Inferences from Litigated Cases
Author(s): Daniel Klerman and Yoon-Ho Alex Lee
Source: *The Journal of Legal Studies*, Vol. 43, No. 2 (June 2014), pp. 209-248
Published by: The University of Chicago Press for The University of Chicago Law School
Stable URL: http://www.jstor.org/stable/10.1086/678236
Accessed: 29/10/2014 12:07

Inferences from Litigated Cases

Daniel Klerman and Yoon-Ho Alex Lee

ABSTRACT

Priest and Klein argued in 1984 that, because of selection effects, the percentage of litigated cases won by plaintiffs will not vary with the legal standard. Many researchers thereafter concluded that one could not make valid inferences about the character of the law from the percentage of cases plaintiffs won, nor could one measure legal change by observing changes in that percentage. This article argues that, even taking selection effects into account, one may be able to make valid inferences from the percentage of plaintiff trial victories, because selection effects are partial. Therefore, although selection mutes changes in the plaintiff trial win rate, it does not make the win rate completely invariant to legal change. This article shows that inferences from litigated cases may be possible under the standard screening and signaling models of settlement, as well as under Priest and Klein's original divergent-expectations model.

1. INTRODUCTION

For almost 3 decades since the publication of Priest and Klein's (1984) influential article on the selection of disputes for litigation,[1] sophisticated

DANIEL KLERMAN is Charles L. and Ramona I. Hilliard Professor of Law and History, Gould School of Law, University of Southern California. YOON-HO ALEX LEE is Assistant Professor of Law, Gould School of Law, University of Southern California. The authors thank Phillip Bond, Andrew Daughety, John Donohue, Lee Epstein, Daniel Friedman, Jonah Gelbach, Gillian Hadfield, William Hubbard, Louis Kaplow, Bart Kosko, Tom Lyon, John Matsusaka, Jennifer Reinganum, Daria Roithmayr, Steven Shavell, Kathryn Spier, Eric Talley, and Donald Wittman; participants at the 2013 Law and Economic Theory Conference, the 2013 Conference on Empirical Legal Studies, the 2014 American Law and Economics Association annual meeting, and DonaldFest (Economics Department, University of California, Santa Cruz, October 27, 2012); and participants at the Tel Aviv Law School law and economics and the University of Southern California's Center for Law and Social Science and Southern California Empirical Legal Studies workshops for their comments and suggestions. The authors also thank Jeffrey Gilbert for his work with Mathe-

[*Journal of Legal Studies*, vol. 43 (June 2014)]
 0047-2530/2014/4302-0008$10.00

empiricists have hesitated to do research on outcomes in litigated cases. Priest and Klein argued that there was a tendency for plaintiffs to win 50 percent of the time regardless of whether the legal standard favored plaintiffs or defendants. To the extent that there were deviations from the 50-percent prediction, they were explained as resulting from asymmetric stakes or other factors. Changes in the legal standard would not have observable effects on plaintiff win rates. Many researchers thereafter concluded that one could not make valid inferences about the character of the law from the percentage of cases that plaintiffs won, nor could one measure legal change by observing changes in that percentage. Those who analyzed plaintiff win rates faced persistent questions about the validity of their findings. In some ways, this reluctance to use win-loss data was salutary, because it encouraged legal empiricists to search for other data—such as accident and crime rates—that are more directly relevant to the impact of law on social welfare.

Given that settlements are the outcome of the parties' choices rather than random factors, it is not surprising that trial data reflect significant selection effects. Nevertheless, Priest and Klein's article is startling in suggesting that selection bias is so strong that a change in the legal standard would result in no observable change in the plaintiff trial win rate. The strength and importance of this claim merit reexamination of its theoretical grounding and of the assumptions necessary to sustain it.

This article argues that, under all standard settlement models and under a wide range of reasonable conditions, one may be able to make valid inferences from the percentage of plaintiff trial victories. It therefore hopes to open up avenues for empirical research that have previously been neglected as unfruitful and to give legitimacy to those who analyzed plaintiff victory rates in spite of Priest and Klein's arguments.

This article is thus consistent with research that finds that trial win rates vary with judicial characteristics, legal standards, and other factors

---

matica and Frederick Boehmke for writing the Stata code that they modified to produce the simulations.

1. According to several citation studies, Priest and Klein (1984) is one of the most influential law articles of all time (Shapiro and Pearse [2012], rank it 61st of all law review articles, and Landes and Posner [1995, p. 838], rank it 28th of all articles in "predicted 'lifetime' citations"). The literature on the selection of disputes, which starts with Priest and Klein's article, is immense. The most recent contributions examine the selection implications of pretrial motions (see Gelbach 2012; Hubbard 2013). There is also a large empirical literature testing the hypothesis (see Kessler, Meites, and Miller 1996; Klerman 2012; Waldfogel 1995, 1998). For theoretical articles refining the theory of the selection of disputes, see Wittman (1985), Friedman and Wittman (2007), Shavell (1996), Hylton (1993), and Eisenberg and Farber (1997).

that affect case strength. For example, Siegelman and Donohue (1995) find that plaintiff win rates in employment discrimination cases vary with the business cycle and are lower in recessions, when litigants with weaker cases would be more likely to sue. While Siegelman and Donohue (1995, p. 427) find selection effects, they conclude that "the settlement process does not produce complete selection." Such partial selection effects are predicted by this article's analysis of asymmetric-information models and the Priest-Klein model. Similarly, Epstein, Landes, and Posner (2013), Eisenberg and Johnson (1991), and Kulik, Perry, and Pepper (2003) find that plaintiff win rates in federal district court vary with case characteristics and/or judicial characteristics such as age, race, and the party of the appointing president. The fact that others, such as Ashenfelter, Eisenberg, and Schwab (1995), do not find that judicial characteristics influence plaintiff win rates is not inconsistent with this article. In fact, if Priest and Klein (1984) are correct that plaintiff trial win rates are invariant to judicial characteristics, a finding that plaintiff trial win rates do not vary would not shed light on whether judicial background matters. Rather, one can infer that judicial characteristics are irrelevant from similar plaintiff victory rates only if selection effects are not so strong that they would mask the impact of judicial characteristics, if they existed.

This article first analyzes selection effects under asymmetric-information models. Priest and Klein's article was based on the divergent-expectations model of suit and settlement pioneered by John Gould, William Landes, and Richard Posner in the 1970s. Starting with seminal articles by P'ng (1983) and Bebchuk (1984), sophisticated economic analysis of litigation has favored models based on asymmetric information. Hylton (1993) and Shavell (1996) analyze selection effects under the screening model and show that the 50-percent prediction of Priest and Klein (1984) does not hold. Waldfogel (1998) tests the relative predictive power of the divergent-expectations and asymmetric-information models. Nevertheless, Hylton, Shavell, and Waldfogel do not examine the relationship between the decision standard and plaintiff trial win rates, nor do they analyze the signaling model. This article shows that, under a wide array of plausible assumptions, both the screening and the signaling models predict that plaintiff trial victory rates will vary with legal standards in the way that one would expect: a more pro-plaintiff legal standard of liability leads to a larger percentage of plaintiff trial victories. Under these asymmetric-information models, when the defendant has superior information, selection means that cases in which the plaintiff's case is weaker are more likely to be litigated. Nevertheless, because a change in the legal standard alters the probability that plaintiffs will prevail in weak as well as strong cases, legal changes may change the plaintiff trial win rate in the expected way. Conversely, when the plaintiff has superior information, it is more likely to litigate stronger cases. Nevertheless, because legal change alters the chances that plaintiffs will win even in stronger cases, the plaintiff trial win rate may vary in predictable ways.

Second, this article examines selection effects under the Priest-Klein divergent-expectations model of settlement. Priest and Klein's 50-percent prediction, while mathematically valid under the assumptions of their model—see Lee and Klerman (2014)—is a limiting result. According to the model, the plaintiff and defendant make unbiased predictions about trial outcomes, but the parties' predictions are not perfectly accurate. When their predictions diverge sufficiently, they cannot settle. As the variance of the parties' prediction errors goes to 0, they fail to settle only the closest cases, and the percentage of plaintiff trial victories converges to 50 percent. Nevertheless, for empirical work, this limiting result, and the more general result that plaintiff win rates do not vary with the legal standard, is not necessarily relevant, because as the variance of the parties' prediction errors goes to 0, the number of litigated cases also goes to 0. Thus, whenever one is doing empirical work on litigated cases, one is necessarily dealing with a situation in which prediction errors are positive. When prediction errors are positive, close cases are more likely to be litigated, but there is also some randomness, which means that any case might be litigated. As a result, the percentage of plaintiff trial victories reflects not just the 50 percent probability that plaintiffs will win close cases but also the full array of factors that influence plaintiff victories in other cases, such as the content of the law and judicial characteristics. An analytic proof shows that, all other things being equal, a pro-plaintiff shift in the law increases the percentage of cases won by plaintiffs. Priest and Klein recognize that the plaintiff win rate will vary substantially with the decision standard if trial rates are high. Reexamination of the model's parameters suggests that the plaintiff win rate will vary substantially even with trial rates below 2 percent. Our analysis of the Priest-Klein model extends Waldfogel (1995), which uses simulations and structural estimation to draw inferences about the decision standard from trial win rates. Unlike Waldfogel (1995), however, we provide an analytic proof of the relationship between the decision standard and plaintiff trial win rates instead of relying on simulations.

To simplify matters, this article focuses on the effect of legal change. Nevertheless, it should be noted that the argument also applies to different decision makers. For example, if one judge is more pro-plaintiff than another, the Priest-Klein hypothesis would assert that parties would take judicial preferences and biases into account when settling, so one would expect that plaintiffs would win 50 percent of the time regardless of which judge heard the case. Priest and Klein would predict a 50 percent plaintiff trial win rate even if cases were randomly assigned to judges, as long as the assignment was sufficiently in advance of trial that the parties could settle after knowing the identity of the judge. In contrast, the analysis in this article suggests that, controlling for other relevant factors, plaintiffs may win a greater fraction of the cases before the more pro-plaintiff judge (see, for example, Kulik, Perry, and Pepper 2003). Similar analysis would apply to different juries, different procedures that might favor the plaintiff or defendant, or different case compositions (see, for example, Siegelman and Donohue 1995). Priest and Klein would predict no effect on trial outcomes, whereas this article predicts that stronger groups of cases or cases facing more pro-plaintiff juries or procedures may produce higher rates of plaintiff trial victories.

The Priest-Klein hypothesis also suggests that one cannot ascertain whether particular factors—such as whether the plaintiff is Hispanic or whether the defendant acted in bad faith—affect verdicts, because parties will take those factors into account when settling. As a result, Priest and Klein would predict that case characteristics would have no effect on the percentage of cases won by plaintiffs, unless they affect the asymmetry of the stakes. For that reason, one could not determine which factors influence jurors or judges by running regressions in which the dependent variable is the case outcome and the independent variables are factors that might influence the outcome. In contrast, this article suggests that, controlling for other relevant factors, such regressions would yield coefficients of the proper sign (see, for example, Eisenberg and Johnson 1991). Selection effects would cause the coefficients to be closer to 0. That is, selection means that regressions will tend to underestimate the magnitude of the coefficients (and their statistical significance). In other words, real-world effects would be stronger than the regression coefficients would otherwise suggest.

Finally, this article examines the effects of legal changes relating to damages, such as imposition of a cap on damages or allowance of a damages multiplier. It concludes that, unlike a shift in the standard of liability, legal changes relating to the calculation of damages produce more ambiguous effects.

Although this article suggests that plaintiff trial win rates can provide useful information about the law, decision makers, and legal decision making, it is important to emphasize that such inferences must be made cautiously. First, nothing in this article casts doubt on Priest and Klein's central insight that litigated cases are a selected, nonrandom sample of all disputes. The characteristics of litigated cases do deviate significantly from those of settled cases (see Klerman 2012). Second, changes in legal rules, decision makers, or case characteristics will produce predictable changes in the percentage of plaintiff trial victories only holding other factors equal. If other factors change—such as the distribution and characteristics of the underlying disputes, levels of uncertainty, the asymmetry of the stakes, or the distribution of information—then the effect of shifting the legal standard on the percentage of plaintiff victories may well be swamped by the effects of these other factors (see Priest 1987). If these other factors are not carefully considered, changes in the percentage of plaintiff judgments may be falsely attributed to changes in the legal standard or differences in decision makers or case characteristics. Third, this article assumes the validity of the standard settlement models. As a result, whatever criticisms that can be mounted against these models will carry over to our analysis.

Sections 2, 3, and 4 analyze the selection implications of the screening model, the signaling model, and the Priest-Klein divergent-expectations model, respectively. Section 5 discusses the effect of changes relating to damages. Section 6 discusses limitations and caveats. Section 7 concludes.

2. THE SCREENING MODEL

The Priest-Klein model has been criticized as lacking proper game-theoretic rigor, because the parties do not take into account the fact that the other side has information about likely trial outcomes and because the parties do not bargain strategically. Around the same time that Priest and Klein wrote their article, economists began applying modern theories of bargaining under asymmetric information to the problem of litigation. Under these models, the uninformed party makes a take-it-or-leave-it offer to the informed party (see Bebchuk 1984).

In this section we show that, under the Bebchuk screening model, the percentage of plaintiff trial victories will vary predictably with the legal standard. In short, a decision standard that is sufficiently more pro-plaintiff will result in a greater percentage of plaintiff victories at trial than a more pro-defendant standard. We describe the model under the assumption that defendants possess private information about the suit and explore the implications for the selection of suits for litigation. As shown in Section A of the online appendix, the results are similar when plaintiffs possess private information.

Section 2.1 first uses a simple discrete distribution to illustrate the effect of changing the legal standard. Section 2.2 then proves more generally that, under a wide range of plausible distributions, a shift in the legal standard changes the plaintiff trial win rate in the predicted way. Finally, Section 2.3 explores the magnitude of the changes by examining selection effects for a beta distribution.

2.1. Illustration of Inferences under the Screening Model with a Discrete Distribution

The easiest way to see the selection implications of the screening model is to assume a very simple discrete distribution. Suppose, for example, that the parties are risk neutral and agree that damages are \$100, each side has litigation costs of \$10 if the case goes to trial, and there are two types of defendants. High-liability defendants lose their cases with a probability of 70 percent, while low-liability defendants lose their cases with a probability of 30 percent. There are equal proportions of each type of defendant in the total population. Defendants know their types, but plaintiffs do not. The plaintiff makes a take-it-or-leave-it offer to the defendant.

Under these assumptions, the highest offer that a high-liability defendant will accept is \$80, which is its \$10 litigation cost plus its expected payout to the plaintiff (\$70 = 70 percent × \$100). The highest offer that a low-liability defendant will accept is \$40, which is its \$10 litigation cost plus its expected payout to the plaintiff (\$30 = 30 percent × \$100). Plausible candidates for the plaintiff's best strategy are to settle with both kinds of defendants (by offering \$40) or to settle with just the high-liability defendant (by offering \$80). Under the parameters assumed here, the plaintiff's payoff is higher when it offers \$80. That offer induces a separating equilibrium in which high-liability defendants settle and low-liability defendants litigate. Since only low-liability defendants litigate, plaintiffs win 30 percent of the cases that make it to trial.

Now assume that a legal change increases the plaintiff's probability of prevailing against both types so that high-liability defendants lose with a probability of 80 percent while low-liability defendants lose with a probability of 40 percent. Now the plaintiff's optimal offer is $90. Again, only low-liability defendants will litigate, but, because of the legal change, the plaintiff now wins 40 percent of litigated cases. Thus, a pro-plaintiff legal change causes the plaintiff trial win rate to increase from 30 percent to 40 percent. There is clearly selection: only low-liability defendants litigate under either legal standard. Nevertheless, the plaintiff trial win rate changes with the legal standard because the legal change increases the probability that the plaintiff prevails against both types.[2]

2.2. A General Proof of Inferences under the Screening Model

The example in Section 2.1 generalizes to continuous distributions of defendant types. The first part of this section follows Shavell (1996). Suppose that defendants know the probability p that they will lose if their case goes to trial, but plaintiffs do not know that probability. Defendants vary in the probability that they will lose, and a defendant for whom the probability of losing is p will be called a defendant of type p. Probability p is distributed on the interval $[\underline{p}, 1]$, where $0 < \underline{p} < 1$.[3] In order to assure that the plaintiff has a credible threat to go to trial against any defendant, it is assumed that $\underline{p} > C_\pi/J$, where C_π is the plaintiff's trial costs and J is the damages that the defendant will pay if the plaintiff prevails.[4] The term C_δ is the defendant's trial costs. Both parties are risk neutral.

2. In the discrete case, what is important is that the legal change increases the probability with which the plaintiff prevails against low-liability types. If the plaintiff's probability of prevailing against high-liability types had remained the same or even dipped a little, the plaintiff's trial victory rate would still rise from 30 percent to 40 percent. Nevertheless, if the probability that the plaintiff prevailed against the high-probability types dipped sufficiently, the separating equilibrium would collapse, and the plaintiff would find it worthwhile to settle with all defendants; in that case it would not be possible to calculate the plaintiff's trial win rate, because there would be no trials.

3. Bebchuk (1984) and Shavell (1996) call the upper and lower bounds a and b. To make the notation easier to remember, we call the lower bound $\underline{p}$. We also assume that the upper bound is 1 and assume that $f(p)$ takes on a positive value up to 1. This simplifies the proofs but is not a significant change since we can always assume that values of $f(p)$ are arbitrarily small after some threshold value of p.

4. Nalebuff (1987) considers the implications of relaxing the requirement that $\underline{p} > C_\pi/J$. In order to assure that it has a credible threat to go to trial against a defendant who does not settle, the plaintiff in Nalebuff's model sometimes needs to increase its settlement offer. This higher settlement offer is necessary when a certain condition, which Nalebuff calls condition 2, is satisfied. If condition 2 is not satisfied, then the analysis and conclusions of the Bebchuk/Shavell model remain valid and, as shown in Section 2.2, a sufficiently more pro-plaintiff legal standard leads to a higher plaintiff trial win rate. If condition 2 is

The probability density of defendant types will depend on the legal standard. Suppose that there are two legal standards. Each legal standard produces a distribution of defendant types represented by the probability density functions $f(p)$ and $g(p)$, distributed on the interval $[\underline{p}, 1]$. Let $F(p)$ and $G(p)$ be the respective cumulative distribution functions. Without loss of generality, we assume that $f(p)$ and $g(p)$ are differentiable and thus continuous, that $f(\underline{p}) = g(\underline{p}) = 0$, and that $F(p)$ and $G(p)$ are strictly increasing over $[\underline{p}, 1]$, which implies that $f(p) > 0$ and $g(p) > 0$ on $(\underline{p}, 1)$. These assumptions simplify the proofs but are not restrictive, because any functions not meeting those criteria can be approximated with any desired precision by functions that do meet those criteria. For example, a discontinuous function with a finite number of discontinuities can be approximated with any desired precision by a continuous function.[5] Similarly, the assumption that $f(p)$ and $g(p)$ share a lower bound $\underline{p}$ is not restrictive, because, if one wanted to consider a function $g(p)$ with a higher lower bound, one could approximate it to any desired degree of precision by assuming that $g(p)$ takes an arbitrarily low value between $\underline{p}$ and the desired lower bound.

Under the screening model of litigation, if the defendant has superior information, the plaintiff makes a take-it-or-leave-it settlement offer x to the defendant. If the defendant accepts, it pays x to the plaintiff, and the case is over. If the defendant rejects the offer, the case goes to trial. A rational defendant of type p accepts the offer if and only if $x \leq pJ + C_\delta$. Equivalently, the defendant settles if and only if $p \geq (x - C_\delta)/J$. So defendants with stronger cases (those that the defendant is likely to win) litigate, and defendants with weaker cases settle. If the case settles, the plaintiff gets the settlement x; if the case goes to trial, the plaintiff's expected recovery is $pJ - C_\pi$. The plaintiff selects its settlement offer to maximize its expected recovery. If defendants are distributed according

---

satisfied, the plaintiff trial win rate is simply C_π/J, which is lower than if the condition is not satisfied and will not vary with the legal standard (unless litigation costs also vary with the litigation standard). The Nalebuff model thus gives different results than the Bebchuck/Shavell model only when the plaintiff trial win rate is relatively low. In addition, if a pro-plaintiff legal change is sufficiently large, condition 2 would no longer be satisfied, and the plaintiff trial win rate would rise. Nalebuff's model does not suggest that inferences are not possible. It will still be true that if one observes an increase in the plaintiff win rate (all else equal), one can infer a pro-plaintiff change in the legal standard. But, unlike the Bebchuk/Shavell model, the Nalebuff model does not imply the converse. Some legal changes (those for which condition 2 is satisfied before and after the change) will not result in changes in the plaintiff trial win rate.

5. This is similar to Shavell (1996, p. 497 n.12): "[T]he discrete case (of interest in its own right) can always be approximated as closely as desired by a continuous density."

to probability density function $f(p)$, then the plaintiff's expected recovery is

$$\int_{\underline{p}}^{(x-C_\delta)/J} (pJ - C_\pi)f(p)dp + \left[1 - F\left(\frac{x - C_\delta}{J}\right)\right]x. \quad (1)$$

The first term in expression (1) represents the plaintiff's expected recovery from litigation, and the second term represents the expected recovery from settlement. If the plaintiff's optimal offer is interior to $[\underline{p}, 1]$, then the optimal offer x^* is determined by the first-order condition

$$1 - F\left(\frac{x - C_\delta}{J}\right) = \frac{[f((x - C_\delta)/J)(C_\pi + C_\delta)]}{J}. \quad (2)$$

Let $P_t(f(p))$ denote the plaintiff trial win rate under the legal standard $f(p)$. Then

$$P_t(f(p)) = \frac{\int_{\underline{p}}^{(x^*-C_\delta)/J} pf(p)dp}{F((x^* - C_\delta)/J)}. \quad (3)$$

This expression can be thought of as the weighted average value of p, the plaintiff's probability of prevailing, for defendants who litigate. If defendants are distributed according to probability density function $g(p)$, expressions (1), (2), and (3) are the same, except one substitutes $g(p)$ and $G(p)$ for $f(p)$ and $F(p)$.

Although it is not strictly necessary, it simplifies the proofs greatly if there is only one settlement offer that satisfies the first-order condition in equation (2). As discussed further below, it is sufficient for that purpose to assume that $f(p)$ and $g(p)$ have strictly increasing hazard rates, $f(p)/[1 - F(p)]$ and $g(p)/[1 - G(p)]$ (see Nalebuff 1987). The assumption of increasing hazard rates is stronger than necessary but not unduly restrictive, as most familiar distribution functions exhibit increasing hazard rates.[6]

We now define what it means for one legal rule to be sufficiently more pro-plaintiff than another. The idea is that a more pro-plaintiff legal rule assigns at least as high and sometimes a higher probability of plaintiff success to any factual situation. That is, if rule g is more pro-plaintiff than rule f, and if a plaintiff with a particular set of facts has

6. Finite distributions with increasing hazard rates include the beta distribution (for $\alpha > 1$, $\beta > 1$), uniform distribution, and rising and falling triangle distributions. More generally, any probability density function that is log concave has an increasing hazard rate (see Bagnoli and Bergstrom 2005). For another article using the screening model and assuming an increasing hazard rate, see Spier (1992, p. 96).

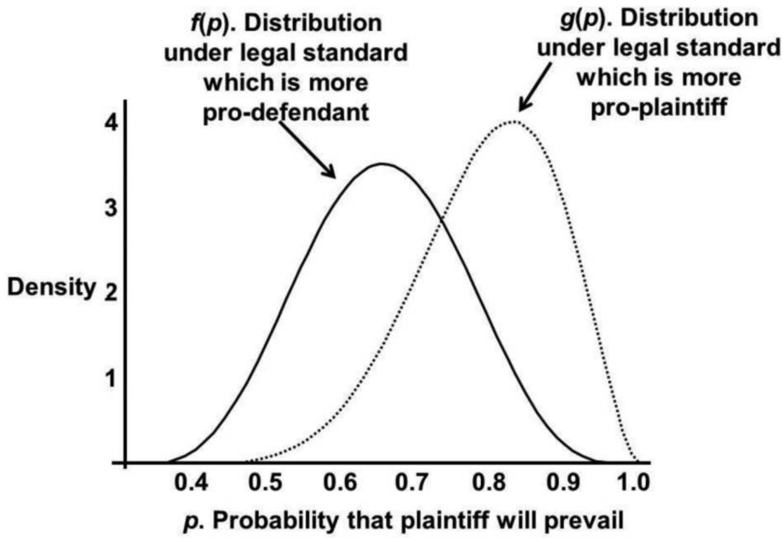

Figure 1. Two distributions of p, the probability that the plaintiff will prevail

probability of success p under rule f, then a plaintiff with the same fact pattern has probability of success $p' \geq p$ under rule g. It is also sensible to think that a more pro-plaintiff rule preserves the ordering of fact patterns. That is, if the probability of plaintiff success under rule f is higher for fact pattern B than for fact pattern A, then the probability of plaintiff success under rule g is also higher for fact pattern B than for fact pattern A. These assumptions suggest that a more pro-plaintiff rule can be represented by a probability distribution that is shifted to the right, as illustrated in Figure 1.

Of course, the shift cannot preserve the shape of the distribution precisely, because if it did, there would be fact patterns for which the plaintiff's probability of success is greater than 100 percent, which is impossible. Therefore, the more pro-plaintiff density needs to be a bit more spread apart and thus lower at lower probabilities and compressed and thus higher at higher probabilities, a phenomenon again illustrated by Figure 1.

This idea of a more pro-plaintiff rule can be captured by the monotone likelihood ratio. That is, $g(p)$ is the probability density function associated with a legal rule that is more pro-plaintiff than the legal rule associated with probability density function $f(p)$ if and only if $g(p_0)/f(p_0)$

$\leq g(p_1)/f(p_1)$ whenever $\underline{p} < p_0 < p_1 < 1$. Equivalently, $g(p)$ is the probability density function associated with a legal rule that is more pro-plaintiff than the legal rule associated with probability density function $f(p)$ if and only if $(d/dp)(g(p)/f(p)) \geq 0$ (assuming that the two functions are not in fact identical). When the monotone likelihood ratio property is satisfied, $f(p) > g(p)$ when p is small, but $g(p) > f(p)$ when p is large, which accords with the idea that under a more pro-plaintiff legal standard, there are fewer cases for which the plaintiff has a low probability of success and more cases for which the plaintiff has a higher probability of success. Many families of distributions have the monotone likelihood ratio property.[7] The fact that $f(p)$ and $g(p)$ have the monotone likelihood ratio property implies that $g(p)$ stochastically dominates $f(p)$. Stochastic dominance by itself, however, does not imply that the monotone likelihood ratio property is satisfied and is not sufficient to ensure that the plaintiff's win rate in litigated cases is higher (see Section B of the online appendix). Because stochastic dominance is not sufficient for the asymmetric-information models, our propositions for both the screening and signaling models are phrased in terms of a sufficiently more pro-plaintiff legal standard.[8]

The proposition to be proved can be stated as follows:

Proposition 1: Inferences under the Bebchuk Screening Model When the Defendant Has Superior Information. When the defendant has the informational advantage, the plaintiff's trial win rate under the screening model is strictly higher under a sufficiently more pro-plaintiff legal standard. Specifically, if (i) $f(p)$ and $g(p)$ are distinct, differentiable probability density functions on the interval $[\underline{p}, 1]$, $f(\underline{p}) = g(\underline{p}) = 0$, and $C_\pi/J < \underline{p} < 1$; (ii) $f(p)$ and $g(p)$ have increasing hazard rates and satisfy the monotone likelihood ratio property on $(\underline{p}, 1)$, $(d/dp)(g(p)/f(p)) \geq 0$; and (iii) $F(p)$ and $G(p)$, the associated cumulative distribution functions, are strictly increasing over $[\underline{p}, 1]$, then there exist unique solutions x_f^* and x_g^* to the first-order conditions for the plaintiff's optimal settlement offer, and

7. These include, for example, the exponential, binomial, Poisson, normal, beta (if $\alpha + \beta$ is constant), uniform, rising triangle, and falling triangle distributions. However, only those with a finite support are relevant for this model.

8. We also investigated the monotone probability ratio, a distribution property intermediate in stringency between stochastic dominance and the monotone likelihood ratio, but were able neither to prove our propositions using that property nor to find a counterexample.

$$P_t(f(p)) = \frac{\int_{\underline{p}}^{p^*{}_f} uf(u)du}{F(p_f^*)} < \frac{\int_{\underline{p}}^{p^*{}_g} ug(u)du}{G(p_g^*)} = P_t(g(p)),$$

where $p_f^* = (x_f^* - C_\delta)/J$ and $p_g^* = (x_g^* - C_\delta)/J$.

Section A.1 of the Appendix contains the proof. Proposition 1 states sufficient conditions for a more pro-plaintiff distribution to result in a higher rate of plaintiff victories at trial. Less restrictive conditions could be formulated, but they would complicate the statement of the proposition and its proof, while adding little insight.[9]

2.3. Empirical Relevance

The fact that a more pro-plaintiff legal standard can result in a greater percentage of plaintiff trial victories suggests that empirical work using the percentage of trial victories as a dependent variable may be possible. Nevertheless, it is also important to verify that the changes in the percentage of plaintiff trial victories can be large enough to be detected empirically. To do that, we solve the model using beta distributions.

The beta distribution is a family of continuous probability distributions on a closed interval.[10] For the interval $[\underline{p}, 1]$, the beta distribution is

$$B(x;\, \alpha,\, \beta) = \frac{(x - \underline{p})^{\alpha-1}(1 - x)^{\beta-1}}{(1 - \underline{p})^{\alpha+\beta-1}\int_0^1 u^{\alpha-1}(1 - u)^{\beta-1}du}.$$

Depending on the parameters α and β, the beta distribution takes various shapes. When $\alpha > 1$ and $\beta > 1$, the distribution is single peaked. If $\alpha = \beta$, then the distribution is symmetrical, and it looks like a normal distribution confined to the interval $[\underline{p}, 1]$. In Figure 1, $f(p)$ illustrates a symmetrical beta distribution with $\alpha = \beta = 5$. If $\alpha > \beta$, the distribution looks like a negatively skewed normal distribution on the interval $[\underline{p}, 1]$. In Figure 1, $g(p)$ is negatively skewed with $\alpha = 7$ and $\beta = 3$. Conversely, if $\alpha < \beta$, the distribution is positively skewed. If the parameters α and β are exchanged, then the resulting distribution is a mirror image. For example, if $\alpha = 3$ and $\beta = 7$, the distribution would be similar to $g(p)$ in Figure 1, except positively skewed. Beta distributions are very

9. For example, the result does not require the monotone likelihood ratio property but only that $g(p)$ stochastically dominate $f(p)$ for all right-truncated distributions. In addition, it is not necessary that the monotone likelihood ratio or increasing hazard rate conditions are satisfied for all p, but rather they need be satisfied only over particular intervals. Similarly, the assumptions of continuity could be relaxed, but then breaks would need to be analyzed separately.

10. We thank Bart Kosko for encouraging us to use the beta distribution.

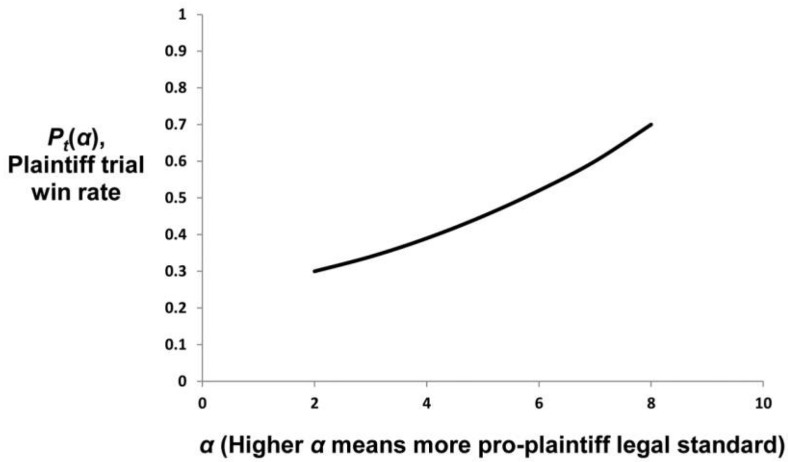

Figure 2. Legal standard (α) and percentage of plaintiff trial victories (p_t): screening model

useful, because they allow one to model the effect of shifting the legal rule or standard by varying the parameters α and β. For example, $\alpha + \beta = 10$, $\alpha > 1$, and $\beta > 1$ define a family of unimodal distributions in which the higher α is, the more pro-plaintiff the legal rule.

Figure 2 shows how the percentage of plaintiff victories varies with the legal standard within the beta distribution family $\alpha + \beta = 10$, $\alpha > 1$, $\beta > 1$ when $C_\pi = C_\delta = .3J$. The figure shows that, under the assumptions used in this section, as the legal standard becomes more pro-plaintiff (as α increases) the percentage of plaintiff trial victories (P_t) also increases. It also shows that the change in the percentage of plaintiff trial victories is potentially large—more than 40 percentage points. Of course, smaller changes in the legal standard will result in smaller changes in the percentage of plaintiff victories, and other distributions may produce larger or smaller changes in the plaintiff trial win rate. Nevertheless, the figure suggests that changes in the percentage of trial victories may be large enough to be observable in data sets of reasonable size.

3. THE SIGNALING MODEL

An alternative asymmetric-information model involves signaling rather than screening. Under the signaling model, the party with superior information makes the offer. Section 3.1 uses a simple discrete distribution to illustrate the intuition behind the effect of changing the legal standard under the signaling model. Section 3.2 shows more generally that when the monotone likelihood ratio property is satisfied, a shift in the legal standard changes the plaintiff trial win rate in the predicted way. Section 3.3 explores the empirical relevance of this finding using beta distributions.

3.1. Illustration of Inferences under the Signaling Model with a Discrete Distribution

The equilibrium and the selection effects of the signaling model are easiest to understand when there are only two types of defendants—low- and high-liability defendants. As in many signaling models, the equilibrium involves mixed strategies. In the situation in which the defendant has superior information about its probability of losing, the defendant offers the plaintiff a settlement equal to the plaintiff's expected recovery net of litigation costs, so the plaintiff is indifferent between accepting and rejecting the settlement offer. In order to make sure that high-liability defendants do not mimic low-liability defendants (and thus get off with low settlement payments), plaintiffs accept all high settlement offers but randomize and reject some low settlement offers. When the probability of rejection is set properly, high-liability defendants are indifferent between low and high offers, so it is incentive compatible for high-liability defendants to make high offers and low-liability defendants to make low offers. The result is similar to the screening model: all trials involve low-liability defendants, because all high-liability defendants settle while some low-liability defendants litigate. Therefore, the plaintiff trial win rate is the probability that the plaintiff prevails against a low-liability defendant. If the law shifts in a pro-plaintiff direction, the probability that the plaintiff prevails against each type increases, but the plaintiff trial win rate is still determined by the probability that the plaintiff prevails against a low-liability defendant. Since that probability is higher under the more pro-plaintiff legal rule, plaintiffs win more of the litigated cases under the more favorable legal standard.

3.2. A General Proof of Inferences under the Signaling Model

This section shows that the analysis of a simple discrete distribution of types (discussed in Section 3.1) remains valid when there is a continuum of types. Since the signaling model was first applied to a situation in which the plaintiff has superior information, that case is discussed here, and the situation in which the defendant has superior information is discussed in Section C of the online appendix.

In the canonical Reinganum and Wilde (1986) model, the parties agree on the probability that the plaintiff will prevail, p, but the plaintiff has private information about damages. The plaintiff offers the defendant a settlement that varies with the plaintiff's private information, and the defendant rejects the offer with a probability that increases with the amount of the settlement offer. Because, in this model, the informational asymmetry concerns damages rather than the probability of plaintiff victory, the implications for the observed rate of plaintiff victories in litigated cases are trivial: the plaintiff will prevail at rate p. If, as assumed above for the screening model, the effect of a legal change is to increase the probability that the plaintiff will prevail for each factual situation, a pro-plaintiff change in the law will increase p and thus increase the observed plaintiff trial victory rate.[11]

The Reinganum-Wilde model can be modified to make its assumptions more similar to that of the other settlement models discussed in this article and to make the selection effects more interesting. The modification also makes the model applicable to a wider range of disputes. As in the Bebchuk screening model, assume that the parties agree on damages J but that the plaintiff has private information about the probability that she will prevail at trial, p. Probability p is distributed on the interval $[\underline{p}, 1]$ according to a cumulative distribution function, $F(p)$. Like Reinganum and Wilde, we assume that this cumulative distribution function is strictly increasing[12] and that the probability is bounded below, $\underline{p} > C_\pi/J$, so as to rule out nuisance suits, where C_π and C_δ are the plaintiff's and defendant's litigation costs, respectively.[13] Let $C = C_\pi + C_\delta$. Both parties are risk neutral.

This modified model is solved in essentially the same way as the original model, and, for that reason, many steps are omitted. A strategy for the plaintiff is a function, $S = s(p)$, that specifies a settlement demand for each possible level of probability of prevailing at trial. A strategy

11. The effect of a legal change that increases damages is discussed in Section 5.

12. The assumption that the cumulative distribution function is strictly increasing (rather than weakly increasing) simplifies the math but does not impair the generality of the results, because any weakly increasing function can be approximated with any desired degree of precision by a strictly increasing function by assuming a very small positive slope over the relevant interval.

13. Reinganum and Wilde (1986, p. 559) make the more restrictive assumption that $\underline{p} \geq C/J$. They do so because they also analyze fee shifting. That more restrictive assumption is not necessary for this article.

for the defendant is a function, $\rho = r(S)$, that specifies the probability that the defendant rejects the demand S. Because the defendant does not know the true probability of the plaintiff's victory, p, he must form some conjectures or beliefs about p on the basis of the settlement demand S. Beliefs $b(S)$ assign a unique type of plaintiff (probability of plaintiff's victory) to each settlement demand.

Following Reinganum and Wilde (1986), we define a separating equilibrium (b^*, r^*, s^*) as follows: (a) given the beliefs b^*, the probability of rejection policy $r^*(\cdot)$ maximizes the defendant's expected wealth; (b) given r^*, the settlement demand policy $s^*(\cdot)$ maximizes the plaintiff's expected wealth; and (c) $b(S) \in [\underline{p}, 1]$ for all S with $b^*(s^*(p)) = p$ for all $p \in [\underline{p}, 1]$—that is, the defendant must assign an existing plaintiff type to every demand S, and the beliefs must be correct for demands that are made in equilibrium. Let $\underline{S} = \underline{p}J + C_\delta$ and $\overline{S} = J + C_\delta$; then the "unique"[14] separating equilibrium is (i) $r^*(S) = 1$ for $S > \overline{S}$, $r^*(S) = 1 - \exp\{-(S - \underline{S})/C\}$ for $S \in [\underline{S}, \overline{S}]$, and $r^*(S) = 0$ for $S < \underline{S}$; (ii) $s^*(p) = pJ + C_\delta$ for $p \in [\underline{p}, 1]$; and (iii) $b^*(S) = 1$ for $S \geq \overline{S}$, $b^*(S) = (S - C_\delta)/J$ for $S \in [\underline{S}, \overline{S}]$, and $b^*(S) = \underline{p}$ for $S \leq \underline{S}$.

This equilibrium can be used to analyze the plaintiff's trial win rate under competing legal standards. As in Section 2, we assume that a legal standard characterized by probability density function $g(p)$ is more pro-plaintiff than one characterized by $f(p)$ if and only if $f(p)$ and $g(p)$ have the monotone likelihood ratio property.[15] Under the modified signaling model, the proposition to be proved can be stated as follows and is proved in Section A.2.

Proposition 2: Inferences under the Reinganum-Wilde Signaling Model When the Plaintiff Has Superior Information. When the plaintiff has the informational advantage, the plaintiff trial win rate under the signaling model is higher under a sufficiently more pro-plaintiff legal standard. Specifically, if (i) $f(p)$ and $g(p)$ are distinct, differentiable probability density functions on the interval $[\underline{p}, 1]$ with $f(\underline{p}) = g(\underline{p}) = 0$ and $\underline{p} > C_\pi/J$; (ii) $f(p)$ and $g(p)$ satisfy the monotone likelihood ratio property on $[\underline{p}, 1]$, $(d/dp)(g(p)/f(p)) \geq 0$; and (iii) $F(p)$ and $G(p)$, the associated cumulative distribution functions, are strictly increasing over $[\underline{p}, 1]$, then

14. As in Reinganum and Wilde (1986), "unique" is in quotation marks because beliefs could take on a range of values. Nevertheless, the settlement offers and their associated probabilities of rejection are truly unique.

15. As with the screening model, it is not sufficient that $g(p)$ stochastically dominate $f(p)$ (see Section D of the online appendix).

$$P_t(f(p)) = \frac{\int_{\underline{p}}^{1} r^*(s^*(p))pf(p)dp}{\int_{\underline{p}}^{1} r^*(s^*(p))f(p)dp} < \frac{\int_{\underline{p}}^{1} r^*(s^*(p))pg(p)dp}{\int_{\underline{p}}^{1} r^*(s^*(p))g(p)dp} = P_t(g(p)).$$

3.3. Empirical Relevance

As with the screening model (see Section 2.3), the empirical relevance of the results under the signaling model can be illustrated using the beta distribution. Figure 3 shows how the plaintiff trial win rate varies with the legal standard within the beta distribution family $\alpha + \beta = 10$, $\alpha > 1$, $\beta > 1$ when $C_\pi = C_\delta = .3J$.[16]

Like the corresponding figure for the screening model (Figure 2), Figure 3 shows that as the legal standard become more pro-plaintiff (as α increases) the percentage of plaintiff trial victories (P_t) also increases. Figure 3 also suggests that legal changes may produce changes in the plaintiff trial win rate large enough to be observable in data sets of reasonable size.

4. PRIEST-KLEIN DIVERGENT-EXPECTATIONS MODEL

Selection effects were initially studied under divergent-expectations models of settlement, and most later work has remained within that paradigm.[17] Under divergent-expectations models, both the plaintiff and the defendant estimate the likelihood of the plaintiff's victory with some error, and those estimates are unbiased. The errors have a mean of zero and are equal in magnitude. That is, unlike asymmetric-information models in which each party has different information, in divergent-expectations models, both sides have the same information or information of the same quality but make different predictions on the basis of that information. Their different predictions may reflect inconsistent priors (Daughety and Reinganum 2012).

So far, this article has focused on the selection implications of newer, asymmetric-information models of settlement, because those models have better game-theoretical grounding and are more widely accepted among sophisticated scholars. Nevertheless, in spite of the introduction of more rigorous models of settlement, Priest and Klein (1984) continues

16. The plaintiff trial win rate in Figure 3 was estimated using a discrete distribution of plaintiff types at .01 intervals. That is, it was assumed that plaintiff types were .30, .31, .32 . . . .98, .99, 1.00. The probability of each plaintiff type was calculated using the beta distribution. The plaintiff trial win rate was calculated for $\alpha \in \{1.01, 2, 3, 4, 5, 6, 7, 8, 8.99\}$. Although the graph looks linear, the relationship is not precisely linear.

17. See the references in note 1.

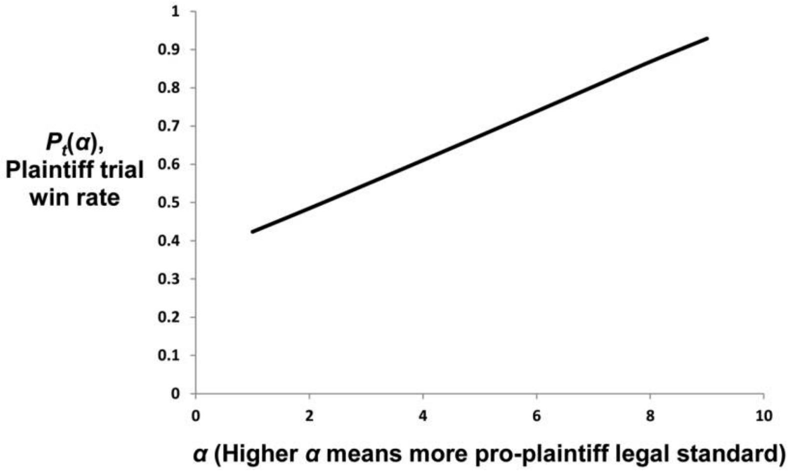

Figure 3. Legal standard (α) and percentage of plaintiff trial victories (p_t): signaling model

to be cited by sophisticated empiricists and in respected peer-reviewed journals (see, for example, Ashenfelter, Eisenberg, and Schwab 1995; Hubbard 2013; Gelbach 2012; Atkinson, Marco, and Turner 2009; Bernardo, Talley, and Welch 2000; Waldfogel 1995; Siegelman and Donohue 1995). More generally, Priest and Klein (1984) is one of the most influential law articles of all time, and its influence has continued to grow as empirical work on law has become more important (compare Shapiro and Pearse [2012] to Shapiro [1996]). Because of the historic importance and continuing influence of the Priest-Klein divergent-expectations model, this section shows that, even under that model, inferences from litigated cases may be possible.

Unlike asymmetric-information models, in which the parties start by estimating the plaintiff's probability of prevailing, Priest and Klein start with a measure of defendant fault, Y', that parties estimate with error. Priest and Klein then derive the parties' estimates of the plaintiff's probability of prevailing from their estimates of the defendant's fault. This setup introduces considerable complexity and meant that Priest and Klein (1984) could not provide an analytic proof of their key finding and Waldfogel (1995) could not derive closed-form expressions for the probability of trial or plaintiff trial victory. Lee and Klerman (2014) provide the first analytic proof.

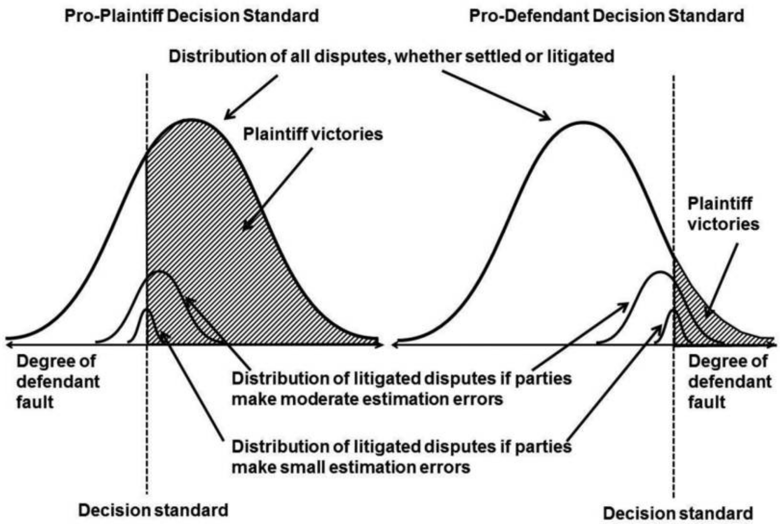

Figure 4. Illustrations of the Priest-Klein model

4.1. Informal Argument

Priest and Klein's argument for the 50-percent prediction is complex, and this is not the place to restate it fully (see generally Priest and Klein 1984; Waldfogel 1995; Lee and Klerman 2014). Figure 4 illustrates both the Priest-Klein hypothesis and the argument in this article.

Under the Priest-Klein model, all cases are arrayed along a horizontal line according to the defendant's degree of fault, with greater fault to the right. In Figure 4, the distribution of defendants is the largest bell curve, and it is the same on the left and right panels. Under this distribution, there are relatively few defendants with extremely low or high fault and higher densities of defendants with intermediate levels of fault. Nevertheless, as we show in a companion paper (Lee and Klerman 2014), the validity of the Priest-Klein hypothesis—that in the limit, the plaintiff trial win rate will be 50 percent—is not dependent on the shape of the distribution of disputes.[18]

18. Lee and Klerman (2014) show that, as long as the parties' error functions and stakes are symmetric and the density of disputes is continuous on an interval containing the decision standard, nonzero at the decision standard, and bounded above, the fraction of litigated cases won by the plaintiff will converge to 50 percent as the standard error of the parties' estimates goes to 0. The shape, symmetry, and slope of the distribution of potential cases are almost entirely irrelevant.

The law (or decision standard) divides defendants into two groups—those who would be found liable if the case went to trial and those who would be found not liable. The left panel of Figure 4 shows a relatively pro-plaintiff decision standard. The shaded area under the largest bell curve shows the proportion of disputes that would result in verdicts for the plaintiff if all disputes went to trial. That shaded area is larger than the unshaded area under the largest bell curve. This suggests that the percentage of plaintiff victories would be greater than 50 percent if all disputes went to trial. In the right panel, the decision standard is farther to the right and thus more favorable to defendants. As a result, the shaded area under the largest bell curve is smaller, which indicates that if all disputes went to trial, the plaintiff would win fewer than 50 percent of cases.

Of course, most disputes settle. The smallest bell curves in both panels of Figure 4 show the distribution of litigated cases if the plaintiff and defendant are very accurate in predicting trial outcomes. When the parties estimate trial outcomes very accurately, they litigate almost exclusively the hardest cases. That is, litigated cases are overwhelmingly those closest to the decision standard. As a result, the distribution of litigated cases is tightly centered around the decision standard, whether the decision standard is favorable to the plaintiff or defendant. As a result, plaintiffs win 50 percent of cases that go to trial. This is the key insight of Priest and Klein (1984).

The intermediate-sized bell curves in both panels of Figure 4 show the distribution of litigated cases if the parties are less accurate in predicting case outcomes. When parties' estimation errors are of intermediate size, they are still more likely to litigate close cases, but they also make more random errors, so the distribution of litigated cases is influenced by the distribution of all cases. Thus, the intermediate-sized bell curves, which represent litigated cases when parties have moderate estimation errors, have shapes that are centered closer to the decision standard than all cases (the largest bell curves) but are somewhat shifted from the decision standard in the direction of the full distribution of all cases. As a result, the shaded areas under the intermediate bell curves indicate that the percentage of litigated cases that plaintiffs win is influenced by the decision standard. When the law is more plaintiff friendly (left panel), the shaded area is larger, and the plaintiff wins more than 50 percent of cases. When the law is more defendant friendly (right panel), the shaded area is smaller, and the plaintiff wins less than 50 percent of cases.

4.2. General Proposition

The conclusions of the informal argument in Section 4.1 can be generalized in the following proposition, which is proved in Section A.3:

Proposition 3: Inferences under the Priest-Klein Model. Under the assumptions of the Priest-Klein model, in which parties predict a case's true merit with errors distributed according to a bivariate distribution with a mean of zero and a positive standard deviation that is the same under the two legal standards and in which the distribution of disputes is log concave with full support over the real line, the plaintiff trial win rate is strictly higher under the more pro-plaintiff legal standard.

Note that the proposition is far more general than the model in Priest and Klein (1984). Although Priest and Klein (1984) assume that the distribution of prediction errors is bivariate normal, proposition 3 assumes only that the bivariate distribution has a mean of zero and a positive standard deviation. In addition, the result is valid even when the stakes are asymmetric. In addition, the distribution of disputes need only be log concave. The class of probability density functions that are log concave is fairly large.[19] Nevertheless, as shown in Section E of the online appendix, the proposition would not hold if the probability density function were only quasi-concave. The assumption that the standard deviation of the parties' estimation errors does not change with the legal standard is restrictive. It could present problems if one examined plaintiff win rates right after a legal change, because new legal rules are likely to be uncertain, so the standard deviation would probably be higher (see Priest 1987). In contrast, if one were analyzing steady-state plaintiff win rates under the two legal standards, the assumption would be less problematic.

4.3. Empirical Relevance

Section 4.2 and Section A.3 show that a pro-plaintiff legal change could lead to an increase in the plaintiff trial win rate under the Priest-Klein model. This section shows that changes in the plaintiff trial win rate may be large enough to be observable in empirical studies of reasonable

19. These include normal distributions, skew normal distributions, generalized normal distributions, exponential distributions, uniform distributions over any convex set, logistic distributions, extreme-value distributions, Laplace distributions, and chi distributions (see generally Bagnoli and Bergstrom 2005; Mohtashami Borzadaran and Mohtashami Borzadaran 2011). Distributions with full support over the real line are most relevant for the Priest-Klein model.

Table 1. Effect of Changing the Decision Standard on the Plaintiff Trial Win Rate

Litigation Costs $((C - S)/J)$	Prediction Errors $(\sigma_{\pi,\delta})$	% Litigated	Decision Standard (Y^*) −1.0	−.5	.0	.5	1.0
.80	1.5	1.4	78.7	65.6	50.2	34.5	21.4
	.5	.7	65.4	57.9	49.9	42.1	34.6
	.1	.1	53.9	51.1	50.2	48.2	46.9
.67	1.5	4.0	79.0	65.8	50.0	34.3	20.9
	.5	1.9	66.1	58.1	49.8	41.8	33.8
	.1	.4	53.8	51.7	49.9	47.9	46.3
.33	1.5	18.2	80.6	66.6	50.0	33.3	19.4
	.5	9.2	68.6	59.6	50.0	40.5	31.4
	.1	2.0	54.6	52.1	50.0	47.9	45.8
If no settlement[a]			84.1	69.1	50.0	30.9	15.9

Note. Results are from 10 million simulated observations (disputes) under the assumption that all cases went to trial. The results of additional simulations can be found in Section F of the online appendix.

[a]Plaintiff trial win rate if all cases went to trial.

sample size. In addition, it tries to explain why Priest and Klein drew such different conclusions from their model.

Like Priest and Klein, we explore the effect of changing the legal standard on the plaintiff trial win rate with simulations. In Table 1, the decision standard Y^* represents the law relevant to the dispute. Higher values of Y^* mean law more favorable to the defendant. In particular, following Priest and Klein, Table 1 assumes that potential disputes are distributed normally, with a mean of zero and a standard deviation of 1. So a decision standard of $Y^* = 0$ implies that half of all potential disputes are meritorious, while a decision standard of $Y^* = 1$ implies that plaintiffs should prevail 15.9 percent of the time if no cases are settled and all cases go to trial.

The percentage of plaintiff trial victories also varies with litigation costs and the accuracy with which parties predict trial outcomes. In the Priest-Klein model, $C = C_\pi + C_\delta$ is the sum of the plaintiff's and defendant's costs of litigating the case through trial, while $S = S_\pi + S_\delta$ is the sum of the plaintiff's and defendant's costs if the case settles. The expression $C - S$ therefore reflects the amount the parties save by settling, and J is the judgment or damages that the plaintiff will receive if it prevails. The expression $(C - S)/J$ is therefore a measure of the cost savings from settlement relative to potential damages. The parties estimate the degree of defendant fault Y' with error. The term $\sigma_{\pi,\delta}$ is the standard deviation of those errors. As expected, the percentage of litigated cases falls as litigation costs increase and as parties increase the accuracy of their estimates.

From Table 1, one can see that the percentage of plaintiff trial victories changes significantly with the legal standard and in the predicted direction. As the decision standard becomes more pro-defendant (as Y^* increases), the percentage of plaintiff trial victories falls. When prediction errors are the smallest ($\sigma_{\pi,\delta} = .1$), the percentage of plaintiff trial victories varies by only a few points. On the other hand, changes in the percentage of plaintiff trial victories are otherwise rather large. For example, for $(C - S)/J = .67$ and $\sigma_{\pi,\delta} = .5$, as the decision standard changes from $-.5$ to 0 to .5, the percentage of plaintiff trial victories falls from 58.1 percent to 49.8 percent to 41.8 percent. These differences are large enough to be detected in data sets of moderate size. For other parameters, especially those involving larger prediction errors, the change in plaintiff win rates is even larger.

Priest and Klein (1984) acknowledged that, if prediction errors are large, the decision standard will affect the percentage of plaintiff trial victories. Nevertheless, they thought that parties were extremely accurate in predicting trial outcomes. They estimated that trial rates were under 2 percent and that litigation costs, as measured by $(C - S)/J$, were .33. From these two estimates, using simulations similar to those in Table 1, they inferred that prediction errors were about .1. That is, they looked at simulations similar to those in Table 1 and noted that if trial rates were 2 percent or less, then prediction errors must be about .1 or smaller. From that they concluded that changes in the decision standard would lead to very small changes in the percentage of trials won by plaintiffs. For example, when $(C - S)/J = .33$ and $\sigma_{\pi,\delta} = .1$, as the decision standard changes from $-.5$ to 0 to .5 the percentage of plaintiff trial victories falls from 52.1 percent to 50.0 percent to 47.9 percent. Such small changes are unlikely to be detectable even with relatively large data sets. For these reasons, Priest and Klein emphasized the limiting result, that as parties' prediction errors go to 0, plaintiffs will win 50 percent of tried cases, and that percentage will not vary with the legal standard.

Unfortunately, Priest and Klein made a simple but consequential mistake. In all of their simulations, they assumed $(C - S)/J = .33$. They justified this assumption by noting that .33 is "the amount of the most common contingency fee in personal injury litigation" (Priest and Klein 1984, p. 22 n.49). Nevertheless, $(C - S)/J = .33$ does not follow from the fact that plaintiffs pay 33 percent in contingent fees for two reasons:

C and S represent the sum of the plaintiff's and defendant's litigation costs, whereas the contingent fee represents only the plaintiff's costs, and to calculate $(C - S)/J$ one needs to know the difference between the cost of litigating and the cost of settling. A contingent-fee percentage does not provide that information. In fact, under a simple 33 percent contingent-fee arrangement, $(C - S)/J = 0$, because the plaintiff would pay 33 percent whether the case was settled or litigated.[20]

The most plausible value of $(C - S)/J$ is hard to determine, because there are very little data on the relative cost of litigation and settlement. The best data come from a RAND study of federal and state litigation in 1985. It found that litigation costs, including the value of parties' time and expenses, were 34 percent of judgments for plaintiffs and 38 percent for defendants (Kakalik and Pace 1986, p. ix). Nevertheless, such numbers are not very helpful, because they represent the weighted average cost of litigated and settled cases, whereas $(C - S)/J$ requires knowing the difference between the cost of litigation and the cost of settlement. Only for medical malpractice does the RAND study distinguish between litigation costs for settled and litigated cases. It reports that settlements and judgments average roughly the same amount and that defendants' litigation costs are 64 percent of judgments for litigated cases and 26 percent for settled cases (Kakalik and Pace 1986, p. 55). This implies that, for defendants, the difference between the cost of litigation and the cost of settlement is 38 percent of the judgment (64 percent − 26 percent). If litigation costs vary in similar fashion for plaintiffs,[21] a plausible value of $(C - S)/J$ would be 76 percent = 2 ×

20. Consider a situation in which the plaintiff has probability p of winning judgment J and hires a lawyer on a 33 percent contingent fee. First, consider settlements. Possible settlements would range from $pJ - C_\pi$ to $pJ + C_\delta$. For simplicity, assume that $C_\pi = C_\delta$ and that settlement is in the middle of that range, pJ, in which case the plaintiff's settlement costs would be $.33pJ$. So $S_\pi = .33pJ$. Now consider cases that go to trial. With probability p, the plaintiff wins J and pays $.33J$ to its attorney. With probability $1 - p$, the plaintiff loses and pays nothing to its attorney. So if the case goes to trial, $C_\pi = .33pJ$. So $(C_\pi - S_\pi)/J = 0$. That, of course, is only the plaintiff's half of $(C - S)/J$, but to calculate the defendant's half, one would need to have information on the defendant's litigation costs, which the contingent fee does not provide. Of course, this simple calculation ignores costs (such as filing fees) that the plaintiff pays and are not covered by the contingent fee. Nevertheless, the point remains that the contingent fee provides little or no information about the parameter of interest, $(C - S)/J$.

21. Litigation and settlement costs for plaintiffs who hire lawyers on a contingent-fee basis are difficult to calculate. As discussed below, if the contingent fee is a fixed percentage (for example, 33 percent or 40 percent), $(C - S)/J$ is approximately 0. This suggests, implausibly, that, for the plaintiff, litigation is no more costly than settlement. Nevertheless, even when the contingent fee is a fixed percentage, the costs to the lawyer (especially the

38 percent. One must multiply the defendants' percentage (38 percent) times 2, because $(C - S)/J$ represents the sum of plaintiffs' and defendants' litigation costs minus the sum of their settlement costs. Since 76 percent is a very rough estimate of $(C - S)/J$, Table 1 includes simulations for both $(C - S)/J = .67$ and $(C - S)/J = .80$ to show the effect of litigation costs a little higher and a little lower than for $(C - S)/J = .76$.

With these parameters, changes in the legal standard produce pronounced changes in plaintiffs' victories. For example, even assuming that the litigation rate is the same as in Priest and Klein's preferred simulations (2.0 percent), if $(C - S)/J = .67$, the most plausible value for prediction errors, according to Table 1, is $\sigma_{\pi,\delta} = .5$. With this value, changes in the legal standard produce large changes in the percentage of tried cases won by plaintiffs. If $(C - S)/J = .80$, then the simulation that produces a percent litigated closest to 2.0 percent is $\sigma_{\pi,\delta} = 1.5$, and changes in the decision standard produce even more pronounced changes in the plaintiff win rate. Simulations with other distributions—such as beta distributions and gamma distributions—confirm that changes in the legal standard produce changes in the plaintiff win rate large enough to be detectable with reasonable sample sizes for many but, of course, not all parameters.

5. CHANGES RELATING TO DAMAGES

Thus far, we have analyzed only the consequence of shifting the liability standard. It is also important to consider the effect of legal changes relating to the calculation of damages.[22] Such changes might include caps on damages, damage multipliers, or the exclusion of damages for

---

lawyer's time) certainly increase if the case proceeds to trial. Since the lawyer is likely to be highly involved in settlement negotiations and is likely to use his or her powers of persuasion to convince the client to accept a settlement offer that would save the lawyer considerable time and expense, looking solely at the fee paid by the client (plaintiff) is misleading. If transaction costs were 0, the lawyer and client would formulate a joint-maximizing settlement strategy, and the actual costs, which are borne primarily by the plaintiff's lawyer, would be what determine settlement. For this reason, it is reasonable to assume that the difference between the cost of litigation and settlement for the plaintiff is roughly equal to the figure calculated above for the defendant. Other reasons that the plaintiff would save money by settling include the value of the plaintiff's time, trial expenses that the contingent-fee contract might require the plaintiff to pay, and the fact that many contingent-fee agreements allow the lawyer to claim a larger share (for example, 40–50 percent) if the case goes to trial.

22. We thank Andrew Daughety and Jennifer Reinganum for encouraging us to look into damages, for pointing out the ambiguous effect of changes relating to damages, and for proving some of those ambiguous effects.

pain and suffering. A shift in the amount of damages will likely change the cost of litigation and settlement. If we assume proportional litigation costs—for instance, that a 50 percent increase in the damages would increase the cost of litigation by 50 percent—then legal changes relating to damages will have no effect on the plaintiff trial win rate. This lack of effect follows from the assumption of risk neutrality. However, it seems more realistic to assume that the cost of litigation will indeed increase with the amount of damages but less than proportionally. We can analyze the effect of such a change in damages by looking at the comparative statics of increasing or decreasing the amount of damages, while fixing all other variables. In this analysis, however, we find that the results are ambiguous and sensitive to model specification.

Proposition 4: Inferences and Changes Relating to Damages. When the standard of liability and litigation costs remain fixed: (i) under the screening and signaling models, an increase (decrease) in damages leads to a higher (lower) trial win rate for the uninformed party; and (ii) under the Priest-Klein model, an increase (decrease) in damages causes the plaintiff trial win rate to converge to (diverge from) the plaintiff trial win rate that would be observed if all disputes were litigated and diverge from (converge to) the limit value of the plaintiff's win rate—that is, if the stakes are symmetric, diverge from (converge to) 50 percent.

Proposition 4 is described in greater detail and proved in Section A.4.[23]

6. CAVEATS

The analysis in Sections 2, 3, and 4 has assumed that the only thing that changes is the legal rule. We began with asymmetric-information models—screening and signaling—and assumed that the distribution of information remains the same after the legal change and that the distribution of types under the two rules has the monotone likelihood ratio property. As discussed in Section 2, the latter assumption is satisfied if the distribution of defendant fault remains the same under both rules, because it is reasonable to believe that, holding defendant fault constant, a more pro-plaintiff rule will increase the probability that a plaintiff will

23. Under the Nalebuff (1987) version of the screening model, discussed in note 4, if damages J increase (decrease) and condition 2 (the credibility constraint) is satisfied, then the plaintiff trial win rate decreases (increases), because the plaintiff trial win rate is C_π/J. If condition 2 is not satisfied, then the effect of changing damages is the same as the standard asymmetric-information models analyzed in proposition 4.

prevail if the case goes to trial for each type. Similarly, under the Priest-Klein model, we assumed that the effect of a legal change is to change the decision standard but not the distribution of defendant fault or the level of uncertainty.

In contrast, if damages, the asymmetry of the stakes, the distribution of information, the level of uncertainty, the cases that are worth bringing suit for, the distribution of fault, or other aspects of predispute behavior change, the analysis in this article may no longer hold. Consider an extreme example. Suppose that, under the original legal standard, potential defendants' behavior results in a distribution of plaintiff probabilities of prevailing that is uniform over the interval $[a, b]$. Further suppose that, as a result of a pro-plaintiff legal change, defendants dramatically reduce their liability-inducing behavior, so that all defendants now act to reduce the probability that they would lose at trial to $a/2$ (see, for example, Bernardo, Talley, and Welch 2000). In this situation, it is clear that the more pro-plaintiff rule cannot result in a higher plaintiff trial win rate.

As a result, inferences from plaintiff win rates are reliable only in circumstances in which potential defendants have not changed their behavior substantially as a result of the new legal rule. One such situation is where a new legal rule was a surprise and was applied retroactively (see Hubbard 2013). In that situation, one could draw valid inferences by studying the cases in which the new rule was applied to behavior that the defendant thought would be governed by the old rule. Similarly, if one is dealing with parties who are not legally sophisticated or where adaptation to the new rule takes time or requires costly investments, one might be able to draw inferences from cases arising out of actions taken for a short period after the legal change.[24] In any case, we suggest that the primary challenge to empirical work using litigated cases to measure legal change is not selective settlement but behavioral responses to the new rule.

The issue of behavioral responses is less serious when analyzing the effect of different legal decision makers. For example, a researcher might use plaintiff trial win rates to determine whether judges appointed by Democratic presidents are more pro-plaintiff than judges appointed by

24. Under the Priest-Klein model, even if prelitigation behavior did not change, one would also need to be concerned about the way a legal change would introduce uncertainty about application of the legal standard and thus increase the standard deviation of the parties' prediction errors. This would also substantially complicate inferences from litigated cases.

Republicans. In this context, one need not be concerned about the possibility that judges appointed by Republicans or Democrats might have different effects on predispute behavior, because judges are assigned randomly well after the primary behavior that triggers a dispute. As a result, controlling for the court, the probability distributions of cases assigned to different judges will be the same whether the judge was appointed by a Republican or Democrat. Thus, inferences from litigated cases about judges (or other legal decision makers, such as jurors) are more likely to be valid than inferences about legal change.

Similarly, if one is analyzing the effect of case characteristics, the only ones that could affect primary behavior are those that the parties could know in advance, such as their own gender or race, or situational facts, such as whether an intentional tort was committed at night. Other case factors, such as whether a car accident victim was a child, are unlikely to be known in advance. Even if potential defendants knew that jurors were more or less likely to find liability when the victim was a child, they would not be able to adjust their primary behavior in a way that selectively reduced the probability of accidents affecting children. As a result, in such situations, inferences from litigated cases about the effect of case characteristics are likely to be reliable because they are unaffected by possible adjustments to primary conduct. In contrast, if one is studying the effect of case characteristics known in advance—such as whether a tort defendant is foreign or domestic—the researcher needs to take into account the possibility of differences in primary behavior. For example, if juries are biased against foreign defendants and foreign defendants are aware of that bias, such defendants might take more precautions than domestic defendants. As a result, the distribution of defendant fault will be different for foreign and domestic corporations. Consequently, even if juries are biased against foreign defendants, win rates might be lower when the defendant is foreign, because such defendants might have taken more precautions. Thus, inferences from plaintiff win rates to jurors' attitudes toward foreign defendants may not be valid.

The analysis of damages in Section 5 complicates empirical work considerably. Many legal changes affect both the criteria for liability and the calculation of damages. Since the effects of changes relating to damages are more ambiguous, the predicted overall effect may be ambiguous as well. Similarly, if particular decision makers are pro-plaintiff or pro-defendant both in the way they decide liability and the way they calculate damages, which seems likely, prediction will be difficult.

It is also important to remember that all of the propositions in this article make assumptions about the distribution of disputes. For example, the propositions about the asymmetric-information models assume that the distributions satisfy the monotone likelihood ratio property, and the proposition about the Priest-Klein model assumes log concavity. Empirical researchers need to consider whether those assumptions are satisfied, and that may be difficult.

7. CONCLUSION

This article has shown that, under the three standard settlement models and a wide array of parameters and distribution functions, the proportion of plaintiff victories at trial will vary in a predictable fashion with the legal standard, legal decision makers, or case characteristics. Under these models and assumptions, a pro-plaintiff legal change will result in a greater percentage of plaintiff trial victories. Judges or jurors biased in favor of ~~defendants~~ plaintiffs will generate more plaintiff victories at trial. Case characteristics—such as particular fact patterns or the age or race of the parties—may influence the rate at which plaintiffs prevail at trial. Therefore, although selection effects exist across all models, we believe that carefully designed empirical studies may permit valid inferences about legal standards, decision makers, and case characteristics. In fact, since selection dampens these effects, the real-world impact of legal, judicial, juror, and case factors is likely to be larger than those measured by researchers. In contrast, the effect of legal changes relating to damages is more ambiguous and complicates empirical work.

APPENDIX

A1. Proof of Inferences under the Screening Model

The key to the proof is to notice that the first-order condition for the plaintiff's optimal offer, equation (2), can be rewritten in terms of the hazard rates,[25] $h_f(p) = f(p)/[1 - F(p)] = 1/k$ and $h_g(p) = g(p)/[1 - G(p)] = 1/k$, where $p = (x - C_\delta)/J$ and $k = (C_\pi + C_\delta)/J$. Lemmas A1–A3 are illustrated in Section G of the online appendix.

Lemma A1. Under the conditions of proposition 1, either $f(1) > 0$ or $f'(1) < 0$ is sufficient to ensure that there exists a unique solution to equation (2).

Proof of Lemma A1. Since $f(\underline{p}) = 0$, the hazard rate starts at 0. Since

25. We thank Jonah Gelbach for suggesting this approach.

the hazard rate is monotonically increasing, it suffices to show that $\lim_{p\to 1} f(p)/[1 - F(p)] = \infty$. This is trivially satisfied if $f(1) > 0$. Otherwise, $f(1) = 0$. Then apply l'Hôpital's rule to the limit: $\lim_{p\to 1} f(p)/[1 - F(p)] = \lim_{p\to 1} f'(p)/[-f(p)]$. Since $f'(1) < 0$ while $f(1) = 0$, the overall limit is infinity. Hence, the hazard rate must approach infinity under either condition. Since the hazard rate starts at 0 and is monotonically increasing toward infinity, it must intersect $y = 1/k$ exactly once. So there exists a unique settlement offer that satisfies the first-order condition, and this offer also satisfies the second-order condition. Q.E.D.

Lemma A2. If $g(p)$ and $f(p)$ have the monotone likelihood ratio property such that $(d/dp)[g(p)/f(p)] \geq 0$ for all $p \in (\underline{p}, 1)$, then $h_g(p) \leq h_f(p)$ for all $p \in (\underline{p}, 1)$.

Proof of Lemma A2. Given p, $q \in (\underline{p}, 1)$ such that $p < q$, we have $g(p)/f(p) \leq g(q)/f(q)$, so $g(p)f(q) \leq f(p)g(q)$. Integrating the expression over all $q \in [p, 1]$, we have $g(p)(1 - F(p)) \leq f(p)(1 - G(p))$. Q.E.D.

Lemma A3. Under the conditions of proposition 1, $p_f^* \leq p_g^*$ and $x_f^* \leq x_g^*$. In other words, the plaintiff's optimal settlement demand is higher under the more pro-plaintiff legal standard.

Proof of Lemma A3. The optimal settlement demand is defined by the first-order condition that the hazard rate equals $1/k$. We have seen that for each pair of probability density functions f and g satisfying the conditions of proposition 1, there are unique probability values p_f^* and p_g^* that satisfy this condition. By lemma A2, however, $h_g(p) \leq h_f(p)$ for all $p \in (\underline{p}, 1)$. Since both hazard rates are strictly increasing, it follows that $p_f^* \leq p_g^*$ and $x_f^* \leq x_g^*$. Q.E.D.

The probability that the plaintiff will prevail in litigated cases under a given legal standard is the average value of p for defendants who litigate, where values of p are weighted by the corresponding probability density function, $f(p)$ or $g(p)$. The monotone likelihood ratio implies that $g(p)$ stochastically dominates $f(p)$, which means, as shown below, that the average value of p, where $p < p_f^*$, is greater under $g(p)$ than $f(p)$. A fortiori, since $p_g^* \geq p_f^*$, the average value of p under $g(p)$ where $p < p_g^*$ is greater than the average value of p under $f(p)$, where $p < p_f^*$.

Proof of Proposition 1. Since $p_f^* \leq p_g^*$, it suffices to show, first, that

$$\frac{\int_{\underline{p}}^{p} u f(u)\,du}{F(p)}$$

is generally increasing in p for any distribution $f(p)$ and, second, that for all p,

$$\frac{\int_{\underline{p}}^{p} uf(u)du}{F(p)} < \frac{\int_{\underline{p}}^{p} ug(u)du}{G(p)}.$$

To see the first, we need to show only that

$$\frac{d}{dp}\left(\frac{\int_{\underline{p}}^{p} uf(u)du}{F(p)}\right) > 0.$$

By the quotient rule, this is true if and only if the numerator of the derivative is positive. In other words,

$$0 < F(p)pf(p) - f(p)\int_{\underline{p}}^{p} uf(u)du = f(p)\int_{\underline{p}}^{p} (p-u)f(u)du.$$

The last term is clearly greater than 0 since the integral is over $u \le p$ and $f(u)$ is everywhere positive in the open interval.

As for the second, integration by parts gives us

$$\frac{\int_{\underline{p}}^{p} uf(u)du}{F(p)} = \frac{pF(p) - \int_{\underline{p}}^{p} F(u)du}{F(p)} = p - \frac{\int_{\underline{p}}^{p} F(u)du}{F(p)}$$

and likewise

$$\frac{\int_{\underline{p}}^{p} ug(u)du}{G(p)} = p - \frac{\int_{\underline{p}}^{p} G(u)du}{G(p)}.$$

It thus suffices to show that

$$\int_{\underline{p}}^{p} G(u)du < \int_{\underline{p}}^{p} F(u)\left(\frac{G(p)}{F(p)}\right)du,$$

which will certainly be true if $G(u) < F(u)(G(p)/F(p))$ for $u < p$, or, alternatively, if $G(p)/F(p)$ is increasing in p. By the quotient rule, it suffices to show in turn that

$$0 < F(p)g(p) - G(p)f(p) = f(p)\left(\int_{\underline{p}}^{p} f(u)\left(\left(\frac{g(p)}{f(p)}\right) - \left(\frac{g(u)}{f(u)}\right)\right)du\right),$$

which is true since $(d/dp)/[g(p)/f(p)] \ge 0$ under the monotone likelihood ratio property and is not always 0 since $f(p)$ and $g(p)$ are distinct functions. Q.E.D.

A2. Proof of Inferences under the Signaling Model

The key to the proof is to recognize that the inequality to be proved is a comparison of the mean value of p under two new probability density functions,

$$f_1(p) = \frac{r^*(s^*(p))f(p)}{\int_{\underline{p}}^1 r^*(s^*(p))f(p)dp}$$

and

$$g_1(p) = \frac{r^*(s^*(p))g(p)}{\int_{\underline{p}}^1 r^*(s^*(p))g(p)dp}.$$

The monotone likelihood ratio condition between $f(p)$ and $g(p)$ implies that $f_1(p)$ and $g_1(p)$ also have that property. This in turn implies that $g_1(p)$ stochastically dominates $f_1(p)$, and the inequality follows.

Proof of Proposition 2. The plaintiff trial win rate is

$$P_t(f(p)) = \frac{\int_{\underline{p}}^1 r^*(s^*(p))pf(p)dp}{\int_{\underline{p}}^1 r^*(s^*(p))f(p)dp}.$$

Note that $r^*(s^*(p)) = 0$ at $\underline{p}$ and is strictly positive at all other points. If we let

$$f_1(p) = \frac{r^*(s^*(p))f(p)}{\int_{\underline{p}}^1 r^*(s^*(p))f(p)dp}$$

and

$$g_1(p) = \frac{r^*\big(s^*(p)\big)g(p)}{\int_{\underline{p}}^1 r^*\big(s^*(p)\big)g(p)dp},$$

$f_1(p)$ and $g_1(p)$ are probability density functions. We can rewrite the inequality as

$$\int_{\underline{p}}^{1} pf_1(p)dp < \int_{\underline{p}}^{1} pg_1(p)dp.$$

Therefore, it suffices to show that $g_1(p)$ stochastically dominates $f_1(p)$. Since the monotone likelihood ratio implies stochastic dominance, it suffices to show that $(d/dp)/[g_1(p)/f_1(p)] \geq 0$. But this is immediate since

$$\frac{d}{dp}\left(\frac{g_1(p)}{f_1(p)}\right) = \left(\frac{\int_{\underline{p}}^1 r^*(s^*(p))f(p)dp}{\int_{\underline{p}}^1 r^*(s^*(p))g(p)dp}\right)\left(\frac{d}{dp}\left(\frac{g(p)}{f(p)}\right)\right) \geq 0$$

by the monotone likelihood ratio property of $g(p)$ and $f(p)$. And since the two distributions are not identical, the inequality between the win rates under the two distributions must be strict. Q.E.D.

A3. Proof of Inferences under the Priest-Klein Model

Proof of Proposition 3. The proof relies on the formalization of the Priest-Klein model by Waldfogel (1995); see also Lee and Klerman (2014). Let $g(Y')$ be the distribution of disputes, and let Y^* be the decision standard. Suppose that the plaintiff estimates case strength (defendant fault) to be $Y'_\pi = Y' + \varepsilon_\pi$ and the defendant estimates case strength to be $Y'_\delta = Y' + \varepsilon_\delta$. Let $f(\varepsilon_\pi, \varepsilon_\delta)$ be the bivariate distribution of errors for the parties, with positive standard deviations σ_π and σ_δ for plaintiff and defendant, respectively. Because we are interested in the dynamics of settlement and litigation in the presence of a change in the legal standard of liability—rather than the limit value of the plaintiff's win rate as σ_π and σ_δ approach 0—we assume that σ_π and σ_δ remain fixed as the standard of liability changes.

We allow for asymmetric stakes by assuming $J_\pi = \alpha J_\delta = \alpha J$, where $\alpha \in ((C-S)/J, \infty)$. If $\alpha \leq (C-S)/J$, no cases will be litigated, because cases go to trial if and only if $\alpha P_\pi - P_\delta > (C-S)/J$. If $F_\pi[x]$ and $F_\delta[x]$ are the cumulative distribution functions for the plaintiff and defendant, respectively, then the trial condition can be rewritten as[26]

$$\alpha F_\pi[Y' + \varepsilon_\pi - Y^*] - F_\delta[Y' + \varepsilon_\delta - Y^*] > \frac{C-S}{J}.$$

Let $P(Y'; Y^*)$ denote the probability that a case of merit Y' goes to trial when the decision standard is Y^*. Then $P(Y'; Y^*)$ can be calculated as

$$\iint_{R(Y'; Y^*)} f(\varepsilon_\pi, \varepsilon_\delta)\, d\varepsilon_\pi d\varepsilon_\delta,$$

where the region of integration is defined implicitly as

26. The model in this proof follows Priest and Klein (1984) and Waldfogel (1995) in that the trial condition is both a necessary and sufficient condition and in that parties estimate Y' only on the basis of signals they receive, without taking into account the distribution of disputes (see Lee and Klerman 2014). Waldfogel (1995) has the expressions inside the brackets normalized by σ_π and σ_δ because he explicitly assumes F to be the cumulative distribution function of the standard normal distribution. Because we are writing the expressions more generally, we need not similarly normalize the expressions.

$$R(Y'; Y^*) = \left\{(\varepsilon_\pi, \varepsilon_\delta) \in R^2 \middle| \alpha F_\pi[Y' + \varepsilon_\pi - Y^*] - F_\delta[Y' + \varepsilon_\delta - Y^*] > \frac{C - S}{J}\right\}.$$

The plaintiff trial win rate is

$$P_t(Y^*) = \frac{\int_{Y^*}^{\infty} P(Y'; Y^*)g(Y')dY'}{\int_{-\infty}^{\infty} P(Y'; Y^*)g(Y')dY'} = \frac{1}{1 + 1/P_{t,\pi/\delta}(Y^*)},$$

where

$$P_{t,\pi/\delta}(Y^*) = \frac{\int_{Y^*}^{\infty} P(Y'; Y^*)g(Y')dY'}{\int_{-\infty}^{Y^*} P(Y'; Y^*)g(Y')dY'}$$

is the ratio of the plaintiff's wins over the defendant's wins. It thus suffices to show that $\partial P_{t,\pi/\delta}(Y^*)/\partial Y^* < 0$. Notice that since $R(Y'; Y^*) = R(Y' - Y^*; 0)$, we must also have $P(Y'; Y^*) = P(Y' - Y^*; 0)$. So if we make a change of variable $z = Y' - Y^*$, then we can rewrite $P(Y'; Y^*)$ as $P(z; 0)$. So

$$P_{t,\pi/\delta}(Y^*) = \frac{\int_{0}^{\infty} P(z; 0)g(z + Y^*)dz}{\int_{-\infty}^{0} P(z; 0)g(z + Y^*)dz}.$$

By the quotient rule and Leibniz's rule, $\partial P_{t,\pi/\delta}(Y^*)/\partial Y^* < 0$ if and only if

$$\left[\int_{-\infty}^{0} P(z; 0)g(z + Y^*)dz\right]\left[\int_{0}^{\infty} P(z; 0)g'(z + Y^*)dz\right] < \left[\int_{-\infty}^{0} P(z; 0)g'(z + Y^*)dz\right]\left[\int_{0}^{\infty} P(z; 0)g(z + Y^*)dz\right].$$

Since $g(Y')$ is log concave, we have $g(Y')g''(Y') \leq [g'(Y')]^2$. This in turn implies that $g_0(Y') = g'(Y')/g(Y')$ is constantly (weakly) decreasing in Y'. Therefore, we must have $g'(z + Y^*) \leq g(z + Y^*)g_0(Y^*)$ for all $z > 0$ and $g'(z + Y^*) \geq g(z + Y^*)g_0(Y^*)$ for all $z < 0$. This indicates that the left-hand side is bounded above by

$$g_0(Y^*)\left[\int_{-\infty}^{0} P(z; 0)g(z + Y^*)dz\right]\left[\int_{0}^{\infty} P(z; 0)g(z + Y^*)dz\right],$$

while the right-hand side is bounded below by it. Moreover, the inequality must be strict since $g_0(Y')$ cannot be a constant function when $g(Y')$ is a probability density function over the entire real line.

Finally, to show that we can apply Leibniz's rule over the infinite interval, we need to show that $|\partial P(z;\ 0)g(z+Y^*)/\partial Y^*| = |P(z;\ 0)g'(z+Y^*)|$ is bounded by an integrable function. Indeed, $|P(z;\ 0)g'(z+Y^*)| \leq |g'(z+Y^*)| = g(z+Y^*)|g_0(z+Y^*)|$, and $g_0(z+Y^*)$ is constantly (weakly) decreasing in Y^*. If $g_0(z+Y^*)$ is always positive over the relevant interval, then $|g'(z+Y^*)| = g'(z+Y^*)$, which must integrate to a finite value since $g(Y')$ is a probability density function. If $g_0(z+Y^*)$ is negative at any point, then the integral of $|g'(z+Y^*)|$ over the relevant interval will be the difference of two integrals of $g'(z+Y^*)$ (over complementary regions), each of which must integrate to a finite value. Hence, $|g'(z+Y^*)|$ is an integrable function, and the result follows. Q.E.D.

A4. Damages

Proof of Proposition 4. Consider first the screening model when the defendant has superior information. The first-order condition for the plaintiff's offer is $1 - F((x - C_\delta)/J) = [f((x - C_\delta)/J)(C_\pi + C_\delta)]/J$. Let $x^*(J)$ be the optimal offer amount for the plaintiff, as a function of J, and let $p^*(J) = (x^*(J) - C_\delta)/J$. Then we can rewrite the first-order condition as $J(1 - F(p^*(J)) = f(p^*(J))(C_\pi + C_\delta)$. Differentiating this equation with respect to J, we get

$$(p^*(J))' = \left[\frac{1 - F(p^*(J))}{C_\pi + C_\delta}\right] \Big/ \left[\frac{1}{k} f(p^*(J)) + f'(p^*(J))\right].$$

This value will be positive as long as $(1/k) f(p^*(J)) + f'(p^*(J)) > 0$. Since $p^*(J) = (x^*(J) - C_\delta)/J$, the question is whether $(1/k) f(p) + f'(p)$ is positive at the point where the graph of $f(p)$ intersects the graph of $(1/k)(1 - F(p))$. The latter graph begins at $1/k$ and is (weakly) decreasing at all times, while, as shown in Section A.1, $f(p)$ starts from 0 and crosses the graph of $(1/k)(1 - F(p))$ exactly once. Therefore, the slope of $f(p)$ at the intersection must be greater than the slope of $(1/k)(1 - F(p))$ at the same point. This means that $f'(p^*(J)) > (1/k)(-f(p^*(J)))$. Hence, $(1/k) f(p^*(J)) + f'(p^*(J))$ must be positive, and $(p^*(J))' > 0$. Therefore, $p^*(J)$ will increase as J increases, and then it is straightforward to show that

$$P_t = \frac{\int_{\underline{p}}^{p^*(J)} p f(p)\, dp}{F(p^*(J))}$$

will also increase.[27] By the same reasoning, we can also show that when the plaintiff has superior information, an increase in J will lead to a higher rate of victory for the defendant.

Now consider the signaling model when the plaintiff has superior information. The plaintiff's win rate can be written as

$$P_t(f(p)) = \frac{\int_{\underline{p}}^{1} r^*(s^*(p))pf(p)dp}{\int_{\underline{p}}^{1} r^*(s^*(p))f(p)dp} = \frac{\int_{\underline{p}}^{1} pL(p, J)dp}{\int_{\underline{p}}^{1} L(p, J)dp},$$

where $L(p, J) = (1 - \exp\{-(p - \underline{p})J/C\})f(p)$. It suffices to show that this is decreasing in J. In other words $dP_t(f(p))/dJ < 0$. Then we need to show that

$$\frac{d}{dJ}\left(\frac{\int_{\underline{p}}^{1} pL(p, J)dp}{\int_{\underline{p}}^{1} L(p, J)dp}\right) = \frac{\int_{\underline{p}}^{1} L(p, J)dp \int_{\underline{p}}^{1} p(\partial L/\partial J)(p, J)dp - \int_{\underline{p}}^{1} pL(p, J)dp \int_{\underline{p}}^{1} (\partial L/\partial J)(p, J)dp}{\left(\int_{\underline{p}}^{1} L(p, J)dp\right)^2}$$

Since $(\partial L/\partial J)(p, J) > 0$ at all times,

$$\int_{\underline{p}}^{1} \frac{\partial L}{\partial J}(p, J)dp > 0$$

and

$$\int_{\underline{p}}^{1} L(p, J)dp > 0.$$

Thus, it suffices to show that

$$\frac{\int_{\underline{p}}^{1} p(\partial L/\partial J)(p, J)dp}{\int_{\underline{p}}^{1} (\partial L/\partial J)(p, J)dp} < \frac{\int_{\underline{p}}^{1} pL(p, J)dp}{\int_{\underline{p}}^{1} L(p, J)dp}.$$

As with the proof for proposition 2, it suffices to show that

$$\frac{L(p, J)}{\int_{\underline{p}}^{1} L(p, J)dp}$$

and

$$\frac{\partial L/\partial J(p, J)}{\int_{\underline{p}}^{1} (\partial L/\partial J)(p, J)dp}$$

have a monotone likelihood ratio property in p. Since the denominators are constant in p, we need only check that

27. Since we assumed $\underline{p} > C_\pi/J$ from the original setup, increasing J will not affect $\underline{p}$ and thus will not affect the likelihood of low-probability litigation. In contrast, a decrease in J might. Therefore, we assume that $\underline{p}$ is sufficiently large.

$$\frac{\partial}{\partial p}\left[\frac{L(p, J)}{\partial L/\partial J(p, J)}\right] > 0.$$

To confirm this, let $(p - \underline{p})/C = q$, and we show that

$$\frac{\partial}{\partial q}\left[\frac{L(q, J)}{\partial L/\partial J(q, J)}\right] > 0.$$

Since

$$\frac{L(q, J)}{\partial L/\partial J(q, J)} = \frac{1 - e^{-qJ}}{qe^{-qJ}},$$

we need to show (using the quotient rule) that $0 < qe^{-qJ}(1 + Je^{-qJ}) + Jqe^{-qJ}(1 - e^{-qJ})$. Since $-qJ \leq 0$, $(1 - e^{-qJ}) \geq 0$, and the rest follows. By a similar proof, we can show that when the defendant has superior information, an increase in J will lead to a higher rate of victory for the plaintiff.

Finally, under the Priest-Klein model, an increase in J (not accompanied by a proportional increase in C and S) will reduce the value of $(C - S)/J$, and will make it more likely for cases to go to trial, all else equal. This means that, if the stakes are symmetric, the rate of the plaintiff's victory will, in the limit, as J goes to infinity, diverge from 50 percent and converge to the plaintiff trial win rate if all disputes are litigated. If the defendant has more at stake, the plaintiff trial win rate will converge to a percentage less than 50 percent as the standard deviation of the parties' estimation errors goes to 0, and an increase in the number of trials caused by an increase in damages will cause the plaintiff trial win rate to diverge from that percentage and converge to the plaintiff trial win rate if all cases are litigated. Conversely, if the plaintiff has more at stake, the plaintiff trial win rate will converge to a percentage greater than 50 percent as the standard deviation of the parties' estimation errors goes to 0, and an increase in the number of trials caused by an increase in damages will cause the plaintiff trial win rate to diverge from that percentage and converge to the plaintiff trial win rate if all cases are litigated.

REFERENCES

Ashenfelter, Orley, Theodore Eisenberg, and Stewart J. Schwab. 1995. Politics and the Judiciary: The Influence of Judicial Background on Case Outcomes. *Journal of Legal Studies* 24:257–81.

Atkinson, Scott E., Alan C. Marco, and John L. Turner. 2009. The Economics of a Centralized Judiciary: Uniformity, Forum Shopping, and the Federal Circuit. *Journal of Law and Economics* 52:411–43.

Bagnoli, Mark, and Ted Bergstrom. 2005. Log-Concave Probability and Its Applications. *Economic Theory* 26:445–69.

Bebchuck, Lucian Arye. 1984. Litigation and Settlement under Imperfect Information. *RAND Journal of Economics* 15:404–15.

Bernardo, Antonio E., Eric Talley, and Ivo Welch. 2000. A Theory of Legal Presumptions. *Journal of Law, Economics, and Organization* 16:1–49.

Daughety, Andrew F., and Jennifer F. Reinganum. 2012. Settlement. Pp. 386–471 in *Procedural Law and Economics,* edited by Christopher W. Sanchirico. 2d ed. Cheltenham: Edward Elgar.

Eisenberg, Theodore, and Henry S. Farber. 1997. The Litigious Plaintiff Hypothesis: Case Selection and Resolution. *RAND Journal of Economics* 28: S92–S112.

Eisenberg, Theodore, and Sheri Lynn Johnson. 1991. The Effects of Intent: Do We Know How Legal Standards Work? *Cornell Law Review* 76:1151–97.

Epstein, Lee, William M. Landes, and Richard A. Posner. 2013. *The Behavior of Federal Judges*. Cambridge, Mass.: Harvard University Press.

Friedman, Daniel, and Donald Wittman. 2007. Litigation with Symmetric Bargaining and Two-Sided Incomplete Information. *Journal of Law, Economics, and Organization* 23:98–126.

Gelbach, Jonah B. 2012. Locking the Doors to Discovery? Assessing the Effects of *Twombly* and *Iqbal* on Access to Discovery. *Yale Law Journal* 121:2273–2345.

Hubbard, William H. J. 2013. Testing for Change in Procedural Standards, with Application to *Bell Atlantic v. Twombly. Journal of Legal Studies* 42:35–68.

Hylton, Keith N. 1993. Asymmetric Information and the Selection of Disputes for Litigation. *Journal of Legal Studies* 22:187–210.

Kakalik, James S., and Nicholas M. Pace. 1986. *Costs and Compensation Paid in Tort Litigation*. Santa Monica, Calif.: Rand Institute for Civil Justice.

Kessler, Daniel, Thomas Meites, and Geoffrey Miller. 1996. Explaining Deviations from the Fifty-Percent Rule: A Multimodal Approach to the Selection of Cases for Litigation. *Journal of Legal Studies* 25:233–59.

Klerman, Daniel. 2012. The Selection of 13th-Century Disputes for Litigation. *Journal of Empirical Legal Studies* 9:320–46.

Kulik, Carol T., Elissa L. Perry, and Molly B. Pepper. 2003. Here Comes the Judge: The Influence of Judge Personal Characteristics on Federal Sexual Harassment Case Outcomes. *Law and Human Behavior* 27:69–86.

Landes, William M., and Richard A. Posner. 1995. Heavily Cited Articles in Law. *Chicago-Kent Law Review* 71:825–40.

Lee, Yoon-Ho Alex, and Daniel Klerman. 2014. The Priest-Klein Hypotheses: Proofs, Generality, and Extensions. Unpublished manuscript. University of Southern California, Gould School of Law, Los Angeles.

Mohtashami Borzadaran, G. R., and H. A. Mohtashami Borzadaran. 2011. Log-Concavity Property for Some Well-Known Distributions. *Surveys in Mathematics and Its Applications* 6:203–19.

Nalebuff, Barry. 1987. Credible Pretrial Negotiation. *RAND Journal of Economics* 18:198–210.

P'ng, I. P. L. 1983. Strategic Behavior in Suit, Settlement, and Trial. *Bell Journal of Economics* 14:539–50.

Priest, George L. 1987. Measuring Legal Change. *Journal of Law, Economics, and Organization* 3:193–225.

Priest, George L., and Benjamin Klein. 1984. The Selection of Disputes for Litigation. *Journal of Legal Studies* 13:1–55.

Reinganum, Jennifer F., and Louis L. Wilde. 1986. Settlement, Litigation, and the Allocation of Litigation Costs. *RAND Journal of Economics* 17:557–66.

Shapiro, Fred R. 1996. The Most-Cited Law Review Articles Revisited. *Chicago-Kent Law Review* 71:751–79.

Shapiro, Fred R., and Michelle Pearse. 2012. The Most-Cited Law Review Articles of All Time. *Michigan Law Review* 110:1484–1520.

Shavell, Steven. 1996. Any Frequency of Plaintiff Victory at Trial Is Possible. *Journal of Legal Studies* 25:493–501.

Siegelman, Peter, and John J. Donohue III. 1995. The Selection of Employment Discrimination Disputes for Litigation: Using Business Cycle Effects to Test the Priest-Klein Hypothesis. *Journal of Legal Studies* 24:427–62.

Spier, Kathryn E. 1992. The Dynamics of Pretrial Negotiation. *Review of Economic Studies* 59:93–108.

Waldfogel, Joel. 1995. The Selection Hypothesis and the Relationship between Trial and Plaintiff Victory. *Journal of Political Economy* 103:229–60.

———. 1998. Reconciling Asymmetric Information and Divergent Expectations Theories of Litigation. *Journal of Law and Economics* 41:451–76.

Wittman, Donald. 1985. Is the Selection of Cases for Trial Biased? *Journal of Legal Studies* 14:185–214.

Correction to *Inferences from Litigated Cases*

Daniel Klerman and Yoon-Ho Alex Lee

Posted February 1, 2018

The first full paragraph on p. 244 should be revised as follows:

> Finally, to show that we can apply Leibniz's Rule over the infinite interval, we need to show there exist $k(z)$ and $h(z)$, integrable over z, such that $|P(z;0)g(z+Y^*)| \leq h(z)$ and $\left|\frac{\partial P(z;0)g(z+Y^*)}{\partial Y^*}\right| \leq k(z)$ for all z and Y^*. Since $g(\cdot)$ is a log-concave probability density function, both $g(\cdot)$ and $|g'(\cdot)|$ must both be bounded above by some constant, say $c \in \boldsymbol{R}$. In particular, $|g'(z+Y^*)|$ cannot spike to infinity at any finite z value since $|g_0(z+Y^*)| = \frac{|g'(z+Y^*)|}{g(z+Y^*)}$ must be (at least weakly) decreasing, and cannot approach infinity in either direction since it must integrate to a finite value. Therefore, $|P(z;0)g(z+Y^*)|$ and $\left|\frac{\partial P(z;0)g(z+Y^*)}{\partial Y^*}\right| = |P(z;0)g'(z+Y^*)|$ are both bounded above by $cP(z;0)$, which is integrable due to Proposition 3 from Lee & Klerman (2016). *Q.E.D.*

We have appended this correction to posted versions of *Inferences from Litigated Cases* because University of Chicago Press policy reserves published errata for substantial errors and does not allow post-publication revision of online appendices.

References

Lee, Yoon-Ho Alex and Daniel Klerman. 2016. The Priest-Klein Hypotheses: Proofs and Generality. International Review of Law & Economics 48:59-76.