TRIAL SELECTION THEORY: A UNIFIED MODEL

Keith N. Hylton[*]

Haizhen Lin[**]

October 2009

Abstract: This paper provides a formal model of the trial selection process that incorporates the Priest-Klein hypothesis and alternative theories of selection. We derive the conditions under which the hypothesis is valid, and examine implications for the relationship between trial outcome uncertainty and litigation. The model suggests a generalization of the hypothesis.

[*] Boston University, School of Law, knhylton@bu.edu.
[**] Indiana University, Kelley School of Business, hzlin@indiana.edu.

I. Introduction

> In *The Common Law*, Oliver Wendell Holmes noted that
>
> [l]egal, like natural divisions, however clear their general outline, will be found on exact scrutiny to end in a penumbra or debatable land. This is the region of the jury, and only cases falling on this doubtful border are likely to be carried far in court.[1]

In spite of this early recognition by a prominent legal theorist, the connection between legal uncertainty and litigation was not examined within a theoretical framework until Priest and Klein (1984). Since then, a substantial literature has developed on the selection of disputes for litigation.

Trial selection theory consists of models that explain or predict the characteristics that distinguish cases that are litigated to judgment from those that settle, and the implications of those characteristics for important trial outcome parameters, such as the plaintiff win rate, and for the development of legal doctrine.[2] The starting point for this literature is the Priest-Klein hypothesis, which holds that the plaintiff win rate will tend toward fifty percent unless the litigants have asymmetric stakes.[3]

Although the Priest-Klein hypothesis is widely cited in the law and economics literature and has been tested empirically, it still lacks a formal treatment.[4] The original Priest-Klein article provides an informal argument, as do later articles testing it.[5] The lack of a formal model makes it difficult to separate important from unimportant assumptions in the Priest-Klein analysis, and to formally separate Priest-Klein analysis from competing theories of trial selection. Moreover, the interpretation of tests of trial

---

[1] Holmes (1881), at 127.
[2] On trial selection and the development of legal doctrine, see Priest (1980), Gennaioli and Shliefer (2007), Miceli (2009).
[3] On the theory of stakes asymmetry and litigation, see Che and Yi (1993).
[4] Waldfogel, 1995, comes closest to providing a formal treatment of the Priest-Klein analysis. However, the fifty percent prediction of the analysis is demonstrated in the Waldfogel article through the use of a simulation rather than a formal proof.
[5] Of the articles testing the Priest-Klein hypothesis or related claims, see Waldfogel, (1995), Waldfogel (1998), Siegelman and Waldfogel (1999), Siegelman and Donohue (1995), Eisenberg (1990), Eisenberg and Heise (2007).

selection theories is also dependent on understanding the formal scope and limitations of the underlying models.[6]

This paper offers a formal model of trial selection in the manner of Priest and Klein. The model incorporates alternative (e.g., asymmetric-information) theories of selection as well. Given that the empirical and theoretical literature on trial selection has been described as a horse race between Priest-Klein (symmetric information) and asymmetric information theories of trial selection,[7] the model in this paper effectively unifies the major trial selection theories under one framework.[8]

The model permits an examination within a single framework of the results in the trial selection literature. For example, the fifty percent prediction of the Priest-Klein analysis is based on a hypothetical distribution of the probability of litigation mapped over an index of the defendant's probability of violating the legal standard (i.e., the guilt level). The fifty percent prediction holds in the limit, it has been said, as the trial rate approaches zero – which may happen because uncertainty over trial outcomes diminishes or because trial becomes more expensive.[9] While this is helpful to the fifty percent prediction in the symmetric information model, it is not a necessary feature.

This model shows that the implications of trial selection theory depend almost entirely on the "censoring function" – i.e., the function describing the probability of litigation conditional on the defendant's guilt level – and the probability distribution of guilt. When both are symmetric the fifty percent prediction holds precisely – and there is no need to talk about limiting conditions. When the guilt distribution is not symmetric we derive a general condition for it ("window property") that supports the tendency toward fifty percent. The fifty percent prediction is also generated in the limit as the censoring function becomes more convex, which occurs as trial outcome uncertainty

---

[6] For a discussion of this issue, see our companion piece to this paper, Hylton and Lin (2009).

[7] Waldfogel, 1998, at 451.

[8] However, as noted in the Hylton and Lin (2009) review, trial selection theories can be based on a larger set of variables than examined in the symmetric and asymmetric information literature. Eisenberg and Farber (1996) develop a theory of selection based on the cost of litigation.

[9] Waldfogel, 1995, at 232-33. Waldfogel's 1995 article, which is perhaps the most elaborate formalization to date of the Priest-Klein model, describes the key Priest-Klein prediction as follows: "The limiting implication of their model is that, with equal stakes to the parties, as the fraction of cases going to trial approaches zero (either because plaintiff or defendant uncertainty about trial outcomes declines or because trial costs increase), plaintiff win rates at trial will approach 50 percent." Waldfogel, at 229-30.

declines at the endpoints of the guilt spectrum relative to the middle – i.e., in Holmes's terms, umbral uncertainty declines relative to penumbral uncertainty.

There is a simple generalization of the Priest-Klein hypothesis in terms of this framework. When the censoring function is skewed right (because trial outcome uncertainty is greater at high levels of guilt than at low levels) the expected win rate will tend toward a level greater than fifty percent; and, conversely, when the function is skewed left, the win rate will tend toward less than fifty percent. Symmetry of the censoring function is a necessary condition for the fifty percent result.

The model also permits us to consider the extent to which informational asymmetry leads to departures from the symmetric information analysis. When the symmetry condition holds for the censoring function, informational asymmetry generates intuitively sensible predictions: e.g., when defendants only are informed, the win rate is less than fifty percent because guilty-and-informed defendants settle disproportionately. However, when the symmetry condition does not hold, then it is no longer clear that informational asymmetry will result in win rates greater or less than fifty percent, consistent with Shavell (1996).

Our approach, which consists of relaxing symmetry conditions for the censoring and distribution functions, can be contrasted with much of the asymmetric information modeling of trial selection, which is based on models of strategic behavior in settlement. However, the implications of the strategic behavior models are highly dependent on their particular assumptions.[10] We avoid dependence on specific strategic behavior assumptions. This model shows that a rich analysis of trial selection theory can be based on general statistical properties of the variables that determine the propensity to litigate.

II. Literature

Trial selection theory,[11] as initially presented in Priest and Klein (1984), builds on the idea, recognized at least since Holmes, that only the most uncertain disputes go all the

---

[10] Compare, for example, the divergent implications of Bebchuk (1984) and Png (1987), which both present strategic behavior models of the settlement process.

[11] We are distinguishing "trial selection theory" from "settlement theory". Trial selection theory generates predictions on important trial outcome parameters, such as the plaintiff win rate. The more general

way to a judgment in litigation without being settled beforehand. According to the Priest-Klein analysis, if litigants have symmetric stakes the win rate for plaintiffs will tend toward fifty percent, like coin tosses. The model assumes litigants have symmetric information and does not explicitly incorporate strategic behavior.

If litigants have asymmetric stakes, the Priest-Klein conjecture holds that the plaintiff win rate may exceed or fall below fifty percent.[12] One of the parties may have a strong desire to litigate in order to persuade or induce the court to change the law or to establish a reputation as a litigator. Priest and Klein introduced empirical evidence to support their hypothesis. Eisenberg (1990) reexamined the empirical evidence and found significant deviations from the fifty percent hypothesis. Waldfogel (1995), in contrast, finds evidence consistent with the Priest-Klein hypothesis; specifically, that plaintiff win rates tend toward fifty percent as the trial rate approaches zero. This finding suggests that the fifty percent win rate is more likely to be observed as legal uncertainty declines, or as trial costs increase.

The trial selection literature has been expanded by the incorporation of strategic behavior and asymmetric information. The first formal model of trial selection under informational asymmetry was offered in Shavell (1996).[13] Building on the screening model of Bebchuk (1984), Shavell concluded that any win rate percentage could be observed, and that there was no clear tendency for the plaintiff win rate to be less than or greater than fifty percent in the context of informational asymmetry. Hylton (2002), building on the signaling model of Png (1983, 1987), argues that win rates will tend to be consistent with the Priest-Klein analysis, and to show predictable deviations from fifty percent when information is asymmetric.

Although the asymmetric information models have provided a rigorous framework for trial selection theory, in contrast to the informal analysis of Priest and Klein, they have failed to generate a consistent set of testable predictions with respect to

---

settlement literature examines settlement incentives. On the settlement literature, see Bebchuk (1984), Daughety and Reinganum (1993), Spier (1992).

[12] For an early critique of the Priest-Klein model, see Wittman (1985). Wittman found that in a more general model there was no tendency toward a fifty percent win rate.

[13] Froeb (1993) presents a model of case selection under asymmetric information that precedes that of Shavell. However, Froeb focuses on the criminal law setting rather than the civil law setting of the Priest-Klein analysis. Hylton (1993) presents an informal analysis of trial selection in the civil context under informational asymmetry.

trial outcome parameters. The asymmetric information models have to be assessed within the context of their particular assumptions.

This paper offers an alternative strategy to modeling trial selection theory that captures both Priest-Klein and asymmetric information models.[14] We pinpoint the distributional assumptions that are needed for the Priest-Klein conjecture to hold, as well as those needed for informational asymmetry models.

III. Model

A. Assumptions

The core component of this model is the familiar Landes-Posner-Gould (LPG) litigation condition: parties choose to litigate rather than settle a dispute if and only if

$$(P_p - P_d) > \gamma \tag{1}$$

where P_p = plaintiff's estimate of the probability of a verdict in his favor, P_d = defendant's estimate of the probability of a verdict in plaintiff's favor; $\gamma = C/J$, where C = the sum of the plaintiff's litigation cost (C_p) and the defendant's litigation cost (C_d), and J = the value of the judgment. I assume that the settlement cost is zero (i.e., the bargaining costs to reach settlement are zero). If the litigation condition (1) holds, the set of mutually beneficial settlement agreements is empty, so the parties choose to litigate.

Each party's predictions is the sum of a rational estimate and an error term: $P_p = P'_p + \varepsilon_p$, $P_d = P'_d + \varepsilon_d$. If Ω_p represents the information set of the plaintiff, and Ω_d the information set of the defendant, $P'_p = E(P_p | \Omega_p)$, $P'_d = E(P_d | \Omega_d)$, $E(\varepsilon_p | \Omega_p) = 0$, $E(\varepsilon_d | \Omega_d) = 0$.

---

[14] These models use the litigants' predictions as the basis of a model settlement and trial selection. However, a trial selection theory can be based on any factor that determines the decision to litigate rather than settle. Eisenberg and Farber (1996) introduce the litigious-plaintiff hypothesis, which holds that win rates can be understood according to the plaintiff's cost of litigation, which varies more for individuals than for corporations. In this paper we will restrict ourselves to the Priest-Klein and asymmetric information models. Those models deliver the clearest implications for trial outcome parameters.

The error terms result from lapses or random shocks in the prediction process. The litigant has all of the information that he would use to reach a prediction of the probability of plaintiff victory, and then something happens that disturbs the true prediction from the rational estimate based on his information set. One can draw an analogy to a production process that experiences a glitch in one of every one thousand runs of the process, resulting in an altered product. The production glitch could be due to a random mechanical failure or to an external shock, such as a lighting strike. In the same sense, a litigant's prediction is based on a set of informational inputs and a cognitive process for converting those inputs into a prediction. But the conversion process is not error-free. A litigant can have a lapse that leads to an error in prediction; or the information needed to make the prediction may exceed the litigant's cognitive capacity at the moment. Alternatively, a random shock – an unforeseen intervention – can disturb the actual prediction from the rational estimate.

Suppose, for example, that both plaintiff and defendant have access to the same information bearing on the defendant's guilt. Both observed the manner in which the defendant, a medical doctor, conducted a test. A rational observer, examining the same information, would predict that the probability that the doctor would be held liable for malpractice is .6. However, the plaintiff may fail to take all of the facts favoring the defendant into account, or mistakenly believe some fact improves his likelihood of success when it does not, and predict that probability he will win is .65.

This sort of fact-based error should be distinguished from judicial error. The rational observer might predict a .6 likelihood of victory on the basis of the law, but the probability of judicial error could lead such an observer to predict a higher likelihood of victory. For example, a rational observer may believe that the plaintiff will be particularly appealing to a jury because of some unfortunate event in his life (e.g., job loss). Even though the law, as applied to the facts, would lead the rational observer to put the probability of plaintiff victory at .6, the possibility of judicial error, in the sense of deviating from the law, might lead the observer to set the probability of victory higher.

Rational predictions regarding the trial outcome will incorporate judicial error probabilities. Let v = probability of a verdict for the plaintiff, w = probability of guilt (equivalently, the probability that the defendant in a legal dispute violated the legal

standard), q_1 = probability that a defendant who has violated the legal standard will be found innocent (type-1 judicial error), and q_2 = probability that a defendant who has not violated the legal standard will be found guilty (type-2 judicial error). Assuming w, q_1, and q_2 are public information, the *rational estimate* (or the objective probability) *of a verdict in the plaintiff's favor*, given public information, is $v = w(1-q_1) + (1-w)q_2$. So that courts are at least minimally accurate, we assume $1-q_1 > q_2$.

We will focus on two information scenarios below. One is the scenario in which each litigant forms a rational estimate based solely on public information of the likelihood of a verdict for the plaintiff $(P'_p = P'_d)$. This is a case of symmetric information, with litigation driven by prediction errors. The other scenario is that of asymmetric information, in which the defendant knows whether or not he complied with the legal standard.

Another important assumption of this model is heteroscedasticity of the prediction error variances. From the perspective of a litigant, the outcome of a dispute is most uncertain when the rational component of the litigants' prediction is equal to fifty percent. This is the case in which the trial outcome is viewed by the litigant as a coin toss. We will therefore assume that the variance of the prediction error term is a function of the rational component of the litigant's prediction, and that the variance reaches a maximum when the rational component is fifty percent and with minima at the endpoints.

Symmetric Heteroscedasticity: $\sigma = \sigma(v) = \sigma(1-v)$ and, for $0 < v < \frac{1}{2}, \sigma'(v) > 0$.

B. The Probability of a Verdict for the Plaintiff, the Probability of Guilt, and the Frequency of Litigation

The probability of a verdict for the plaintiff (v) can be viewed as an index of case quality. If the distribution of v is uniform, then the probability of a verdict for the plaintiff is the same in all disputes; or equivalently, all claims have the same quality. If the v distribution is skewed right, then there is a tendency for legal disputes to involve guilty defendants. We will assume that v is governed by the probability density function

$h(v)$. Since v is a linear function of the probability of guilt, the probability density function for v will be related to the density for the probability of guilt w.

To simplify, we will start with the assumption that courts are error free, so that the probability of a verdict for the plaintiff is the same as the probability of guilt ($v = w$). Thus, $h(v)$ describes both the distribution of guilt and the distribution of the probability of a plaintiff verdict.

If all disputes were litigated to a judgment, the average plaintiff win rate would be determined by the distribution of guilt (or case quality) in the population. Thus, if all disputes were litigated the *expected plaintiff win rate* would be the same as the *expected population level of guilt*, which is

$$E(v) \equiv \mu = \int_0^1 vh(v)dv \tag{2}$$

Of course, not all cases are litigated. And the cases that are litigated are not necessarily a random sample from the population of disputes. For this reason it is necessary to consider the factors that influence litigation.

The probability of litigation conditional on the guilt level is $f = prob((P_p - P_d) > \gamma)$ which can be expressed as $f = prob(\varepsilon_p - \varepsilon_d > \gamma - \Delta)$, where $\Delta = P'_p - P'_d$. We assume that the error difference $\varepsilon_p - \varepsilon_d$ is generated by a truncated normal distribution with mean zero and variance σ^2, where $\varepsilon_p - \varepsilon_d \in [-1-\Delta, 1-\Delta]$. The variance of the error difference can be decomposed $\sigma^2 = \sigma^2_p + \sigma^2_d - 2\rho$.

Given these assumptions, the probability of litigation conditional on the guilt level can be expressed as

$$f = 1 - G, \tag{3}$$

where G, the probability of settlement conditional on the guilt level, is

$$G(\gamma - \Delta; -1 - \Delta, 1 - \Delta) = \frac{\Phi(\frac{\gamma - \Delta}{\sigma}) - \Phi(\frac{-1 - \Delta}{\sigma})}{\Phi(\frac{1 - \Delta}{\sigma}) - \Phi(\frac{-1 - \Delta}{\sigma})} \quad . \tag{4}$$

There are some immediate implications from this setup. As the degree of uncertainty regarding the probability of liability (σ) increases, the probability of litigation rises, a basic result of the Priest-Klein model. The frequency of litigation function f combines features from several models of litigation. As the cost of litigation rises relative to the judgment (γ increases) the probability of litigation falls (Landes-Posner-Gould). Over-optimism generates litigation. A negative correlation between prediction errors (ρ), consistent with litigant over-optimism (Shavell, 1982), reduces σ which in turn increases litigation.

C. Symmetric Information Model: Preliminary Results

The symmetric information case ($\Delta = 0$) has been associated with the Priest-Klein analysis, according to which litigation is driven by uncertainty and the plaintiff win rate tends toward fifty percent (assuming symmetric stakes). In this part, we will examine the basic components of that analysis.

The Priest-Klein conjecture is based on what should be observed within the sample of cases that have been litigated – i.e., the ex post plaintiff win rate. Within the sample of litigated cases, according to the conjecture, one will observe a plaintiff win rate that tends toward fifty percent, irrespective of the underlying distribution of guilt. We will focus on a formal construction of this argument.

Given the assumption of no judicial error, the expected plaintiff win rate is equal to the expected guilt level conditional on litigation

$$E(v \mid lit) \equiv \hat{\mu} = \int_0^1 \frac{vf(v)h(v)dv}{\int_0^1 f(v)h(v)dv} \quad . \tag{5}$$

For simplicity, we will refer to this measure as the plaintiff win rate. An alternative way of measuring the plaintiff success would focus on the expected plaintiff win rate within the set of litigated and settled disputes: $\int_0^1 vf(v)h(v)dv$. However, the selection literature has focused on the plaintiff win rate within litigated disputes, and this is a sensible decision given the great difficulty in finding reliable statistics on settled disputes (Priest and Klein, 1984).

The following proposition is implied by (5):

Proposition 1: *The expected population level of guilt is equal to the sum of the plaintiff win rate, multiplied by the trial rate, and the expected level of guilt within settled cases, multiplied by the settlement rate.*

Proof: This follows from a straightforward decomposition for the expected population win rate:

$$E(v) = \left(\int_0^1 f(v)h(v)dv\right)\left(\frac{\int_0^1 vf(v)h(v)dv}{\int_0^1 f(v)h(v)dv}\right) + \left(\int_0^1 G(v)h(v)dv\right)\left(\frac{\int_0^1 vG(v)h(v)dv}{\int_0^1 G(v)h(v)dv}\right) \blacksquare \quad (6)$$

Proposition 1 clarifies the relationship between the quality of claims and the frequencies of trial and settlement under the symmetric information model. Assume, for example, that the Priest-Klein conjecture holds, so that the plaintiff win rate is fifty percent. If, in addition, the expected guilt level in the population is .75 and the trial rate is .25, the expected guilt level within the sample of settled cases would be .83.

Proposition 1 also clarifies how little information is conveyed by the observed plaintiff win rate, and how much information must be obtained to fully understand the distribution of case quality within the samples of settled and litigated disputes. Knowing the plaintiff win rate tells us little about the average quality of all disputes. However, knowing the plaintiff win rate, the rate at which cases settle, and the average probability

of prevailing within the sample of settled disputes would allow us to infer the average quality of all claims.

It should be clear that the Priest-Klein conjecture depends on there being a "pyramid of disputes" (Felstiner, et al. 1980) with trial occurring within a fraction of disputes at the top. The conjecture holds that the plaintiff win rate is invariant to the underlying distribution of guilt (or case quality). This claim becomes more plausible as the trial rate diminishes.

Let the trial rate be represented by:

$$\theta \equiv \int_0^1 f(v)h(v)dv = 1 - \int_0^1 G(v)h(v)dv . \tag{7}$$

Let the average quality of settled disputes be represented by

$$\tilde{\mu} \equiv \frac{\int_0^1 vG(v)h(v)dv}{\int_0^1 G(v)h(v)dv} .$$

Then the expected guilt level decomposition (6) can be written as:

$$\mu = \theta\hat{\mu} + (1-\theta)\tilde{\mu} \tag{8}$$

With basic terms defined, we will focus on constructing the Priest-Klein argument within this model. The first result, which is an important component of the symmetric information analysis, shows the relationship between the frequency of litigation and the guilt level.

Proposition 2: The probability of litigation conditional on the guilt level, f, reaches its maximum when the guilt level (the probability of a verdict for the plaintiff) is equal to fifty percent.

Proof: We need to show that the f is symmetric around ½ and it is strictly increasing for $0 \le v \le$ ½. Combining these two properties, the proof is done.

Step 1: Symmetry property:

By definition, $f(\sigma(v)) = 1 - G(\sigma(v))$. Because $\sigma(v) = \sigma(1-v)$, we have

$$f(v) = 1 - G(\sigma(v))$$
$$= 1 - G(\sigma(1-v))$$
$$= f(\sigma(1-v)) = f(1-v)$$

Step 2: Monotonically increasing property:

First, we will show that the frequency of settlement, $G(\sigma)$, is strictly decreasing in σ. In other words,

$$\frac{\partial G(\sigma)}{\partial \sigma} < 0 \text{ where } G(\gamma; -1, 1) = \frac{\Phi(\frac{\gamma}{\sigma}) - \Phi(\frac{-1}{\sigma})}{\Phi(\frac{1}{\sigma}) - \Phi(\frac{-1}{\sigma})}.$$

Let $\lambda(\varepsilon; x_0) = \dfrac{-(x_0 + \varepsilon)\phi(x_0 + \varepsilon) + (x_0)\phi(x_0)}{\Phi(x_0 + \varepsilon) - \Phi(x_0)}$ where $x_0 = \dfrac{-1}{\sigma}$. The sign of $\dfrac{\partial G(\sigma)}{\partial \sigma}$ is the same as

$$\frac{\partial G(\sigma)}{\partial \sigma} \propto \left\{ \frac{[-(\frac{\gamma}{\sigma})\phi(\frac{\gamma}{\sigma}) + (\frac{-1}{\sigma})\phi(\frac{-1}{\sigma})]}{\Phi(\frac{\gamma}{\sigma}) - \Phi(\frac{-1}{\sigma})} - \frac{[-(\frac{1}{\sigma})\phi(\frac{1}{\sigma}) + (\frac{-1}{\sigma})\phi(\frac{-1}{\sigma})]}{\Phi(\frac{1}{\sigma}) - \Phi(\frac{-1}{\sigma})} \right\}$$

$$= \lambda(\varepsilon = \frac{\gamma + 1}{\sigma}; x_0) - \lambda(\varepsilon = \frac{2}{\sigma}; x_0)$$

The above equation can be considered as the difference of two points along the curve of $\lambda(\varepsilon; x_0)$ and the sign of $\dfrac{\partial G(\sigma)}{\partial \sigma}$ will depend on the shape of $\lambda(\varepsilon; x_0)$. We need to check when γ moves between the range of $[0,1]$, how does the value of $\lambda(\varepsilon = \dfrac{\gamma + 1}{\sigma}; x_0)$ change? We have

$$\frac{\partial \lambda(\varepsilon; x_0)}{\partial \gamma} = \frac{1}{\sigma} \frac{\ell(\varepsilon; x_0)\phi(x_0 + \varepsilon)}{(\Phi(x_0 + \varepsilon) - \Phi(x_0))^2}$$

where $\ell(\varepsilon; x_0) = ((x_0 + \varepsilon)^2 - 1)(\Phi(x_0 + \varepsilon) - \Phi(x_0)) + (x_0 + \varepsilon)\phi(x_0 + \varepsilon) - x_0\phi(x_0)$. To analyze the sign of $\dfrac{\partial \lambda(\varepsilon; x_0)}{\partial \gamma}$, it is equivalent to study the sign of $\ell(\varepsilon; x_0)$. Note that $\ell(\varepsilon; x_0)$ is increasing in γ because for any $\gamma \in [0,1]$

$$\frac{\partial \ell(\varepsilon; x_0)}{\partial \gamma} = \frac{2}{\sigma}(x_0 + \varepsilon)(\Phi(x_0 + \varepsilon) - \Phi(x_0))$$

$$= \frac{2}{\sigma}(\frac{\gamma}{\sigma})(\Phi(\frac{\gamma}{\sigma}) - \Phi(\frac{-1}{\sigma})) > 0$$

Now we show here that $\ell(\varepsilon; x_0)$ evaluated at the left-end point ($\varepsilon = \frac{1}{\sigma}$ or $\gamma = 0$) is

negative and $\ell(\varepsilon; x_0)$ evaluated at the right-end point ($\varepsilon = \frac{2}{\sigma}$ or $\gamma = 1$) is positive.

It is easy to show that at the left-end point

$$\ell^L = \ell(\varepsilon; x_0)\big|_{\varepsilon = \frac{1}{\sigma}, x_0 = -\frac{1}{\sigma}} = \Phi(-\frac{1}{\sigma}) - \Phi(0) + \frac{1}{\sigma}\phi(\frac{1}{\sigma}) < 0 \text{ because } \ell^L \text{ is increasing in } \sigma \text{ and it}$$

reaches zero while σ goes to infinity. Similarly, we can show that at the right-end

point $\ell^R = \ell(\varepsilon; x_0)\big|_{\varepsilon = \frac{2}{\sigma}, x_0 = -\frac{1}{\sigma}} = (\frac{1}{\sigma^2} - 1)(2\Phi(\frac{1}{\sigma}) - 1) + \frac{2}{\sigma}\phi(\frac{1}{\sigma}) > 0 \text{ because } \ell^R \text{ is decreasing in}$

σ and it reaches zero while σ goes to infinity.

Based on the above arguments, we can conclude that $\lambda(\varepsilon; x_0)$ is first decreasing and then

increasing in γ.

Now we show that $\lambda(\varepsilon = \frac{1}{\sigma}; x_0) = \lambda(\varepsilon = \frac{2}{\sigma}; x_0)$, which means that γ has the same value

evaluated at its left-end point and right-end point.

To sum up, $\lambda(\varepsilon; x_0)$ is first decreasing and then increasing in γ, and λ has the

same value at its left-end point and right-end point. As a result, for $\forall \gamma \in (0, 1]$ or

$\forall \varepsilon \in (\frac{1}{\sigma}, \frac{2}{\sigma}]$, $\lambda(\varepsilon = \frac{\gamma + 1}{\sigma}; x_0) < \lambda(\varepsilon = \frac{2}{\sigma}; x_0)$. This proves that $\frac{\partial G}{\partial \sigma} < 0$.

The final step of the argument is to show what the foregoing implies for the

frequency of litigation function. For any $0 \le v \le 1/2$, we have

$$\frac{\partial f(v)}{\partial v} = \frac{\partial (1 - G(v))}{\partial v} = -\frac{\partial G(v)}{\partial v} = -\frac{\partial G(\sigma)}{\partial \sigma}\frac{\partial \sigma(v)}{\partial v} > 0. \blacksquare$$

Although the frequency of litigation conditional on the guilt level is symmetric

about ½ and has a maximum at that point, the conditional probability of litigation

function   does   not   constitute   a   probability   density   over v,   given

that $\int_0^1 f(v)dv = 1 - \int_0^1 G(v)dv < 1$. Since f is not a density over v, the better description of

the Proposition 2 is that f is stochastically dominant when $v = \frac{1}{2}$.

The next step in formalizing the Priest-Klein conjecture is to identify the set of special cases in which the fifty percent claim holds with precision.

Proposition 3: *The plaintiff win rate is equal to fifty percent for any symmetric distribution of guilt.*

Proof: Note that

$$\int_0^1 vf(v)h(v)dv = \int_0^1 (1-v)f(1-v)h(1-v)dv$$

$$= \int_0^1 (1-v)f(v)h(v)dv = \int_0^1 f(v)h(v)dv - \int_0^1 vf(v)h(v)dv$$

The first equality is by integration by substitution; the second is by the symmetry properties of $f(v)$ and $h(v)$. From the above equation, we get

$$\int_0^1 vf(v)h(v)dv = \frac{1}{2}\int_0^1 f(v)h(v)dv$$

so that

$$E(v\,|\,lit) \equiv \hat{\mu} = \frac{\int_0^1 vf(v)h(v)dv}{\int_0^1 f(v)h(v)dv} = \frac{1}{2}.\ \blacksquare$$

Proposition 3 is a stronger result than the Priest-Klein conjecture. The Priest-Klein conjecture predicts a *tendency* of the plaintiff win rate toward fifty percent, no matter what form the underlying distribution of case quality takes. The reason is that cases of high and low quality are censored out of the final litigation sample by the settlement process. Our third proposition shows that the censoring process yields a precise fifty percent outcome in the case of a symmetric distribution of case quality. Moreover, there is no need for the trial rate to be small (as argued in Waldfogel, 1995) in

order to get the fifty percent prediction – the prediction is invariant to the trial rate in the case of symmetry.

The symmetric case quality distribution includes the normal (truncated) and uniform as special cases.[15] The result is not surprising, given that the average guilt level in the symmetric distribution case is fifty percent. We have only shown that when the guilt distribution is symmetric, the plaintiff win rate is not biased away from the population average guilt level as a result of the settlement process. Since the conditional probability of litigation given guilt level does not itself form a probability density over the guilt level, it is not immediately obvious that the settlement-censoring process should preserve the population average even in the symmetric guilt distribution case.

If the underlying distribution of case quality is not symmetrical, the plaintiff win rate is not necessarily fifty percent. The question then becomes under what conditions the censoring process causes the win rate to tend toward fifty percent.

In the more general setting, the Priest-Klein conjecture has two components, or implications. One is that the plaintiff win rate will be closer to fifty percent than is the expected population guilt level. The second is that as the censoring process becomes more severe, in the sense that the conditional litigation function f becomes more convex, the tendency of the plaintiff win rate toward ½ will become more pronounced or reliable. One might think of the first component as a weak requirement, but it is not. As we show below, the first implication is not more likely to be satisfied than is the second. We will take up the two implications in order.

First, we consider whether the plaintiff win rate will tend to be closer to fifty percent than is the expected population guilt level. In other words, if all disputes were litigated to a final judgment, the plaintiff win rate would always be the same as the expected population guilt level. Since not all disputes litigate to a final judgment, the win rate will differ from the population average guilt level. But will it move closer to fifty percent?

---

[15] Waldfogel, 1995, at 232, notes that his model assumes that the "distribution of filed cases' underlying quality is standard normal," which means that guilt is assumed to normally distributed in his study. Given Proposition 3, the fifty percent should be observed in any model that assumes a normal or symmetric distribution of case quality.

Proposition 4: The plaintiff win rate will be closer to fifty percent than is the expected population level of guilt, that is,

$$\left| \hat{\mu} - \frac{1}{2} \right| < \left| \mu - \frac{1}{2} \right|$$

if and only if $\tilde{\mu} <(>) \mu <(>) \dfrac{1}{2} + \dfrac{1}{2}(1-\theta)\tilde{\mu}$.

Proof: $\left(\hat{\mu} - \dfrac{1}{2} \right)^2 - \left(\mu - \dfrac{1}{2} \right)^2 < 0$ holds if and only if $\left(\hat{\mu} - \mu \right)\left(\hat{\mu} - (1-\mu) \right) < 0$.

Let $A = \displaystyle\int_0^1 v f(v) h(v) dv$, which is the numerator of $\hat{\mu}$. Integrating by parts

$$A = \left. H(1-G)v \right]_0^1 + \int_0^1 Hgv\,dv - \int_0^1 H(1-G)dv$$

After applying integration by parts again

$$A = \left. Hv \right]_0^1 + \int_0^1 vGh(v)dv - \int_0^1 H(v)dv \quad ,$$

or, equivalently,

$$A = 1 + \int_0^1 vGh(v)dv - \int_0^1 H(v)dv$$

It follows that

$$\hat{\mu} = \frac{1 + \displaystyle\int_0^1 vGh(v)dv - \int_0^1 H(v)dv}{\displaystyle\int_0^1 f(v)h(v)dv}$$

Since $\mu = 1 - \displaystyle\int_0^1 H(v)dv$, we have

$$\hat{\mu} - \mu = \frac{\mu \displaystyle\int_0^1 vGh(v)dv - \int_0^1 vGh(v)dv}{\displaystyle\int_0^1 f(v)h(v)dv} \quad . \tag{9}$$

From the same derivation it follows that

$$\hat{\mu} - (1 - \mu) = \frac{\mu - \int\limits_0^1 vGh(v)dv - \int\limits_0^1 H(v)dv}{\int\limits_0^1 f(v)h(v)dv} \tag{10}$$

One case in which the inequality $(\hat{\mu} - \mu)(\hat{\mu} - (1 - \mu)) < 0$ holds is when $\hat{\mu} > \mu$ and $(\hat{\mu} - (1 - \mu)) < 0$. From (9), $\hat{\mu} > \mu$ will hold when

$$\mu \int\limits_0^1 vGh(v)dv - \int\limits_0^1 vGh(v)dv > 0,$$

or equivalently when $\mu > \tilde{\mu}$. The other inequality follows from (8) and setting (10) less than zero. ∎

Given the relationship among μ, $\tilde{\mu}$, and $\hat{\mu}$ implied by (8), there is an equivalent set of conditions governing the relationship between μ and $\hat{\mu}$. For example, the conditions $\mu > \tilde{\mu}$ and $\mu < \frac{1}{2} + \frac{1}{2}(1 - \theta)\tilde{\mu}$ imply $\mu < \hat{\mu}$ and $\mu < 1 - \theta\hat{\mu}$. Without imposing additional conditions not imposed in Proposition 4, there is no reason to believe that these parameter constraints are likely to be satisfied in general.

The question remains whether the first implication of the Priest-Klein conjecture (Proposition 4) is likely to hold for a diverse set of guilt distributions. To examine this, we take a more formal approach below. Specifically, we examine the conditions that must be imposed on the distribution of guilt to satisfy the first and second implications of the Priest-Klein conjecture, given the symmetry property of the conditional probability of litigation function.

D. Implications of Priest-Klein Conjecture

Taking advantage of the symmetry of f, we will examine the properties that the population distribution of guilt must have in order for the first and second implications of

the Priest-Klein conjecture to hold. Recall that the first implication is that the plaintiff win rate is closer to fifty percent than is the population guilt level. We derive a condition under which this holds.

Theorem 1: Let v be a random variable with the density function h(v): $[0,1] \rightarrow$ $[0, +\infty)$. Then for any $f(v):[0,1] \rightarrow [0,+\infty)$ such that f is increasing on the interval $[0, \frac{1}{2}]$ and symmetric around $\frac{1}{2}$, i.e., $f(\frac{1}{2} - v) = f(\frac{1}{2} + v)$, for any $v \in [0, \frac{1}{2}]$,

$$\left| \frac{\int_0^1 vf(v)h(v)dv}{\int_0^1 f(v)h(v)dv} - \frac{1}{2} \right| \leq \left| \int_0^1 vh(v)dv - \frac{1}{2} \right| \tag{11}$$

if and only if for any $\alpha \in (0, \frac{1}{2}]$,

$$\left| \frac{\int_{\frac{1}{2}-\alpha}^{\frac{1}{2}+\alpha} vh(v)dv}{\int_{\frac{1}{2}-\alpha}^{\frac{1}{2}+\alpha} h(v)dv} - \frac{1}{2} \right| \leq \left| \int_0^1 vh(v)dv - \frac{1}{2} \right| \tag{12}$$

Before we begin the proof, we first see two definitions and one lemma.

Definition 1: A function $\chi_\alpha : [0,1] \rightarrow [0, +\infty)$ is called the characteristic function of the interval $[\frac{1}{2} - \alpha, \frac{1}{2} + \alpha]$ where $\alpha \in (0, \frac{1}{2}]$ if it is 1 on interval $[\frac{1}{2} - \alpha, \frac{1}{2} + \alpha]$ and zero otherwise.

Definition 2: A function f is called a symmetric simple function if there exist a finite increasing sequence of numbers $\alpha_n \in [0, \frac{1}{2}]$ $(n = 1, 2, 3, \ldots N)$ where $\alpha_0 = 0$ and $\alpha_N = \frac{1}{2}$ and a finite non-increasing sequence of nonnegative numbers M_n $(n = 0, 1, 2, \ldots, N)$ such that $f = M_n$ on the interval $[\frac{1}{2} - \alpha_n, \frac{1}{2} - \alpha_{n-1}]$ and $[\frac{1}{2} + \alpha_{n-1}, \frac{1}{2} + \alpha_n]$ for $n = 1, 2, 3, \ldots, N$.

Remark: based on the two definitions above, it is not hard to see any symmetric simple function f can be expressed as a linear combination of characteristic functions,

i.e., $f = \sum_{n=1}^{N-1}(M_n - M_{n+1})\chi_{\alpha_n} + M_N$ or $f = \sum_{n=1}^{N}(M_n - M_{n+1})\chi_{\alpha_n}$ if we define $M_{n+1} = 0$.

In particular, if $N = 2$, $f = (M_1 - M_2)\chi_{\alpha_1} + M_2$.

Lemma 1. For any function $f : [0,1] \to [0,+\infty)$ which is increasing on the interval $[0,\frac{1}{2}]$ and symmetric around $\frac{1}{2}$, there exists a sequence $\{f_n\}$ of symmetric simple functions such that $\lim_{n\to\infty} f_n(v) = f(v)$ almost everywhere on $[0,1]$ where $f_n(v) \le f_{n+1}(v)$ for $n = 1,2,3,\ldots$

We omit the proof of Lemma1 here since it simply follows the idea that any real function can be almost everywhere pointwisely approximated by the simple functions and the proof of that can be found in every real analysis book.

Proof of Theorem 1: Since (1.2) has the exactly the same form as

$$\left| \frac{\int_0^1 v\chi_\alpha(v)h(v)dv}{\int_0^1 \chi_\alpha(v)h(v)dv} - \frac{1}{2} \right| \le \left| \int_0^1 vh(v)dv - \frac{1}{2} \right|,$$

Taking $f = \chi_\alpha$ in (11) simply gives us the necessity.

As to the sufficiency, (12) gives us that (11) holds for the characteristic functions χ_α and by the linearity, (11) also holds for any characteristic function multiplied by a constant. Then since any symmetric simple function is a linear combination of characteristic functions, it is natural for us to expect (11) holds for any symmetric simple function f. Indeed, without lost of generality, to prove that we may assume $N = 2$. The proofs of other cases follow completely the same idea.

Now, $f = (M_1 - M_2)\chi_{\alpha_1} + M_2$. From above, we already have

$$\left| \frac{\int_0^1 v(M_1 - M_2)\chi_{\alpha_1}(v)h(v)dv}{\int_0^1 (M_1 - M_2)\chi_{\alpha_1}(v)h(v)dv} - \frac{1}{2} \right| \le \left| \int_0^1 vh(v)dv - \frac{1}{2} \right|,$$

i.e.,

$$-\left|\int_0^1 vh(v)dv-\frac{1}{2}\right|\le\frac{\int_0^1 v(M_1-M_2)\chi_{\alpha_1}(v)h(v)dv-\frac{1}{2}\int_0^1(M_1-M_2)\chi_{\alpha_1}(v)h(v)dv}{\int_0^1(M_1-M_2)\chi_{\alpha_1}(v)h(v)dv}$$ (13)

$$\le\left|\int_0^1 vh(v)dv-\frac{1}{2}\right|$$

and also

$$\left|\frac{\int_0^1 vM_2h(v)dv}{\int_0^1 M_2h(v)dv}-\frac{1}{2}\right|\le\left|\int_0^1 vh(v)dv-\frac{1}{2}\right|$$

i.e.,

$$-\left|\int_0^1 vh(v)dv-\frac{1}{2}\right|\le\frac{\int_0^1 vM_2h(v)dv-\frac{1}{2}\int_0^1 M_2h(v)dv}{\int_0^1 M_2h(v)dv}\le\left|\int_0^1 vh(v)dv-\frac{1}{2}\right|.$$ (14)

Then

$$\frac{\int_0^1 vf(v)h(v)dv}{\int_0^1 f(v)h(v)dv}-\frac{1}{2}=\frac{\int_0^1 v(M_1-M_2)\chi_{\alpha_1}(v)h(v)dv+\int_0^1 vM_2h(v)dv}{\int_0^1(M_1-M_2)\chi_{\alpha_1}(v)h(v)dv+\int_0^1 M_2h(v)dv}-\frac{1}{2}$$

$$=\frac{\{\int_0^1 v(M_1-M_2)\chi_{\alpha_1}(v)h(v)dv-\frac{1}{2}\int_0^1(M_1-M_2)\chi_{\alpha_1}(v)h(v)dv\}+\{\int_0^1 vM_2h(v)dv-\frac{1}{2}\int_0^1 M_2h(v)dv\}}{\int_0^1(M_1-M_2)\chi_{\alpha_1}(v)h(v)dv+\int_0^1 M_2h(v)dv}$$

By (13), (14), it is easy to see

$$-\left|\int_0^1 vh(v)dv-\frac{1}{2}\right|\le\frac{\int_0^1 vf(v)h(v)dv}{\int_0^1 f(v)h(v)dv}-\frac{1}{2}\le\left|\int_0^1 vh(v)dv-\frac{1}{2}\right|.$$

Therefore, (11) holds for any symmetric simple function.

Our final step is to use Lemma1 to extend the result of symmetric simple functions to their limits. For any function $f(v):[0,1]\to[0,+\infty)$ such that f is increasing on the interval $[0,\frac{1}{2}]$ and symmetric around $\frac{1}{2}$, by Lemma1, there exists a sequence $\{f_n\}$ of symmetric simple functions where $f_n(v)$ increases to $f(v)$ almost everywhere.

Then by the Lebesgue Dominance Convergence Theorem, we have

$$\int_0^1 vf(v)h(v)dv = \lim_{n\to\infty}\int_0^1 vf_n(v)h(v)dv \text{ and } \int_0^1 f(v)h(v)dv = \lim_{n\to\infty}\int_0^1 f_n(v)h(v)dv. \text{ Since}$$

$f_n(v)$ increases to $f(v)$, $\int_0^1 f_n(v)h(v)dv$ also increases to $\int_0^1 f(v)h(v)dv$. So $\int_0^1 f(v)h(v)dv$

should be greater than zero so that $\left| \dfrac{\int_0^1 vf(v)h(v)dv}{\int_0^1 f(v)h(v)dv} - \dfrac{1}{2} \right|$ is well-defined and from above,

$$\left| \frac{\int_0^1 vf(v)h(v)dv}{\int_0^1 f(v)h(v)dv} - \frac{1}{2} \right| = \lim_{n\to\infty} \left| \frac{\int_0^1 vf_n(v)h(v)dv}{\int_0^1 f_n(v)h(v)dv} - \frac{1}{2} \right| \le \left| \int_0^1 vh(v)dv - \frac{1}{2} \right|$$

which completes the proof.■

Theorem 1 means that a necessary and sufficient condition for the first implication of the Priest-Klein conjecture – that the win rate is closer to fifty percent than is the population level of guilt – is that the distribution of guilt have the following property: within any window around the fifty percent guilt level, the conditional mean given you are within the window is at least as close to fifty percent as is the mean of the distribution.

Theorem 1 is intuitive in view of the effect of the settlement-censoring process under symmetric information. The effect of settlement censoring is to exclude or under-weight cases in which the guilt level is close to one or close to zero. If the population distribution of guilt is such that the mean is at least as close to fifty percent within any window around ½ than for the whole distribution, then the censoring that takes place under symmetric information will always move the win rate closer to fifty percent than is the population guilt level. Thus, the first implication of Priest-Klein conjecture holds when (and only when) the population distribution of guilt satisfies the *window property* in (12).

The window property is satisfied by a diverse set of distributions. Certainly any symmetric distribution will satisfy the window property. Moreover, a large set of

asymmetric distributions also satisfy it. We tried several simulations with the Beta distribution and could not find a parameter set that failed to satisfy the window property.

The following example illustrates a case in which Theorem 1 does not hold. The distribution of guilt in this example does not satisfy the window property.

Example: Let $h(v) = v + 15(2^{13})(v - \frac{1}{2})^{14}$ describe the population distribution of guilt, and let the function $f(v) = -(v - \frac{1}{2})^2 + .25$ approximate the conditional probability of litigation given the guilt level. In the model, the conditional probability of litigation function is determined by the underlying truncated normal. We have chosen a simple function that is symmetric about $\frac{1}{2}$ to aid the illustration. The population distribution of guilt has the following shape:
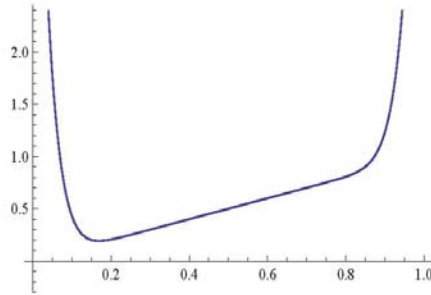

Figure 1: A hypothetical guilt distribution

For this example,

$$\left| \mu - \frac{1}{2} \right| = \left| E(v) - \frac{1}{2} \right| = \frac{1}{12}$$

$$\left| \hat{\mu} - \frac{1}{2} \right| = \left| \frac{\int_0^1 vf(v)h(v)dv}{\int_0^1 f(v)h(v)dv} - \frac{1}{2} \right| = \left| \frac{\frac{117}{2040}}{\frac{5}{51}} - \frac{1}{2} \right| = \frac{17}{200} > \frac{1}{12} \; ,$$

so the win rate conditional on litigation is further away from fifty percent than is the population average level of guilt.

Recall that the second implication of the Priest-Klein conjecture is that as the censoring process becomes more severe, in the sense that the conditional probability of litigation function f becomes more convex, the plaintiff win rate will tend more reliably toward fifty percent. While Proposition 2 is a statement about the first order stochastic dominance, the second implication is really a statement about second order stochastic dominance. In other words, as the conditional litigation function becomes more dominant in the second order sense, the tendency of the plaintiff win rate to fifty percent becomes more certain or reliable.

Theorem 2: *Let v be a random variable with the density function*

$h(v):[0,1] \to [0,+\infty)$. *The function $f(v):[0,1] \to [0,+\infty)$ is increasing on the*

interval $[0,\frac{1}{2}]$ *and symmetric around* $\frac{1}{2}$, *i.e.,* $f(\frac{1}{2}-v) = f(\frac{1}{2}+v)$ *for any* $v \in [0,\frac{1}{2}]$.

For any $\varepsilon > 0$, define $f_\varepsilon:[0,1] \to [0,+\infty)$ as $f_\varepsilon(v) = \frac{1}{\varepsilon} f(\frac{1}{2} - \frac{\frac{1}{2}-v}{\varepsilon})$ *if* $\left|\frac{1}{2}-v\right| \le \frac{\varepsilon}{2}$

and $f_\varepsilon(v) = 0$ if $\frac{\varepsilon}{2} < \left|\frac{1}{2}-v\right| \le \frac{1}{2}$. *Define $g_\varepsilon(x):[0,1] \to [0,1]$ be* $g_\varepsilon(x) = \frac{f_\varepsilon(x)}{Max\{f_\varepsilon(x)\}}$, *then*

$$\lim_{\varepsilon \to 0} \frac{\int_0^1 v g_\varepsilon(v) h(v) dv}{\int_0^1 g_\varepsilon(v) h(v) dv} = \frac{1}{2}.$$

In particular, if $E(V) \neq \frac{1}{2}$, $\left| \frac{\int_0^1 v g_\varepsilon(v) h(v) dv}{\int_0^1 g_\varepsilon(v) h(v) dv} - \frac{1}{2} \right| \le \left| \int_0^1 v h(v) dv - \frac{1}{2} \right|$ *when ε is small enough.*

Proof: First notice that the boundedness of $f(v)$ guarantees the finiteness of the

integrals. Next to show $\lim\limits_{\varepsilon \to 0} \dfrac{\int_0^1 vg_\varepsilon(v)h(v)dv}{\int_0^1 g_\varepsilon(v)h(v)dv} = \dfrac{1}{2}$, it is enough to

prove $\lim\limits_{\varepsilon \to 0} \dfrac{\int_0^1 vf_\varepsilon(v)h(v)dv}{\int_0^1 f_\varepsilon(v)h(v)dv} = \dfrac{1}{2}$.

Now we construct an instrument function $F(v)$ which extends the domain to $(-\infty, +\infty)$. Note that $F(v)$ is exactly the same as $f(v)$, except for the extended domain for the demonstration of the following proof. Define $F(v):(-\infty, +\infty) \to [0, +\infty)$ as $F(v) = f(v)$ for $v \in [0,1]$ and $F(v) = 0$ otherwise. Then define $G(v) = F(\dfrac{1}{2} - v)$ and

thus $G(v) = 0$ if $v \notin [-\dfrac{1}{2}, \dfrac{1}{2}]$ and the function $G_\varepsilon(v) = \dfrac{1}{\varepsilon}G(\dfrac{v}{\varepsilon}) = f_\varepsilon(\dfrac{1}{2} - v)$ for $v \in [-\dfrac{1}{2}, \dfrac{1}{2}]$ and equals 0 otherwise. Then,

$$\frac{\int_0^1 vf_\varepsilon(v)h(v)dv}{\int_0^1 f_\varepsilon(v)h(v)dv} = \frac{\int_{-\frac{1}{2}}^{\frac{1}{2}}(\frac{1}{2}-v)f_\varepsilon(\frac{1}{2}-v)h(\frac{1}{2}-v)dv}{\int_{-\frac{1}{2}}^{\frac{1}{2}} f_\varepsilon(\frac{1}{2}-v)h(\frac{1}{2}-v)dv} = \frac{\int_{-\infty}^{\infty}(\frac{1}{2}-v)G_\varepsilon(v)h(\frac{1}{2}-v)dv}{\int_{-\infty}^{\infty} G_\varepsilon(v)h(\frac{1}{2}-v)dv}$$

$$= \frac{1}{2} - \frac{\int_{-\infty}^{\infty} vG_\varepsilon(v)h(\frac{1}{2}-v)dv}{\int_{-\infty}^{\infty} G_\varepsilon(v)h(\frac{1}{2}-v)dv}$$

To prove the final result, now it is enough to show $\lim\limits_{\varepsilon \to 0} \dfrac{\int_{-\infty}^{\infty} vG_\varepsilon(v)h(\frac{1}{2}-v)dv}{\int_{-\infty}^{\infty} G_\varepsilon(v)h(\frac{1}{2}-v)dv} = 0$.

Let $z = \dfrac{v}{\varepsilon}$, by changing variables,

$$\frac{\int v G_\varepsilon(v) h(x-v) dv}{\int G_\varepsilon(v) h(x-v) dv} = \frac{\int v \frac{1}{\varepsilon} G(\frac{v}{\varepsilon}) h(x-v) dv}{\int \frac{1}{\varepsilon} G(\frac{v}{\varepsilon}) h(x-v) dv} = \frac{\int \varepsilon z G(z) h(x-\varepsilon z) dz}{\int G(z) h(x-\varepsilon z) dz} = \varepsilon \frac{\int z G(z) h(x-\varepsilon z) dz}{\int G(z) h(x-\varepsilon z) dz}$$

Since $G(v) = 0$ outside $[-\frac{1}{2}, \frac{1}{2}]$, $\left| \dfrac{\int z G(z) h(x-\varepsilon z) dz}{\int G(z) h(x-\varepsilon z) dz} \right| = \left| \dfrac{\int_{-\frac{1}{2}}^{\frac{1}{2}} z G(z) h(x-\varepsilon z) dz}{\int_{-\frac{1}{2}}^{\frac{1}{2}} G(z) h(x-\varepsilon z) dz} \right| \leq \dfrac{1}{2}.$

Therefore,

$$\lim_{\varepsilon \to 0} \frac{\int_{-\infty}^{\infty} v G_\varepsilon(v) h(\frac{1}{2}-v) dv}{\int_{-\infty}^{\infty} G_\varepsilon(v) h(\frac{1}{2}-v) dv} = \lim_{\varepsilon \to 0} \varepsilon \frac{\int z G(z) h(x-\varepsilon z) dz}{\int G(z) h(x-\varepsilon z) dz} = 0,$$

which completes the proof.■

The intuition for Theorem 2 is straightforward. As the conditional probability of litigation function becomes more convex, it will approach the shape of a spike at the fifty percent guilt level. The rest of the argument can be drawn on the reasoning for the familiar Chebyshev inequality. However, intuition based on the Chebyshev inequality is not technically appropriate in this case. The Chebyshev inequality addresses a relationship between a realization of a random variable and its mean, as the variance of the underlying distribution collapses. In contrast, this is a model of a censoring process applied to a random variable (the guilt level). The "distribution" that "collapses" in this case is the conditional probability function describing the censoring process. Rather than a case where the variance of a specific probability distribution collapses as in the Chebyshev result, this is instead a case where the censoring process becomes more convex around a particular conditioning value.

The second theorem provides an alternative way of looking at the argument that the fifty percent prediction of Priest and Klein holds in the limit as the trial rate goes to

zero (Waldfogel, 1995). The "in the limit" result can be viewed as a statement about the convexity of the censoring function f. That convexity, in turn, is determined by the convexity of the prediction error variance $\sigma(v)$. Thus, if we use $\sigma''(v)/\sigma(v)$ as a measure of the convexity of the error variance, the limiting result depends on $\sigma''(v)/\sigma(v)$ increasing. As the convexity of $\sigma(v)$ increases, the win rate approaches fifty percent.

In terms of implications for litigation, the second theorem implies that the fifty percent result holds as relative uncertainty lessens at the extremes of the guilt probability scale, relative to the center. Greater convexity translates into cases settling with greater frequency when the guilt status of the defendant is close to one or close to zero. This is more likely to be true when the law is clear, which makes it easier to determine whether a violation occurred, or the facts of the dispute are clear.

The following simple numerical example illustrates the effects of greater convexity in the censoring function.

Example: Consider two censoring processes, with the second being more convex than the first. Assume the connection between the guilt of the defendant, the distribution of guilt, and the probability of litigation given guilt is as follows:

v	¼	½	¾
$Prob(lit \mid v)$	.2	.6	.2
$Prob(v = v_o)$	.3	.3	.4

Under these assumptions,

$$E(v) = (\tfrac{1}{4})(.3) + (\tfrac{3}{4})(.4) + (\tfrac{1}{2})(.3) = .52$$

$$Prob(lit) = (.3)(.2) + (.4)(.2) + (.3)(.6) = .32$$

$$E(v \mid lit) = \left(\frac{1}{4}\right)\left(\frac{(.2)(.3)}{.32}\right) + \left(\frac{1}{2}\right)\left(\frac{(.3)(.6)}{.32}\right) + \left(\frac{3}{4}\right)\left(\frac{(.2)(.4)}{.32}\right) = .515$$

Now consider an alternative censoring process that is more convex in the sense described earlier.

v	¼	½	¾
$Prob(lit \mid v)$	.1	.7	.1
$Prob(v = v_o)$	.3	.3	.4

For this case, we have

$$E(v) = (¼)\,(.3) + (¾)(.4) + (½)(.3) = .52$$

$$Prob(lit) = (.3)(.1) + (.4)(.1) + (.3)(.7) = .28$$

$$E(v \mid lit) = \left(\frac{1}{4}\right)\left(\frac{(.1)(.3)}{.28}\right) + \left(\frac{1}{2}\right)\left(\frac{(.7)(.3)}{.28}\right) + \left(\frac{3}{4}\right)\left(\frac{(.1)(.4)}{.28}\right) = .509$$

The trial rate could also approach zero as the cost of litigation increases relative to the judgment (γ increases) (Waldfogel, 1995). This is equivalent to shifting the censoring function f down, which will generally result in the disputes at the extreme of the guilt distribution settling. However, Theorem 1 and the counterexample imply that this effect will not necessarily lead to a movement toward fifty percent. If the window property is not satisfied by the guilt distribution, a shift downward of the censoring function may not result in a win rate closer to fifty percent.[16]

E. Generalization of Conjecture

Much of the foregoing has been an effort to understand the conditions under which the Priest-Klein conjecture may be valid. The core assumption is that the variance of the difference between the parties' trial outcome prediction errors is a function of the level of guilt (heteroscedasticity) and that this function is symmetric about the midpoint of the range of guilt levels. The prediction errors can be viewed as random shocks that cause the litigant's subjective prediction of the probability of a verdict for the plaintiff to differ from the rational estimate based on the litigant's information set. Since parties are

---

[16] Of course, if the underlying distribution of guilt is symmetric (which would satisfy the window property of Theorem 1), then shifting the censoring function down will lead to the fifty percent result. Waldfogel (1995, at 232) assumes the distribution of guilt is normal, which may explain his finding. We impose no functional assumptions on the distribution of guilt.

assumed to have symmetric information, the rational prediction of the guilt level is the same for both parties and is equal to the true guilt level. When the litigant's information generates a rational prediction equal to the most uncertain guilt level, fifty percent, his subjective prediction is most sensitive to a random shock.

Of course, the variance symmetry property need not hold. It could be that the variance of the prediction error difference is greater when guilt levels are small than when they are large (or the converse). Consider, for example, the simple case of failing to look both ways before crossing an intersection in a car. A failure to look is obviously negligence, and the prediction of guilt would be relatively insensitive to information-processing errors. On the other hand, looking both ways may have been held reasonable care in a previous case, but may not necessarily be so in any other case with more complicated facts. The actual prediction would be relatively more sensitive to prediction error than in the clear non-compliance scenario (failing to look).

The possibility that the prediction error variance might be asymmetric suggests a generalization on the Priest-Klein conjecture. If the assumption of symmetric heterscedasticity is replaced with an assumption of asymmetric heteroscedasticity, then there should be no clear tendency for the plaintiff win rate to move toward fifty percent, even under the window property of Theorem 1. Indeed, if the direction of the asymmetry (or skewness) is right (left), the model presented here would suggest a win rate that is greater than (less than) fifty percent. The easiest way to see this generalization of the Priest-Klein conjecture is to consider the case where the distribution of guilt is symmetric.

Proposition 5: *Suppose the prediction error difference is asymmetrically hesteroscedastic, with a left skew. Then the plaintiff win rate will be less than fifty percent for any symmetric guilt distribution. Conversely, if the prediction error difference has a right skew, the plaintiff win rate will be greater than fifty percent for any symmetric guilt distribution.*

Proof: Assume the following form of asymmetric heteroskedasticity:

(1) $\sigma(v) \geq \sigma(1-v)$ and $\sigma'(v) > 0$ for $0 < v < \frac{1}{2}$.

(2) $\sigma(v) \le \sigma(1-v)$ and $\sigma'(v) < 0$ for $\frac{1}{2} < v < 1$.

This implies

$$f(v) \ge f(1-v) \text{ for } 0 < v < \tfrac{1}{2}, \text{ and}$$

$$f(v) \ge f(1-v) \text{ for } \tfrac{1}{2} < v < 1,$$

which means that the conditional probability of litigation function, graphed over v, is skewed toward the left. Following the proof of Proposition 2

$$\int_0^1 vf(v)h(v)dv = \int_0^1 (1-v)f(1-v)h(1-v)dv$$

$$= \int_0^1 (1-v)f(1-v)h(v)dv$$

$$\le \int_0^1 (1-v)f(v)h(v)dv$$

The second equality follows from the symmetry of $h(v)$. The inequality results from the left skewness of f (induced by $\sigma(v)$). From the above inequality, we get

$$\int_0^1 vf(v)h(v)dv \le \frac{1}{2}\int_0^1 f(v)h(v)dv$$

so that

$$\hat{\mu} \le \frac{1}{2}.$$

If we reverse the inequalities in (1) and (2), the same argument leads to the conclusion that $\hat{\mu} \ge \frac{1}{2}$. ∎

Proposition 5 illuminates some limitations on the Priest-Klein conjecture. Recall that we described the conjecture as having two key implications: first, that the plaintiff win rate would be closer to fifty percent than is the population guilt level, and, second, that as the censoring function became more convex, the tendency of the win rate to equal fifty percent would become more reliable. It should be clear that in the presence of asymmetric heteroscedasticity, the first implication is not valid. Theorem 1 shows that for any population distribution of guilt that satisfies what we have called the window property, the win rate will be closer to fifty percent than is the expected population guilt level. That result is inconsistent with Proposition 5. The symmetric distribution of guilt

obviously satisfies the window property of Theorem 1. That the settlement process leads to a win rate different from fifty percent shows that the first implication of the Priest-Klein conjecture is not valid when the symmetric heteroscedasticity assumption is replaced by asymmetric heteroscedasticity.

The second implication of the Priest-Klein conjecture no longer continues to hold, though the deviation from fifty percent is vanishing in the limit. As the censoring process becomes more severe (convex), the expected win rate conditional on litigation will necessarily move toward fifty percent. However, the plaintiff win rate will be biased in the direction in which the prediction error variance is skewed.

F. Extension to Judicial Error Setting and Asymmetric Information

To simplify the analysis we assumed in the previous part that courts operate without error. Under that assumption, the probability of a verdict for the plaintiff is the same as the probability that the defendant is guilty. If courts make mistakes, the probability of a verdict for the plaintiff will differ from the probability of guilt. Error distorts the link between case quality and merit.

When courts make mistakes, the foregoing analysis remains intact, though the notation is more complicated. The distribution of the probability of a verdict in favor of the plaintiff will now differ from the population guilt distribution. But since the two are linearly related, the two distributions will also be related. If the density for the probability of guilt is $k(w)$, then the density for the probability of a plaintiff verdict is

$$h(v) = \frac{1}{(1-q_1-q_2)} k(\frac{v-q_2}{(1-q_1-q_2)}) ,$$

over the interval q_2 and $1-q_1-q_2$. None of the results in the previous part is dependent on the assumption that v runs from 0 to 1, and they can be modified where necessary to incorporate the new limits.

For the asymmetric information setting examined next, there are substantial differences between the model with judicial error and the model without error. The most

basic implications are the same between these two models, but the details are so different that it would be difficult to analyze the model with error by reinterpreting the model without error. Because the error-free model is a special case of the judicial error model, we will examine the model with error in the remaining parts.

As a general matter, $P'_p = w_p(1-q_{1p}) + (1-w_p)q_{2p}$, where $w_p = E(w| \Omega_p)$, $q_{1p} = E(q_1| \Omega_p)$, $q_{2p} = E(q_2| \Omega_p)$; and, similarly, $P'_d = w_d(1-q_{1d}) + (1-w_d)q_{2d.}$ For each litigant, the rational prediction of the probability of guilt and the prediction of the likelihood of error will both depend on the information available to him. In the case of the uninformed litigant, his rational predictions are equal to the objective probability based on public information. For the informed litigant, his estimate of w is equal to 1 in the case of non-compliance by the defendant, or 0 in the case of compliance by the defendant.

Thus, in a dispute involving an informed-innocent defendant and an uninformed plaintiff, $P'_p = v$ and $P'_d = q_2$, and the actual predictions are:

$$P_p = v + \varepsilon_p \tag{15}$$

$$P_d = q_2 + \varepsilon_d \tag{16}$$

As in the previous part, we will assume litigation is determined by condition (1) from the LPG model. However, the LPG condition requires additional justification in a model that allows for asymmetric information. In asymmetric information models the settlement decision is not governed solely by the LPG condition. Indeed, litigation occurs more frequently under Bebchuk's screening model than would be implied by the LPG model. The reason is that the uninformed plaintiff trades off the cost of litigating for the gain of getting a larger share of the settlement surplus.

In spite of this, there are two justifications for using the LPG model to describe litigation here. The first is Occam's razor. This model assumes that error in litigants' predictions of the probability of plaintiff victory plays a substantial role in generating litigation. In the absence of such error, settlement decisions would separate the guilty from the innocent, unless the parties behaved strategically (Bebchuk, 1984). With error

in predictions, the defendant's willingness to settle does not reveal his type. Strategic behavior is therefore not a necessary feature of the litigation model.

Second, the LPG condition is broadly consistent with both screening and signaling models of litigation. In signaling models LPG condition (1) still separates the conditions in which litigation may occur from those in which it will not (Png, 1983; Png 1987, Hylton 2002). In other words, the LPG condition provides a lower bound on the settlement zone. In Bebchuk's screening model, the LPG condition is a necessary but not sufficient condition for settlement – that is, it provides an upper bound on the settlement zone.[17] The LPG model permits us to capture in broad form the implications of both types of asymmetric information model without having to commit to one of them.

We consider two asymmetric information scenarios: where the defendant has the informational advantage and where the plaintiff has the informational advantage. When the defendant has the informational advantage, the frequency of litigation will depend on the defendant's type. If the plaintiff is uninformed and the defendant is innocent, $P'_p = w(1-q_1) + (1-w)q_2$, $P'_d = q_2$; thus, $\Delta_I = w(1-q_1-q_2)$.

In the asymmetric information setting w should be understood as the rational prediction of the probability of guilt given information publicly available about the parties and the facts of the dispute. In the previous section of this paper, where we considered symmetric information, the plaintiff and the defendant chose the same estimate for w. In this part, only the uninformed party chooses w as his estimate of the probability of guilt. Having chosen w, the uninformed party's estimate of the likelihood of a verdict for the plaintiff is equal to $v = w(1-q_1)+(1-w)q_2$.

Let f_I be the probability of litigation for the innocent-defendant, uninformed plaintiff pairing, given guilt level w. The frequency of litigation is

$$f_I = 1 - G_I \tag{17}$$

where the probability of settlement, is

---

[17] One way to think of the difference between the models is by drawing an analogy to price setting. In Bebchuk's model the uninformed litigant screens by setting the settlement amount. He sets the price in order to grab some of the surplus, which results in the standard monopoly-pricing inefficiency. In the signaling model the informed actors set prices in a way that price discriminates.

$$G_I(\gamma - \Delta_I; -1 - \Delta_I, 1 - \Delta_I) = \frac{\Phi(\frac{\gamma - \Delta_I}{\sigma}) - \Phi(\frac{-1 - \Delta_I}{\sigma})}{\Phi(\frac{1 - \Delta_I}{\sigma}) - \Phi(\frac{-1 - \Delta_I}{\sigma})} \quad . \tag{18}$$

It should be clear that litigation can occur even though the defendant's guilt level is low. The plaintiff is uninformed, so makes his decision on whether to litigate on the population average level of guilt.

If the defendant is guilty, $P'_p = w(1-q_1) + (1-w)q_2$, $P'_d = 1-q_1$, and $\Delta_G = - (1-w)(1-q_1-q_2)$. Let f_G be the probability of litigation for the guilty-defendant, uninformed plaintiff pairing, given guilt level w.

$$f_G = 1 - G_G \tag{19}$$

where the probability of settlement, is

$$G_G(\gamma - \Delta_G; -1 - \Delta_G, 1 - \Delta_G) = \frac{\Phi(\frac{\gamma - \Delta_G}{\sigma}) - \Phi(\frac{-1 - \Delta_G}{\sigma})}{\Phi(\frac{1 - \Delta_G}{\sigma}) - \Phi(\frac{-1 - \Delta_G}{\sigma})} \tag{20}$$

Proposition 6: *For any* $\Delta_1 > \Delta_2$,

$$G(\gamma - \Delta_1; -1 - \Delta_1, 1 - \Delta_1) < G(\gamma - \Delta_2; -1 - \Delta_2, 1 - \Delta_2) \quad .$$

Proof: This is to show that $G(\gamma - \Delta; -1 - \Delta, 1 - \Delta)$ is decreasing in Δ, or increasing in $(-\Delta)$, namely $\frac{\partial G}{\partial(-\Delta)} > 0$.

Let $a = \frac{-1 - \Delta}{\sigma}, b = \frac{\gamma - \Delta}{\sigma}, c = \frac{1 - \Delta}{\sigma}$, we have

$$\frac{\partial G}{\partial(-\Delta)} = \frac{(\phi(b) - \phi(a))(\Phi(c) - \Phi(a)) - (\phi(c) - \phi(a))(\Phi(b) - \Phi(a))}{(\Phi(c) - \Phi(a))^2}$$

Let $x = \Phi(t)$, $y = \phi(t)$, we have $\dfrac{dy}{dx} = \dfrac{-t\phi(t)}{\phi(t)} = -t < 0$ and $\dfrac{d^2 y}{dx^2} = \dfrac{-1}{\phi(t)} < 0$

For any value of t, the curve of y with regard to x is strictly concave. On the curve, we can identify three points where $A(\Phi(a), \phi(a)), B(\Phi(b), \phi(b)), C(\Phi(c), \phi(c))$. Since $a < b < c$, we have $\Phi(a) < \Phi(b) < \Phi(c)$ so that point A is to the left of B and point B is to the left of C. Given the curve is strictly concave, $slope(AB) > slope(AC)$, namely

$\dfrac{(\phi(b) - \phi(a))}{(\Phi(b) - \Phi(a))} > \dfrac{(\phi(c) - \phi(a))}{(\Phi(c) - \Phi(a))}$. This proves that $\dfrac{\partial G}{\partial(-\Delta)} > 0$. ∎

It follows that $f_I > f_G$, since it is a special case where $\Delta_I = w(1 - q_1 - q_2)$ and $\Delta_G = -(1 - w)(1 - q_1 - q_2)$. Since $\Delta_I > \Delta_G$, we have $G_I(\gamma - \Delta_I; -1 - \Delta_I, 1 - \Delta_I) < G_G(\gamma - \Delta_G; -1 - \Delta_G, 1 - \Delta_G)$; and it follows that $f_I > f_G$. The reason $f_I > f_G$ is that guilty informed defendants settle their cases at a higher rate than do the innocent informed defendants. Since guilty defendants are more likely to be held liable than are the innocent, the potential payoff from litigating is lower for the guilty.

As in the previous parts we will examine the plaintiff win rate. When the defendant has the informational advantage, the win rate is

$$\hat{\mu} = \dfrac{\displaystyle\int_0^1 [w f_G (1 - q_1) + (1 - w) f_I q_2] k(w) dw}{\displaystyle\int_0^1 [w f_G + (1 - w) f_I] k(w) dw} \tag{21}$$

Following the approach of the previous part, we can derive the key implications for trial selection by examining the setting in which the guilt distribution is symmetric. If the fifty percent prediction fails to hold in the symmetric guilt distribution case, then it will clearly not hold in the more general analysis.

Proposition 7: *Assume the guilt distribution is symmetric. When the defendant (plaintiff) has the informational advantage, the plaintiff win rate will be less than (greater than) or equal to fifty percent, as long as error rates satisfy reasonable assumptions.*

Proof: Letting Ψ represent the denominator, (21) can be rewritten as:

$$\hat{\mu} = (1-q_1)\left(\frac{\int_0^1 wf_{wG}k(w)dw}{\int_0^1 f_{wG}k(w)dw}\right)\left(\frac{\int_0^1 f_{wG}k(w)dw}{\Psi}\right) + q_2\left(\frac{\int_0^1 (1-w)f_{wI}k(w)dw}{\int_0^1 f_{wI}k(w)dw}\right)\left(\frac{\int_0^1 f_{wI}k(w)dw}{\Psi}\right)$$

The symmetry of f_I and f_G are established by the first step of the proof of Proposition 2. Assuming the symmetry of the guilt distribution and taking advantage of Proposition 3, this simplifies to:

$$\hat{\mu} = (1-q_1)\left(\frac{1}{2}\right)\left(\frac{\int_0^1 f_G k(w)dw}{\Psi}\right) + q_2\left(\frac{1}{2}\right)\left(\frac{\int_0^1 f_I k(w)dw}{\Psi}\right) \tag{22}$$

Since $\int_0^1 wf_G k(w)dw < \int_0^1 (1-w)f_I k(w)dw$ and $\int_0^1 wf_G k(w)dw < 2\int_0^1 f_I k(w)dw$,

$$\frac{\int_0^1 f_G k(w)dw}{\Psi} < 1 \tag{23}$$

and

$$\frac{\int_0^1 f_I k(w)dw}{\Psi} > 1. \tag{24}$$

Rewrite (22) as

$$\hat{\mu} = (1-q_1)\left(\frac{1}{2}\right)\tilde{\theta} + q_2\left(\frac{1}{2}\right)\hat{\theta}.$$

Where $\tilde{\theta}$ is equal to the left hand side of (23) and $\hat{\theta}$ is equal to the left hand side of (24). Moreover, it is straightforward to show that $\tilde{\theta} < \hat{\theta}, \tilde{\theta} < 1, \hat{\theta} > 1$, and $\tilde{\theta} + \hat{\theta} = 2$. The question is whether $\hat{\mu} \leq \frac{1}{2}$. To simplify the remaining argument, assume the judicial error terms q_1 and q_2 are close in value, so that $q_2 = q_1 + \varepsilon$. Then $2\hat{\mu} \leq 1$ if and only if $(1-q_1)\tilde{\theta} \leq 1 - (q_1 + \varepsilon)\hat{\theta}$. Since $\tilde{\theta} + \hat{\theta} = 2$, this is equivalent to $0 \leq (1-2q_1)(1-\tilde{\theta}) - \varepsilon(2-\tilde{\theta})$. For ε sufficiently close to zero and $q_1 < \frac{1}{2}$, this holds with a strict inequality. Indeed, for reasonable assumptions on the judicial error terms, the condition holds. It is violated when $q_2 > q_1$ (ε large) and q_1 is close to $\frac{1}{2}$, which are conditions under which judicial decisions are essentially coin tosses. ∎

It follows that even if the window property of Theorem 1 holds, the plaintiff win rate will not tend toward fifty percent. However, there are directional tendencies. Where the defendant has the informational advantage, the win rate will tend toward a level that is less than or equal to fifty percent, and conversely where the plaintiff has the informational advantage.

G. Asymmetric Stakes

We have to this point ignored the issue of asymmetric stakes, which was a part of the original Priest-Klein analysis. The asymmetric stakes theory was offered by Priest and Klein as an explanation for areas of litigation, such as medical malpractice, in which plaintiff win rates were consistently below the fifty percent prediction of their model. The model presented in the previous parts generates plaintiff win rates that deviate from fifty percent without the additional assumption of asymmetric stakes. However, the

model can be extended to incorporate asymmetric stakes.[18] In the remainder of this paper we will discuss the application of this model to trial win rate data.

III. Discussion and Implications

We have identified three main settings in which to examine the tendency of plaintiff win rates toward fifty percent, as hypothesized by Priest and Klein. The first is where the censoring function (probability of litigation conditional on the guilt level) is symmetric over the range of guilt levels with a maximum at fifty percent. The second is where the censoring function has a maximum at fifty percent but is asymmetrical. The third is where the litigants have asymmetric information. There is a straightforward generalization of the Priest-Klein hypothesis: (1) under symmetric information and uncertainty, the plaintiff win rate tends toward fifty percent; (2) under symmetric information and asymmetric uncertainty, the win rate tends toward the pole with greatest uncertainty; and (3) under asymmetric information and symmetric uncertainty, the win rate favors the party with the information advantage

The symmetric censoring function corresponds to the original Priest-Klein analysis. This is a setting that can be described, to use Holmes's language, as one of *penumbral uncertainty*. The uncertainty around the trial verdict is greatest when the defendant's compliance is in a region of uncertainty between clear non-compliance and clear compliance. Uncertainty at the edges is minimal. Consider a medical malpractice case. Under the law of torts, a doctor's negligence is determined by his compliance with custom of the medical profession. If the doctor complies with the custom he is not negligent, and conversely. Uncertainty is minimal at the endpoints of compliance and noncompliance. Litigation is most likely to arise when the fact of compliance is unclear. If both doctor and patient have symmetric information with respect to the facts, the conditions of the Priest-Klein analysis will hold. Symmetric information might be observed when the doctor is charged with negligence in a course of conduct that is observable by the patient – for example, the diagnosis of a physical ailment.

---

[18] Let the plaintiff's payoff from litigation be the sum of the damage award J and the "stakes gain" G. The defendant loses J and the "stakes loss" L. The condition for litigation then becomes $(P_p–P_d)(J+L) + P_p(G–L) > C$, which can then be treated in a manner similar to (1).

The second case identified by the model of this paper is one of symmetric information where the level of uncertainty is greatest midway between the endpoints of compliance and noncompliance, but uncertainty is greater at one endpoint than at the other. To use Holmes's terms, this is the setting of *asymmetric (or one-sided) umbral uncertainty*. This case was not identified in the Priest-Klein analysis, and has not been examined in previous analyses of trial selection. Consider the road crossing example. Failing to look both ways is negligence. On the other hand, looking both ways may have been deemed reasonable care in a previous case, but may not be under more complicated facts. In other words, while there is a set of acts that will always be deemed a failure to comply, it may be impossible to identify a set of acts that will always be deemed compliance. The umbral uncertainty associated with compliance may be greater than that associated with noncompliance. The win rate will tend toward the node with greatest umbral uncertainty. Thus, if the uncertainty surrounding the court's decision in the case of compliance is greater than in the case of noncompliance, the win rate will tend toward a level less than fifty percent.

As an alternative example consider malpractice. Many courts have replaced local standards of medical custom with national standards. In a setting where the litigants expect that the local standard to be replaced by a national standard, there would be one-sided umbral uncertainty, with the uncertainty associated with compliance greater than that associated with noncompliance. The uncertainty would be legal rather than factual, but the difference is not important in this example. Legal uncertainty simply means that the given facts of the case might generate a different legal outcome than expected under a particular view of the law. This is no different from saying that given a clear rule, factual uncertainty implies that the clear legal rule might generate a different outcome than expected under a particular view of the facts.

The third set of cases identified by this model is informational asymmetry, a topic that the literature has explored. Medical malpractice offers many scenarios of informational asymmetry. If the patient is put under anesthesia, he will not know what procedures were performed on him, while the doctor will know. Any negligence claim brought against the doctor will involve a patient who is ignorant of the facts bearing on the doctor's compliance with medical custom during the period he was anesthetized.

Doctors who know they violated the custom will tend to settle, while doctors who know that they did not will litigate to judgment. The result will be a tendency of the plaintiff win rate below fifty percent.

This model does not incorporate the asymmetric stakes argument of Priest and Klein, though it is not difficult to extend the model to do so. For Priest and Klein, asymmetric stakes provided an argument that explained plaintiff win rates that consistently deviated from fifty percent. However, we do not have a need to employ the asymmetric stakes argument, given that plaintiff win rates that deviate from fifty percent are generated on the basis of the uncertainty captured within this model.

There are difficult questions associated with the empirical application of this model. Any attempt to empirically test this model would have to distinguish cases of penumbral uncertainty from asymmetric umbral uncertainty, and in addition distinguish cases of informational asymmetry. Since penumbral uncertainty is a characteristic of all litigated disputes, the analyst would have to try to identify the cases where the effects of umbral uncertainty or informational asymmetry are likely to outweigh those of penumbral uncertainty. Similarly, any attempt to use the model to provide a positive theory of trial outcome statistics must distinguish the various types of uncertainty that generate plaintiff win rate patterns that diverge from the fifty percent prediction.

Table 1 provides a summary of the salient results from empirical studies of trial selection. The studies summarized in the table are those of Priest and Klein (1984), Eisenberg (1990), Kessler, Meites, and Miller (1996), and Waldfogel (1995). Priest and Klein is the only article in the table that uses a sample based on state court trials. The remaining three studies use samples from federal court. The distinction could be important in examining the evidence of trial selection. To bring a lawsuit in federal court, the plaintiff's case must raise a federal question (e.g., an issue under a federal statute) or there must be "diversity of citizenship" between the plaintiff and defendant, meaning that the parties are not from the same state. For the ordinary personal injury torts shown in the first two rows of Table 1, a large number of the lawsuits must be based on diversity, which means that the amount in controversy had to cross a minimum financial threshold. The diversity requirements impart some degree of selection immediately in the federal sample.

In any event, if there are selection biases embedded in the federal sample, they are not obvious to the naked eye. The most obvious impression is that the fifty percent hypothesis of Priest and Klein does not hold generally. This paper's model implies that the fifty percent win rate prediction will hold under rather special conditions: no (one-sided) informational asymmetry, no asymmetric umbral uncertainty, and the underlying guilt distribution satisfies the conditions of Theorem 1.

TABLE 1:

Summary of Plaintiff Win Rates in Four Studies

	Priest-Klein (1984)	Eisenberg (1990)	Kessler et al. (1996)	Waldfogel (1995)
Personal injury torts (non-traffic)	.51[*] (3,045)	.46 (3,808)	.27 (97)	.15 (639)
Personal injury torts (traffic)	.47 (9,987)	.60 (3,261)	.17 (35)	.34 (337)
Product liability	.43 (477)	.25 (3,255)	.41 (83)	.01 (243)
Medical malpractice	.40 (202)	.38 (697)	.00 (6)	.19 (143)
Employment discrimination		.21 (7,165)	.15 (448)	.14 (666)
Antitrust		.43 (586)	.26 (31)	.32 (98)
Years	1959-1979	1978-1985	1982-1987	1984-1987
Court	State (Cook County)	Federal	Federal (7th Circuit)	Federal (S.D.N.Y.)

Notes: *Based on common carrier, injury on property, street hazard, and dramshop categories in Table 7 of Priest and Klein (1984).

The results that seem closest to the fifty percent prediction are the first two cells from the Priest-Klein study, showing plaintiff win rates for "non-traffic torts" and for "traffic torts",[19] and the first cell from the Eisenberg study, showing the win rate from "non-traffic torts". However, even within their respective rows, the win rates from the other studies diverge widely from the fifty percent prediction, though most of them are below fifty percent. Waldfogel (1995) finds evidence that the fifty percent hypothesis holds in samples in which the trial rate is relatively low. If the data in the first two rows were disaggregated by court or by year, the fifty percent result might be observed in specific courts or years in which trials were less frequent. In addition, the first row (personal injury torts, non-traffic) aggregates different areas of litigation (premises liability, worker injury, etc.), some of which may fail to satisfy the requirements of the fifty percent prediction.

The more consistent patterns appear in the third through sixth rows. The product liability, medical malpractice, employment discrimination, and antitrust categories show a consistent pattern of plaintiff win rates less than fifty percent.

The model in this paper incorporates two uncertainty based explanations for the win rates below fifty percent. One is defendant-sided informational asymmetry: if the defendant has an informational advantage over the plaintiff, then the plaintiff win rate will tend to be less than fifty percent. The other explanation is compliance-centered umbral uncertainty: if there is greater uncertainty associated with compliance than with noncompliance, win rates will tend toward the level associated with the compliance endpoint.

The consistent low win rates for medical malpractice litigation invite an explanation based on informational asymmetry. Informational asymmetry with respect to facts is a recognized feature of medical malpractice litigation.

---

[19] The personal injury torts (non-traffic) cell from Priest and Klein (the first cell in Table 1) excludes traffic-based torts, medical malpractice, product liability, and worker injury torts. Priest and Klein treated the last three categories as anomalous because of asymmetric stakes. We think informational asymmetry is largely responsible for the results reported by Priest and Klein in these categories. Priest and Klein reported high win rate for worker injury cases. They could not explain the result. The most plausible explanation is that the worker injury lawsuits, which are brought by workers against non-employers (e.g., property owners), reflect selection based on asymmetric information. Most weak worker lawsuits will be selected into workers compensation. Workers who have relatively strong claims (say, because of an absence of any contributory negligence) will select into the tort system by suing non-employers.

The low product liability win rates could also be explained by the asymmetric information theory. Liability for defective design is determined by the risk-utility test, which compares the incremental risk and the incremental utility of the challenged design relative to some feasible safer alternative. The test may give an informational advantage to the defendant, provided the defendant knows more than the plaintiff about the incremental risk and the range of feasible alternatives, which seems plausible. But this is unlikely to be a complete explanation for the low win rates. The risk and utility features of many products are easily observable or at least discoverable early in a trial.[20] In addition, some courts have shifted the burden of proof on the risk-utility where evidence is entirely in the hands of the defendant.[21] Given this, compliance-centered umbral uncertainty provides an alternative explanation for the low win rates in product liability litigation. In cases where the product cannot be made entirely free of risk, there is no absolutely safe feasible alternative; the manufacturer has to trade off some risks for others. In these cases, there is likely to be a great deal of uncertainty over the extent of compliance. A firm that concludes, after a review of the risk tradeoffs, that its product is relatively safe will still face substantial uncertainty over whether a court would find that the design was defective.

Employment discrimination and antitrust share the same features as product liability. There is a plausible argument in both areas that the defendant has an informational advantage, more so in the discrimination case than in the antitrust case. The defendant in an employment discrimination action knows more about the efficiency justifications for its decisions than will the plaintiff. But employment discrimination is an area where compliance-centered umbral uncertainty is likely to exist. An employer can design a test for screening potential employees with the purpose of avoiding a discriminatory impact, and still be unsure that it would not lose in a discrimination lawsuit later brought on the basis of the test outcome.

We have not attempted to conduct an empirical test of the selection pressures identified in this article. However, such a test should start with identifying areas in which informational asymmetry or one-sided umbral uncertainty are likely to dominate the

---

[20] Of course, the discovery process may not fully reveal the private information of the defendant, see Hay (1995). Discovery may be insufficient to change the dispute into a symmetric information case.

[21] See, e.g., Barker v. Lull Engineering Co., 573 P.2d 443 (Cal. 1978).

penumbral uncertainty emphasized by Priest and Klein. This requires an examination of the relevant legal tests that will determine the outcome of litigation. The simplest types of litigation to examine will be those in which the legal test examines the conduct of only one party. For example, in medical malpractice cases, the legal test focuses on the doctor alone; contributory negligence is almost never an issue. Within the set of tests that focus on the conduct of one actor, the relative influence of informational asymmetry and penumbral uncertainty can be assessed.

IV. Conclusion

We have offered a model that unifies existing trial selection theories and introduces a new theory as well. The model identifies conditions under which the fifty percent prediction holds, and also suggests that the conditions under which it holds are unlikely to be observed generally. The symmetric information version of our model, which incorporates the Priest-Klein analysis, generates the fifty percent prediction of the Priest-Klein analysis. However, it also predicts win rates that deviate from fifty percent when litigation uncertainty is greater at one end of the guilty spectrum than at the other. The general prevalence of below fifty percent win rate statistics in litigation may be explained by the simple fact that trial outcome uncertainty surrounding compliance is often greater than uncertainty surrounding noncompliance.

References

Lucian Arye Bebchuk, Litigation and Settlement under Imperfect Information, Rand Journal of Economics, vol.15 (3), 404-415 (1984)

Yeon-Koo Che and Jong Goo Yi, The Role of Precedents in Repeated Litigation, Journal of Law, Economics, and Organization, 9(2), 399—424, 1993

Andrew F. Daughety and Jennifer F. Reinganum, Endogenous Sequencing in Models of Settlement and Litigation, Journal of Law, Economics, & Organization, vol. 9, 314-348 (1993).

Theodore Eisenberg, Testing the Selection Effect: A New Theoretical Framework with Empirical Tests, Journal of Legal Studies, vol. 19 (2), 337-358 (1990).

Theodore Eisenberg and Michael Heise, Plaintiphobia in State Courts? An Empirical Study of State Court Trials on Appeal, Journal of Legal Studies, vol. 38, 121-155 (2009).

Theodore Eisenberg and Henry S. Farber, The Litigious Plaintiff Hypothesis: Case Selection and Resolution, Rand Journal of Economics, vol. 28, S92- S112 (1997).

W.L. Felstiner, R.L. Abel, A. Sarat, The emergence and transformation of disputes: Naming, blaming, claiming …, Law and Society Review, Vol. 15, 630-649 (1980-1981).

Luke Froeb, Adverse Selection of Cases for Trial, 13 International Review of Law and Economics 317-324 (1993).

Nicola Gennaioli and Andrei Shleifer, A. The Evolution of Common Law. Journal of Political Economy, Vol. 115, 43-68 (2007).

John P. Gould, The Economics of Legal Conflicts, 2 J. Legal Stud. 279 (1973).

Bruce L. Hay, Effort, Information, Settlement, Trial, Journal of Legal Studies, vol.24, 29-62 (1995).

Oliver Wendell Holmes, Jr. The Common Law (Boston: Little, Brown & Co., 1881).

Keith N. Hylton, An Asymmetric Information Model of Litigation, International Review of Law and Economics, vol.22, 153-175 (2002).

Keith N. Hylton, Asymmetric Information and the Selection of Disputes for Litigation, Journal of Legal Studies, vol. 22, 187-210, (1993).

Keith N. Hylton and Haizhen Lin, Trial Selection Theory and Evidence: A Review, forthcoming in Encyclopedia of Law and Economics: Volume X: Procedural
Law and Economics, edited by Chris Sanchirico, Edward Elgar Publishing, 2010.

Daniel Kessler, Thomas Meites, Geoffrey Miller, Explaining Deviations from the Fifty-Percent Rule: A Multimodal Approach to the Selection of Cases for Litigation, Journal of Legal Studies, vol. 25, 233-259 (1996).

William M. Landes, An Economic Analysis of the Courts, 14 J. Law & Econ. 61 (1971).

Thomas J. Miceli, Legal Change and the Social Value of Lawsuits, University of Connecticut Economics Department Working Paper 2008-34, available at
http://www.econ.uconn.edu/working/2008-34.pdf.

Richard A. Posner, An Economic Approach to Legal Procedure and Judicial Administration, 2 J. Legal Stud. 399 (1973).

George Priest, The Common Law Process and the Selection of Efficient Rules, Journal of Legal Studies, vol.6, 65-82 (1977).

George Priest, The Selection of Disputes for Litigation, Journal of Legal Studies, vol.13, 1-55, (1984).

George Priest, Selective Characteristics of Litigation, Journal of Legal Studies, vol.9, 399-421 (1980).

Steven Shavell, Any Frequency of Plaintiff Victory at Trial is Possible, Journal of Legal Studies, vol.25, 493-501 (1996).

Peter Siegelman 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, vol. 24(2), pages 427-62, June, 1995.

Peter Siegelman and Joel Waldfogel, Toward a Taxonomy of Disputes: New Evidence through the Prism of the Priest/Klein Model, 28 Journal of Legal Studies, 101-130 (1999).

Kathryn E. Spier The Dynamics of Pretrial Negotiation, Review of Economic Studies, vol. 59, 93-108 (1992).

Joel Waldfogel, Reconciling Asymmetric Information and Divergent-Expectations Theories of Litigation, Journal of Law and Economics, vol. 41, 451-476 (1998).

Joel Waldfogel, The Selection Hypothesis and the Relationship between Trial and Plaintiff Victory, 103 Journal of Political Economy, 229-260 (Apr., 1995).

Donald Wittman, Is the Selection of Cases for Trial Biased? Journal of Legal Studies, vol. 14, 185-214 (1985).