



The Priest-Klein hypotheses: Proofs and generality



Yoon-Ho Alex Lee*, Daniel Klerman

USC Gould School of Law, United States

ARTICLE INFO

Article history:

Received 18 June 2015

Received in revised form 22 March 2016

Accepted 16 June 2016

Available online 1 July 2016

JEL classification:

D81

K

K4

K41

ABSTRACT

Priest and Klein's 1984 article, "The Selection of Disputes for Litigation," famously hypothesized a "tendency toward 50 percent plaintiff victories" among litigated cases. Despite the article's enduring influence, its results have never been formally proved, and doubts remain about their meaning, validity, and generality. This article makes two main contributions. First, it distinguishes six hypotheses plausibly attributable to Priest and Klein. Second, it mathematically proves or disproves the hypotheses under a generalized version of Priest and Klein's model. The Fifty-Percent Limit Hypothesis and three other hypotheses attributable to Priest and Klein (1984) are mathematically well-founded and true under the assumptions made by Priest and Klein. In fact, they are true under a wider array of assumptions. More specifically, the Trial Selection Hypothesis, Fifty-Percent Limit Hypothesis, Asymmetric Stakes Hypothesis, and Irrelevance of Dispute Distribution Hypothesis are true for any distribution of disputes that is bounded, strictly positive, and continuous. The Fifty-Percent Bias Hypothesis is true when the parties are very accurate in estimating case outcomes, but only sometimes true when they are less accurate. As shown in Klerman and Lee (2014), the No Inferences Hypothesis is false.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Priest and Klein's 1984 article, "The Selection of Disputes for Litigation," famously hypothesized that there will be a "tendency toward 50 percent plaintiff victories" among litigated cases (p. 20). Their article has been one of the most influential legal publications, and its influence is growing as empirical work on law has become more common. Compare Shapiro and Pearce (2012) to Shapiro (1996). Even with the introduction of asymmetric information models of settlement, Priest and Klein's article continues to be cited by sophisticated scholars and in respected peer-reviewed journals (see Prescott and Spier, 2016; Hubbard, 2013; Gelbach, 2012; Atkinson et al., 2009; Yildiz, 2004; Yildiz, 2003; Bernardo et al., 2000; Waldfogel, 1995; Siegelman and Donohue, 1995). In addition, Waldfogel (1998) found greater empirical support for the Priest-Klein model than for asymmetric information models. But see Daughety and Reinganum (2012, pp. 439–440).

Nevertheless, despite the passage of more than thirty years since the publication of Priest and Klein's original article, their results have never been rigorously proved, and doubts remain about the assumptions needed to sustain their conclusions. In their article, Priest and Klein supported their main claims with simulations using normal distributions and an informal, graphical argument. As a result, the precise statement and scope of their claims have never been entirely clear. Others have discussed Priest and Klein's claims in more mathematically grounded terms. For example, Waldfogel (1995) formalized Priest and Klein's model, following carefully their original set-up and notation. Shavell (1996, p. 499, n. 20) set out "two key steps in the proof," including one "not supplied by Priest and Klein." But neither Waldfogel nor Shavell supplied proofs. Part of the challenge is that the plaintiff trial win rate involves triple integrals over a region of integration that is defined in terms of functions without closed forms (Waldfogel, 1995, p. 237). Although Hylton and Lin (2012) also formalize and prove some of Priest and Klein's claims, they do so using a model substantially different from, and in many ways less general than, Priest and Klein's.¹ This article provides the first set of rigorous proofs of Priest and Klein's claims, while remaining faithful to Priest and Klein's original set-up.

* Corresponding author.

E-mail addresses: alee@law.usc.edu (Y.-H.A. Lee), dklerman@law.usc.edu (D. Klerman).

¹ Hylton and Lin's model differs from Priest and Klein's in that Priest and Klein assume that case merit is measured by a real number, $y \in (-\infty, \infty)$, where the plaintiff prevails if $y > y^*$, and where y^* represents the decision standard or legal rule. In contrast, Hylton and Lin assume that case merit is a probability, and therefore, their model is more similar to Wittman (1985).

Understanding selection is fundamental to empirical analysis of decided cases. Because most cases settle, decided cases are a non-random subset of all cases. Hence, it is essential that scholars understand the processes by which cases settle or go to trial. While there has been some research on selection when one party is perfectly informed and the other party is completely uninformed, [Shavell \(1996\)](#), [Klerman and Lee \(2014\)](#), there has been little investigation of the selection implication of two-sided asymmetric information models other than Priest and Klein's original paper and [Friedman and Wittman \(2007\)](#). Further investigation of such models is warranted by the fact that [Waldfogel \(1998\)](#) found greater empirical support for the Priest-Klein model than for one-sided asymmetric information models. But see [Daughety and Reinganum \(2012, pp. 439–440\)](#). Many lawyers and scholars find Priest and Klein's main result—that close cases are more likely to be litigated—more plausible than the selection implications of one-sided asymmetric information models, which suggest that extreme cases are more likely to be litigated ([Klerman and Lee, 2014](#)). For these reasons, it is valuable to see if Priest and Klein's results have mathematical foundation.

Before setting out the formal analysis, however, it is helpful to distinguish six different hypotheses, all of which are plausibly attributable to [Priest and Klein \(1984\)](#):

The Trial Selection Hypothesis. “[D]isputes selected for litigation (as opposed to settlement) will constitute neither a random nor a representative sample of the set of all disputes” (p. 4). This proposition is probably the most important contribution of Priest and Klein's article.

The Fifty-Percent Limit Hypothesis. “[A]s the parties' error diminishes and the litigation rates declines, the proportion of plaintiff victories will approach 50 percent” (p. 19). “In the limit, the proportion of victories will approach 50 percent exactly” (p. 18). This hypothesis is often called the Priest-Klein hypothesis.

The Fifty-Percent Bias Hypothesis. Regardless of the legal standard, the plaintiff trial win rate will exhibit “a strong bias toward... fifty percent” as compared to the percentage of cases plaintiff would have won if all cases went to trial (pp. 5 and 23). This is plausibly a statement away from the limit: that is, the plaintiff trial win rate will be *closer* to fifty percent than the plaintiff win rate that would be observed if all cases went to trial.

The Asymmetric Stakes Hypothesis. If the defendant would lose more from an adverse judgment than the plaintiff would gain, then the plaintiff will win less than fifty percent of the litigated cases. Conversely, if the plaintiff has more to gain, then the plaintiff will win more than Fifty-Percent (see pp. 24–26). This hypothesis is most plausibly, like the Fifty-Percent Limit Hypothesis, a statement about the limit percentage of plaintiff victories as the parties become increasingly accurate in predicting trial outcomes.

The Irrelevance of the Dispute Distribution Hypothesis. The plaintiff trial win rate will be “unrelated... to the shape of the distribution of disputes” (pp. 19 and 22). This hypothesis is about the plaintiff trial win rate in the limit as the parties become increasingly accurate in predicting trial outcomes. This hypothesis is closely related to the Fifty-Percent Limit Hypothesis, but also more fundamental and more general, because it also applies when the stakes are unequal.

The No Inferences Hypothesis. Because selection effects are so strong, no inferences can be made about the law or legal decisionmakers from the plaintiff trial win rate. Rather, “the proportion of observed plaintiff victories will tend to remain constant over time regardless of changes in the underlying standards applied” (p. 31).

This paper explores the mathematical validity of each of these hypotheses, except the No Inferences Hypothesis. [Klerman and Lee \(2014\)](#) showed that the No Inferences Hypothesis is generally false under Priest and Klein's original model as well as under the canonical asymmetric information models. Instead, under all standard litigation models and a plausible set of assumptions, information about the legal standard and decisionmakers may be inferred from the plaintiff trial win rate. Because the No Inferences Hypothesis is analyzed extensively in [Klerman and Lee \(2014\)](#), we do not discuss it further in this article.²

We also note that Priest and Klein's model shares many features with the literature on global games (see [Morris and Shin, 2006](#)).³ As a result, the analysis in this article may open up additional avenues of research lying at the intersection of law and economics and the global games literature.

The rest of the article proceeds as follows. Section 2 discusses two different ways of interpreting Priest and Klein's model. We note that the model lends itself to either a non-common-priors interpretation or a common-prior interpretation, and relate this discussion to the game theory literature. Section 2 also discusses the absence of communication in the Priest-Klein model, a feature of the model that is often criticized by game theorists. Section 3 begins the formalization of Priest and Klein's model. This section assumes familiarity with [Priest and Klein \(1984\)](#) and is consistent with [Waldfogel \(1995\)](#). Although there have been other attempts to formalize Priest and Klein's model (see [Wittman 1985](#); [Hylton and Lin, 2012](#)), Waldfogel offers the formalization that is most faithful to the model in Priest and Klein's original article (see [Hylton and Lin, 2012, n. 5](#)). Nevertheless, because of the complexity of Priest and Klein's model, we begin by presenting a simple model in which disputes are assumed to be distributed uniformly across the real line. This simplification helps provide intuition. In Section 4, we allow for a more general distribution of disputes. Our second model is quite close to Priest and Klein's original model. In this model, the plaintiff trial win rate takes into account the underlying distribution of disputes, which is no longer assumed to be improper uniform, but the parties continue to estimate the likelihood that the plaintiff will prevail at trial as if the underlying distribution were improper uniform. Section 4 also presents a third model, in which the parties' estimates take into account the underlying distribution of disputes, which may take any distribution that is continuous, strictly positive, and bounded above. Section 5 concludes.

Our main findings can be summarized as follows. The Trial Selection Hypothesis, the Fifty-Percent Limit Hypothesis and the Asymmetric Stakes Hypothesis are valid under all three models under a wide range of assumptions. The Irrelevance of Dispute Distribution Hypothesis

² [Klerman and Lee \(2014\)](#) discusses the sufficient conditions under which one can make inferences under what Section 3 *infra* calls Model 2. The validity of the No Inferences Hypothesis under what Section 3 *infra* calls Model 3 is discussed in [Lee and Klerman \(2015\)](#). That article shows that if the standard of deviation of the errors is sufficiently low, a pro-plaintiff shift in the law may result in a decrease in the plaintiff trial win rate under Model 3.

³ See *infra* Section 3.

is valid under the second and third models, and is not meaningful under the first model, because that model assumes a single distribution of disputes (improper uniform). The Fifty-Percent Bias Hypothesis is valid under the second model, the one most faithful to Priest and Klein's original model, but only for a more limited set of dispute distributions.

2. Priors and communication in the Priest-Klein model

Priest and Klein's model can be interpreted in two different ways: either as a two-sided asymmetric information model with common priors (also known as consistent priors) or as a model with non-common priors (also known as inconsistent priors). The difference between the two interpretations comes down to whether the parties have the same beliefs about the distribution of disputes. Under the common-prior interpretation, the parties have the same beliefs (priors) about the distribution of disputes, but each receives a private signal about the merits of the particular case in dispute. Under the non-common-priors interpretation, the parties have different beliefs (priors) about the distribution of disputes, those beliefs are common knowledge, and the parties do not receive any additional information (private or public) about the particular case they are litigating.

In the modern literature on settlement and litigation, the common-prior asymmetric-information approach is more common (see [Bebchuk, 1984](#); [Reinganum and Wilde 1986](#); [Schweizer, 1989](#)). In recent years, however, the non-common-priors approach has received increasing attention ([Spier and Prescott, 2016](#); [Watanabe, 2005](#); [Yildiz, 2004](#); see generally [Morris, 1995](#)). One benefit of the non-common-priors approach is that it “allows us to focus on differences in beliefs without getting drowned in the informational issues.” [Yildiz \(2004, p. 237\)](#). In addition, it provides justification for the divergent expectations models of suit and settlement that dominated the literature until the mid-1980s. [Posner \(1973\)](#); [Shavell \(1982\)](#). In fact, the Priest-Klein model is usually classified as a divergent expectations model, and several scholars have stated that it should be interpreted as a model with non-common priors ([Daughety and Reinganum, 2012, pp. 439–440](#); [Yildiz, 2004, pp. 223–224](#)).

For the purposes of presenting our results in Sections 3 and 4, we adopt the common-prior, two-sided asymmetric information interpretation of Priest and Klein's model. Nevertheless, because the non-common-priors approach is interesting and yields remarkably similar results, we discuss it at various points as well.⁴

Priest and Klein's model has been criticized for not allowing the parties to communicate with each other. Such communication would presumably cause the parties to revise their assessments of the case and alter their settlement behavior ([Hay and Spier, 1998](#)). Before considering this criticism, we note that this “no communication” feature is not unique to Priest and Klein's model or to divergent expectations models. Indeed, the same criticism applies to some asymmetric information models as well. For example, in the standard screening model, [Bebchuk \(1984\)](#), the uninformed party makes a take-it-or-leave-it offer, and the informed party does not communicate with the uninformed party except to accept or reject the uninformed party's offer. In [Bebchuk's](#) model, the offer by the uninformed party obviously conveys no information, and acceptance or rejection by the informed party only partially reveals the informed party's type and does not cause any meaningful revision of the uninformed party's assessment of the case, because acceptance ends the case and rejection leads to trial without further opportunity to settle. Similarly, the parties in [Friedman and Wittman's \(2007\)](#) two-sided asymmetric information model use the Chatterjee-Samuelson mechanism, under which the parties submit offers to a machine or neutral third-party, but do not communicate with each other or revise their case assessments for settlement purposes based on the other party's offer.⁵

The “no communication” criticism has two components. First, one might think that if the parties could communicate, the sharing of information might lead them to agree on the merits of the case and thus always settle. This criticism is grounded in [Aumann \(1976\)](#), which proved that if the parties have common priors and common knowledge of posteriors, the posteriors must be equal, and they cannot “agree to disagree.” Because Aumann's argument presupposes common priors, it does not apply to the non-common-priors interpretation of the Priest-Klein model. As applied to the common-prior interpretation, Aumann's argument requires that communication between the parties result in common knowledge of their posteriors. But whether one can assume the existence of an incentive-compatible mechanism that would induce both parties to truthfully reveal their posteriors is unclear.

This consideration of incentive-compatible mechanisms leads to the second component of the “no communication” criticism.” Why doesn't the Priest-Klein model specify a mechanism, such as take-it-or-leave-it offers, under which the parties would, at least partially, reveal their information to each other? Although we do not presume to explain Priest and Klein's modeling choices, we follow this aspect of their set-up for several reasons. First, Priest and Klein assume that the parties always settle if the plaintiff's reservation price is lower than the defendant's reservation price. This is equivalent to assuming an *ex post* efficient bargaining.⁶ This assumption is justifiable as long as there exists a bargaining mechanism that guarantees *ex post* efficiency. Although mechanism design research has shown that, in many bargaining situations, achieving *ex post* efficiency is impossible without outside subsidies,⁷ (see [Myerson and Satterthwaite, 1983](#); [Spier, 1994](#)), this impossibility theorem does not apply to the common-prior interpretation of the Priest and Klein's model. Myerson and Satterthwaite's theorem assumes that parties' valuations are not correlated, while in Priest and Klein's model the plaintiff's and defendant's signals and valuations are correlated. Instead, [McAfee and Reny \(1992\)](#) show that, with correlated valuations, *ex post* efficiency is attainable without external subsidy (although a budget balancer who breaks even on average may be necessary) (see also [Gelbach, 2016](#)). Hence, according to [McAfee and Reny \(1992\)](#), it is possible to design a mechanism that implements Priest and Klein's model with *ex post* efficiency. Second, under the non-common-priors interpretation of the Priest-Klein model, the parties have complete information, and it is well known that *ex post* efficiency is attainable ([Rubinstein, 1982](#); [Spier and Prescott, 2016](#)). Third, under the common-prior interpretation, modeling particular mechanisms, such as take-it-or-leave-it offers or the Chatterjee-Samuelson mechanism, leads to rather complicated mathematics

⁴ See *infra* note 9 and Section 3.

⁵ Of course, there are litigation models in which at least one party learns from the other's offers (see, e.g. [Reinganum and Wilde, 1986](#); [Nalebuff, 1987](#); [Spier, 1992](#)).

⁶ By *ex post* efficient bargaining, we mean that the parties settle whenever it is mutually beneficial to do so given their subjective assessments of case merit. The parties would, however, still fail to settle if they were sufficiently optimistic relative to each other, in which case the plaintiff's reservation value would be higher than the defendant's. Complete efficiency would require that the parties always settle, because litigation is a negative-sum game. Complete efficiency would be attained if the parties could, as [Aumann \(1976\)](#) suggests, reach agreement in their assessments of case merit. In that situation, every case would settle.

⁷ As is typical, the Myerson-Satterthwaite model assumes individual rationality and incentive compatibility.

because there is two-sided asymmetric information (Daughety and Reinganum, 2012, p. 440). We think it worthwhile to work through the implications of such mechanisms and do so elsewhere. For example, Lee and Klerman (2016) show that when Priest and Klein's model is coupled with take-it-or-leave-it offers or the Chatterjee-Samuelson bargaining mechanism, the Fifty-Percent Limit Hypothesis sometimes remains valid. Nevertheless, given their complexity, we think it best to analyze the selection implications of bargaining mechanisms in a separate article. Fourth, because Priest and Klein's famous paper assumed *ex post* efficiency, we think it worthwhile to employ that assumption here so as to show which of the hypotheses in that paper are valid under a model similar in all important respects to their original.

3. A simple model with an improper uniform distribution of disputes

According to Priest and Klein's model, the merits of a case, which are determined by the case facts, are represented by a random variable Y , which takes on a real number. Y is distributed according to a probability density function, $g_Y(y)$. We shall call this the *distribution of disputes*.

The legal standard is given by $y^* \in \mathbf{R}$. If the realization of the case merit, y ,⁸ exceeds y^* , then the plaintiff prevails. If it falls below or is equal to y^* , then the defendant prevails. For example, in a negligence case, Y might be the efficient level of precaution expenditures minus defendant's actual precaution, in which case $y^* = 0$. It should be noted that the discontinuous, step-function relationship between case merit and the plaintiff's probability of prevailing is a critical feature that distinguishes Priest and Klein's model from canonical asymmetric information models. In most asymmetric information models, party type and thus case merit is represented as the probability that the plaintiff will prevail and can take on a wide range of values between zero and one. In contrast, in Priest and Klein's model, although case merit is represented as any real number, those real numbers map onto only two probabilities—zero percent or one hundred percent.

The plaintiff and the defendant do not observe y directly, but each party instead estimates y with some error. This error may be due to a number of factors, such as errors in interpreting the law, errors in interpreting his or her case facts, or incomplete information concerning either. These errors can be interpreted as resulting from noisy signals.⁹ The plaintiff receives a signal according to $Y_p = Y + \varepsilon_p$, where ε_p is distributed normally with mean zero and standard deviation, σ . The defendant likewise receives signal $Y_d = Y + \varepsilon_d$, where ε_d is distributed normally with mean zero and standard deviation, σ . Priest and Klein assume that ε_p , ε_d , and Y are independently distributed.¹⁰

Although Priest and Klein predicted that their Fifty-Percent predictions would hold for a wide range of dispute distributions, for simulation purposes, they used a standard normal distribution for $g_Y(y)$. For the purposes of our simplified exposition in this section, we assume instead that Y is distributed uniformly over the entire \mathbf{R} . In other words, $g_Y(y) = 1$ for all $y \in \mathbf{R}$. We will call this Model 1.

Although the assumption of uniform distribution over the entire \mathbf{R} contradicts the standard view that densities must integrate to one, its mathematical properties have been worked out by Hartigan (1983) and DeGroot (2004), and it is also used in the global games literature (see Morris and Shin, 2006). Indeed, the global games literature shares a number of features with Priest and Klein's model. Under both models, parties receive noisy signals about the state of the world and decide on their actions based on the inferences they make without updating or communication of signals. Morris and Shin show that as the standard deviation of the noise approaches zero, the parties behave as if the underlying distribution is uniform. Morris and Shin thus motivate this “improper prior” by arguing that it can be seen as a limiting case *either* as the prior distribution becomes diffuse *or* as the standard deviation of the noise becomes small.¹¹ The prior distribution would become more diffuse, if, for example, g_Y is in the normal family and one considers the limiting situation when the standard deviation becomes arbitrarily large. The equivalence between results under the improper uniform prior and results under a general distribution as the standard deviation of the noise becomes small is proven in Section 4.

Note that given the improper and uniform prior, the plaintiff's posterior is that cases are distributed normally with mean, y_p , and standard deviation, σ . The defendant's posterior likewise has mean, y_d , and standard deviation, σ (see DeGroot, 2004, p. 191).

Next, given the plaintiff's posterior, the plaintiff computes the (subjective) conditional probability, P_p , that he will prevail. In other words, the plaintiff estimates the probability that case merit lies above y^* , given his realized signal, y_p . Thus, $P_p \equiv \Pr(Y > y^* | Y_p = y_p)$. At this point, let $V = \frac{Y - y_p}{\sigma}$, the random variable representing the normalized difference between Y and the signal y_p . Since the plaintiff prevails if Y is greater than y^* , the plaintiff's conditional probability of prevailing at trial is the probability that $V > \frac{y^* - y_p}{\sigma}$, or equivalently, the probability that $V < \frac{y_p - y^*}{\sigma}$. Thus, we can write the plaintiff's conditional probability as $P_p = \Phi\left(\frac{y_p - y^*}{\sigma}\right)$, where Φ is the cumulative distribution function of the standard normal distribution. Note that as y_p goes up, the plaintiff believes he is more likely to prevail. When $y_p = y^*$, the plaintiff

⁸ Priest and Klein (1984) denote realized case merit Y' , but, we use y to be more consistent with modern notation. Similarly, Priest and Klein denote the plaintiff's and defendant's point estimates of realized case merit, \hat{Y}_p and \hat{Y}_d , but we denote them as y_p and y_d . In order to harmonize the Priest-Klein with more modern asymmetric approaches, we treat y_p and y_d as signals rather than estimates.

⁹ As noted in the Introduction, the interpretation in the text assumes common priors and asymmetric information. Another interpretation would be that the parties' different estimates reflect different priors. Under this interpretation, the parties do not know the distribution of disputes, $g_Y(y)$. Instead, the plaintiff's prior is that the disputes are normally distributed with mean y_p and standard deviation σ , and the defendant's prior is that they are normally distributed with mean y_d and standard deviation σ , and each party knows the other's prior. For a model of this type, see Spier and Prescott (2016). With one exception, the math and results are the same under the non-common-priors approach as under the common-prior approach taken in the text. The exception relates to the calculation of conditional probabilities when the true distribution of disputes is not uniform (see Section 3). Some of the terminology used in the text and propositions would need to be altered to reflect the non-common-priors approach. For example, one would need to refer to the accuracy of the parties' priors rather than the accuracy of their signals.

¹⁰ Although we maintain these assumptions, our proof generalizes to non-normal distributions of errors and where ε_p and ε_d are not independent from one another (see Proof of Proposition 3 in Appendix A).

¹¹ Morris and Shin's finding in the global games literature that a general prior behaves as if it were an improper uniform distribution is similar to Priest and Klein's intuition and to our main finding. Nevertheless, Morris and Shin's results do not translate directly to Priest and Klein's model, because Morris and Shin have, in their key integrand, a probability density function, which is by definition integrable. This means they can immediately apply the Dominated Convergence Theorem to take the limit under the integral. By contrast, our “litigation probability function” is not a probability density function, so its integrability must be established (see Proof of Proposition 3 in Appendix A).

believes he has a fifty-percent chance of prevailing. The defendant's conditional probability, his estimate of the probability that the plaintiff will prevail given his signal, is likewise given by $P_d \equiv \Pr(Y > y^* | Y_d = y_d) = \Phi\left(\frac{y_d - y^*}{\sigma}\right)$.

Priest and Klein assume that the parties go to trial¹² if (and only if) $P_p J - C_p + S_p > P_d J + C_d - S_d$, where $J > 0$ is the damages that the defendant pays the plaintiff if the case is litigated and the plaintiff prevails, C_p and C_d are litigation costs for the plaintiff and the defendant, respectively, and S_p and S_d are settlement costs for the plaintiff and the defendant, respectively.¹³ This condition for litigation makes sense, because settlement can only happen if both parties perceive the payoffs to settlement to be higher than the payoffs to litigation. The litigation condition can be rewritten as $P_p - P_d > \frac{C_p - S_p}{J}$, where $C = C_p + C_d$ and $S = S_p + S_d$. This last inequality is known as the Landes-Posner-Gould condition for litigation, after the three scholars who formulated it. We define $K \equiv \frac{C - S}{J}$ for further notational simplicity. Priest and Klein simulate their results with $K = 1/3$. We assume $0 < K \leq 1$. Priest and Klein assume that the plaintiff always has a credible threat to go to trial and thus can litigate or settle even when $P_p J < C_p$. We retain this assumption for the paper, but discuss two ways to address plaintiff threat credibility in [Appendix A](#).

As mentioned already, Priest and Klein are silent about how the parties bargain to arrive at a settlement. Technically, the Landes-Posner-Gould condition should be seen as a sufficient condition for litigation, rather than a necessary one: that is, litigation might happen even if there is a range of settlement amounts that would be in their perceived mutual interest, because parties, in bargaining strategically, might not be able to agree on the settlement amount. Modern mechanism design research has shown that strategic bargaining frequently produces *ex post* inefficiency, although introduction of a third-party can also help eliminate such inefficiency (see [Myerson and Satterthwaite, 1983](#); but see [McAfee and Reny, 1992](#)). Nevertheless, [Priest and Klein \(1984\)](#) and others using the divergent expectations model have assumed that the Landes-Posner-Gould condition is necessary as well as sufficient for litigation, and, for the reasons set out in [Section 2](#), we retain this assumption as well.

Let $\Pi_\sigma(y; y^*)$ denote the objective probability that a dispute of merit y goes to trial when the decision standard is y^* , and where the parties predict case merit with errors ε_p and ε_d that are distributed with mean zero and the common standard deviation σ . We call $\Pi_\sigma(y; y^*)$ the “litigation probability function.” Based on our analysis above, $\Pi_\sigma(y; y^*)$ can be written as the probability that

$$P_p - P_d = \Phi\left(\frac{y_p - y^*}{\sigma}\right) - \Phi\left(\frac{y_d - y^*}{\sigma}\right) > K.$$

This expression is equivalent to

$$P_p - P_d = \Phi\left(\frac{y + \varepsilon_p - y^*}{\sigma}\right) - \Phi\left(\frac{y + \varepsilon_d - y^*}{\sigma}\right) > K. \quad (1)$$

Thus,

$$\Pi_\sigma(y; y^*) = \iint_{R_\sigma(y; y^*)} \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^2 e^{-\frac{\varepsilon_p^2}{2\sigma^2}} e^{-\frac{\varepsilon_d^2}{2\sigma^2}} d\varepsilon_p d\varepsilon_d,$$

where

$$R_\sigma(y; y^*) = \left\{ (\varepsilon_p, \varepsilon_d) \in \mathbb{R}^2 \mid \Phi\left(\frac{y + \varepsilon_p - y^*}{\sigma}\right) - \Phi\left(\frac{y + \varepsilon_d - y^*}{\sigma}\right) > K \right\}$$

is the *litigation set*. Therefore, $\Pi_\sigma(y; y^*)$ can be expressed as a double integral over a region of integration that is defined by the inequality.¹⁴

[Fig. 1](#) plots the litigation probability function, $\Pi_\sigma(y; y^*)$, for large and small for σ , when $y^* = 1$. Note that because $\Pi_\sigma(y; y^*)$ represents the probability of litigation at each y , the graph is bounded below by 0 and above by 1. Importantly, $\Pi_\sigma(y; y^*)$ is not a probability density function, because it does not always integrate to one.

These graphs provide the three key intuitions for all of the results in this article. First, note that litigation is most likely when the dispute is close. That is, the closer a dispute, y , is to the decision standard, y^* , the greater the probability that the dispute will go to trial. This result follows from the normal distribution of the parties' errors. When the true value of the dispute, y , is far from y^* , the parties' signals are also likely to be far from y^* . Although the parties' signals will differ from each other, since each party's estimate of the plaintiff's probability of prevailing is the area under the portion of a normal distribution centered at their signal that exceeds y^* , the difference in the parties' estimates will be small, because it will represent the difference in the (left or right) tails of the distributions. Conversely, when the true value of the disputes, y , is close to y^* , the parties' signals are also likely to be close to y^* , and the difference in the parties' estimates will be large, because it will represent differences in the broad centers of the distributions. Priest and Klein illustrate this effect in [Fig. 6](#) of their article.

Second, note that $\Pi_\sigma(y; y^*)$ is symmetric around y^* . This also follows from the symmetry of the normal distribution of the parties' errors, although the precise reasoning is best seen in the Proof of Proposition 1. The symmetry of $\Pi_\sigma(y; y^*)$ around y^* is important, because it means that, all other things being equal, litigation is equally likely when the plaintiff will prevail (i.e., $y > y^*$) and when the defendant will prevail (i.e., $y < y^*$). This provides the key intuition for the idea that plaintiff is likely to win fifty percent of the litigated cases (and lose the other fifty percent).

¹² Priest and Klein and much of the later literature assume that “litigate” and “go to trial” are synonymous, because they assume that all cases either settle or go to trial. More recent work explores the fact that many cases are resolved by motions to dismiss or summary judgment. [Gelbach \(2012\)](#); [Hubbard \(2013\)](#). Cases resolved by such motions are litigated, but did not go to trial. This article, however, retains the simplifying assumption that all litigated cases go to trial. The term “disputes” or “all disputes” means both cases that settle and cases that are litigated.

¹³ We are assuming symmetric stakes here. Asymmetric stakes are discussed below.

¹⁴ The inequality defining the region of integration, $R_\sigma(y; y^*)$, can also be written in an explicit form as $R_\sigma(y; y^*) = \left\{ (\varepsilon_p, \varepsilon_d) \in \mathbb{R}^2 \mid \sigma\Phi^{-1}\left(K + \Phi\left(\frac{y + \varepsilon_d - y^*}{\sigma}\right)\right) - (y - y^*) < \varepsilon_p \right\}$. Nevertheless, we retain the implicit form because it shortens the proofs.

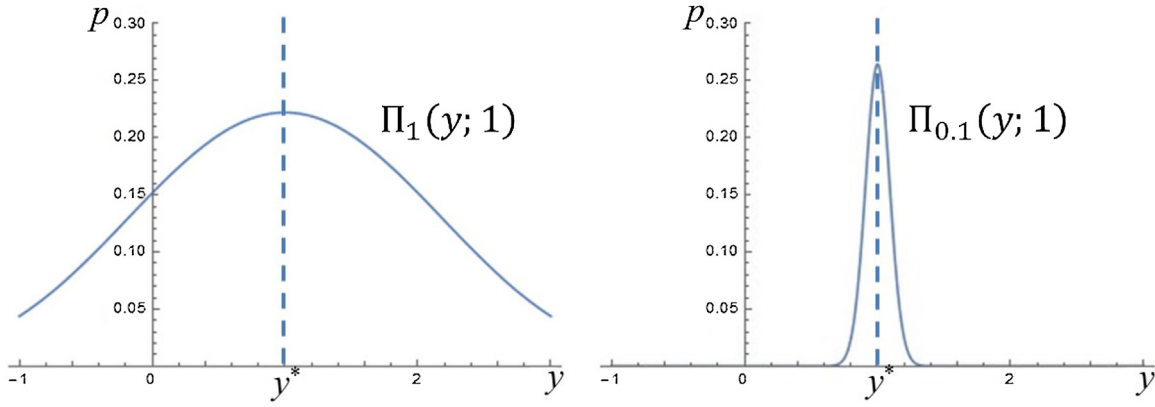


Fig. 1. Litigation Probability Functions ($\sigma = 1$ (Left), $\sigma = 0.1$ (Right); $y^* = 1$; $K = \frac{1}{3}$).

Third, and finally, note that as the parties become more accurate in their estimates (as σ becomes smaller), the distribution of $\Pi_\sigma(y; y^*)$ becomes very tight around y^* . That is, only disputes very close to the decision standard have any significant probability of being litigated. This will be important for the next section, where the assumption of a uniform distribution of disputes is dropped.

Using the probability of litigation, $\Pi_\sigma(y; y^*)$, we can calculate the plaintiff trial win rate:

$$W_\sigma(y^*) = \frac{\int_{y^*}^{\infty} \Pi_\sigma(y; y^*) dy}{\int_{-\infty}^{\infty} \Pi_\sigma(y; y^*) dy}. \quad (2)$$

The numerator is the fraction of litigated cases that are decided in favor of the plaintiff, and the denominator is the fraction of all cases that are litigated.¹⁵ Since $\Pi_\sigma(y; y^*)$ is a double integral, the plaintiff trial win rate is a fraction in which the numerator and denominator are both triple integrals.

Then, given the symmetry of $\Pi_\sigma(y; y^*)$ around y^* we have the following result.

Proposition 1. (The Fifty-Percent Hypothesis under the Improper Uniform Distribution of Disputes). *Regardless of the accuracy of the parties' signals, if the stakes are symmetric, the plaintiff always wins fifty percent of all litigated cases. That is, for all $\sigma > 0$, $W_\sigma(y^*) = 1/2$.*

Proof. We begin with a change of variables that normalizes case merit and the parties' signals. This will turn out to be useful for all the other results as well. Let $u = \frac{y + \varepsilon_p - y^*}{\sigma}$, $v = \frac{y + \varepsilon_d - y^*}{\sigma}$, and $z = \frac{y - y^*}{\sigma}$. Equivalently, $y = \sigma z + y^*$. Rewriting the litigation probability function in terms of the new variables, we have

$$\Pi_\sigma(y; y^*) = \Pi_\sigma(\sigma z + y^*; y^*) = \frac{1}{2\pi} \int \int_{R(u,v)} e^{-\frac{(u-z)^2}{2}} e^{-\frac{(v-z)^2}{2}} du dv$$

where $R(u, v) \equiv R_\sigma(y; y^*) = \{(u, v) | \Phi(u) - \Phi(v) > K\}$. Note that prior to normalization, $\Pi_\sigma(y; y^*)$ was a double integral of a fixed bivariate distribution over a region of integration in the $\varepsilon_p \varepsilon_d$ -plane that depended on two parameters: y and σ . After the variable changes,

$$\Pi_\sigma(y; y^*) = \Pi_\sigma(\sigma z + y^*; y^*) = \frac{1}{2\pi} \int \int_{R(u,v)} e^{-\frac{(u-z)^2}{2}} e^{-\frac{(v-z)^2}{2}} du dv$$

is a double integral of a bivariate distribution over a region of integration in the uv -plane that depends on only one parameter: z . As a result, the value of $\Pi_\sigma(\sigma z + y^*; y^*)$ is independent with respect to the choice of σ and y^* . In other words, $\Pi_\sigma(\sigma z + y^*; y^*) = \Pi_1(z; 0)$ for all $\sigma > 0$ and $y^* \in \mathbf{R}$. Thus, we can define $\Pi(z) \equiv \Pi_1(z; 0) = \Pi_\sigma(\sigma z + y^*; y^*)$ for all $\sigma > 0$ and y^* . For each z , $\Pi(z)$ is an integral over $R(u, v)$ of a bivariate normal distribution centered at $(u, v) = (z, z)$.

The litigation probability function lets us compute the plaintiff trial win rate in terms of z . Under Eq. (2), we have

$$W_\sigma(y^*) = \frac{\int_0^\infty \Pi(z) dz}{\int_{-\infty}^\infty \Pi(z) dz}. \quad (3)$$

At this point, as long as $\Pi(z)$ is integrable (a result we show in the proof of Proposition 3), the fifty-percent plaintiff trial win rates follows from the symmetry of $\Pi(z)$ around $z = 0$. This is true because symmetry of $\Pi(z)$ would imply that $\int_0^\infty \Pi(z) dz = \int_{-\infty}^0 \Pi(z) dz$, so that the denominator would equal $2 \int_0^\infty \Pi(z) dz$. To prove the symmetry of $\Pi(z)$ around $z = 0$, we need only show that for all $z > 0$, $\Pi(z) = \Pi(-z)$. This is easy. Since Φ is a cumulative distribution function of the standard normal distribution (which is itself symmetric

¹⁵ Note that the assumption of an improper uniform distribution raises the possibility that neither the numerator nor the denominator may converge, since $\Pi_\sigma(y; y^*)$ is not a probability density function. It is a function that takes on a value between 0 and 1 for each y . Nevertheless, this function does integrate to a finite value—a result that turns out to be necessary for the rest of the analysis (see Proof of Proposition 3 in Appendix A).

around 0), we have $\Phi(x) = 1 - \Phi(-x)$ for all $x \in \mathbf{R}$. Therefore, $\Phi(u) - \Phi(v) = (1 - \Phi(-u)) - (1 - \Phi(-v)) = \Phi(-v) - \Phi(-u)$. This means that $R(u, v) = R(-v, -u)$. Therefore,

$$\begin{aligned}\Pi(z) &= \frac{1}{2\pi} \iint_{R(u,v)} e^{-\frac{(u-z)^2}{2}} e^{-\frac{(v-z)^2}{2}} dudv = \frac{1}{2\pi} \iint_{R(-v,-u)} e^{-\frac{(u-z)^2}{2}} e^{-\frac{(v-z)^2}{2}} dudv = \frac{1}{2\pi} \iint_{R(u,v)} e^{-\frac{(-v-z)^2}{2}} e^{-\frac{(-u-z)^2}{2}} (-dv)(-du) \\ &= \frac{1}{2\pi} \iint_{R(u,v)} e^{-\frac{(u+z)^2}{2}} e^{-\frac{(v+z)^2}{2}} dudv = \frac{1}{2\pi} \iint_{R(u,v)} e^{-\frac{(u-(-z))^2}{2}} e^{-\frac{(v-(-z))^2}{2}} dudv = \Pi(-z). \square\end{aligned}$$

Before moving on, we make three observations. First, [Proposition 1](#) is a stronger result than the Fifty-Percent Limit Hypothesis in that the plaintiff trial win rate is fifty-percent for all $\sigma > 0$, not just in the limit. This is due to our assumption of the improper uniform distribution of disputes, which is symmetric around y^* (or around any y , for that matter). We show in the next section that if the improper uniform distribution is replaced with a “well-behaved” (continuous, bounded, strictly positive) distribution of disputes, which may not be symmetric around y^* , the result will hold only in the limit as σ approaches zero. Second, the Fifty-Percent Bias Hypothesis is not relevant to the model with the improper uniform distribution of disputes, because the probability that the plaintiff will prevail if all disputes were litigated is a fraction for which the numerator and the denominator are both infinite—the integral of a uniform distribution over the real line or a portion of the real line truncated at the left by y^* . Third, the Irrelevance of the Distribution Dispute Hypothesis is inapplicable in the case of the improper uniform distribution, because this model assumes only one distribution of disputes.

Now we move on to discuss asymmetric stakes. Priest and Klein allow the parties to have asymmetric stakes and suggest there will be a deviation from fifty-percent plaintiff trial victories in such cases. For example, the defendant might have more at stake in cases involving product liability, where an adverse judgment would damage the defendant’s reputation or could be used against it in other cases. Conversely, but less commonly, the plaintiff would have more at stake in cases alleging patent infringement, where a judgment invalidating the patent would bar suits against other alleged infringers. As [Waldfoegel \(1995\)](#) points out, asymmetric stakes can be formalized by assuming that the plaintiff would win αJ if it prevailed and the defendant would lose J if the plaintiff won. If $\alpha < 1$, this indicates that the defendant has more at stake in the litigation than the plaintiff, and vice versa. The trial condition now becomes $\alpha P_p - P_d > K$.¹⁶ Note, however, that when $\alpha \leq K$, no disputes will go to trial because the trial condition is never satisfied given that $P_p \leq 1$. Therefore, we will always assume $\alpha > K$ for the remainder of our analysis.

We can then follow the same approach as when the stakes were symmetric. First, let $\Pi_{\alpha,\sigma}(y; y^*)$ denote the probability that a dispute of merit y goes to trial when the stakes are asymmetric. By analogous reasoning, we obtain

$$\Pi_{\alpha,\sigma}(y; y^*) = \iint_{R_{\alpha,\sigma}(y; y^*)} \left(\frac{1}{\sigma\sqrt{2\pi}} \right)^2 e^{-\frac{\varepsilon_p^2}{2\sigma^2}} e^{-\frac{\varepsilon_d^2}{2\sigma^2}} d\varepsilon_p d\varepsilon_d,$$

where

$$R_{\alpha,\sigma}(y; y^*) = \left\{ (\varepsilon_p, \varepsilon_d) \mid \alpha \Phi \left(\frac{y + \varepsilon_p - y^*}{\sigma} \right) - \Phi \left(\frac{y + \varepsilon_d - y^*}{\sigma} \right) > K \right\}$$

is the litigation set. The plaintiff trial win rate becomes

$$W_{\alpha,\sigma}(y^*) = \frac{\int_{y^*}^{\infty} \Pi_{\alpha,\sigma}(y; y^*) dy}{\int_{-\infty}^{\infty} \Pi_{\alpha,\sigma}(y; y^*) dy}. \quad (4)$$

Importantly, when $\alpha \neq 1$, the litigation probability function is no longer symmetric around y^* . Instead, if the plaintiff has less at stake than the defendant, the litigation probability will peak before y^* ; and if the plaintiff has more at stake than the defendant, the litigation probability function will peak after y^* . [Fig. 2](#) depicts litigation probability functions for $\alpha = 0.5$ and $\alpha = 1.5$.

Under this set-up, we have the following result.

Proposition 2. (The Asymmetric Stakes Hypothesis under the Improper Uniform Distribution of Disputes). *Regardless of the accuracy of the parties’ signals, the party with more at stake wins more of the litigated cases. That is, for all $\sigma > 0$, $W_{\alpha,\sigma}(y^*) > 1/2$ for $\alpha > 1$ and $W_{\alpha,\sigma}(y^*) < 1/2$ for $\alpha < 1$.*

We include the proof in [Appendix A](#), but the main strategy is to show that when $\alpha \neq 1$, the graphs of $\Pi_{\alpha,\sigma}(y; y^*)$ are indeed skewed as in [Fig. 2](#). Consequently, when the defendant has more at stake, more litigated disputes are to the left of y^* (the left graph), and so the defendant wins a greater fraction of litigated cases. Conversely, when the plaintiff has more at stake, more litigated cases are to the right of y^* (the right graph), and so the plaintiff wins a greater fraction of litigated cases.

[Fig. 3](#) illustrates the plaintiff trial win rate as α varies. The result is not strictly monotonic. Instead, simulation shows a slight dip in the beginning near K . In addition, as discussed further in [Appendix A](#), there is a discontinuity at $\alpha = 1 + K$, where the plaintiff trial win rate jumps to one.

¹⁶ Note that under Waldfoegel’s formalization, K does not vary with α . We retain this assumption for our analyses. Allowing K to vary with α will not affect any of the outcomes of [Proposition 2](#) or [Proposition 5](#), which are all statements about a given α .

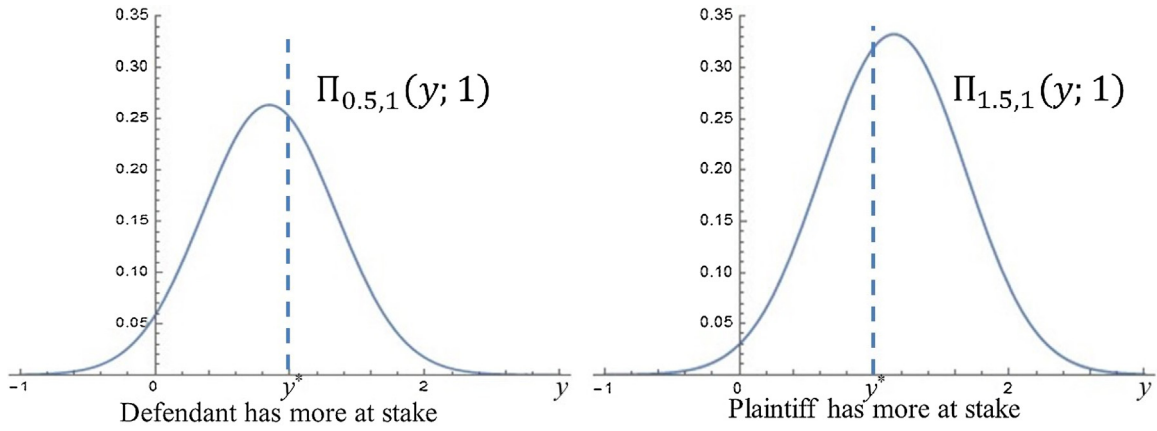


Fig. 2. Litigation Probability Functions with Asymmetric Stakes ($\alpha = 0.5$ (Left), $\alpha = 1.5$ (Right); $\sigma = 1$; $y^* = 1$; $K = \frac{1}{3}$).

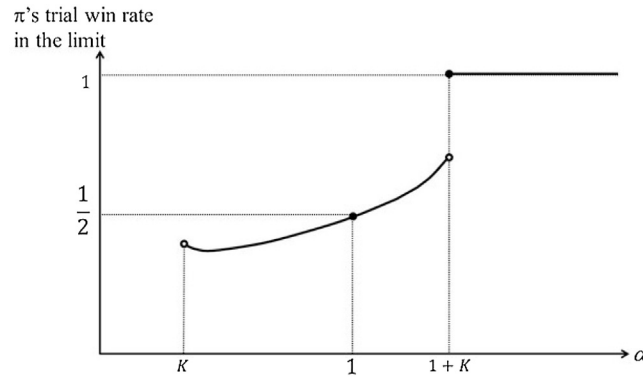


Fig. 3. The Plaintiff Trial Win Rate (in the Limit) for Asymmetric Stakes ($K = \frac{1}{3}$).

4. The general case

In this section, we retain most of the set-up of the model as discussed in Section 3 but relax the assumption that Y is distributed uniformly across the entire real line. Instead, we assume g_Y is a well-behaved probability density function in the following sense.

Assumption A1. (Well-Behaved Dispute Distribution). $g_Y(y)$ is a probability density function that is strictly positive, continuous, and bounded above.¹⁷

A1 assures, for example, that $g_Y(y)$ has full support and that it does not contain any probably mass functions. Using a well-behaved g_Y rather than the improper uniform distribution affects the analysis in two ways. First, the distribution affects the plaintiff trial win rate. $W_{\alpha,\sigma}(y^*)$ will no longer be $\frac{\int_{y^*}^{\infty} \Pi_{\alpha,\sigma}(y; y^*) dy}{\int_{-\infty}^{\infty} \Pi_{\alpha,\sigma}(y; y^*) dy}$, but instead, the integrand must be the litigation probability function weighted and thus multiplied by the distribution of disputes. Hence, we have

$$W_{\alpha,\sigma}(y^*) = \frac{\int_{y^*}^{\infty} \Pi_{\alpha,\sigma}(y; y^*) g_Y(y) dy}{\int_{-\infty}^{\infty} \Pi_{\alpha,\sigma}(y; y^*) g_Y(y) dy} = \frac{\int_0^{\infty} \Pi_{\alpha}(z) g_Y(\sigma z + y^*) dz}{\int_{-\infty}^{\infty} \Pi_{\alpha}(z) g_Y(\sigma z + y^*) dz}. \quad (5)$$

Fig. 4 shows how $\Pi_{\alpha,\sigma}(y; y^*)$ and g_Y interact for different σ values when stakes are symmetric, $\alpha = 1$.¹⁸ The graphs on the left show the distribution of *all* disputes (the solid line) and the probability of litigation for each value of y (the dotted line). The graphs on the right show the distribution of *litigated* disputes, which is derived by multiplying the density of the dispute distribution times the probability of litigation. The top right panel shows the result when $\sigma = 0.5$. The bottom right panel shows the result when $\sigma = 0.2$. The plaintiff trial win rate, $W_{\alpha,\sigma}(y^*)$, is the area under the product graph (the right panel) for $y > y^*$ divided by the area under the entire product graph.

These graphs provide the key intuition for the general model. As the parties become more accurate in predicting outcomes (i.e., as σ gets smaller), the distribution of litigated disputes becomes more symmetric around the decision standard, y^* . This follows from the fact, discussed in the previous section, that the litigation probability function, $\Pi_{\alpha,\sigma}(y; y^*)$, is symmetric and becomes tighter and tighter as the parties become more accurate. Since the distribution of litigated disputes is the litigation probability function, $\Pi_{\alpha,\sigma}(y; y^*)$, times the distribution of all disputes, $g_Y(y)$, as the litigation probability function becomes tighter and tighter, the distribution of litigated disputes also

¹⁷ In fact, all of the propositions are true if the distribution is bounded everywhere, but strictly positive and continuous only at y^* . It can be zero or discontinuous elsewhere.

¹⁸ The graphs also assume Assumption A2a, discussed later. Assumptions A2b would not fundamentally change the graphs. Although the litigation probability would no longer be symmetric, it would become increasingly symmetric as σ approaches zero.

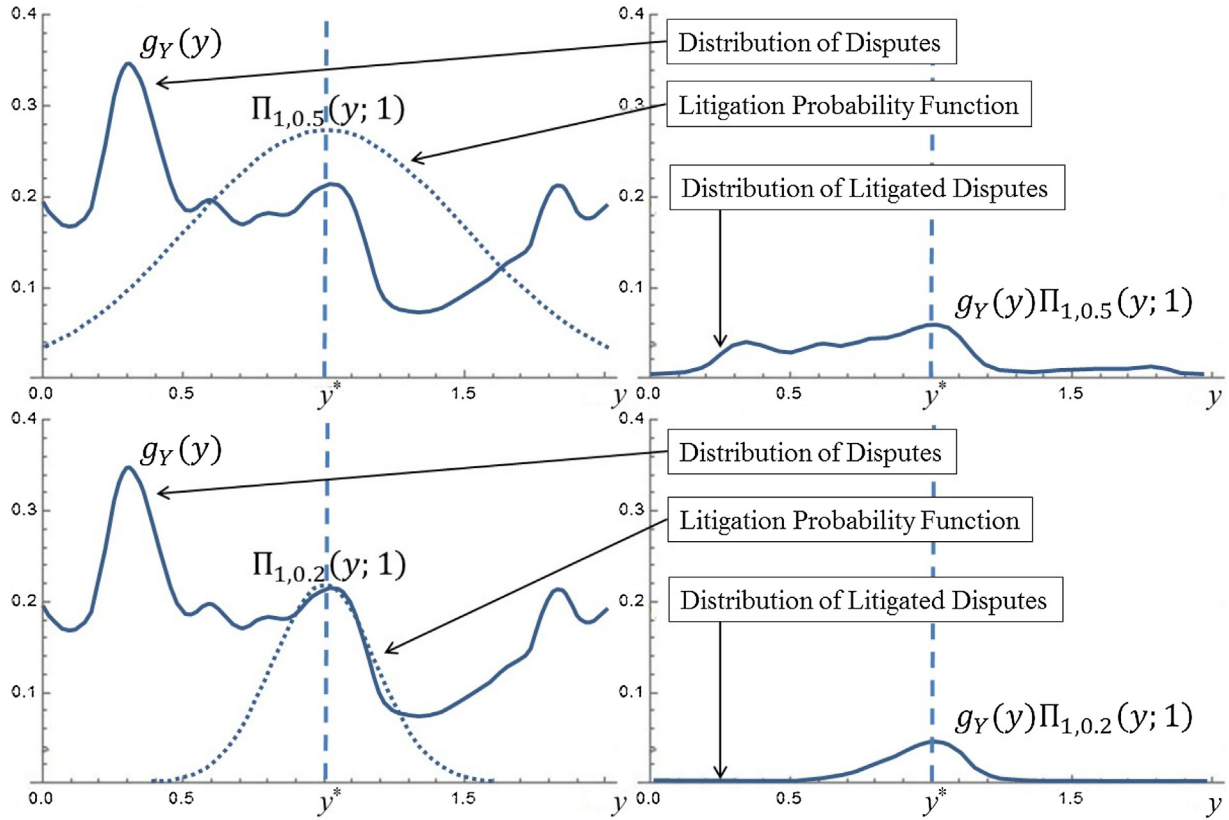


Fig. 4. Distribution of Disputes and Litigated Cases ($\sigma = 0.5$ (top), $\sigma = 0.2$ (Bottom); $\alpha = 1$; $y^* = 1$; $K = \frac{1}{3}$).

becomes tighter and more symmetric. The fact that only disputes close to the decision standard will be litigated means that the distribution of all disputes is nearly irrelevant. It simply does not matter how disputes far from the decision standard are distributed, because they will, for all practical purposes, never be litigated. As the parties' estimates become increasingly accurate (i.e., as σ goes to zero), more and more disputes become "far" from y^* , and the fraction of disputes that matter can become arbitrarily small. That is, all that matters are disputes that are arbitrarily close to y^* . As long as the distribution of all disputes is continuous at y^* , the difference in the density of disputes to the left and right of y^* can be made arbitrarily small. As a result, the density of the distribution of disputes—loosely speaking—will look flat near y^* . This is why models with a well-behaved distribution of disputes in this section produce results that are nearly identical to models with a uniform improper distribution, discussed in the previous section.

Note that the plaintiff trial win rate, $W_{\alpha,\sigma}(y^*)$, does initially depend on the shape of the distribution of disputes, g_Y . But as Fig. 4 shows, as σ becomes small, the graphs of the litigated cases become increasingly symmetric around y^* . Thus, it makes sense that this ratio converges to fifty percent as σ approaches zero.¹⁹ In Proposition 3, we show that the term $g_Y(\sigma z + y^*)$ disappears from $W_{\alpha,\sigma}(y^*)$ in the limit as σ approaches 0, and therefore the limit value of the plaintiff trial win rate will depend only on the shape of $\Pi_{\alpha,\sigma}(y; y^*)$. Put differently, in the limit, g_Y may be assumed to equal 1 for all y , which allows us to appeal to Proposition 1 and Proposition 2 for the main results.

There is a second way that g_Y can affect the analysis. It could affect how each party calculates the probability that the plaintiff will prevail at trial. Under the improper uniform distribution, the parties' estimates of the plaintiff's probability of prevailing were $P_p = \Phi\left(\frac{y_p - y^*}{\sigma}\right)$ and $P_d = \Phi\left(\frac{y_d - y^*}{\sigma}\right)$. We designate this assumption as A2a.

Assumption A2a. (Conditional Probability of Plaintiff Victory with Laplacian Beliefs).

$$P_p \equiv \Pr(Y > y^* | Y_p = y_p) = \Phi\left(\frac{y_p - y^*}{\sigma}\right) \text{ and } P_d \equiv \Pr(Y > y^* | Y_d = y_d) = \Phi\left(\frac{y_d - y^*}{\sigma}\right).$$

Assumption A2a is consistent with Priest and Klein (1984) and Waldfogel (1995). In addition, other divergent expectations models have made similar assumptions (see Wittman, 1985). Nevertheless, it is important to note that, under Assumption A2a, the parties construct their subjective probabilities as if the distribution of disputes were flat. This may be a reasonable assumption if g_Y is not known by the parties. After all, the fact that we incorporate the well-behaved distribution into the plaintiff win rate does not imply that the parties know it. Morris and Shin in their work on global games justify an assumption similar in spirit to Assumption A2a by calling this type of belief

¹⁹ Nonetheless, the fifty-percent result is still not obvious, because, as the graph becomes increasingly symmetric, both the area to the right of y^* and the entire area under the curve approach zero. In other words, as σ approaches zero, both the numerator and the denominator of $W_{\alpha,\sigma}(y^*)$ will approach zero. Meanwhile, because both the numerator and the denominator involve a triple integral over the region of integration defined by an inequality, L'Hopital's rule cannot readily be used to simplify the limit. For this reason, it is not clear that the limit will be exactly fifty percent when $\alpha = 1$, and an analysis of the behavior at the limit is warranted (see the Appendix A).

“Laplacian” (Morris and Shin, 2006, p. 58), following Laplace’s suggestion that one should apply a uniform distribution to unknown events from the “principle of insufficient reason.”²⁰

On the other hand, if they were fully informed about g_Y , rational parties would take into account that knowledge when calculating the probability that the plaintiff would prevail. For example, suppose the plaintiff receives a particular signal y_p , but also knows that according to g_Y very few cases lie above y_p but a great many cases lie below y_p . Then it would be rational for the plaintiff to assume that it is more likely that true case merit lies below y_p . For this reason, the plaintiff’s subjective probability that he will win at trial will not simply be $P_p = \Phi\left(\frac{y_p - y^*}{\sigma}\right)$, as it was when g_Y was uniform improper. Rather, possible values of Y will have to be weighted by the distribution of disputes, g_Y . Therefore,

$$P_p = \Pr(Y > y^* | Y_p = y_p, g_Y) = \frac{\int_{y^*}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y_p - y)^2}{2\sigma^2}} g_Y(y) dy}{\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y_p - y)^2}{2\sigma^2}} g_Y(y) dy} = \frac{\int_{y^*}^{\infty} e^{-\frac{(y_p - y)^2}{2\sigma^2}} g_Y(y) dy}{\int_{-\infty}^{\infty} e^{-\frac{(y_p - y)^2}{2\sigma^2}} g_Y(y) dy}. \quad (6)$$

The defendant will similarly take into account the distribution of disputes when calculating its estimate of the probability that the plaintiff prevails. We designate this alternative assumption as A2b.

Assumption A2b. (Conditional Probability of Plaintiff Victory with Knowledge of g_Y).

$$P_p \equiv \Pr(Y > y^* | Y_p = y_p, g_Y) = \frac{\int_{y^*}^{\infty} e^{-\frac{(y_p - y)^2}{2\sigma^2}} g_Y(y) dy}{\int_{-\infty}^{\infty} e^{-\frac{(y_p - y)^2}{2\sigma^2}} g_Y(y) dy} \text{ and } P_d \equiv \Pr(Y > y^* | Y_d = y_d, g_Y) = \frac{\int_{y^*}^{\infty} e^{-\frac{(y_d - y)^2}{2\sigma^2}} g_Y(y) dy}{\int_{-\infty}^{\infty} e^{-\frac{(y_d - y)^2}{2\sigma^2}} g_Y(y) dy}.$$

Because A2a and A2b are conflicting assumptions, they generate two different models. We present the results under both assumptions for four reasons. First, both are empirically plausible. If parties lack good information about the distribution of disputes, then A2a is more realistic. If parties have good knowledge about the distribution of disputes, then A2b is more realistic. Second, both are of theoretical interest. A2a is more faithful to Priest & Klein’s original model, but A2b is more consistent with other asymmetric information models of settlement in that parties calculate their subjective conditional probabilities using accurate information about the distribution of disputes, g_Y . Third, A2a best reflects the non-common-priors interpretation of the Priest-Klein model. If the parties’ different estimates reflect not different signals, but rather different assumptions (priors) about the distribution of disputes, where both assume that the distribution of disputes is normally distributed with standard deviation σ (but different means, y_p and y_d), then A2a is the correct statement of the conditional probabilities. Fourth, proving results under A2b requires first proving them under A2a.

In spite of the differences between Assumptions A2a and A2b, all results pertaining to limit values of the plaintiff win rate (as σ approaches zero) coincide. This is driven by the fact that, in the limit, the underlying distribution of disputes behaves as if it were flat. Indeed, note that the conditional probabilities in A2b reduce to those in A2a if we assume $g_Y(y) = 1$. Therefore, it is only away from the limit that the plaintiff trial win rates under the two models may diverge.

To simplify the propositions, we define two models:

Model 2 (Well-Behaved Dispute Distribution with Laplacian Beliefs). *Assumptions A1 and A2a.*

Model 3 (Well-Behaved Dispute Distribution with Knowledge of g_Y). *Assumptions A1 and A2b.*

With these assumptions, we have the following results.

Proposition 3. (The Irrelevance of the Dispute Distribution Hypothesis under Well-Behaved Dispute Distribution). *Under both Model 2 (well-behaved dispute distribution with Laplacian beliefs) and Model 3 (well-behaved dispute distribution with knowledge of g_Y), the limit value of the plaintiff trial win rate as the parties become increasingly accurate in estimating the case merit is independent of the distribution of disputes because the plaintiff trial win rate converges to the rate that would obtain under the improper uniform distribution of disputes. More specifically,*

$$\lim_{\sigma \rightarrow 0^+} W_{\alpha, \sigma}(y^*) = \begin{cases} \frac{\int_0^{\infty} \Pi_{\alpha}(z) dz}{\int_{-\infty}^{\infty} \Pi_{\alpha}(z) dz} & \text{for } K < \alpha < 1 + K \\ 1 & \text{for } \alpha \geq 1 + K \end{cases}.$$

Note that $W_{\alpha, \sigma}(y^*)$ is undefined at $\sigma = 0$. This is because when $\sigma = 0$, each party knows whether it will win or lose with one-hundred-percent certainty and no disputes will go to trial. Proposition 3 is therefore a statement about the limit value of a function at a point at which it is undefined.

Note that $W_{\alpha, \sigma}(y^*)$ is undefined at $\sigma = 0$ when $K < \alpha \leq 1 + K$ because no disputes will go to trial. In this case, both sides will know with one-hundred-percent certainty whether the plaintiff will win or lose. Therefore, $P_p = P_d = 0$ or 1, and the condition for litigation ($\alpha P_p - P_d > K$) can never be met. For these cases, Proposition 3 is a statement about the limit value of a function at a point at which it is undefined. By

²⁰ Hartigan (1983, p. 2) notes that “Laplace, following Bernoulli (1713) used the principle of insufficient reason which specifies that probabilities of two events will be equal if we have no reason to believe them different. An early user of this principle was Thomas Bayes (1763), who apologetically postulated that a binominal parameter p was uniformly distributed if nothing were known about it.” If one extends this idea that “the probabilities of two events will be equal if we have no reason to believe them different” to continuous distributions, then the result is a uniform distribution.

contrast, when $\alpha > 1 + K$, trial is possible even when the parties completely agree on the case outcome, because plaintiff has far more to gain from a trial victory than defendant would lose, so the defendant would not be willing to pay enough to convince the plaintiff to settle. It is because those certain-to-win cases (for plaintiffs) will always go to trial when $\alpha > 1 + K$ that the plaintiff trial win rate in the limit is inevitably 1 for such cases.

We provide a very brief sketch of the proof here under Assumption A2a when $K < \alpha < 1 + K$. A modified argument applies for the scenario with Assumption A2b, and the full proofs for both cases (as well as when $\alpha \geq 1 + K$) are included in [Appendix A](#). Recall that

$$W_{\alpha,\sigma}(y^*) = \frac{\int_0^\infty \Pi_\alpha(z) g_Y(\sigma z + y^*) dz}{\int_{-\infty}^\infty \Pi_\alpha(z) g_Y(\sigma z + y^*) dz}.$$

For the main result, we are done if we can take the limits under the integral by applying Lebesgue's Dominated Convergence Theorem. In that case, because $\Pi_\alpha(z)$ is independent of σ (in the uv -coordinate space) and $g_Y(y)$ is continuous and strictly positive, we have $\lim_{\sigma \rightarrow 0^+} g_Y(\sigma z + y^*) = g_Y(y^*) \neq 0$ and therefore,

$$\lim_{\sigma \rightarrow 0^+} W_{\alpha,\sigma}(y^*) = \lim_{\sigma \rightarrow 0^+} \frac{\int_0^\infty \Pi_\alpha(z) g_Y(\sigma z + y^*) dz}{\int_{-\infty}^\infty \Pi_\alpha(z) g_Y(\sigma z + y^*) dz} = \frac{\int_0^\infty \lim_{\sigma \rightarrow 0^+} \Pi_\alpha(z) g_Y(\sigma z + y^*) dz}{\int_{-\infty}^\infty \lim_{\sigma \rightarrow 0^+} \Pi_\alpha(z) g_Y(\sigma z + y^*) dz} = \frac{\int_0^\infty \Pi_\alpha(z) g_Y(y^*) dz}{\int_{-\infty}^\infty \Pi_\alpha(z) g_Y(y^*) dz} = \frac{\int_0^\infty \Pi_\alpha(z) dz}{\int_{-\infty}^\infty \Pi_\alpha(z) dz}.$$

Application of Lebesgue's Dominated Convergence Theorem, however, requires existence of a Lebesgue-integrable function that dominates $\Pi_\alpha(z) g_Y(\sigma z + y^*)$. Since $g_Y(\sigma z + y^*)$ is bounded above by Assumption A1, we need only show $\Pi_\alpha(z)$ is integrable—that is, both $\int_0^\infty \Pi_\alpha(z) dz$ and $\int_{-\infty}^\infty \Pi_\alpha(z) dz$ converge.

As noted already, for each z , $\Pi_\alpha(z)$ is an integral of a bivariate normal distribution centered at $(u, v) = (z, z)$ and evaluated over $R_\alpha(u, v)$. Meanwhile, it is easy to show that, when $K < \alpha < 1 + K$, $R_\alpha(u, v)$, as a region in the uv -space, lies entirely below some horizontal asymptote and to the right of some vertical asymptote. This means that, as z moves away from 0 (in either direction), the center of the distribution, too, will move farther away from $R_\alpha(u, v)$. Therefore, according to bivariate Chebyshev's Inequality, $\Pi_\alpha(z)$ must eventually decrease at the speed of z^{-2} (or faster). Hence, $\int_{-\infty}^\infty \Pi_\alpha(z) dz$ must converge, and [Proposition 3](#) follows.

Furthermore, because [Proposition 3](#) establishes that in the limit, the case of a well-behaved distribution reduces to the case of the improper uniform distribution, we can appeal to the results of [Proposition 1](#) and [Proposition 2](#) for the limit cases with a well-behaved distribution. Thus, we immediately get the following two results.

Proposition 4. (The Fifty-Percent Limit Hypothesis under Well-Behaved Dispute Distribution). *Under both Model 2 and Model 3, if the stakes are symmetric, the plaintiff trial win rate converges to fifty percent as the parties become increasingly accurate in estimating the case merits. That is, $\lim_{\sigma \rightarrow 0^+} W_\sigma(y^*) = 1/2$.*

Proposition 5. (The Asymmetric Stakes Hypothesis under Well-Behaved Dispute Distribution). *Under both Model 2 and Model 3, as the parties become increasingly accurate in estimating the case merits, the plaintiff trial win rate converges to less than fifty-percent if the defendant has more at stake, and converges to more than fifty percent if the plaintiff has more at stake. That is, $\lim_{\sigma \rightarrow 0^+} W_{\alpha,\sigma}(y^*) > 1/2$ for $\alpha > 1$ and $\lim_{\sigma \rightarrow 0^+} W_{\alpha,\sigma}(y^*) < 1/2$ for $\alpha < 1$.*

Because [Proposition 5](#) follows directly from [Proposition 2](#), note that even when g_Y is a well-behaved distribution satisfying Assumption A1, [Fig. 3](#) still accurately depicts the limit value of the plaintiff trial win rate for disputes with asymmetric stakes.

Thus far we have limited our analysis to hypotheses relating to the limit as parties become increasingly accurate in predicting trial outcomes. Our analysis, however, also offers some insight into the Fifty-Percent Bias Hypothesis. That hypothesis says that the plaintiff trial win rate will be closer to fifty percent than the percentage of cases that plaintiff would have won if all cases had been litigated and none had settled.

Note first that this hypothesis is *a priori* plausible only if the following two conditions are met: the stakes are symmetric ($\alpha = 1$), and the plaintiff trial win rate if all cases were litigated is not itself fifty percent. If the first condition (symmetric stakes) is not satisfied, [Proposition 5](#) tells us that the limit value of the plaintiff trial win rate will not itself be fifty percent, and therefore, the Fifty-Percent Bias Hypothesis will sometimes be false for sufficiently small values of σ . On the other hand, if the first condition (symmetric stakes) is satisfied, the plaintiff trial win rate will converge to fifty percent, and thus it is reasonable to think that the plaintiff trial win rate in the real world (where most cases settle) will be closer to its limit value than the plaintiff trial win rate in a counterfactual world where no cases settle. The second condition is necessary because if the plaintiff's trial win rate when no cases are settled happens to be fifty percent, it is impossible for the plaintiff trial win rates to be closer to fifty percent. Meanwhile, if these two conditions are satisfied, as a corollary to the Fifty-Percent Limit Hypothesis, we can conclude that for σ values that are sufficiently small, the Fifty-Percent Bias Hypothesis must be true.

On the other hand, if σ is sufficiently high, the Fifty-Percent Bias Hypothesis will not be generally true unless we make more restrictive assumptions about g_Y . We show that under Model 2,²¹ if g_Y is symmetric (not necessarily around y^*) and is logarithmically concave—conditions which are satisfied, for example, by normal distributions but also others as well—the Fifty-Percent Bias Hypothesis is true away from the limit as well. These are sufficient conditions, rather than necessary conditions. This is clear since a small perturbation on the distribution of disputes would be unlikely to thwart the overall selection bias. Nevertheless, we also show that neither symmetry nor logarithmically concave cumulative distribution by itself is sufficient (see the [Appendix A](#) for proofs).

Proposition 6. (The Fifty-Percent Bias Hypothesis under Well-Behaved Dispute Distribution). *Under both Model 2 and Model 3, when the parties' stakes are symmetric, the Fifty-Percent Bias Hypothesis will be true for sufficiently small values of σ as long as $\int_{y^*}^\infty g_Y(y) dy \neq 1/2$. In*

²¹ We do not have results for the Fifty-Percent Bias Hypothesis for general σ under Model 3.

other words, as σ approaches zero, the plaintiff trial win rate will (eventually) be closer to fifty percent than the would-be plaintiff win rate if all disputes were to go to trial. For general values of $\sigma > 0$, the Fifty-Percent Bias Hypothesis will be true under Model 2, if g_Y is symmetric and logarithmically concave, $\alpha = 1$, and $\int_{y^*}^{\infty} g_Y(y) dy \neq 1/2$.

5. Conclusion

This paper provides a rigorous analysis of Priest and Klein's conclusions about the selection of disputes for litigation. It distinguishes several hypotheses plausibly attributable to Priest and Klein, and proves or disproves them. We conclude that several of the hypotheses attributable to Priest and Klein (1984) are mathematically well-founded and true under the assumptions made by Priest and Klein. More specifically, under Priest and Klein's original model, the Trial Selection Hypothesis, Fifty-Percent Limit Hypothesis, Asymmetric Stakes Hypothesis, and Irrelevance of Dispute Distribution Hypothesis are true for any distribution of disputes that is bounded, strictly positive and continuous. The Fifty-Percent Bias Hypothesis is true when the parties are very accurate in estimating case outcomes, but only sometimes true when parties are less accurate. Finally, we have also shown that even under the modified model in which the parties make inferences regarding the probability that the plaintiff would prevail by taking into account the underlying distribution of disputes, all of the limit results from the original model go through.

Acknowledgements

The authors are especially grateful to Ken Alexander, Scott Baker and Jonah Gelbach, who were extraordinarily generous with their feedback and advice on multiple drafts. The authors also thanks Andrew Daughety, Dan Erman, Richard Holden, Bart Kosko, Jennifer Reinganum, Kathryn Spier, Neelesh Tiruvilumala, as well as the participants at the Twenty-Fourth Annual Meeting of the American Law and Economics Association, the 2014 NBER Summer Institute in Law and Economics, Pepperdine University School of Law Faculty Workshop, and the UCLA Law and Economics Workshop for their helpful comments and suggestions.

Appendix A.

This Appendix A contains proofs not included in the main text as well as some additional results referenced in the main text. We prove all the results beginning with Proposition 2, and then turn to discuss the credibility of the plaintiff's threat to go to trial. We begin with Lemma A1, which describes the general shape of $R_\alpha(u, v)$ for various values of α .

Lemma A1. (The Shape of $R_\alpha(u, v)$)

- $R_\alpha(u, v)$ is non-empty if and only if $\alpha > K$.
- For all $\alpha > K$, $R_\alpha(u, v)$ is bounded left by a vertical asymptote at $u = \Phi^{-1}\left(\frac{K}{\alpha}\right)$.
- For all $\alpha < 1 + K$, $R_\alpha(u, v)$ is bounded above by a horizontal asymptote $v = \Phi^{-1}(\alpha - K)$.
- When $\alpha \geq 1 + K$, $R_\alpha(u, v)$ is not bounded above, and if $\alpha > 1 + K$, it is characterized by a region to the right of an increasing curve that has two vertical asymptotes, at $u = \Phi^{-1}\left(\frac{K}{\alpha}\right)$ and $u = \Phi^{-1}\left(\frac{1+K}{\alpha}\right)$.
- When $\alpha = 1 + K$, then the boundary of $R_\alpha(u, v)$ in the first quadrant is characterized by an asymptote with a slope of 1, and the boundary approaches it monotonically.

Proof of Lemma A1. If $\alpha \leq K$, $R_\alpha(u, v)$ is empty since $\alpha\Phi[u] - \Phi[v] < \alpha \leq K$ for all (u, v) . On the other hand, if $\alpha > K$, for high enough u and low enough v , we can find some value such that $\alpha\Phi[u] - \Phi[v] > K$. Meanwhile, if $\alpha\Phi[u] - \Phi[v] > K$, then $\alpha\Phi[u] > K$, and thus $R_\alpha(u, v)$ is bounded left by $u = \Phi^{-1}\left(\frac{K}{\alpha}\right)$. Intuitively, this vertical bound implies that if the defendant believes that the plaintiff will prevail with probability zero, then as long as the plaintiff draws a signal less than a certain value, the parties will not go to trial because there will not be enough disagreement to induce litigation. If $\alpha < 1 + K$, $\Phi[v] < \alpha\Phi[u] - K \leq \alpha - K$. Thus, $R_\alpha(u, v)$ is bounded above by $v = \Phi^{-1}(\alpha - K)$. Similarly, this horizontal bound implies that if the plaintiff believes that he will prevail with probability one, then as long as the defendant draws a signal above a certain value, the case will settle. Meanwhile, if $\alpha \geq 1 + K$, then the boundary curve $\alpha\Phi[u] - \Phi[v] = K$ must continually increase as u increases. Furthermore, if $\alpha > 1 + K$, then the inequality must hold for all value of $u > \Phi^{-1}\left(\frac{1+K}{\alpha}\right)$ and all values of v .

Now suppose $\alpha = 1 + K$. Note first that $u > v$ for all $(u, v) \in R_\alpha(u, v)$. This is because $(1 + K)\Phi[u] - \Phi[v] > K$ implies $\Phi[u] - \Phi[v] > K(1 - \Phi[u]) > 0$. Therefore, $R_\alpha(u, v)$ must lie strictly under the line $v = u$, and the boundary cannot have a slope greater than 1 in the limit. Now it suffices to show that, given any point $(u, v) \in R_\alpha(u, v)$, we must also have $(u + x, v + x) \in R_\alpha(u, v)$ for all $x \geq 0$. This means $R_\alpha(u, v)$ must wholly contain its own translation in the direction of $v = u$. This will ensure that the boundary will have slope at least 1 at all points. To see this, we show that all points in $R_\alpha(u, v)$ will be properly contained in $R_\alpha(u, v)$ when translated upward by x . This amounts to showing the following: if $(1 + K)\Phi[u] - \Phi[v] \geq K$, then $(1 + K)\Phi[u + x] - \Phi[v + x] \geq K$, where $x \geq 0$. To show this, we rewrite the statement as follows: if $\left(\frac{1 - \Phi[v]}{1 - \Phi[u]}\right) \geq K + 1$, then $\left(\frac{1 - \Phi[v + x]}{1 - \Phi[u + x]}\right) \geq K + 1$. It is then sufficient to show that $\left(\frac{1 - \Phi[v + x]}{1 - \Phi[u + x]}\right) \geq \left(\frac{1 - \Phi[v]}{1 - \Phi[u]}\right)$, or simply that $\frac{1 - \Phi[v + x]}{1 - \Phi[u + x]}$ is increasing in x . Under the quotient rule, this will be true as long as

$$(1 - \Phi[u + x])(-f(v + x)) - (1 - \Phi[v + x])(-f(u + x)) \geq 0,$$

where f is the standard normal density function. This can be rewritten as $\frac{f(u + x)}{(1 - \Phi[u + x])} \geq \frac{f(v + x)}{(1 - \Phi[v + x])}$. The last result is true (for $u > v$) when $f(x)$ has an increasing hazard rate, which is true for the standard normal density function. \square

Fig. 5 illustrates the regions of integration. The left figure shows the region of integration for $\alpha = 1$. Note that the region of integration, $R_1(u, v)$, is always bounded by a horizontal asymptote above and by a vertical asymptote on the left. In other words, the region of integration can be contained by a translated fourth quadrant. As discussed in Lemma A1, $R_\alpha(u, v)$ will always satisfy this property whenever $K < \alpha <$

$$K = \frac{1}{3}).$$

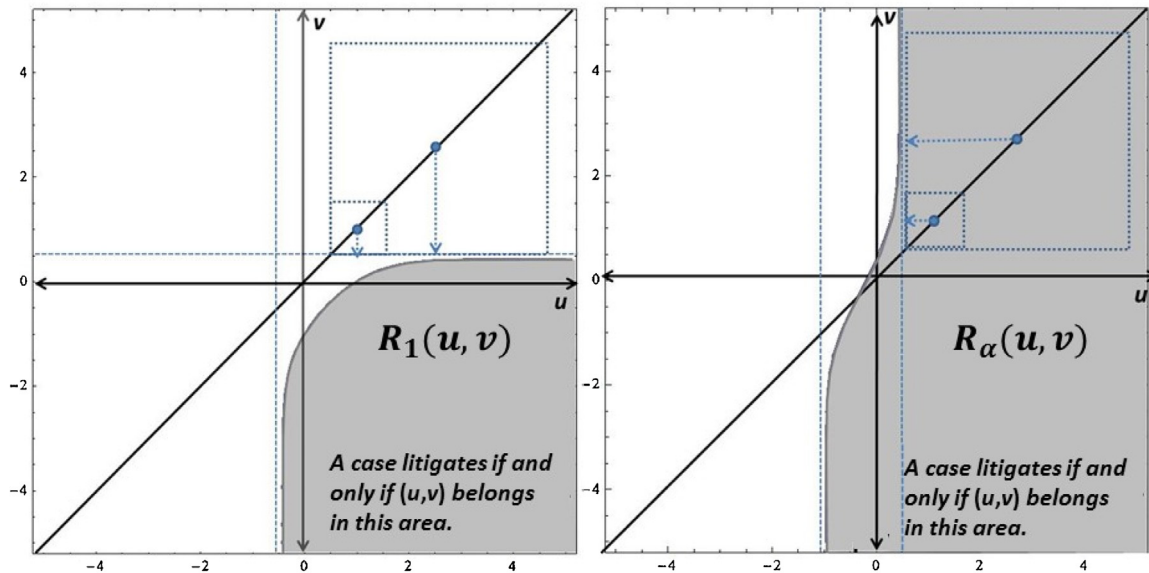


Fig. 5. Litigation Sets ($\alpha = 1$ (Left) and $\alpha > 1 + K$ (Right); $K = \frac{1}{3}$).

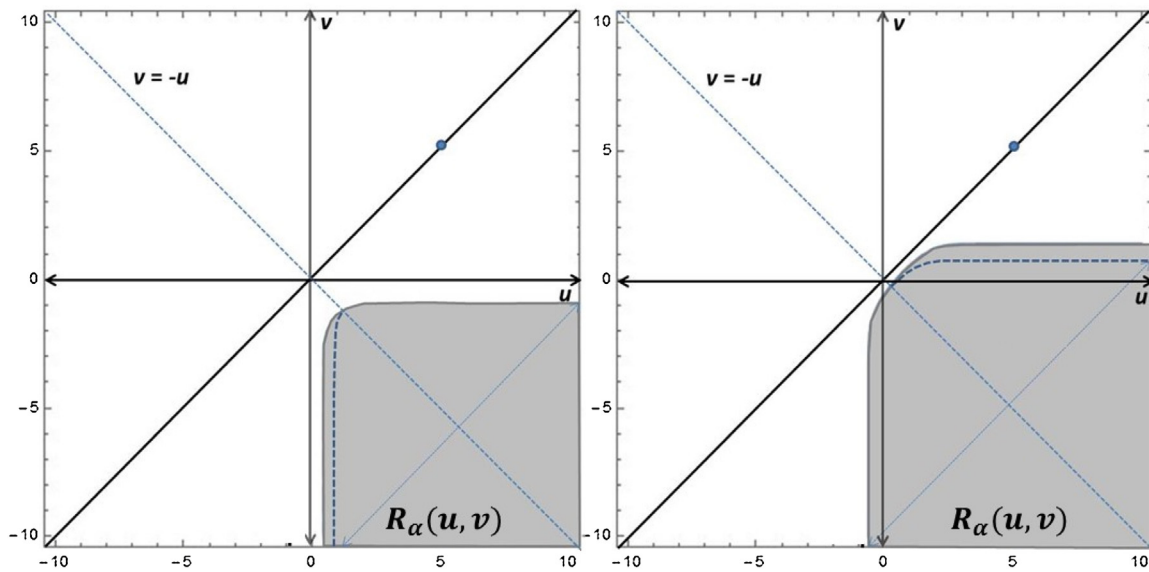


Fig. 6. Litigation Sets ($K < \alpha < 1$ (Left) and $1 < \alpha < 1 + K$ (Right); $K = \frac{1}{3}$).

$1 + K$. In contrast, as illustrated by the right figure, for $\alpha > 1 + K$, $R_\alpha(u, v)$ will not be bounded by any horizontal asymptote, but will be characterized by two vertical asymptotes.

Proof of Proposition 2. Note that $\int_0^\infty \Pi_\alpha(z) dz$ is evaluated by centering the bivariate distribution of errors at each point along the line $v = u$ in the first quadrant and integrating it over the shaded area. $\int_{-\infty}^\infty \Pi_\alpha(z) dz$ is evaluated by carrying out the similar integration along the entire line $v = u$. To show that the win rate is greater or less than fifty-percent, it suffices then to compare $\int_0^\infty \Pi_\alpha(z) dz$ to $\int_{-\infty}^0 \Pi_\alpha(z) dz$. This can be done by looking at symmetry or asymmetry of $R_\alpha(u, v)$ across the line $v = -u$.

For $\alpha > 1 + K$, the result is obvious since $R_\alpha(u, v)$ shows that $\Pi_\alpha(z)$ approaches 1 as z approaches infinity, but will approach 0 as z approaches negative infinity. A similar argument applies when $\alpha = 1 + K$.

Now assume $K < \alpha < 1 + K$. In this case, as noted already, the region of integration, $R_\alpha(u, v)$, is always bounded by a horizontal asymptote above and by a vertical asymptote on the left. As shown in Fig. 6, when $K < \alpha < 1$, the top boundary of $R_\alpha(u, v)$, when reflected across the line $v = -u$ is bounded left by the left boundary of $R_\alpha(u, v)$. The reflection of the top boundary is represented by the dotted line that lies inside $R_\alpha(u, v)$ and is nearly vertical. This shows that $R_\alpha(u, v)$ is composed of a subregion that is symmetric around $v = -u$ and a separate subregion that lies entirely below $v = -u$ (the side more favorable toward the plaintiff). Conversely, when $1 < \alpha < 1 + K$, the region of integration will be biased towards the first-quadrant. As shown in the right graph of Fig. 6, this time the left boundary of $R_\alpha(u, v)$, when reflected across the line $v = -u$ is bounded above by the top boundary of $R_\alpha(u, v)$, indicating more litigated cases for the defendant.

We show this more formally. Note that the horizontal asymptote, $\Phi^{-1}(\alpha - K)$, occurs at a value that is more positive than the vertical asymptote, $\Phi^{-1}(\frac{K}{\alpha})$, is negative if and only $\alpha > 1$. In other words, $|\Phi^{-1}(\alpha - K)|$ is greater (less) than $|\Phi^{-1}(\frac{K}{\alpha})|$ if $\alpha > 1$ ($\alpha < 1$). This is obvious since $\Phi(-\Phi^{-1}(\frac{K}{\alpha})) = 1 - \Phi(\Phi^{-1}(\frac{K}{\alpha})) = (1 - \frac{K}{\alpha})$ and $\Phi(\Phi^{-1}(\alpha - K)) = \alpha(1 - \frac{K}{\alpha})$.

Suppose $\alpha > 1$. To show that the limit value of the plaintiff trial win rate is greater than $\frac{1}{2}$, it suffices to show that $R_\alpha(u, v)$ is skewed in the direction of the first quadrant in the following sense. Take $S_\alpha(u, v) = R_\alpha(u, v) \cap \{(u, v) | v < -u\}$. Then $S_\alpha(u, v)$ is the portion of $R_\alpha(u, v)$ that lies below the line $v = -u$. Let $S'_\alpha(u, v) = \{(-v, -u) | (u, v) \in S_\alpha(u, v)\}$. $S'_\alpha(u, v)$ is the reflection of $S_\alpha(u, v)$ over the line $v = -u$. We show that $S'_\alpha(u, v) \subset R_\alpha(u, v)$. This means that the entire region of integration that lies below $v = -u$ can be reflected across the line $v = -u$ and that reflection will be properly contained in $R_\alpha(u, v)$. Since $R_\alpha(u, v)$ contains additional regions above the line $v = -u$ (because the horizontal asymptote is more positive than the vertical asymptote is negative in this case), this shows that $R_\alpha(u, v)$ is skewed in the direction of the first quadrant, and this is sufficient to show that the plaintiff win rate will be higher than fifty percent. To see this, take away the portion that is symmetric, which is $S_\alpha(u, v) \cup S'_\alpha(u, v)$. What is left must lie entirely on the side of $v > -u$, and hence closer to the plaintiff's win side than the defendant's win side. To show $S'_\alpha(u, v) \subset R_\alpha(u, v)$, we need to show the following: If (i) $\alpha > 1$, (ii) $\alpha\Phi[u] - \Phi[v] > K$, and (iii) $v < -u$, then $\alpha\Phi[-v] - \Phi[-u] > K$. In turn, it suffices to show that if (i) $\alpha > 1$ and (ii) $\Phi[v] < \Phi[-u]$, then $\alpha\Phi[-v] - \Phi[-u] > \alpha\Phi[u] - \Phi[v]$, since the last expression will be greater than K . But this last inequality rearranges as follows:

$$\alpha\Phi[-v] - \Phi[-u] > \alpha\Phi[u] - \Phi[v] \Leftrightarrow \alpha(1 - \Phi[v]) - \Phi[-u] > \alpha(1 - \Phi[-u]) - \Phi[v] \Leftrightarrow (\alpha - 1)(\Phi[-u] - \Phi[v]) > 0,$$

which is immediate since $\alpha > 1$ and $-u > v$ for $(u, v) \in S_\alpha(u, v)$. Therefore, the limit value of the plaintiff's win rate is greater than fifty percent. Notice by the same logic that the last inequality also holds when we have $\alpha < 1$ and $v > -u$. This shows that when $\alpha < 1$, $R_\alpha(u, v)$ is skewed in the direction of the third quadrant, and hence the limit value of the plaintiff win rate is lower than $\frac{1}{2}$. \square

Proof of Proposition 3. We first show the results for Model 2 when $\alpha \in (K, 1 + K)$. Given the sketch of the proof included in the main text, we need only show that $\int_0^\infty \Pi_\alpha(z) dz$ and $\int_{-\infty}^0 \Pi_\alpha(z) dz$ are finite. Lemma A1 tells us that $R_\alpha(u, v)$ is properly contained by a translated fourth quadrant in the uv -plane with the origin at $(\Phi^{-1}(\frac{K}{\alpha}), \Phi^{-1}(\alpha - K))$, which in turn is contained by a similar quadrant with the origin at $(\Phi^{-1}(\frac{K}{\alpha}), |\Phi^{-1}(\alpha - K)|)$. This set is obviously contained in all of \mathbf{R}^2 . Moreover, for $z > |\Phi^{-1}(\alpha - K)| > 0$, the point (z, z) is at least $z - |\Phi^{-1}(\alpha - K)| = |z| - |\Phi^{-1}(\alpha - K)|$ away from $R_\alpha(u, v)$. For $z < \Phi^{-1}(\frac{K}{\alpha})$, (z, z) is at least $|z - \Phi^{-1}(\frac{K}{\alpha})| = |z| - |\Phi^{-1}(\frac{K}{\alpha})|$ away from $R_\alpha(u, v)$. This means that if we let $c = \max\{|\Phi^{-1}(\frac{K}{\alpha})|, |\Phi^{-1}(\alpha - K)|\} > 0$ and $d = \min\{|\Phi^{-1}(\frac{K}{\alpha})|, |\Phi^{-1}(\alpha - K)|\} > 0$, then for $K < \alpha < 1 + K$, $R_\alpha(u, v)$ is properly contained in $M_\alpha(u, v; z)$ for every $z \in \mathbf{R}$, where

$$M_\alpha(u, v; z) = \begin{cases} \mathbf{R}^2 & \text{for } |z| < c, \\ \{(u, v) | |u - z| > |z| - d \text{ or } |v - z| > |z| - d\} & \text{for } |z| \geq c. \end{cases}$$

It follows that for $K < \alpha < 1 + K$ and for each $z \in \mathbf{R}$,

$$\Pi_\alpha(z) < \Pi_\alpha^U(z) \equiv \frac{1}{2\pi} \iint_{M_\alpha(u, v; z)} e^{-\frac{(u-z)^2}{2}} e^{-\frac{(v-z)^2}{2}} dudv.$$

It now suffices to show that $\int_0^\infty \Pi_\alpha^U(z) dz$ and $\int_{-\infty}^0 \Pi_\alpha^U(z) dz$ are finite. But notice that

$$\int_0^\infty \Pi_\alpha^U(z) dz = \int_0^c \Pi_\alpha^U(z) dz + \int_c^\infty \Pi_\alpha^U(z) dz = c + \int_c^\infty \Pi_\alpha^U(z) dz.$$

Meanwhile, for each $z > c$, $\Pi_\alpha^U(z) \leq 1 - \Pr(|u - z| < |z| - d \text{ and } |v - z| < |z| - d)$. By bivariate Chebyshev's Inequality, $\Pr(|u - z| < |z| - d \text{ and } |v - z| < |z| - d) \geq 1 - \frac{1 + \sqrt{1 + \text{Cor}(u, v)^2}}{(|z| - d)^2}$.

Therefore,

$$\Pi_\alpha^U(z) \leq \frac{1 + \sqrt{1 + \text{Cor}(u, v)^2}}{(|z| - d)^2},$$

which is quadratic in z in the denominator and therefore integrates to a finite value over $z \in [c, \infty)$. The integral over $z \in (-\infty, 0]$ can likewise be shown to be finite.²²

Now we show the result for Model 3. Although the region of integration is no longer independent of σ , constructing a Lebesgue-integrable dominating function does not require the *actual* region of integration to be independent of σ , but only that the region of integration, for sufficiently small values of σ , can be *contained* in another region of integration that is in fact independent of σ .

As explained in the main text, the trial condition will be determined by the following inequality:

$$\alpha \left(\frac{\int_{y^*}^\infty e^{-\frac{(yp-y)^2}{2\sigma^2}} g_Y(y) dy}{\int_{-\infty}^\infty e^{-\frac{(yp-y)^2}{2\sigma^2}} g_Y(y) dy} \right) - \left(\frac{\int_{y^*}^\infty e^{-\frac{(yp-y)^2}{2\sigma^2}} g_Y(y) dy}{\int_{-\infty}^\infty e^{-\frac{(yp-y)^2}{2\sigma^2}} g_Y(y) dy} \right) > K.$$

²² Under our assumption, of course, $\text{Cor}(u, v) = 0$. But note that the proof goes through even if $\text{Cor}(u, v) \neq 0$.

For a given $\sigma > 0$, we use the same change of variables: $u = \frac{y_p - y^*}{\sigma} = \frac{y + \varepsilon_p - y^*}{\sigma}$, $v = \frac{y + \varepsilon_d - y^*}{\sigma}$. Then by setting the dummy variable ω appropriately, we can rewrite $R_{\alpha, \sigma}(y; y^*)$ as

$$\{(u, v) | \alpha \varphi_p(u, \sigma) - \varphi_d(v, \beta \sigma) > K\} = R_{\alpha, \sigma}(u, v; y^*)$$

where

$$\varphi_p(u, \sigma) = \frac{\int_{-\infty}^u e^{-\frac{\omega^2}{2}} g_Y(y^* + \sigma(u - \omega)) d\omega}{\int_{-\infty}^{\infty} e^{-\frac{\omega^2}{2}} g_Y(y^* + \sigma(u - \omega)) d\omega} \text{ and } \varphi_d(v, \sigma) = \frac{\int_{-\infty}^v e^{-\frac{\omega^2}{2}} g_Y(y^* + \sigma(v - \omega)) d\omega}{\int_{-\infty}^{\infty} e^{-\frac{\omega^2}{2}} g_Y(y^* + \sigma(v - \omega)) d\omega}.$$

Therefore, for each $\sigma > 0$,

$$\Pi_{\alpha, \sigma}(y; y^*) = \frac{1}{2\pi} \iint_{R_{\alpha, \sigma}(u, v; y^*)} e^{-\frac{(u-z)^2}{2}} e^{-\frac{(v-z)^2}{2}} dudv.$$

Now the region of integration in the uv -plane will continue to depend on σ and y^* . But still, we have

$$\frac{\int_{y^*}^{\infty} \Pi_{\alpha, \sigma}(y; y^*) g_Y(y) dy}{\int_{-\infty}^{\infty} \Pi_{\alpha, \sigma}(y; y^*) g_Y(y) dy} = \frac{\int_0^{\infty} \Pi_{\alpha, \sigma}(\sigma z + y^*; y^*) g_Y(\sigma z + y^*) dz}{\int_{-\infty}^{\infty} \Pi_{\alpha, \sigma}(\sigma z + y^*; y^*) g_Y(\sigma z + y^*) dz}.$$

Notice that once we assume σ to be sufficiently small (say $\sigma < \bar{\sigma}$), there will be an absolute lower limit \underline{u} and an absolute upper limit \bar{v} such that, for each σ such that $0 < \sigma < \bar{\sigma}$ and for each $y^* \in \mathbf{R}$, $R_{\alpha, \sigma}(u, v; y^*)$ is bounded above by $v = \bar{v}$ and bounded on the left side by $u = \underline{u}$. In other words, for sufficiently small σ , the boundary of $R_{\alpha, \sigma}(u, v; y^*)$ will not go to infinity (negative infinity) even as u approaches

infinity (v approaches negative infinity). But this is obvious since $\frac{\int_{-\infty}^u e^{-\frac{\omega^2}{2}} g_Y(y^* + \sigma(u - \omega)) d\omega}{\int_{-\infty}^{\infty} e^{-\frac{\omega^2}{2}} g_Y(y^* + \sigma(u - \omega)) d\omega}$ and $\frac{\int_{-\infty}^v e^{-\frac{\omega^2}{2}} g_Y(y^* + \sigma(v - \omega)) d\omega}{\int_{-\infty}^{\infty} e^{-\frac{\omega^2}{2}} g_Y(y^* + \sigma(v - \omega)) d\omega}$, purely as functions

defined in terms of variable σ , are continuous in σ at $\sigma = 0$, and we know the behavior of the boundaries of $R_{\alpha, 0}(u, v; y^*) = R_{\alpha}(u, v)$.

Therefore, once we assume σ to be sufficiently small,

$$\Pi_{\alpha, \sigma}(\sigma z + y^*; y^*) g_Y(\sigma z + y^*) < \frac{\bar{g}}{2\pi} \iint_{\{(u, v) | u \geq \underline{u}, v \leq \bar{v}\}} e^{-\frac{(u-z)^2}{2}} e^{-\frac{(v-z)^2}{2}} dudv,$$

where \bar{g} is the upper bound for g_Y . This value will eventually decrease at least as fast as $|z|^{-2}$ as z approaches infinity (or negative infinity). This is because, as we have done in Proposition 3, we can apply the bounds (based on Chebyshev's inequality) to the region $\{(u, v) | u \geq \underline{u}, v \leq \bar{v}\}$. Consequently, we can still take the limits inside the integral. Then

$$\lim_{\sigma \rightarrow 0^+} \frac{\int_0^{\infty} \Pi_{\alpha, \sigma}(\sigma z + y^*; y^*) g_Y(\sigma z + y^*) dz}{\int_{-\infty}^{\infty} \Pi_{\alpha, \sigma}(\sigma z + y^*; y^*) g_Y(\sigma z + y^*) dz} = \frac{\int_0^{\infty} \lim_{\sigma \rightarrow 0^+} \Pi_{\alpha, \sigma}(\sigma z + y^*; y^*) dz}{\int_{-\infty}^{\infty} \lim_{\sigma \rightarrow 0^+} \Pi_{\alpha, \sigma}(\sigma z + y^*; y^*) dz}$$

since g_Y is continuous and strictly positive everywhere. Meanwhile

$$\lim_{\sigma \rightarrow 0^+} \Pi_{\alpha, \sigma}(\sigma z + y^*; y^*) = \lim_{\sigma \rightarrow 0^+} \frac{1}{2\pi} \iint_{R_{\alpha, \sigma}(u, v; y^*)} e^{-\frac{(u-z)^2}{2}} e^{-\frac{(v-z)^2}{2}} dudv = \frac{1}{2\pi} \iint \lim_{\sigma \rightarrow 0^+} R_{\alpha, \sigma}(u, v; y^*) e^{-\frac{(u-z)^2}{2}} e^{-\frac{(v-z)^2}{2}} dudv.$$

Since $\lim_{\sigma \rightarrow 0^+} R_{\alpha, \sigma}(u, v; y^*) = \left\{ (u, v) | \alpha \left(\frac{\lim_{\sigma \rightarrow 0^+} \int_{-\infty}^u e^{-\frac{\omega^2}{2}} g_Y(y^* + \sigma(w - \omega)) d\omega}{\lim_{\sigma \rightarrow 0^+} \int_{-\infty}^{\infty} e^{-\frac{\omega^2}{2}} g_Y(y^* + \sigma(w - \omega)) d\omega} \right) - \left(\frac{\lim_{\sigma \rightarrow 0^+} \int_{-\infty}^v e^{-\frac{\omega^2}{2}} g_Y(y^* + \sigma(w - \omega)) d\omega}{\lim_{\sigma \rightarrow 0^+} \int_{-\infty}^{\infty} e^{-\frac{\omega^2}{2}} g_Y(y^* + \sigma(w - \omega)) d\omega} \right) > K \right\}$, and since each inte-

grand is bounded above by $e^{-\frac{\omega^2}{2}} \bar{g}$, which is Lebesgue-integrable, we can once again take the limits inside the integral and factor out $g_Y(y^*)$. And therefore,

$$\lim_{\sigma \rightarrow 0^+} R_{\alpha, \sigma}(u, v; y^*) = \left\{ (u, v) | \alpha \left(\frac{\int_{-\infty}^u e^{-\frac{\omega^2}{2}} d\omega}{\int_{-\infty}^{\infty} e^{-\frac{\omega^2}{2}} d\omega} \right) - \left(\frac{\int_{-\infty}^v e^{-\frac{\omega^2}{2}} d\omega}{\int_{-\infty}^{\infty} e^{-\frac{\omega^2}{2}} d\omega} \right) > K \right\} = \{(u, v) | \alpha \Phi[u] - \Phi[v] > K\} = R_{\alpha}(u, v).$$

Therefore,

$$\lim_{\sigma \rightarrow 0^+} \Pi_{\alpha, \sigma}(\sigma z + y^*; y^*) = \frac{1}{2\pi} \iint_{R_{\alpha}(u, v)} e^{-\frac{(u-z)^2}{2}} e^{-\frac{(v-z)^2}{2}} dudv = \Pi(z)$$

and

$$\lim_{\sigma \rightarrow 0^+} \frac{\int_{-\infty}^{\infty} \Pi_{\alpha, \sigma}(\sigma z + y^*; y^*) g_Y(\sigma z + y^*) dz}{\int_{-\infty}^{\infty} \Pi_{\alpha, \sigma}(\sigma z + y^*; y^*) g_Y(\sigma z + y^*) dz} = \frac{\int_0^{\infty} \lim_{\sigma \rightarrow 0^+} \Pi_{\alpha, \sigma}(\sigma z + y^*; y^*) dz}{\int_{-\infty}^{\infty} \lim_{\sigma \rightarrow 0^+} \Pi_{\alpha, \sigma}(\sigma z + y^*; y^*) dz} = \frac{\int_0^{\infty} \Pi(z) dz}{\int_{-\infty}^{\infty} \Pi(z) dz}.$$

We now show the result for Model 2 when $\alpha > 1 + K$. Note that

$$\frac{\int_0^{\infty} \Pi_{\alpha}(z) g_Y(\sigma z + y^*) dz}{\int_{-\infty}^{\infty} \Pi_{\alpha}(z) g_Y(\sigma z + y^*) dz} = \frac{\int_0^{\infty} \Pi_{\alpha}(z) g_Y(\sigma z + y^*) dz}{\int_{-\infty}^0 \Pi_{\alpha}(z) g_Y(\sigma z + y^*) dz + \int_0^{\infty} \Pi_{\alpha}(z) g_Y(\sigma z + y^*) dz}.$$

Here $\lim_{\sigma \rightarrow 0^+} \int_{-\infty}^0 \Pi_{\alpha}(z) g_Y(\sigma z + y^*) dz = \int_{-\infty}^0 \Pi_{\alpha}(z) g_Y(y^*) dz$ as before, since $R_{\alpha}(u, v)$ is bounded by a left asymptote. Meanwhile, we cannot apply Lebesgue's Dominated Convergence Theorem to $\int_0^{\infty} \Pi_{\alpha}(z) g_Y(\sigma z + y^*) dz$ because $R_{\alpha}(u, v)$ contains all of the increasing large rectangles. Instead, we write

$$\int_0^{\infty} \Pi_{\alpha}(z) g_Y(\sigma z + y^*) dz = \int_0^{\infty} (1 - \Gamma_{\alpha}(z)) g_Y(\sigma z + y^*) dz = \int_0^{\infty} g_Y(\sigma z + y^*) dz - \int_0^{\infty} \Gamma_{\alpha}(z) g_Y(\sigma z + y^*) dz$$

where $\Gamma_{\alpha}(z) = \frac{1}{2\pi} \iint_{R_{\alpha}^C(u, v)} e^{-\frac{(u-z)^2}{2}} e^{-\frac{(v-z)^2}{2}} du dv$ and $R_{\alpha}^C(u, v)$ is the complement of $R_{\alpha}(u, v)$. Then we can apply Lebesgue's Dominated Convergence Theorem to $\int_0^{\infty} \Gamma_{\alpha}(z) g_Y(\sigma z + y^*) dz$ since the integral over the complement of $R_{\alpha}(u, v)$ can now be bounded above. Therefore,

$$\begin{aligned} & \lim_{\sigma \rightarrow 0^+} \frac{\int_0^{\infty} \Pi_{\alpha}(z) g_Y(\sigma z + y^*) dz}{\int_{-\infty}^{\infty} \Pi_{\alpha}(z) g_Y(\sigma z + y^*) dz} \\ &= \lim_{\sigma \rightarrow 0^+} \frac{\int_0^{\infty} g_Y(\sigma z + y^*) dz - \int_0^{\infty} \Gamma_{\alpha}(z) g_Y(\sigma z + y^*) dz}{\int_{-\infty}^0 \Pi_{\alpha}(z) g_Y(\sigma z + y^*) dz + \int_0^{\infty} g_Y(\sigma z + y^*) dz - \int_0^{\infty} \Gamma_{\alpha}(z) g_Y(\sigma z + y^*) dz} \\ &= \frac{\left(\lim_{\sigma \rightarrow 0^+} \int_0^{\infty} g_Y(\sigma z + y^*) dz \right) - \left(\lim_{\sigma \rightarrow 0^+} \int_0^{\infty} \Gamma_{\alpha}(z) g_Y(\sigma z + y^*) dz \right)}{\left(\lim_{\sigma \rightarrow 0^+} \int_0^{\infty} g_Y(\sigma z + y^*) dz \right) + \left(\lim_{\sigma \rightarrow 0^+} \int_{-\infty}^0 \Pi_{\alpha}(z) g_Y(\sigma z + y^*) dz \right) - \left(\lim_{\sigma \rightarrow 0^+} \int_0^{\infty} \Gamma_{\alpha}(z) g_Y(\sigma z + y^*) dz \right)} \\ &= \frac{\left(\lim_{\sigma \rightarrow 0^+} \int_0^{\infty} g_Y(\sigma z + y^*) dz \right) - g_Y(y^*) \int_0^{\infty} \Gamma_{\alpha}(z) dz}{\left(\lim_{\sigma \rightarrow 0^+} \int_0^{\infty} g_Y(\sigma z + y^*) dz \right) + g_Y(y^*) \int_{-\infty}^0 \Pi_{\alpha}(z) dz - g_Y(y^*) \int_0^{\infty} \Gamma_{\alpha}(z) dz} \end{aligned}$$

Since $\lim_{\sigma \rightarrow 0^+} \int_0^{\infty} g_Y(\sigma z + y^*) dz = \lim_{\sigma \rightarrow 0^+} \frac{1 - G_Y(y^*)}{\sigma} = \infty$, while all other terms are finite, the limit value of the plaintiff win rate is 1.

When $\alpha = 1 + K$, Lemma A1 tells us that the boundary of $R_{\alpha}(u, v)$ in the first quadrant is characterized by an asymptote with a slope of 1, and the boundary approaches it monotonically. Note further that from Lemma A1, we must have $\Pi_{\alpha}(z) \geq \Pi_{\alpha}(0) > 0$ for all z . This is because $\Pi_{\alpha}(0)$ is equivalent to taking the double integral centered at (z, z) over a region that corresponds to $R_{\alpha}(u, v)$ translated by (z, z) , which would be wholly contained in the original $R_{\alpha}(u, v)$. Therefore,

$$\int_0^{\infty} \Pi_{\alpha}(z) g_Y(\sigma z + y^*) dz \geq \int_0^{\infty} \Pi_{\alpha}(0) g_Y(\sigma z + y^*) dz = \Pi_{\alpha}(0) \int_0^{\infty} g_Y(\sigma z + y^*) dz.$$

As shown above, $\lim_{\sigma \rightarrow 0^+} \int_0^{\infty} g_Y(\sigma z + y^*) dz = \infty$. Thus, the rest follows as before, and the limit value is 1.

Finally, the logic from the first half of the proof applies to lead to the same results (for $\alpha > 1 + K$) under Model 3. \square

Proof of Proposition 4. Proposition 4 follows directly from Proposition 1 and Proposition 3. \square

Proof of Proposition 5. Proposition 5 follows directly from Proposition 2 and Proposition 3. \square

Proof of Proposition 6. Without loss of generality, assume $g_Y(y)$ is centered around 0. Since $g_Y(y)$ is log-concave, (i) it is also single-peaked (including the possibility of a flat-top) and (ii) its cumulative distribution, $G_Y(y)$, is also log-concave. Without loss of generality, suppose $y^* < 0$. In that case, $\int_{y^*}^{\infty} g_Y(y) dy > \frac{1}{2}$. Then we need to show that $|W_{\sigma}(y^*) - 1/2| < \int_{y^*}^{\infty} g_Y(y) dy - \frac{1}{2}$. First notice $W_{\sigma}(y^*) > \frac{1}{2}$. This can be seen as follows. $W_{\sigma}(y^*)$ over all of $g_Y(y)$ is greater than $W_{\sigma}(y^*)$ over a symmetric image of $g_Y(y)|_{y \leq y^*}$ around y^* . The latter graph lies completely under $g_Y(y)$ because $g_Y(y)$ is symmetric and single-peaked. And clearly, $W_{\sigma}(y^*) = \frac{1}{2}$ over a symmetric image of $g_Y(x)|_{x \geq y^*}$ around y^* . Thus, we need only show $\int_{y^*}^{\infty} g_Y(y) dx > W_{\sigma}(y^*)$. This is equivalent to showing

$$\frac{\int_{y^*}^{\infty} g_Y(y) dy}{\int_{-\infty}^{\infty} g_Y(y) dy} = \int_{y^*}^{\infty} g_Y(y) dx > W_{\sigma}(y^*) = \frac{\int_{y^*}^{\infty} \Pi_{\sigma}(y; y^*) g_Y(y) dy}{\int_{-\infty}^{\infty} \Pi_{\sigma}(y; y^*) g_Y(y) dy}.$$

This inequality rearranges to

$$\frac{\int_{-\infty}^{y^*} \Pi_{\sigma}(y; y^*) g_Y(y) dy}{\int_{-\infty}^{y^*} g_Y(y) dy} > \frac{\int_{y^*}^{\infty} \Pi_{\sigma}(y; y^*) g_Y(y) dy}{\int_{y^*}^{\infty} g_Y(y) dy}.$$

By changing the variable to $y = y^* - w$ for the left-side integral and $y = y^* + w$ for the right-side integral and recognizing that $\Pi_{\sigma}(y; y^*)$ is symmetric around y^* , we have

$$\int_0^{\infty} \Pi_{\sigma}(y^* - w; y^*) \left(\frac{g_Y(y^* - w)}{\int_0^{\infty} g_Y(y^* - w) dw} \right) dw = \int_0^{\infty} \Pi_{\sigma}(y^* + w; y^*) \left(\frac{g_Y(y^* + w)}{\int_0^{\infty} g_Y(y^* + w) dw} \right) dw > \int_0^{\infty} \Pi_{\sigma}(y^* + w; y^*) \left(\frac{g_Y(y^* + w)}{\int_0^{\infty} g_Y(y^* + w) dw} \right) dw.$$

Since $\Pi_{\sigma}(y^* + w; y^*)$ is strictly decreasing in w , we need only show $\left(\frac{g_Y(y^* + w)}{\int_0^{\infty} g_Y(y^* + w) dw} \right)$ as a probability density function defined over $[0, \infty)$

first-order stochastically dominates $\left(\frac{g_Y(y^* - w)}{\int_0^{\infty} g_Y(y^* - w) dw} \right)$. Equivalently, we need to show $\frac{\int_0^w g_Y(y^* - w) dw}{\int_0^{\infty} g_Y(y^* - w) dw} > \frac{\int_0^w g_Y(y^* + w) dw}{\int_0^{\infty} g_Y(y^* + w) dw}$ for all $w > 0$, or

$$\frac{G_Y(y^*) - G_Y(y^* - w)}{G_Y(y^*)} > \frac{G_Y(y^* + w) - G_Y(y^*)}{1 - G_Y(y^*)} = \frac{(1 - G_Y(-y^* - w)) - (1 - G_Y(-y^*))}{G_Y(-y^*)} = \frac{G_Y(-y^*) - G_Y(-y^* - w)}{G_Y(-y^*)}$$

since $1 - G_Y(x) = G_Y(-x)$ for all x . Since $-y^* > y^*$, it now suffices to show that $\frac{G_Y(y^*) - G_Y(y^* - w)}{G_Y(y^*)} = 1 - \frac{G_Y(y^* - w)}{G_Y(y^*)}$ is decreasing in y^* , or that $\frac{G_Y(y^* - w)}{G_Y(y^*)}$ is increasing in y^* for each $w > 0$. Under the Quotient Rule, this is true if

$$G_Y(y^*) g_Y(y^* - w) - g_Y(y^*) G_Y(y^* - w) = G_Y(y^*) G_Y(y^* - w) \left(\frac{g_Y(y^* - w)}{G_Y(y^* - w)} - \frac{g_Y(y^*)}{G_Y(y^*)} \right) > 0$$

which holds if $\frac{g_Y(y^*)}{G_Y(y^*)}$ is decreasing in y^* , or put differently, if $G_Y(y^*) g'_Y(y^*) < (g_Y(y^*))^2$. But this last inequality holds since $G_Y(y)$ is strictly log-concave (which must be true since $g_Y(y)$ is log-concave).

Meanwhile, for a general log-concave function or a single-peaked function that is not symmetric, the inequality generally will not hold for all σ at the mean (that is, for Y^* such that $G_Y(Y^*) = \frac{1}{2}$) since $W_{\sigma}(Y^*)$ will not always equal $1/2$. For a symmetric counter example, consider $g_Y(y)$, which equals zero everywhere but takes on the value of 1 on $[y^* - \frac{1}{2}, y^*]$ and $[y^* + B - \frac{1}{2}, y^* + B]$, where $B > \frac{1}{2}$. This function is clearly symmetric around $y = y^* + \frac{B}{2}$. Then $\int_{y^*}^{\infty} g_Y(y) dy = 1/2$. But in this case, plaintiff trial win rate is

$$\frac{\int_{y^* - \frac{1}{2}}^{y^*} \Pi_{\sigma}(y; y^*) dy}{\int_{y^* - \frac{1}{2}}^{y^*} \Pi_{\sigma}(y; y^*) dy + \int_{y^* + B - \frac{1}{2}}^{y^* + B} \Pi_{\sigma}(y; y^*) dy}.$$

By Chebyshev's Inequality, the numerator can get arbitrarily small as B increases, while the denominator maintains a certain minimum value. Hence the plaintiff trial win rate can get arbitrarily small, even though the win rate if all cases had gone to trial would have been $1/2$. Although this counter-example violates Assumption A1, it is possible to create a continuous $g_Y(y)$ which can arbitrarily closely approximate this counter-example. \square

Next, we discuss Priest and Klein's assumption that the plaintiff always has a credible threat to go to trial. Whether relaxing this assumption makes a difference depends on how plaintiff credibility is incorporated into the model. We sketch two potential ways to address this concern. Under both, the results proved in this article remain valid. Nevertheless, both approaches have problems, and a more in-depth analysis, such as [Nalebuff \(1987\)](#), is warranted.

One approach is to assume away the existence of non-credible threats by specifying that even the weakest plaintiff has probability of prevailing C_p/J . Although this is an unrealistic assumption, it is an approach taken in most asymmetric information models of litigation (see, e.g., [Bebchuk, 1984](#); [Reinganum and Wilde, 1986](#); [Shavell, 1996](#)). For example, suppose that if $y < y^*$, the plaintiff's probability of prevailing is C_p/J rather than zero, and if $y > y^*$, the plaintiff's probability of prevailing is $1 - C_d/J$ rather than 1. Then all of the results in the paper remain valid if $C_p = C_d$. The assumption that if $y > y^*$, the plaintiff's probability of prevailing is $1 - C_d/J$, is necessary to ensure that the defendant, in turn, has a credible threat to defend. If damages are certain (as in this and most other litigation models), then a rational defendant would default if the probability of prevailing were higher than $1 - C_d/J$ (see [Hubbard, 2016](#)). The assumption that $C_p = C_d$ simply means that plaintiff's and defendant's litigation costs are equal, which is roughly true.

An alternative approach to non-credible threats is to assume that potential disputes for which plaintiff does not have a credible threat (i.e., $P_p < C_p/J$) exist, but that plaintiffs does not bring those cases, so they neither settle nor litigate. Under this approach, as long as $K = \frac{C-S}{J} \geq C_p/J$, the plaintiff trial win rate will be unaffected, because cases in which the plaintiff's threat is not credible would never have gone to trial anyway. The litigation condition, $P_p - P_d > \frac{C-S}{J}$ already implies $P_p > K$. So, cases for which the credibility constraint is binding have no effect on plaintiff trial win rates. It is reasonable to assume that $K = \frac{C-S}{J} \geq C_p/J$, because settlement costs, S , are generally thought to be much less than half the cost of litigation, C , and plaintiff's litigation costs, C_p , are roughly half of total litigation costs, C . While this approach is appealing, it has two drawbacks. First, a rational plaintiff might bring cases not satisfying the credibility constraint, if it knew that all such cases would settle, as assumed in the model. The plaintiff's decision to sue will depend, in part, on whether the defendant would go to trial when a positive settlement for less than defendant's litigation costs would be in defendant's interest. Second, if the plaintiff did not bring such cases, the defendant would rationally take that into account when estimating the probability that the plaintiff will prevail. This updating would require extensive modification of the model.

References

- Atkinson, Scott, et al., 2009. *The economics of a centralized judiciary: uniformity, forum shopping, and the federal circuit*. *J. Law Econ.* 52, 411–443.
- Aumann, Robert J., 1976. *Agreeing to disagree*. *Ann. Stat.* 4, 1236–1239.
- Bebchuk, Lucian Arye, 1984. *Litigation and settlement under imperfect information*. *Rand J. Econ.* 15, 404–415.
- Bernardo, Antonio, Talley, Eric, Welch, Ivo, 2000. *A theory of legal presumptions*. *J. Law Econ. Org.* 15, 1–48.
- Daughety, Andrew F., Reinganum, Jennifer F., 2012. *Settlement*. In: Sanchirico, Christopher W. (Ed.), *Procedural Law and Economics*. Edward Elgar, Cheltenham, pp. 386–471.
- DeGroot, Morris H., 2004. *Optimal Statistical Decisions*. Wiley-Intersciences.
- Friedman, Daniel, Wittman, Donald, 2007. *Litigation with symmetric bargaining and two-sided incomplete information*. *J. Law Econ. Org.* 23, 98–126.
- Gelbach, Jonah, 2012. *Locking the doors to discovery? Assessing the effects of Twombly and Iqbal on access to discovery*. *Yale Law J.* 121, 2270–2346.
- Gelbach, Jonah, 2016. *Can simple mechanism design results be used to implement the proportionality standard in discovery?* *J. Inst. Theor. Econ.* 172, 200–221.
- Hartigan, J.A., 1983. *Bayes Theory*. Springer-Verlag.
- Hay, Bruce L., Spier, Kathryn E., 1998. In: Newman, Peter (Ed.), *New Palgrave Dictionary of Economics and the Law*. Macmillan, pp 3. 442–451.
- Hubbard, William, 2013. *Testing for change in procedural standards, with application to Bell Atlantic v. Twombly*. *J. Legal Stud.* 42, 35–68.
- Hubbard, William, 2016. *Sinking costs to force or deter settlement*. *J. Law Econ. Org.* 32 (forthcoming).
- Hylton, Keith N., Lin, Haizhen, 2012. *A formalization and extension of the Priest-Klein hypothesis*. Unpublished., <http://dx.doi.org/10.2139/ssrn.1864728>, Available at SSRN: <http://ssrn.com/abstract=1864728>.
- Klerman, Daniel, Lee, Yoon-Ho Alex, 2014. *Inferences from litigated cases*. *J. Legal Stud.* 43, 209–248.
- Lee, Yoon-Ho Alex, Klerman, Daniel, 2015. *Updating Priest and Klein*. Unpublished, Available at SSRN: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2619856.
- Lee, Yoon-Ho Alex, Klerman, Daniel, 2016. *Settlement and selection with correlated two-sided asymmetric information*. Unpublished (in preparation).
- McAfee, R. Preston, Reny, Philip J., 1992. *Correlated information and mechanism design*. *Econometrica* 60, 395–421.
- Morris, Stephen, Shin, Hyun Song, 2006. *Global games: theory and applications*. In: Dewatripont, M., Hansen, L., Turnovsky, S. (Eds.), *Advances in Economics and Econometrics, the Eighth World Congress*. Cambridge University Press.
- Morris, Stephen, 1995. *The common prior assumption in economic theory*. *Econ. Philos.* 11, 227–253.
- Myerson, Roger B., Satterthwaite, Mark A., 1983. *Efficient mechanisms for bilateral trading*. *J. Econ. Theory* 29, 265–281.
- Nalebuff, Barry, 1987. *Credible pretrial negotiation*. *Rand J. Econ.* 18, 198–210.
- Posner, Richard A., 1973. *An economic approach to legal procedure and judicial administration*. *J. Legal Stud.* 2, 399–458.
- Prescott, J.J., Spier, Kathryn E., 2016. *A comprehensive theory of civil settlement*. *N.Y.U. Law Rev.* 91, 59–143.
- Priest, George, Klein, Benjamin, 1984. *The selection of disputes for litigation*. *J. Legal Stud.* 13, 1–55.
- Reinganum, Jennifer F., Wilde, Louis L., 1986. *Settlement, litigation, and the allocation of litigation costs*. *Rand J. Econ.* 17, 557–566.
- Rubinstein, Ariel, 1982. *Perfect equilibrium in a bargaining model*. *Econometrica* 50, 97–109.
- Schweizer, Urs., 1989. *Litigation and settlement under two-sided incomplete information*. *Rev. Econ. Stud.* 56, 163–177.
- Shapiro, Fred, Pearce, Michelle, 2012. *The most-cited law review articles of all time*. *Mich. Law Rev.* 110, 1483–1520.
- Shapiro, Fred, 1996. *The most-cited law review articles revisited*. *Chicago-Kent Law Rev.* 71, 751–779.
- Shavell, Steven, 1982. *Suit, settlement and trial: a theoretical analysis under alternative methods for the allocation of legal costs*. *J. Legal Stud.* 11, 55–81.
- Shavell, Steven, 1996. *Any frequency of plaintiff victory at trial is possible*. *J. Legal Stud.* 25, 493–501.
- Siegelman, Peter, Donohue III, John J., 1995. *The selection of employment discrimination suits for litigation: using business cycle effects to test the Priest-Klein Hypothesis*. *J. Legal Stud.* 24, 427–462.
- Spier, Kathryn E., Prescott, J.J., 2016. *Contracting on litigation*. Unpublished., <http://dx.doi.org/10.2139/ssrn.2765033>, Available at SSRN: <http://ssrn.com/abstract=2765033>.
- Spier, Kathryn E., 1992. *The dynamics of pretrial negotiation*. *Rev. Econ. Stud.* 59, 93–108.
- Spier, Kathryn E., 1994. *Pretrial bargaining and the design of fee-shifting rules*. *Rand J. Econ.* 25, 197–214.
- Waldfogel, Joel., 1995. *The selection hypothesis and the relationship between trial and plaintiff victory*. *J. Polit. Econ.* 103, 229–260.
- Waldfogel, Joel, 1998. *Reconciling asymmetric information and divergent expectations theories of litigation*. *J. Law Econ.* 41, 451–476.
- Watanabe, Yasutora, 2005. *Learning and bargaining in dispute resolution: theory and evidence from medical malpractice litigation*. Northwestern University Mimeo.
- Wittman, Donald, 1985. *Is the selection of cases for trial biased?* *J. Legal Stud.* 14, 185–214.
- Yildiz, Muhamet, 2003. *Bargaining without a common prior: an immediate agreement theorem*. *Econometrica* 71, 793–811.
- Yildiz, Muhamet, 2004. *Waiting to persuade*. *Q. J. Econ.* 119, 223–248.