A FORMALIZATION AND EXTENSION OF THE PRIEST-KLEIN HYPOTHESIS

Boston University School of Law Working Paper No. 11-28
(June 14, 2011)

Keith N. Hylton
Boston University School of Law

Haizhen Lin
Indiana University, Kelley School of Business

This paper can be downloaded without charge at:

A FORMALIZATION AND EXTENSION OF
THE PRIEST-KLEIN HYPOTHESIS

Keith N. Hylton*
Haizhen Lin**

April 2011

Abstract: This paper provides a formal model of the Priest-Klein trial selection hypothesis, and extends the hypothesis as well. We derive the conditions under which the hypothesis is valid, and examine implications for the relationship between trial outcome uncertainty and litigation.

* Boston University, School of Law, knhylton@bu.edu. The material in this paper was originally part of an earlier working paper titled “Trial Selection Theory: A Unified Model”. This paper focuses on formalizing the Priest-Klein hypothesis, while the revised version of the initial working paper extends the model to incorporate asymmetric information and asymmetric stakes theories of trial selection.

** Indiana University, Kelley School of Business, hzlin@indiana.edu.
I. Introduction

In *The Common Law*, Oliver Wendell Holmes noted that

> [I]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).2 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.3 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.4

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.5 The original Priest-Klein article provides an informal argument, as do later articles testing it.6 The lack of a formal model makes it difficult to separate important from unimportant assumptions in the Priest-Klein analysis, and to formally distinguish Priest-Klein from competing theories of trial selection. The interpretation of tests of trial selection theories is also dependent on understanding the scope and limitations of the underlying models.

---

1 Holmes (1881), at 127.
2 Baxter (1980) anticipates the core of the Priest-Klein argument but does not examine the theory at the same level of detail.
4 On the theory of stakes asymmetry and litigation, see Che and Yi (1993).
5 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.
This paper offers a formal model of trial selection in the manner of Priest and Klein.\textsuperscript{7} We use the model to examine within a single framework the results in the trial selection literature and to offer new results. 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.\textsuperscript{8} While this is helpful to the fifty percent prediction in the symmetric information model, we find that it is not a necessary feature.

We show 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 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.

We provide 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 censoring

\textsuperscript{7} The Priest-Klein model focuses on the expectations of litigants, which we focus on in this paper. We recognize that trial selection theories can be based on a larger set of variables than examined in the symmetric information and asymmetric information literature. Eisenberg and Farber (1996) develop a theory of selection based on the cost of litigation.

\textsuperscript{8} 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.
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.\footnote{By considering a skewed censoring function, we are able to propose a new explanation for plaintiff win rates that consistently deviate from fifty percent. Explanations of win rate need no longer be a "horse race" between Priest-Klein and asymmetric information models (Waldfogel, 1998, at 451).}

Our approach, which consists of relaxing symmetry conditions for the censoring and guilt distribution functions, can be contrasted with the asymmetric information modeling of trial selection, which is based on models of strategic behavior in settlement. Both approaches can account for deviations from the fifty percent prediction. However, the implications of the strategic behavior models are highly dependent on their particular assumptions.\footnote{Compare, for example, the divergent implications of Bebchuk (1984) and Png (1987), which both present strategic behavior models of the settlement process.} We avoid dependence on specific strategic behavior assumptions. A rich analysis of trial selection theory can be based on general statistical properties of the variables that determine the propensity to litigate. We argue that patterns in trial win rate data can be explained largely on the basis of the generalized version of the Priest-Klein model in this paper.

II. Literature

Trial selection theory,\footnote{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 settlement literature examines settlement incentives. On the settlement literature, see Bebchuk (1984), Daughety and Reinganum (1993), Spier (1992).} as initially presented by Priest and Klein, builds on the idea, recognized at least since Holmes, that only the most uncertain disputes go all the 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.\footnote{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.} 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, found support for the hypothesis; specifically, that plaintiff win rates tend toward fifty percent as the trial rate approaches zero.

The trial selection literature has been expanded by the incorporation of strategic behavior and asymmetric information. Shavell (1996), building on the screening model of Bebchuk (1984), concluded that any plaintiff win rate percentage could be observed, and that there was no clear tendency for the win rate to be less than or greater than fifty percent in the presence of informational asymmetry. Hylton (2002), building on the signaling model of Png (1983, 1987), argued that win rates will tend to be consistent with the Priest-Klein conjecture, and to show predictable deviations from fifty percent when information is asymmetric. Although the asymmetric information models provide a rigorous alternative to Priest-Klein, they have not yielded a consistent set of testable predictions with respect to trial outcome parameters.

This paper offers a model of trial selection that is entirely within the Priest-Klein symmetric information assumptions. However, the model is also capable of explaining plaintiff win rates that deviate from fifty percent without having to resort to the hypothesis of asymmetric stakes. In addition, the model provides an approach to understanding deviations from fifty percent that does not require modeling incentives under asymmetric information. Because of these features, this model offers a sparse framework capable of accounting for much of the observed patterns in plaintiff win rates.

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

\[(P_p - P_d) > \gamma\]

---

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.

14 Both symmetric and asymmetric information 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.
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 plaintiff’s and defendant’s litigation costs, and \( J \) = the value of the judgment. I assume that the settlement cost (i.e., bargaining cost to reach settlement) is zero.

Each party’s prediction is the sum of a rational estimate and an idiosyncratic error term: \( P_p = \nu + \epsilon_p \), \( P_d = \nu + \epsilon_d \). The rational prediction \( \nu \) is equal to the objective probability of a verdict for the plaintiff. Since we assume that there is no judicial error, the objective probability of a verdict for the plaintiff is equal to the probability of guilt. The information sets of the plaintiff and defendant are the same in this model, so that the rational predictions are the same too.

The error terms result from lapses or shocks in the prediction process. 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.

The prediction error variances are assumed to be heteroscedastic (Wittman, 1985). 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. We will therefore assume that the variance of the prediction error 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 \( 0 < v < \frac{1}{2} \), \( \sigma'(v) > 0 \).

---

15 In the presence of judicial error, the probability of guilt and the probability of a verdict for the plaintiff would differ. Specifically, if \( w = \) probability of guilt, \( q_1 = \) 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); \( \nu = w(1-q_1) + (1-w)q_2 \). The model presented in the text can be modified easily to allow for judicial error, but doing so would complicate the discussion without seriously modifying the results.

16 This sort of fact-based error should be distinguished from judicial error, which would be reflected in the rational component of the prediction \( (P') \). 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.
B. Probability of Plaintiff Victory, Probability of Guilt, and Frequency of Litigation

The probability of a verdict for the plaintiff ($v$) is also 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 – i.e., all claims have the same quality. If the distribution is skewed right, then legal disputes tend to involve guilty defendants. We assume that $v$ is governed by the probability density function $h(v)$. Since we assume that courts are error free, $h(v)$ 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 (case quality) in the population. Thus, when all disputes are litigated the expected plaintiff win rate is the same as the expected population level of guilt, which is

$$E(v) = \mu = \int_{0}^{1} vh(v) dv$$

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 $f = \text{prob}((P_p - P_d) > \gamma)$, which can be expressed as $f = \text{prob}(\varepsilon_p - \varepsilon_d > \gamma)$. We assume that the error difference $\varepsilon_p - \varepsilon_d$ is generated by a truncated normal distribution with mean zero and variance $\sigma^2$, where $\varepsilon_p - \varepsilon_d \in [-1, 1]$. The variance of the error difference is $\sigma^2 = \sigma_p^2 + \sigma_d^2 - 2\rho$.

Given these assumptions, the probability of litigation conditional on the guilt level can be expressed as $f = 1 - G$, where

$$G(\gamma, \sigma) = \frac{\Phi\left(\frac{\gamma}{\sigma}\right) - \Phi\left(\frac{-1}{\sigma}\right)}{\Phi\left(\frac{1}{\sigma}\right) - \Phi\left(\frac{-1}{\sigma}\right)}.$$

is the probability of settlement conditional on the guilt level. We will refer to the probability of litigation conditional on the guilt level in some contexts as the conditional probability of litigation function, the frequency of litigation function, or as the censoring function.
This setup incorporates earlier results from the literature. As the degree of uncertainty regarding liability (\(\sigma\)) increases, the probability of litigation rises (Priest-Klein, 1984); as the cost of litigation rises relative to the judgment (\(\gamma\) increases) the probability of litigation falls (Landes-Posner-Gould); over-optimism (a negative correlation of prediction errors, \(\rho\), reduces \(\sigma\)) generates litigation (Shavell, 1982).

The symmetric information case studied here 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. The Priest-Klein conjecture holds that within the sample of litigated cases, 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 here.

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

\[
E(v|lit) \equiv \hat{\mu} = \frac{\int f(v)h(v)dv}{\int f(v)dv}.
\]                      (4)

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: \(\int f(v)h(v)dv\). However, the selection literature has focused on the plaintiff win rate within litigated disputes, and this is a sensible decision given the difficulty of finding reliable statistics on settled disputes. The following proposition is implied by the win rate expression (4).

**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:
\[ E(v) = \left( \int_0^1 f(v) h(v) dv \right) \left( \int_0^1 v f(v) h(v) dv \right) + \left( \int_0^1 G(v) h(v) dv \right) \left( \int_0^1 v G(v) h(v) dv \right) \]  \hspace{1cm} (5)

Proposition 1 clarifies how little information is conveyed by the plaintiff win rate, and how much information must be obtained to understand the distribution of case quality within the samples of settled and litigated disputes. Knowing the numerical value of the plaintiff win rate tells us little about the average quality of disputes. However, knowing the plaintiff win rate, the trial rate, and the average probability of prevailing within the sample of settled disputes would allow us to infer the average quality of all claims.

Let the trial rate be represented by:

\[ \theta = \int_0^1 f(v) h(v) dv = 1 - \int_0^1 G(v) h(v) dv. \]  \hspace{1cm} (6)

Let the average quality of settled disputes be represented by

\[ \hat{\mu} = \frac{\int_0^1 v G(v) h(v) dv}{\int_0^1 G(v) h(v) dv}. \]

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

\[ \mu = \theta \hat{\mu} + (1 - \theta) \bar{\mu} \]  \hspace{1cm} (7)

With basic terms defined, we will focus on constructing the Priest-Klein argument. The first result 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 \( \frac{1}{2} \) and it is strictly increasing for \( 0 \leq v \leq \frac{1}{2} \). 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 \( \sigma \), i.e.,

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

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

\[
\frac{\partial G(\sigma)}{\partial \sigma} \approx \left\{ \frac{-(\frac{\gamma}{\sigma})\phi(x_0 + \varepsilon) + (\frac{1}{\sigma})\phi(-1)}{\Phi(x_0 + \varepsilon) - \Phi(x_0)} - \frac{(-1)\phi(x_0 + \varepsilon) + (\frac{1}{\sigma})\phi(-1)}{\Phi(x_0 + \varepsilon) - \Phi(x_0)} \right\}
\]

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

The above equation can be considered as the difference of two points along the curve of \( \lambda(\varepsilon; x_0) \) and the sign of \( \frac{\partial G(\sigma)}{\partial \sigma} \) will depend on the shape of \( \lambda(\varepsilon; x_0) \). We need to check when \( \gamma \) moves between the range of [0,1], how the value of \( \lambda(\varepsilon = \frac{\gamma + 1}{\sigma}; x_0) \) changes. 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 \( \frac{\partial \lambda(\varepsilon; x_0)}{\partial \gamma} \), it is equivalent to study the sign of \( \ell(\varepsilon; x_0) \). Note that \( \ell(\varepsilon; x_0) \) is increasing in \( \gamma \) because for any \( \gamma \in [0,1] \)

\[
\frac{\partial \ell(\varepsilon; x_0)}{\partial \gamma} = \frac{2}{\sigma} (x_0 + \varepsilon)(\Phi(x_0 + \varepsilon) - \Phi(x_0)) = \frac{2}{\sigma} (\frac{\gamma}{\sigma})(\Phi(\frac{\gamma}{\sigma}) - \Phi(-1)) > 0
\]

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

\[ \ell^L = \ell(\varepsilon; x_0) \bigg|_{\varepsilon = \frac{1}{\sigma}, x_0 = -\frac{1}{\sigma}} = \Phi(-\frac{1}{\sigma}) - \Phi(0) + \frac{1}{\sigma} \phi(\frac{1}{\sigma}) < 0 \]

because \( \ell^L \) is increasing in \( \sigma \) and it reaches zero while \( \sigma \) goes to infinity. Similarly, we can show that at the right-end point

\[ \ell^R = \ell(\varepsilon; x_0) \bigg|_{\varepsilon = \frac{2}{\sigma}, x_0 = -\frac{1}{\sigma}} = \left(\frac{1}{\sigma^2} - 1\right)(2\Phi(\frac{1}{\sigma}) - 1) + \frac{2}{\sigma} \phi(\frac{1}{\sigma}) > 0 \]

because \( \ell^R \) is decreasing in \( \sigma \) and it reaches zero while \( \sigma \) goes to infinity. Based on the above arguments, we can conclude that \( \lambda(\varepsilon; x_0) \) is first decreasing and then increasing in \( \gamma \). Moreover, it is straightforward to show that at \( x_0 = -\frac{1}{\sigma} \), \( \lambda(\varepsilon = \frac{1}{\sigma}; x_0) = \lambda(\varepsilon = \frac{2}{\sigma}; x_0) \), which means that \( \gamma \) 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 \( \gamma \), and \( \lambda \) has the same value at its left-end point and right-end point. As a result, for \( \forall \gamma \in (0,1] \) or

\[ \forall \varepsilon \in \left(\frac{1}{\sigma}, \frac{2}{\sigma}\right], \lambda(\varepsilon = \frac{\gamma + 1}{\sigma}; x_0) < \lambda(\varepsilon = \frac{1}{\sigma}; x_0) \] 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 \( 0 \leq \nu \leq 1/2 \), we have

\[ \frac{\partial f(\nu)}{\partial \nu} = \frac{\partial (1 - G(\nu))}{\partial \nu} = - \frac{\partial G(\nu)}{\partial \nu} = - \frac{\partial G(\sigma)}{\partial \sigma} \frac{\partial \sigma(\nu)}{\partial \nu} > 0. \]

Although the frequency of litigation conditional on the guilt level is symmetric about \( 1/2 \) and has a maximum at that point, it does not constitute a probability density over \( \nu \), given that \( \int_0^1 f(\nu) d\nu = 1 - \int_0^1 G(\nu) d\nu < 1 \). Since \( f \) is not a density, we will refer to it at times as the censoring function.\(^{17}\)

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.

\(^{17}\) If the censoring function were a density, Proposition 2 would imply stochastic dominance (first order) when the guilt level is fifty percent. While technically inappropriate, this might be a useful way of thinking about the result and its implications for trial outcome parameters.
Proposition 3: The plaintiff win rate is equal to fifty percent for any symmetric distribution of guilt.

Proof: 
\[
\int_0^1 vf(v)h(v)dv = \int_0^1 (1-v)f(1-v)h(1-v)dv \\
= \int_0^1 (1-v)f(v)h(v)dv = \int_0^1 f(v)h(v)dv - \int_0^1 vf(v)h(v)dv
\]

The first equality is due to 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
\[
E(v|lit) \equiv \mu = \frac{\int_0^1 vf(v)h(v)dv}{\int_0^1 f(v)h(v)dv} = \frac{1}{2}.
\]

Proposition 3 is a stronger result than the Priest-Klein conjecture. Priest-Klein predicts a tendency of the plaintiff win rate toward fifty percent, no matter what form the underlying distribution of case quality takes – because 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 when there is a symmetric case quality distribution. There is no need for the trial rate to be small, as commonly stated (Waldfogel, 1995), in order to get the fifty percent prediction – the result is invariant to the trial rate in the case of symmetry.

The symmetric case quality distribution includes the normal and uniform as special cases.\(^\text{18}\) The result is not surprising, given that the average guilt level in the symmetric case is fifty percent. We have only shown that when the guilt distribution is

\(^{18}\) 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 be normally distributed in his study. Given Proposition 3, the fifty percent should be observed in any model simulation that assumes a normal or symmetric distribution of case quality.
symmetric, the plaintiff win rate is not biased away from the population average guilt level as a result of the settlement process. Since the censoring (conditional probability of litigation) function does not itself form a probability density over the guilt level, it is not immediately obvious that the settlement-censoring process would preserve the population average even in the symmetric guilt distribution case.

If the 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 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 censoring function \( f \) (conditional probability of litigation function) becomes more convex, the tendency of the plaintiff win rate toward fifty percent will become more pronounced or reliable. We will take up the two implications in order. As a preliminary matter, the first implication is equivalent to the following statement.

Proposition 4: The plaintiff win rate will be closer to fifty percent than is the expected population level of guilt, that is \( |\hat{\mu} - \frac{1}{2}| < |\mu - \frac{1}{2}| \), if and only if

\[
\hat{\mu} < (>) \mu < (>) \frac{1}{2} + (\frac{1}{2})(1 - \theta)\mu.
\]

Proof: Let \( A = \int_0^1 vf(v)h(v)dv \), which is the numerator of \( \hat{\mu} \). Integrating by parts yields \( A = H(1 - G)v|_0^1 + \int_0^1 Hgv dv - \int_0^1 H(1 - G)dv \), and integrating by parts again yields \( A = 1 + \int_0^1 vGh(v)dv - \int_0^1 H(v)dv \). It follows that

\[
\hat{\mu} = \frac{1 + \int_0^1 vGh(v)dv - \int_0^1 H(v)dv}{\int_0^1 f(v)h(v)dv}.
\]
Since $\mu = 1 - \int_0^1 H(v) dv$, we have

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

From the same derivation it follows that

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

$(\hat{\mu} - \frac{1}{2})^2 - (\mu - \frac{1}{2})^2 < 0$ is equivalent to $(\hat{\mu} - \mu)(\hat{\mu} - (1 - \mu)) < 0$, and the result follows straightforwardly from this observation. 

Given the relationship among $\mu$, $\hat{\mu}$, and $\hat{\mu}$ in (7), there is an equivalent set of conditions governing the relationship between $\mu$ and $\hat{\mu}$. For example, the conditions $\mu > \hat{\mu}$ and $\mu < \frac{1}{2} + (\frac{1}{2})(1-\theta)\hat{\mu}$ imply $\mu < \hat{\mu}$ and $\mu < 1 - \theta\hat{\mu}$. The upshot is that 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 guilt distribution to satisfy the first and second implications of the Priest-Klein conjecture, given the symmetry property of the censoring function $f$.

### D. Examining Implications of Priest-Klein Conjecture

Taking advantage of the symmetry of $f$, we will examine the properties that the distribution of guilt must have in order for the first and second implications of the Priest-Klein conjecture to hold. Our first result addresses the first implication.

**Theorem 1:** Let $v$ be a random variable with the density function $h(v)$: $[0,1] \rightarrow \mathbb{R}$.
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., \( f\left(\frac{1}{2} - v\right) = f\left(\frac{1}{2} + v\right) \), for any \( v \in [0, \frac{1}{2}] \),

\[
\left| \int_0^1 vf(v)h(v) dv - \frac{1}{2} \right| \leq \left| \int_0^1 vh(v) dv - \frac{1}{2} \right| \tag{10}
\]

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

\[
\left| \int_{\frac{1}{2} - \alpha}^{\frac{1}{2} + \alpha} vh(v) dv - \frac{1}{2} \right| \leq \left| \int_{\frac{1}{2} - \alpha}^{\frac{1}{2} + \alpha} h(v) dv \right| \tag{11}
\]

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.,

\[
 f = \sum_{n=1}^{N-1} (M_n - M_{n+1}) \chi_{\alpha_n} + M_N \quad \text{or} \quad f = \sum_{n=1}^{N} (M_n - M_{n+1}) \chi_{\alpha_n} \quad \text{if we define} \quad M_{N+1} = 0.
\]

In particular, if \( N = 2 \), \( f = (M_1 - M_2) \chi_{\alpha_1} + M_2 \).
Lemma 1. For any function $f : [0, 1] \rightarrow [0, +\infty)$ 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 $f_n(v) \leq f_{n+1}(v)$ for $n = 1, 2, 3, \ldots$

We omit the proof of Lemma 1 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

$$\frac{\int_0^1 v\chi_{\alpha}(v)h(v)dv}{\int_0^1 \chi_{\alpha}(v)h(v)dv} - \frac{1}{2} \leq \frac{\int_0^1 vh(v)dv - \frac{1}{2}}{\int_0^1 \chi_{\alpha}(v)h(v)dv},$$

Taking $f = \chi_{\alpha}$ in (10) simply gives us the necessity. As to the sufficiency, (11) gives us that (10) holds for the characteristic functions $\chi_{\alpha}$ and by the linearity, (10) 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 (10) 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, $f = (M_1 - M_2)\chi_{a_1} + M_2$. From above, we already have

$$\frac{\int_0^1 v(M_1 - M_2)\chi_{a_1}(v)h(v)dv}{\int_0^1 (M_1 - M_2)\chi_{a_1}(v)h(v)dv} - \frac{1}{2} \leq \frac{\int_0^1 vh(v)dv - \frac{1}{2}}{\int_0^1 (M_1 - M_2)\chi_{a_1}(v)h(v)dv},$$

i.e.,

$$- \frac{\int_0^1 vh(v)dv - \frac{1}{2}}{\int_0^1 (M_1 - M_2)\chi_{a_1}(v)h(v)dv} \leq \frac{\int_0^1 v(M_1 - M_2)\chi_{a_1}(v)h(v)dv - \frac{1}{2} \int_0^1 (M_1 - M_2)\chi_{a_1}(v)h(v)dv}{\int_0^1 (M_1 - M_2)\chi_{a_1}(v)h(v)dv}$$

$$\leq \frac{\int_0^1 vh(v)dv - \frac{1}{2}}{\int_0^1 (M_1 - M_2)\chi_{a_1}(v)h(v)dv}$$

and also
\[
\left| \int_0^1 vM_2 h(v) dv - \frac{1}{2} \right| \leq \left| \int_0^1 vh(v) dv - \frac{1}{2} \right|
\]

i.e.,
\[
-\left| \int_0^1 vh(v) dv - \frac{1}{2} \right| \leq \frac{1}{2} \int_0^1 vM_2 h(v) dv - \frac{1}{2} \int_0^1 M_2 h(v) dv \leq \left| \int_0^1 vh(v) dv - \frac{1}{2} \right| \quad (13)
\]

Then
\[
\frac{\int_0^1 vf(v)h(v) dv - \frac{1}{2}}{\int_0^1 f(v)h(v) dv} - \frac{1}{2} = \frac{\int_0^1 v(M_1 - M_2) \chi_\alpha(v) h(v) dv + \int_0^1 vM_2 h(v) dv - \frac{1}{2}}{\int_0^1 (M_1 - M_2) \chi_\alpha(v) h(v) dv + \int_0^1 M_2 h(v) dv - \frac{1}{2}}
\]
\[
= \frac{\left\{ \int_0^1 v(M_1 - M_2) \chi_\alpha (v) h(v) dv - \frac{1}{2} \int_0^1 v(M_1 - M_2) \chi_\alpha (v) h(v) dv \right\} + \left\{ \int_0^1 vM_2 h(v) dv - \frac{1}{2} \int_0^1 M_2 h(v) dv \right\}}{\int_0^1 (M_1 - M_2) \chi_\alpha (v) h(v) dv + \int_0^1 M_2 h(v) dv}
\]

By (12), (13), it is easy to see
\[
-\left| \int_0^1 vh(v) dv - \frac{1}{2} \right| \leq \frac{\int_0^1 vf(v)h(v) dv}{\int_0^1 f(v)h(v) dv} - \frac{1}{2} \leq \left| \int_0^1 vh(v) dv - \frac{1}{2} \right|
\]

Therefore, (10) holds for any symmetric simple function.

Our final step is to use Lemma 1 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 Lemma 1, 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
\[
\int_0^1 vf(v)h(v) dv = \lim_{n \to \infty} \int_0^1 vf_n(v)h(v) dv \quad \text{and} \quad \int_0^1 f(v)h(v) dv = \lim_{n \to \infty} \int_0^1 f_n(v)h(v) dv.
\]

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 \( \left| \int_0^1 vf(v)h(v) dv - \frac{1}{2} \right| \) is well-defined and from above,
which completes the proof. ■

Thus, 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. We will refer to this as the window property.

Theorem 1 is partly intuitive in view of the effect of the settlement-censoring process under symmetric information. Settlement censoring excludes or under-weights disputes 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 \( \frac{1}{2} \) than for the whole distribution, then the censoring that takes place under symmetric information will move the win rate closer to fifty percent than is the population guilt level. That this is also a necessary condition for the Priest-Klein conjecture is less intuitive.

The window property is satisfied by a diverse set of distributions. Certainly any symmetric distribution will satisfy the property. Moreover, a large set of asymmetric distributions also satisfy it. We tried 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, because the guilt distribution 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 \( f(v) = -(v - \frac{1}{2})^2 + .25 \) approximate the censoring function. The population distribution of guilt has the following shape:

![Figure 1: A hypothetical guilt distribution](image)

In this case, \( |\mu - \frac{1}{2}| = \frac{1}{12} \), and

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

so the plaintiff win rate 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 censoring function (conditional probability of litigation function) \( f \) becomes more convex, the plaintiff win rate will tend more reliably toward fifty percent. We explore this next.\(^{19}\)

\(^{19}\) While Proposition 2 can be treated (though inappropriately and only for intuition) as a statement about the first order stochastic dominance, the second implication can be treated (in the same sense) as 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., \( f\left(\frac{1}{2} - v\right) = f\left(\frac{1}{2} + v\right) \) for any \( v \in [0, \frac{1}{2}] \).

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

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

\[
\left|\int_{0}^{1} v g_\varepsilon(v) h(v) dv - \frac{1}{2}\right| \leq \int_{0}^{1} vh(v) dv - \frac{1}{2} \quad \text{when} \quad \varepsilon \quad \text{is small enough.}
\]

Proof: First, the boundedness of \( f(v) \) guarantees the finiteness of the integrals.

Next to show \( \lim_{\varepsilon \to 0} \frac{\int_{0}^{1} v g_\varepsilon(v) h(v) dv}{\int_{0}^{1} g_\varepsilon(v) h(v) dv} = \frac{1}{2} \), it is enough to prove \( \lim_{\varepsilon \to 0} \frac{\int_{0}^{1} v f_\varepsilon(v) h(v) dv}{\int_{0}^{1} f_\varepsilon(v) h(v) dv} = \frac{1}{2} \).

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 \( G(v) = F\left(\frac{1}{2} - v\right) \) and

thus \( G(v) = 0 \) if \( v \not\in [-\frac{1}{2}, \frac{1}{2}] \) and the function \( G_\varepsilon(v) = \frac{1}{\varepsilon} G\left(\frac{v}{\varepsilon}\right) = f_\varepsilon\left(\frac{1}{2} - v\right) \) for \( v \in [-\frac{1}{2}, \frac{1}{2}] \) and equals 0 otherwise. Then,
To prove the final result, now it is enough to show \[
\lim_{\varepsilon \to 0} \frac{\int_{-\infty}^{\infty} v G_\varepsilon(v) (\frac{1}{2} - v) dv}{\int_{-\infty}^{\infty} G_\varepsilon(v) (\frac{1}{2} - v) dv} = 0.
\]

Let \( z = \frac{v}{\varepsilon} \), by changing variables,

\[
\int v G_\varepsilon(v) h(x - v) dv = \int \frac{1}{\varepsilon} G\left(\frac{v}{\varepsilon}\right) h(x - v) dv = \int \varepsilon z G(z) h(x - \varepsilon z) dz = \int z G(z) h(x - \varepsilon z) dz.
\]

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

Therefore,

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

which completes the proof. ■

As the censoring function becomes more convex, it will approach the shape of a spike at the fifty percent guilt level. The rest of the intuition for the second proposition
can be drawn from the reasoning for the familiar Chebyshev inequality, though the comparison is a loose one. 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 “distribution” that “collapses” in this case is the censoring function (i.e., conditional probability function describing the censoring process).

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. The “in the limit” result translates to a statement about the convexity of the censoring function $f$. That convexity, in turn, increases with the convexity of the prediction error variance function $\sigma(v)$. Thus, if we use $\sigma''(v)/\sigma'(v)$ as a measure of the convexity of the prediction error variance, the limiting result depends on $\sigma''(v)/\sigma'(v)$ increasing in absolute value.

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 distribution, relative to the center. Greater convexity translates into cases settling with greater frequency as the guilt status of the defendant moves toward one or zero. This is more likely to be true when the law is clear or the facts of the dispute are clear.

The trial rate could also approach zero as the cost of litigation increases relative to the judgment ($\gamma$ increases). This is equivalent to shifting the censoring function $f$ down, which will 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.\(^{20}\)

\textit{E. Generalization of Conjecture}

\(^{20}\) 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.
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 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 of the Priest-Klein conjecture. If the assumption of symmetric heteroscedasticity 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 heteroscedastic, 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 heteroscedasticity:

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

This implies \( f(v) \geq f(1-v) \) for \( 0 < v < \frac{1}{2} \), and \( f(v) \geq f(1-v) \) for \( \frac{1}{2} < v < 1 \), which means that the censoring function (graphed over \( v \)) is skewed left. Following the proof of Proposition 2

\[
\int_0^1 vf(v)h(v)dv = \int_0^1 (1-v)f(1-v)h(1-v)dv \\
= \int_0^1 (1-v)f(v)h(v)dv \leq \int_0^1 (1-v)f(v)h(v)dv
\]

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

\[
\int_0^1 vf(v)h(v)dv \leq \frac{1}{2} \int_0^1 f(v)h(v)dv
\]

so that \( \hat{\mu} \leq \frac{1}{2} \). If we reverse the inequalities in (1) and (2), the same argument leads to the conclusion that \( \hat{\mu} \geq \frac{1}{2} \). ■

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. 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.

IV. Discussion and Implications

We have identified two 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.

This implies a straightforward generalization of the Priest-Klein hypothesis: (1) under symmetric information and symmetric 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.

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.

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. 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 scenario that we have not incorporated into this model is that of informational asymmetry, a topic that the literature has explored. Medical malpractice probably 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. The informational asymmetry case has been explored in the literature, but we will not rely on the predictions from that literature in this analysis.

Another scenario that this model does not incorporate is 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. 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 would have to 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 (substantial) 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

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal injury</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>torts (non-traffic)</td>
<td>.51* (3,045)</td>
<td>.46 (3,808)</td>
<td>.27 (97)</td>
<td>.15 (639)</td>
</tr>
<tr>
<td>Personal injury</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>torts (traffic)</td>
<td>.47 (9,987)</td>
<td>.60 (3,261)</td>
<td>.17 (35)</td>
<td>.34 (337)</td>
</tr>
<tr>
<td>Product liability</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.43 (477)</td>
<td>.25 (3,255)</td>
<td>.41 (83)</td>
<td>.01 (243)</td>
</tr>
<tr>
<td>Medical malpractice</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.40 (202)</td>
<td>.38 (697)</td>
<td>.00 (6)</td>
<td>.19 (143)</td>
</tr>
<tr>
<td>Employment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>discrimination</td>
<td></td>
<td>.21 (7,165)</td>
<td>.15 (448)</td>
<td>.14 (666)</td>
</tr>
<tr>
<td>Antitrust</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.43 (586)</td>
<td>.26 (31)</td>
<td>.32 (98)</td>
<td></td>
</tr>
<tr>
<td>Years Court</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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”,\textsuperscript{21} 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 a new uncertainty based explanation for the win rates below fifty percent: 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 (i.e., low plaintiff win rates).

The low win rates observed in products liability litigation have been attributed to asymmetric information,\textsuperscript{22} 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

\textsuperscript{21} 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. Priest and Klein reported a high win rate for worker injury cases, a result they could not explain using their model. 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 private information. Weak worker lawsuits will tend to 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.

\textsuperscript{22} 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.
discoverable early in a trial.\textsuperscript{23} In addition, some courts have shifted the burden of proof on the risk-utility test where evidence is entirely in the hands of the defendant.\textsuperscript{24} 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.\textsuperscript{25}

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.

Even medical malpractice win rates may be due in part to compliance-centered uncertainty. Informational asymmetry is effectively eliminated in some malpractice cases by the presence of experts on both sides of the case.

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 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

\textsuperscript{23} 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.


\textsuperscript{25} For example, consider the cases where alternative designs generate equivalent but different types of risks. One famous class of such cases involves the choice between “X frames” and hard shells in cars, see Dawson v. Chrysler Corp., 630 F.2d 950 (3d Cir. 1980).
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.

V. Conclusion

We have presented a model that formalizes the Priest-Klein analysis and extends it as well. The model identifies conditions under which Priest-Klein’s fifty percent plaintiff win rate prediction holds, and also suggests that the conditions under which it holds are unlikely to be observed generally. In addition to generating the fifty percent prediction, the model also predicts win rates that deviate from fifty percent when litigation uncertainty is greater at one end of the guilt spectrum than at the other. The general prevalence of plaintiff win rates that are less than fifty percent in litigation data may be explained by the simple fact that trial outcome uncertainty surrounding compliance is often greater than uncertainty surrounding noncompliance.
References


