{\theta}_+(b_1, b_2, c)$. Then $\mathbb{E}[g|\rho_1, \rho_2, \tilde{\theta}] \leq 0$ for an arbitrary $\tilde{\theta} \in (\hat{\theta}_+(b_1, b_2, c), \hat{\theta})$, where

$$\mathbb{E}[g|\rho_1, \rho_2, \tilde{\theta}] = \int_0^1 g(s, \tilde{\theta})f(s|\tilde{\theta})ds + b_1 \int_0^1 (1 - \rho_1(s|\tilde{\theta}))g(s, \tilde{\theta})f(s|\tilde{\theta})ds + b_2 \int_0^1 (1 - \rho_2(s|\tilde{\theta}))g(s, \tilde{\theta})f(s|\tilde{\theta})ds.$$

Note that both parties never disclose the same s in equilibrium, because (ρ_1, ρ_2) are best response strategies and therefore $\rho_i \in \mathcal{S}_i^{\hat{\theta}}$ for $i \in \{1, 2\}$ (see the proof of Lemma 1). So the term associated with c disappears.

Let us compare $\mathbb{E}[g|\rho_1, \rho_2, \tilde{\theta}]$ and $E[g|0, h(\tilde{\theta}), \tilde{\theta}]$ term by term. First, $\int_0^1 (1 - \rho_2(s, \tilde{\theta}))g(s, \tilde{\theta})f(s|\tilde{\theta})ds \geq \int_0^{h(\tilde{\theta})} g(s, \tilde{\theta})f(s|\tilde{\theta})ds$. Also, by Lemma 1, $\rho_1(s, \theta) = 0$ if $g(s, \theta) > 0$. Hence,

$$\int_0^1 (1 - \rho_1(s|\tilde{\theta}))g(s, \tilde{\theta})f(s|\tilde{\theta})ds \geq \int_0^1 g(s, \tilde{\theta})f(s|\tilde{\theta})ds.$$

Therefore, $\mathbb{E}[g|\rho_1, \rho_2, \tilde{\theta}] > 0$ if $E[g|0, h(\tilde{\theta}), \tilde{\theta}] > 0$. Since $\tilde{\theta} > \hat{\theta}_+(b_1, b_2, c)$, $E[g|0, h(\tilde{\theta}), \tilde{\theta}] > 0$, so $\mathbb{E}[g|\rho_1, \rho_2, \tilde{\theta}] > 0$ - a contradiction. Hence, $\hat{\theta} \leq \hat{\theta}_+(b_1, b_2, c)$. A symmetric argument proves that $\hat{\theta} \geq \hat{\theta}_-(b_1, b_2, c)$.

Turning now to part (ii), let us without loss of generality focus on the case in which only party 1 obtains a legal advice. Choose any $\hat{\theta} \in [\hat{\theta}^*(b_1, b_2), \hat{\theta}_+(b_1, b_2, c)]$. 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 $\mathbb{E}[g|0, h(\hat{\theta}), \theta]$ if $\theta < \hat{\theta}$ and $\mathbb{E}[g|h(\theta), h(\hat{\theta}), \theta]$ if $\theta > \hat{\theta}$. But $\mathbb{E}[g|0, h(\hat{\theta}), \theta] < 0$ and $\mathbb{E}[g|h(\theta), h(\hat{\theta}), \theta] > 0$ because $\hat{\theta} \in [\hat{\theta}^*(b_1, b_2), \hat{\theta}_+(b_1, b_2, c)]$. Hence, the judge's cutoff strategy is optimal. The proof of the converse is analogous to that for part (i) and is therefore omitted. \blacksquare

PROOF OF PROPOSITION 5: The objective function of $[WP]$ can be rewritten as follows:

$$p_{11} \int_{s>h(\theta)} g(s, \theta) f(s|\theta) l(\theta) ds d\theta + p_{10} \int_{\theta>\hat{\theta}} \int_{s>h(\theta)} g(s, \theta) f(s|\theta) l(\theta) ds d\theta +$$

$$p_{01} \left\{ \int_{\theta>\hat{\theta}} \int_0^1 g(s, \theta) f(s|\theta) l(\theta) ds d\theta + \int_{\theta\leq\hat{\theta}} \int_{s>h(\theta)} g(s, \theta) f(s|\theta) l(\theta) ds d\theta \right\} + p_{00} \int_{\theta>\hat{\theta}} \int_0^1 g(s, \theta) f(s|\theta) l(\theta) ds d\theta,$$

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

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

$$-p_{00} \int_0^1 g(s, \hat{\theta}) f(s|\hat{\theta}) ds l(\hat{\theta}) - p_{10} \int_{h(\hat{\theta})}^1 g(s, \hat{\theta}) f(s|\hat{\theta}) ds l(\hat{\theta}) - p_{01} \int_0^{h(\hat{\theta})} g(s, \hat{\theta}) f(s|\hat{\theta}) ds l(\hat{\theta})$$

$$\stackrel{\geq}{\leq} 0 \quad \text{if} \quad \hat{\theta} \stackrel{\leq}{\geq} \hat{\theta}^* \left(\frac{p_{10}}{p_{00}}, \frac{p_{01}}{p_{00}}, \frac{p_{11}}{p_{00}} \right),$$

The inequalities follow from the definition of $\hat{\theta}^* \left(\frac{p_{10}}{p_{00}}, \frac{p_{01}}{p_{00}} \right)$. So the objective function of $[WP]$ attains its maximum at $\hat{\theta}^* \left(\frac{p_{10}}{p_{00}}, \frac{p_{01}}{p_{00}} \right)$. \blacksquare

PROOF OF PROPOSITION 6: To prove (i), let $\bar{\theta}^* := \hat{\theta}^* \left(\frac{p_{10}}{p_{00}}, \frac{p_{01}}{p_{00}} \right)$. Also, let

$$B := \left(\frac{p_{10}}{p_{00}} \right) \int_{h(\bar{\theta}^*)}^1 g(s, \bar{\theta}^*) f(s|\bar{\theta}^*) ds + \left(\frac{p_{01}}{p_{00}} \right) \int_0^{h(\bar{\theta}^*)} g(s, \bar{\theta}^*) f(s|\bar{\theta}^*) ds.$$

If $B > 0$, set $b_1 = \hat{b}_1 \equiv \frac{B}{\int_{h(\bar{\theta}^*)}^1 g(s, \bar{\theta}^*) f(s|\bar{\theta}^*) ds}$. Since $p_{01} > 0$, we have $0 < \hat{b}_1 < \frac{p_{10}}{p_{00}}$. Furthermore,

$$\mathbb{E}[g|h(\bar{\theta}^*), h(\bar{\theta}^*), \bar{\theta}^*; \hat{b}_1, 0, 0] = \int_0^1 g(s, \bar{\theta}^*) f(s|\bar{\theta}^*) ds + \hat{b}_1 \int_{h(\bar{\theta}^*)}^1 g(s, \bar{\theta}^*) f(s|\bar{\theta}^*) ds$$

$$= \int_0^1 g(s, \bar{\theta}^*) f(s|\bar{\theta}^*) ds + B = \mathbb{E}[g|h(\bar{\theta}^*), h(\bar{\theta}^*), \bar{\theta}^*; \frac{p_{10}}{p_{00}}, \frac{p_{01}}{p_{00}}, \frac{p_{11}}{p_{00}}].$$

This, together with Lemma A3, proves that

$$\hat{\theta}^*(\hat{b}_1, 0) = \bar{\theta}^* = \hat{\theta}^* \left(\frac{p_{10}}{p_{00}}, \frac{p_{01}}{p_{00}} \right),$$

as was to be shown. The case with $B < 0$ is treated symmetrically with $b_1 = 0$.

We now prove (ii). Without loss of generality, assume $B > 0$. Consider any inference rule $(b_1, 0, 0)$ with $b_1 > 0$. Clearly, $\mathbb{E}[g|0, h(\theta), \theta; b_1, 0, 0] > 0$ whenever $\mathbb{E}[g|0, h(\theta), \theta; \frac{p_{10}}{p_{00}}, \frac{p_{01}}{p_{00}}, \frac{p_{11}}{p_{00}}] \geq 0$. Hence, $\hat{\theta}_+(b_1, 0, 0) < \hat{\theta}_+(\frac{p_{10}}{p_{00}}, \frac{p_{01}}{p_{00}}, \frac{p_{11}}{p_{00}})$.

Recall that $\hat{\theta}^*(\frac{p_{10}}{p_{00}}, \frac{p_{01}}{p_{00}}) = \hat{\theta}^*(b_1, 0) \geq \hat{\theta}_-(\frac{p_{10}}{p_{00}}, \frac{p_{01}}{p_{00}}, \frac{p_{11}}{p_{00}})$ and that $\hat{\theta}_+(b_1, 0, 0) - \hat{\theta}_-(b_1, 0, 0)$ monotonically converges to zero as b_1 gets small. Hence, there exists b_1 such that $(b_1, 0, 0)$ satisfies (5) in the statement of the Proposition. The case with $B < 0$ can be treated symmetrically. Finally, if $B = 0$, then $\hat{\theta}_-(0, 0, 0) = \hat{\theta}_+(0, 0, 0) = \hat{\theta}^*(0, 0) = \hat{\theta}^*(\frac{p_{10}}{p_{00}}, \frac{p_{01}}{p_{00}})$, so (5) is satisfied with the inference rule $(0, 0, 0)$. \blacksquare

PROOF OF PROPOSITION 8: Given the strategies adopted by parties 1 and 2, the judge's ruling strategy is rational under her Bayes-consistent beliefs. In particular, the symmetry between the two parties implies that the judge's threshold of $\frac{1}{2}$ 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 for party 1's incentives for 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

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

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 $\theta > h^{-1}(s)$. As a result, the judge rules for party 1 if and only if $\theta < h^{-1}(1)$. So, party 1's payoff from nondisclosure is

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

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

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

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. \blacksquare

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 it 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 $\theta < \max\{h^{-1}(s), \frac{1}{2}\}$. Thus, party 1's expected payoff when he hires a lawyer after observing s is

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

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 (6), (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, indeed, 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

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

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

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

Subtracting (20) from (19) gives

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

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 non-disclosure. 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 a 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 a 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 s.t. $g(s, \theta) > 0$. 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

- Bundy, S.M., and Elhauge, E. (1991), "Do Lawyers Improve the Adversary System? A General Theory of Litigation Advice and Its Regulation," *California Law Review*, 79, 313-420.
- Bundy, S.M., and Elhauge, E. (1993). "Knowledge about Legal Sanctions," *Michigan Law Review*, 92, 261-335.
- Crawford, V., and J. Sobel (1982). "Strategic Information Transmission" *Econometrica*, 50, 1431-1451.
- Daughety, A.F., and Reinganum, J.F. (2000a). "Appealing Judgments," *Rand Journal of Economics*, 31, 502-525.
- Daughety, A.F., and Reinganum, J.F. (2000b). "On the Economics of Trials: Adversarial Process, Evidence and Equilibrium Bias," *Journal of Law, Economics, and Organization*, 16, 365-394.
- Dewatripont, M., and Tirole, J (1999). " Advocates," *Journal of Political Economy*, 107, 1-39.
- Frank, J. (1973), *Courts on Trial: Myth and Reality in American Justice*, Princeton, Princeton University Press.
- Froeb, L.M., and Kobayashi, B.H. (1996), "Naive, Biased, yet Bayesian: Can Juries Interpret Selectively Produced Evidence?" *Journal of Law, Economics, and Organization*, 12, 257-276.
- Froeb, L.M., and Kobayashi, B.H. (2001), "Evidence Production in Adversarial vs. Inquisitorial Regimes," *Economics Letters*, 70, 267-272.
- Grossman, S. (1981), "The Informational Role of Warranties and Private Disclosure about Product Quality," *Journal of Law and Economics*, 24, pp. 461-483.
- Hay, B.L., and Spier, K.E. (1997), "Burdens of Proof in Civil Litigation: An Economic Perspective," *Journal of Legal Studies*, 26, 413-431.
- Kaplow, L., and Shavell, S. (1989), "Legal Advice about Information to Present in Litigation: Its Effects and Social Desirability," *Harvard Law Review*, 102, 567-614.
- Krishna, V., and J. Morgan (2001), "A Model Of Expertise," *The Quarterly Journal of Economics*, 116, 747-775.
- Levy, G. (2005), "Careerist Judges," *Rand Journal Economics*, 36, 275-297.
- Lewis, T.R., and Poitevin, M. (1997), "Disclosure of Information in Regulatory Proceedings" *Journal of Law, Economics, and Organization*, 13, 50-73.
- Lipman, B.L., and Seppi, D.J. (1995), "Robust Inference in Communication Games with Partial Provability," *Journal of Economic Theory*, 66, 370-405.
- Luban, D.J. (1983), "The Adversary System Excuse," in *The Good Lawyer: Lawyers' Roles and Lawyers' Ethics*, (D.J. Luban eds.), New Jersey: Rowman & Allanheld.

- Milgrom, P. (1981), "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 12, 380-391.
- Milgrom, P., and Roberts, J. (1986), "Relying on the Information of Interested Parties," *Rand Journal Economics*, 17, 18-32.
- Myerson, R. (1981), "Optimal Auction Design," *Mathematics of Operations Research*, 6, 58-73.
- Posner, R.A. (1999), "An Economic Approach to the Law of Evidence," *Stanford Law Review*, 51, 1477-1546.
- Riley, J., and Samuelson, W. (1981), "Optimal Auctions," *American Economic Review*, 71, 381-392.
- Rhode, D. (1985), "Ethical Perspectives on Legal Practice," *Stanford Law Review*, 37, 589-614.
- Sanchirico, C.W. (1997), "A Burden of Proof in Civil Litigation: A Simple Model of Mechanism Design" *International Review of Law and Economics*, 17, 431-447.
- Sanchirico, C.W. (1998), "Enforcement by Hearing: An Integrated Model of Evidence Production," *USC Law School Olin Working Paper* No. 98-19.
- Sanchirico, C.W. (2001), "Relying on the Information of Interested — and Potentially — Dishonest Parties" *American Law and Economics Review*, 3, 320-357.
- Shin, H.S. (1994), "The Burden of Proof in a Game of Persuasion," *Journal of Economic Theory*, 64, 253-264.
- Shin, H.S. (1998), "Adversarial and Inquisitorial Procedure in Arbitration," *Rand Journal of Economics*, 29, 378-405.
- Sobel, J. (1998), "Disclosure of Evidence and Resolution of Disputes: Who Should Bear the Burden of Proof?" in *Game Theoretical Models of Bargaining*, A. Roth (ed.) Cambridge: Cambridge University Press.
- Spurrer, R.L. (1980), *Inexpensive justice: Self-representation in the small claims court*, Kennikat Press.
- Thompson, W. (1989), "Are Juries Competent to Evaluate Statistical Evidence?" *Law and Contemporary Problems*, 52, 9-41.
- Tribe, L.H. (1971), "Trial by Mathematics: Precision and Ritual in the Legal Process," *Harvard Law Review*, 84, 1329-1393.
- Wells, G. (1992), "Naked Statistical Evidence of Liability: Is Subjective Probability Enough?" *Journal of Personality and Social Psychology*, 62, 739-752.