USC Center for Law, Economics & Organization
CLEO

Research Paper No. C03-8

USC Law and Economics

Working Paper No. 03-8

Private Information, Self-Serving Biases, and Optimal Settlement Mechanisms: Theory and Evidence

Seth Seabury
The Rand Corporation

Eric Talley
USC Law School

Private Information, Self-Serving Biases, and Optimal Settlement Mechanisms: Theory and Evidence

Seth Seabury* and Eric Talley†

PRELIMINARY

April 2003

Abstract

The law and economics literature on suit and settlement has tended to focus on two alternative conceptual models. On the one hand, the "optimism" model of pre-trial negotiation attempts to explain settlment failure as an artifact of unfounded optimisim by one or both parties. The idea that bargaining agents can adopt such non-rational biases receives support from experimental evidence. On the other hand, the "private information" model of pre-trial bargaining portrays settlement failures as an artifact of strategic information rent extraction. It finds support in some experimental evidence as well. This paper presents (for the first time) a mechanism-design approach for studying suit and settlement in the presence of both optimism and two-sided private information. We use a parameterization of our framework to generate testable comparative statics that distinguish between the two competing models, and then test these predictions using data from civil jury trials before and after the limitation on non-economic medical malpractice damages introduced by California legislation during the 1970s. Our (preliminary) results appear to be most consistent with the optimism model rather than the information model.

1 Introduction

Although litigation steals the lion's share of public attention within the civil justice system, settlement is (and always has been) one of the most important

*Associate Economist, RAND Corporation, Santa Monica, CA. Email: *seabury@rand.org*

†Professor of Law, USC Law School and Visiting Fellow, RAND Corporation, Santa Monica, CA. Email: *etalley@law.usc.edu*. Many thanks to Jennifer Arlen, Linda Babcock, Colin Camerer, Amy Farmer, Steve Garber, Bentley MacLeod, Ed McCaffery, Nick Pace, Jeff Rachlinsky, and seminar participants at the University of Arkansas Business School and at the 2003 USC/JEBO Conference in honor of Richard Day for helpful comments and discussions. All errors are ours.

phenomena in civil litigation. Indeed, by recent estimates, more than ninety-five percent of all civil disputes settle before seeing their full blown day in court.[1] The role, therefore, that settlement plays in the overall administration of civil liability systems is significant, growing, and critical for ongoing policy debates. Moreover, any attempt to study and interpret the law as it is reflected in litigated cases must remain mindful of the selection processes that generate litigation in the first place.

Nevertheless, the existing theoretical and empirical literature on settlement is relatively ununified, due in large part to what have emerged as two competing theories of why cases settle – or perhaps more correctly, fail to do so. One significant branch of the literature (e.g., Priest & Klein (1985)) employs what has become known as an "optimism" model of bargaining to explain litigation, in which certain litigants fail to settle disputes because they are unjustifiably sanguine about their prospects in court. If the parties' relative degrees of optimism are sufficiently severe, this theory posits, the plausible range of mutually-beneficial settlements shrinks to nothing, and the case will go to court. A litigation model built on an optimism framework is thought to predict that tried cases will result in plaintiff victories around fifty percent of the time, although though recent work (e.g., Eisenberg & Farber (1997); Klerman (2001); Siegelman & Waldfogel (1999)) demonstrates that this prediction rests on some special distributional assumptions as well as assumptions about the relative stakes involved. Nevertheless, notwithstanding the precariousness of the 50% hypothesis, the Priest/Klein model has been a favored framework for empirical testing, and in fact draws some descriptive support from experimental research demonstrating that litigants (and even their attorneys) tend to skew the facts in their favor (Babcock et al. (1995, 1997); Babcock & Loewenstein (1997)).

An alternative strand within the literature, however, portrays litigation not as an artifact of cognitive bias toward optimism, but rather as a signal of private information. Under this approach (see, e.g., Bebchuk (1984); Reinganum & Wilde (1986)), plaintiffs and defendants have private information about some important element of their case (such as liability or damages), and this information is unobservable to the other side, and only verifiable in court. Demonstrating a willingness to litigate, under this approach, is the only way to signal that one's case is relatively strong. (See, e.g., Spier 1992). Under this theory, plaintiffs may prevail with virtually any probability (Shavell 1996).

Both of these approaches to settlement have garnered significant attention among law and economics scholars. However, they have (for the most part) remained largely segregated from one another; there is currently very little work that explicitly attempts to unify these two approaches in a single, general, and empirically testable model of bargaining (with the exceptions of Farmer & Pecorino (2003) and Waldfogel (1998) discussed below). Perhaps this absence of work is due, at least in part, to the fact that — outside of certain special contexts — settlements are typically private agreements, not readily observable to outsiders (and even pro-actively protected by the settling parties). As such,

[1] See Kakkalik et al (1990).

most attempts to study suit and settlement concentrate on the special context, such as class actions, publicized settlements in sports arbitration, or historical civil procedures in which settlement and verdicts are both observable.[2] None of these settings, however, is particularly germane to the vast majority of civil litigation that occurs in today's courtrooms.

This paper attempts to fill that partial void in two ways. First, we generalize the existing theoretical framework for generating predictions about litigation selection effects that incorporates both the optimism and information model of settlement. Instead of positing a stylized, extensive form game bargaining game (as do Waldfogel (1998) and Farmer & Pecorino (2002, 2003)), we instead focus broadly on the set of bargaining outcomes that can be supported with *direct revelation mechanisms*, and more specifically on optimal bargaining mechanisms (which may mutate along with the underlying legal rule). Our approach is therefore capable of capturing a much more general set of contexts, where private information can be two sided (rather than one sided) and cognitive biases can take on many forms (such as additive, multiplicative, and so forth).

Second, we use a parameterization of this approach to generate specific comparative statics about how litigated cases will change after a shock to the underlying legal doctrine at play. We then use this predictive theory to inform hypotheses that we test using a large data set of jury trials occurring in San Francisco County, CA and Cook County, IL between 1960 and 1999. In particular, we test the optimism model against the information model by examining how noneconomic damage caps imposed on medical malpractice cases in California in 1975 affected the win rates of plaintiffs who litigated after the enactment of the cap. Using difference-in-difference estimates and utilizing three separate controls (non-medical malpractice cases in San Francisco and medical malpractice and non-medical malpractice cases from Cook County), we find (preliminary) evidence that is consistent with the optimism model. Specifically, our results fail to reject the hypothesis that the optimism model best describes pre-trial settlement behavior.

To date, there is (to our knowledge) only one previous attempt to combine both optimism and information asymmetries in a unified theoretical account pretrial bargaining. Farmer and Pecorino (2003) add optimism to a take-it-or-leave it pretrial bargaining model of Bebchuk (1984) to assess whether such biases are beneficial. Their approach, however, is limited to bargaining problems with (1) one-sided information asymmetries (2) about the probability of legal liability, within (3) a simple take-it-or-leave-it bargaining framework. This paper extends and generalizes their analysis to direct revelation mechanisms under two-sided asymmetric information. While many of their results are robust to an environment with an optimal bargaining mechanism, some are not.

From an empirical perspective, our approach is perhaps most akin to Waldfogel (1998), who attempts to distinguish between the information and optimism models using data from litigated cases in federal court, and finds evidence

[2] One exception to this is Sieg (2000) who estimates a one-sided asymmetric information game using closed claim data on medical malpractice cases in Florida.

consistent with ours. Waldfogel's approach, however, depends on predicting aggregate win rates of cases based on a relatively stylized bargaining model of one sided asymmetric information and additive biases. It is therefore not likely to be robust to variations in the theoretical framework that generates those predictions.[3] Our approach, in contrast, attempts to estimate a parameterization of a much more general model, which in turn allows us to be more confident about the our theoretical predictions. In addition, rather than relying on aggregate levels of plaintiff win rates to test the competing models, we instead focus on measuring the effects of a perturbation to the legal system that would change the characteristics of litigated cases in alternative directions. This allows us to avoid hinging our conclusions on aggregate win rates, which are themselves inherently indeterminate and unreliable in either the optimism or the information model.

Our analysis proceeds as follows. Section 2 of the paper presents a general framework for studying suit and settlement in a mechanism design framework. We characterize the qualities of an incentive compatible, individually rational bargaining mechanism when parties possess both private information and optimistic biases. We then characterize an optimal bargaining mechanism within that framework. Section 3 studies a plausible parameterization of the general model in order to generate testable comparative statics on plaintiff win rates when shocks occur to various observable parameters. Section 4 then uses these predictions to test between the information and the optimism model by focusing on the advent of the MICRA legislation in California. As noted above, we find (preliminarily) that the effects on plaintiff win rates are more consistent with the optimism model than they are with the information model. Section 5 concludes.

2 A Mechanism-Design Approach to Settlement

In this section, we develop and then study a framework for analyzing settlement through the lens of a direct revelation mechanism. Unlike previous studies, however (e.g., Stole (1992); Spier (1994); Talley (1995)), we augment the traditional mechanism design approach – which usually assumes only information asymmetry – with the possibility of cognitive bias.

2.1 Framework

Consider a plaintiff (denoted π) and a defendant (denoted Δ) who are contemplating litigation over civil law dispute, such as a medical malpractice claim.

[3] For example, Waldfogel posits a bargaining game in which defendants alone possess private information about the likelihood of liability. This leads him to predict that the information model would yield plaintiff win rates far below 50%. If, however, one flips the model to assume that it is the plaintiff who possesses private information, his prediction would reverse itself. If both possess private information, Waldfogel's approach yields no prediction whatsoever.

Both parties – or at least their lawyers who control bargaining[4] – are assumed for simplicity to be risk neutral, and the costs of litigating (rather than settling) are assumed exogenous, and given by c_π and c_Δ, respectively. Within this setting, the bargaining behavior of each side may turn on private information, cognitive bias, or possibly both, as described below.

Consider first the possibility that each party may potentially possess private information about the expected judgment should the case go to trial. We assume that the expected judgment to be given by $J(x, y, \mathbf{z})$, where x and y denote the plaintiff's and defendant's private information, respectively, about their side of the case, and $\mathbf{z}$ represents a vector of doctrinal parameters and publicly observable characteristics. We shall assume that $J_1(x, y, \mathbf{z}) > 0$, $J_{11}(x, y, \mathbf{z}) \geq 0$, $J_2(x, y, \mathbf{z}) < 0$, and $J_{22}(x, y, \mathbf{z}) \leq 0$, so that a high value of x or y is tantamount to having an objectively 'strong' case (from each side's alternative perspective).[5] We assume that x is distributed according to a commonly-known CDF $F(x)$ with strictly positive density $f(x)$ over the interval $\mathbf{X} \equiv [\underline{x}, \overline{x}]$, and similarly that y is distributed according to a commonly-known CDF $G(y)$ with strictly positive density $g(y)$ over the interval $\mathbf{Y} \equiv [\underline{y}, \overline{y}]$. The distributions of $F(x)$ and $G(y)$ are assumed independent, and $F(.)/f(.)$ and $G(.)/g(.)$ are assumed strictly increasing over $\mathbf{X}$ and $\mathbf{Y}$, respectively. (In the limiting case where private information is no longer at issue, these distributions collapse to a single mass point).

In addition to the information component outlined above, we also assume that each of the parties possibly harbors a bias about her expectations of the case. In particular, the plaintiff's subjective expectation of the final judgment for any given pair of information values (x, y) and observable environmental parameters $\mathbf{z}$ is $J(x, y, \mathbf{z}) + \varepsilon_\pi(x, y, \mathbf{z})$, where $\varepsilon_\pi(x, y, \mathbf{z})$ represents the plaintiff's bias about the case. If $\varepsilon_\pi(.) > 0$, the plaintiff harbors a self-serving bias; if $\varepsilon_\pi(.) = 0$, the plaintiff is unbiased. (We restrict attention to non-negative bias in light of the relatively strong experimental results on the issue[6]). Although the plaintiff's bias may be increasing or decreasing in her type, we assume that $\frac{\partial \varepsilon_\pi}{\partial x} > -\frac{\partial J}{\partial x}$, so as to rule out the perverse case where the plaintiff's perceived judgment is actually decreasing in the strength of her case. Similarly, the defendant's subjective expectation of the final judgment for any given (x, y) and observables $\mathbf{z}$ is $J(x, y, \mathbf{z}) - \varepsilon_\Delta(x, y, \mathbf{z})$,where $\varepsilon_\Delta(x, y, \mathbf{z})$ represents the defendant's bias about the case. If $\varepsilon_\Delta(.) > 0$, the defendant is optimistic; if $\varepsilon_\Delta(.) = 0$, the defendant is unbiased. We similarly assume here that $\frac{\partial \varepsilon_\Delta}{\partial y} < -\frac{\partial J}{\partial y}$, again to rule out the possibility that the defendant's perceived judgement is decreasing

[4] Our paper does not explore any agency cost between the attorney and the client. While that is certainly worth study, our focus is on asking whether the information or optimism outcome can better explain the characteristics of litigation. We have no clear priors on how such an agency cost (or lack thereof) between the attorney and client would affect the answer to our question.

[5] The cross-partial $J_{12}(.)$ could be either positive or negative without affecting our general results. However, in a later parameterization of the model for testing, we will restrict attention to judgment functions such that $|J_{12}|$ is sufficiently "small."

[6] See, e.g., Babcock et al. (1997); Babcock et al. (1997).

in her type.[7] (For notational ease, we suppress $\mathbf{z}$ in the theoretical discussion below, rescussitating it again in the empirical section of the paper). To distinguish the effect of bias from that of private information, we assume that the structure of the bias functions, $\varepsilon_\pi(x,y)$ and $\varepsilon_\Delta(x,y)$, is common knowledge before bargaining commences.[8]

2.2 Reservation Utilities

Under the above specifications, it is possible to describe the expected payoffs that the parties would receive in the absence of bargaining. The plaintiff's reservation payoff is given by:

$$E_y\left[J(x,y)+\varepsilon_\pi(x,y)-c_\pi\right] \tag{1}$$

The defendant's reservation payoff is given by:

$$E_x\left[-J(x,y)+\varepsilon_\Delta(x,y)-c_\Delta\right] \tag{2}$$

Thus, if pre-trial negotiation is to be individually rational for these players, it had better produce an expected payoff for each that exceeds those given in the above expressions. We shall also assume that $J(\underline{x},\overline{y})-c_\pi \geq 0$, and thus suit is always credible, even when the weakest, unbiased plaintiff comes up against the strongest defendant. (In later drafts of this paper, we will explore implications of relaxing this assumption).

2.3 Settlement Mechanism

Given the underlying structure of the litigation game, the revelation principle is clearly applicable, and thus the outcome of any extensive form bargaining game can be represented through a direct revelation mechanism. Consider, therefore, a bargaining mechanism of the form $\{p(\xi,\theta), t(\xi,\theta)\}$, where:

- ξ and θ represent the reported types of the plaintiff (whose true type is x) and the defendant (whose true type is y), respectively;
- $p(.,.)$ denotes the probability of settlement (and thus $(1-p(.,.))$ denotes the probability of trial); and
- $t(.,.)$ denotes the settlement payment made by defendant to plaintiff when a settlement occurs.

[7] We also assume that $\frac{\partial^2\varepsilon_\pi(x)}{\partial y^2} \geq 0$, $\frac{\partial^2\varepsilon_\Delta(y)}{\partial y^2} \geq 0$, and that the cross partials of both are relatively "small."

[8] Note that the functional form of the bias is deliberately kept as general as possible. This generality allows for numerous functional forms. For example, the bias function could be a simple additive or multiplicative transformation of the judgment function; or alternatively, it could be specified as a type of "coding error" that occurs when the parties attempt to perceive their private information.

Under such a mechanism, we first derive the parties' respective *net gains* from participating, over and above their reservation utilities. For the plaintiff whose type is x but who reports ξ, the expected *net gain* from participation is[9]:

$$\begin{aligned} r_{\pi}(\xi|x) &= \underbrace{E_y\left\{p(\xi,y)\,t(\xi,y)+(1-p(\xi,y))\left(J(x,y)+\varepsilon_{\pi}(x,y)-c_{\pi}\right)\right\}}_{\text{Gross Payoff}} \\ &\quad \underbrace{-E_y\left\{J(x,y)+\varepsilon_{\pi}(x,y)-c_{\pi}\right\}}_{\text{Reservation Payoff}} \qquad (3) \end{aligned}$$

Similarly, for the defendant whose type is y but who reports θ, the expected net gain from participation is:

$$\begin{aligned} r_{\Delta}(\theta|y) &= E_x\left\{-p(x,\theta)\,t(x,\theta)-(1-p(x,\theta))\left(J(x,y)-\varepsilon_{\Delta}(x,y)+c_{\Delta}\right)\right\} \\ &\quad +E_x\left\{J(x,y)-\varepsilon_{\Delta}(x,y)+c_{\Delta}\right\} \qquad (4) \end{aligned}$$

If the parties report their information truthfully (that is, $\xi = x$ and $\theta = y$), then the above expressions simplify to read:

$$\begin{aligned} r_{\pi}(x|x) &= E_y\left\{p(x,y)\,t(x,y)\right\}-E_y\left\{p(x,y)\left(J(x,y)+\varepsilon_{\Delta}(x,y)-c_{\pi}\right)\right\} \qquad (5) \\ r_{\Delta}(y|y) &= -E_x\left\{p(x,y)\,t(x,y)\right\}+E_x\left\{p(x,y)\left(J(x,y)-\varepsilon_{\Delta}(x,y)+c_{\Delta}\right)\right\} \end{aligned}$$

(In what follows we will slightly abuse notation and use $r_{\pi}(x)$ and $r_{\Delta}(y)$ to denote $r_{\pi}(x|x)$ and $r_{\Delta}(y|y)$, respectively). Consequently, under truth telling, the expected sum of the parties' *perceived* ex ante gains from settlement is:

$$S_1 = E_{xy}\left\{r_{\pi}(x)+r_{\Delta}(y)\right\} = E_{x,y}\left\{p(x,y)\left(\underbrace{(c_{\pi}+c_{\Delta})}_{\text{Costs Saved}}-\underbrace{(\varepsilon_{\pi}(x,y)+\varepsilon_{\Delta}(x,y))}_{\text{Bias Premium Lost}}\right)\right\} \qquad (6)$$

Note that the above sum incorporates within it the parties' biases about their expected judgment. If this bias is an artifact of transitory (or "hot") preferences, then it may be misleading to characterize the above expression as an appropriate measure of social surplus. In particular, a social planner might prefer to maximize a "paternalistic" social welfare function, consisting of the parties' *actual* (rather than perceived) ex ante gains from settlement, or:

$$S_2 = E_{xy}\left\{r_{\pi}(x)+r_{\Delta}(y)\,|\varepsilon^{\pi}(.)=\varepsilon^{\Delta}(.)=0\right\} = E_{x,y}\left\{p(x,y)\cdot\underbrace{(c_{\pi}+c_{\Delta})}_{\text{Costs Saved}}\right\} \qquad (7)$$

The two above expressions are similar in that they the fact that settlement saves the parties the joint litigation costs they would otherwise have to expend in litigation. Nevertheless, there is an extra term in (6) representing an expected

[9] These expressions are derived formally in the appendix.

loss that bargaining visits on parties who (irrationally) envision a more generous litigation outcome than they will actually receive.

Because our chief enterprise is to explore the characteristics of an "optimal" mechanism, a well-specified problem requires that we commit to some measure of total surplus. Rather than gravitating to one of the above two formulations, we shall simply assume that an optimal mechanism attempts to maximize some convex combination of the two above measures, or $\gamma S_1 + (1-\gamma) S_2$, where $\gamma \in [0,1]$ denotes a free parameter that measures the amount of deference that the optimal bargaining mechanism gives to the parties' cognitive biases. In practice, γ might reflect any number of considerations. For instance, it might (as alluded to above) designate the weight that a benevolent social planner would place on gratifying the parties' temporary biases. More concretely, γ might capture the extent to which the overall structure of the legal system (such as professional attorneys and procedural rules) acts to 'de-bias' the litigants, inducing them to accept a bargaining procedure that comes closer to maximizing the clients' actual (rather than perceived) joint welfare.[10]

Given this definition of optimality, we study the characteristics of the mechanism that maximizes this measure of surplus, subject to "incentive compatibility" constraints (i.e., that the parties truthfully reveal their types) and "individual rationality" constraints (i.e., that parties expect to receive at least their *perceived* reservation values under the mechanism). Consequently, we posit the following maximization problem for the mechanism designer:

$$\begin{array}{ll} Max_{\{p,t\}} & E_{x,y}\left\{p(x,y)\left[(c_\pi + c_\Delta) - \gamma\left(\varepsilon_\pi(x,y) + \varepsilon_\Delta(x,y)\right)\right]\right\} \\ & \text{subject to} \\ (IC_\pi) & x \in \arg\max_\xi \left\{r_\pi(\xi|x)\right\}; \\ (IC_\Delta) & y \in \arg\max_\theta \left\{r_\Delta(\theta|y)\right\}; \\ (IR_\pi) & r_\pi(x) \geq 0; \\ (IR_\Delta) & r_\Delta(y) \geq 0. \end{array} \qquad (*)$$

2.3.1 Incentive compatibility

As a starting point for studying this problem, consider first what restrictions incentive compatibility imposes. Using a relatively standard approach from the literature,[11] and assuming that the rent functions are almost always continuously differentiable, the following lemma follows (whose proof is omitted):

Lemma 1: A bargaining mechanism (p,t) is incentive compatible for π and Δ

[10] There is some evidence that attorneys can play this role. While still possibly somewhat subject to self-serving biases, attorneys are paid to anticipate counterarguments of the other side. In experimental settings at least, it has been shown that being forced to articulate such counterarguments has a debiasing effect. See, e.g., Babcock & Loewenstein (1997).

[11] See, e.g., Guesnerie & Laffont (1984).

if and only if the following conditions hold:

$$\begin{array}{ll} (a) & r'_{\pi}(x) = -E_y\left\{p(x,y)\cdot\left(J_1(x,y)+\varepsilon_{\pi_1}(x,y)\right)\right\} < 0 \\ (b) & r'_{\Delta}(y) = -E_x\left\{p(x,y)\cdot\left(-J_2(x,y)+\varepsilon_{\Delta_2}(x,y)\right)\right\} < 0 \\ (c) & E_y\left\{p_1(x,y)\cdot\left(J_1(x,y)+\varepsilon_{\pi_1}(x,y)\right)\right\} \leq 0 \\ (d) & E_x\left\{p_2(x,y)\cdot\left(-J_2(x,y)+\varepsilon_{\Delta_2}(x,y)\right)\right\} \leq 0 \end{array}$$

The four conditions above consist of first order "envelope" conditions ((a) and (b)) and second order "monotonicity" conditions ((c) and (d)), and are common in the literature.[12] Note from the envelope conditions that expected information rents of both parties are strictly decreasing in their types. This is consistent with one's intuition that, since it is the litigants with the weak cases who have the greatest incentive (and ability) to misrepresent their types. The monotonicity conditions ensure that

2.3.2 Individual Rationality:

Constraining ourselves to incentive compatible mechanisms (as characterized above), consider now the role played by individual rationality: that is, what incentive compatible mechanisms also induce self-interested parties to participate in the mechanism? As noted above, the parties' information rents are decreasing in their types, so that the strong plaintiff ($x = \overline{x}$) and the strong defendant ($y = \overline{y}$) are the ones who have the least to gain from settlement.

Integrating out the plaintiff's and defendant's first order conditions from these "minimum rent" types of plaintiff and defendant yields the following alternative set of expressions for $r_\pi(x)$ and $r_\Delta(y)$:

$$r_\pi(x) = r_\pi(\overline{x}) + \int_x^{\overline{x}} E_y\left\{p(\xi,y)\cdot\left(J_1(x,y)+\varepsilon_{\pi_1}(x,y)\right)\right\}\cdot d\xi \tag{8}$$

$$r_\Delta(y) = r_\Delta(\overline{y}) + \int_y^{\overline{y}} E_x\left\{p(x,\theta)\cdot\left(-J_2(x,y)+\varepsilon_{\Delta_2}(x,y)\right)\right\} d\theta \tag{9}$$

Equating (8) and (9) to the definitions of $r_\pi(x)$ and $r_\Delta(y)$ yields the following lemma (whose proof is in the appendix).

Lemma 2: Consider a trading rule $p : \mathbf{X}\times\mathbf{Y} \to [0,1]$. There exists a transfer function $t : \mathbf{X}\times\mathbf{Y} \to \mathbb{R}^+$ such that (p,t) is incentive and compatible and individually rational if and only if:

$$\begin{array}{ll} (a) & E_y\left\{p_1(x,y)\cdot\left(J_1(x,y)+\varepsilon_{\pi_1}(x,y)\right)\right\} \leq 0, \\ (b) & E_x\left\{p_2(x,y)\cdot\left(-J_2(x,y)+\varepsilon_{\Delta_2}(x,y)\right)\right\} \leq 0, \text{ and} \\ (c) & E_{x,y}\left\{p(x,y)\cdot\phi(x,y,c_\pi,c_\Delta)\right\} \geq 0, \text{ where} \end{array}$$

$$\begin{aligned} \phi(.) &= (c_\pi + c_\Delta) - \left(\varepsilon_\pi(x,y)+\varepsilon_\Delta(x,y)\right) \\ &\quad - \left[\left(J_1(x,y)+\varepsilon_{\pi_1}(x,y)\right)\left(\frac{F(x)}{f(x)}\right) + \left(-J_2(x,y)+\varepsilon_{\Delta_2}(x,y)\right)\left(\frac{G(y)}{g(y)}\right)\right]. \end{aligned}$$

[12] See, e.g., Fudenberg & Tirole (1991).

The procedure we shall follow to identify the optimal mechanism will solve a 'relaxed' problem in which constraints (a) and (b) are ignored, and the resulting solution will be checked to verify that they are satisfied.

2.4 Ex post efficiency

Before commencing with this endeavor, however, it is important to consider whether "ex post" efficiency is attainable. An important observation with this model is that, unlike in a pure asymmetric information model, with self serving biases it may be optimal not to settle. In particular, when viewed from the perspective of maximizing the parties' perceived payoffs (i.e., $\gamma = 1$), litigation turns out to be optimal for those values of x and y such that

$$c_\pi + c_\Delta < \gamma \cdot \left(\varepsilon_\pi \left(x, y\right) + \varepsilon_\Delta \left(x, y\right)\right) \tag{10}$$

in such situations, setting $p\left(x, y\right) = 0$ is first-best efficient.

However, when there are values of (x, y) such that (10) is not satisfied, or when adopting a more paternalistic setting (that is $\gamma < 1$), settlement becomes more desirable. Here, the ex post efficient mechanism would set the probability of settlement so that:

$$p\left(x, y\right) = \begin{cases} 1 & \text{if } c_\pi + c_\Delta \geq \gamma\left(\varepsilon_\pi\left(x, y\right) + \varepsilon_\Delta\left(x, y\right)\right) \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

Analysis of this rule leads to the following proposition (whose proof can be found in the appendix):

Proposition 1: *Ex post efficiency is attainable (if at all) under a direct revelation bargaining mechanism* $\{p, t\}$ *under the following conditions:*

1. *If* $\frac{c_\pi + c_\Delta}{\gamma} \leq \inf\left\{\left(\varepsilon_\pi\left(x, y\right) + \varepsilon_\Delta\left(x, y\right)\right)\right\},$ *ex post efficiency always prescribes litigation, which is always attainable under a direct revelation mechanism;*
2. *If* $\inf\left\{\varepsilon_\pi\left(x, y\right) + \varepsilon_\Delta\left(x, y\right)\right\} < \frac{c_\pi + c_\Delta}{\gamma} < \sup\left\{\varepsilon_\pi\left(x, y\right) + \varepsilon_\Delta\left(x, y\right)\right\},$ *ex post efficiency prescribes settlement iff* $\varepsilon_\pi\left(x, y\right) + \varepsilon_\Delta\left(x, y\right) < \frac{c_\pi + c_\Delta}{\gamma}$*, and is never attainable under a direct revelation mechanism;*
3. *If* $\sup\left\{\varepsilon_\pi\left(x, y\right) + \varepsilon_\Delta\left(x, y\right)\right\} < \frac{c_\pi + c_\Delta}{\gamma}$*, ex post efficiency prescribes* $p\left(x, y\right) = 1 \ \forall\left(x, y\right)$*, which is attainable under a direct revelation mechanism only if* $c_\pi + c_\Delta$ *is sufficiently high.*

The intuition that underlies Proposition 1 is relatively simple. It states that the only circumstance in which ex post efficiency is always possible occurs when the effects of optimism become of first order importance. In particular, when $\frac{c_\pi + c_\Delta}{\gamma} \leq \inf\left\{\varepsilon_\pi\left(x, y\right) + \varepsilon_\Delta\left(x, y\right)\right\},$ optimism is so overpowering relative to litigation costs that it is always socially optimal to bring suit. In this instance, the null mechanism prescribing no settlement at all is universally incentive compatible and individually rational. In all other situations, ex post efficiency is generally unattainable, except in the case where litigation costs grow so large as to swamp the effects of both optimism and information rent extraction. In this situation, all cases will settle.

Significantly, the only two situations in which ex post efficiency is attainable under a direct mechanism prescribe either universal litigation or universal settlement. Neither of these scenarios seems to be terribly plausible as an empirical proposition, given the statistical fact that some disputes settle while some do not. Consequently, we would posit that the lion's share of disputes (and in particular, the marginal dispute) likely reside within either case (2) from the proposition above, or a version of case (3) where litigation costs are insufficiently high to induce universal settlement. Either way, ex post efficiency would not be feasible, which implies that we must consider the characteristics of incentive efficient mechanisms, to which we now turn.

2.5 Incentive (Interim) Efficiency

Assuming ex post efficiency is not feasible, then, our chief question becomes that of deriving the characteristics of an optimal second best mechanism. The programing problem, then, is to maximize expected surplus $\gamma S_1 + (1-\gamma) S_2$, subject to the incentive compatibility and individual rationality constraints identified above. As noted in Lemma 2 above, the optimization problem posited above can now be simplified as the following "relaxed" problem.[13]

$$\begin{array}{l} Max_{p(x,y)} \;\; E_{x,y}\left\{p\left(x,y\right)\left(\left(c_{\pi}+c_{\Delta}\right)-\gamma\left(\varepsilon_{\pi}\left(x,y\right)+\varepsilon_{\Delta}\left(x,y\right)\right)\right)\right\} \\ \quad \text{subject to} \\ \quad E_{x,y}\left\{p\left(x,y\right)\cdot\phi\left(x,y,c_{\pi},c_{\Delta}\right)\right\}\geq 0 \end{array} \qquad (**)$$

Analysis of this program leads to the following proposition (whose proof can be found in the appendix).

Proposition 2: *There exists a unique scalar multiplier $\lambda > 0$ such that the solution to the optimal settlement mechanism program (**) consists of a settlement function $p^*(x,y)$ and transfer function $t^*(x,y)$ such that:*

$$p^*\left(x,y\right)=\begin{cases} 1 \quad \Leftrightarrow c_{\pi}+c_{\Delta}\geq & \frac{(\gamma+\lambda)}{(1+\lambda)}\left(\varepsilon_{\pi}\left(x,y\right)+\varepsilon_{\Delta}\left(x,y\right)\right) \\ & +\frac{\lambda}{(1+\lambda)}\left(J_1\left(x,y\right)+\varepsilon_{\pi_1}\left(x,y\right)\right)\left(\frac{F(x)}{f(x)}\right) \\ & +\frac{\lambda}{(1+\lambda)}\left(-J_2\left(x,y\right)+\varepsilon_{\Delta_2}\left(x,y\right)\right)\left(\frac{G(y)}{g(y)}\right) \\ 0 \quad \text{Otherwise} \end{cases}$$

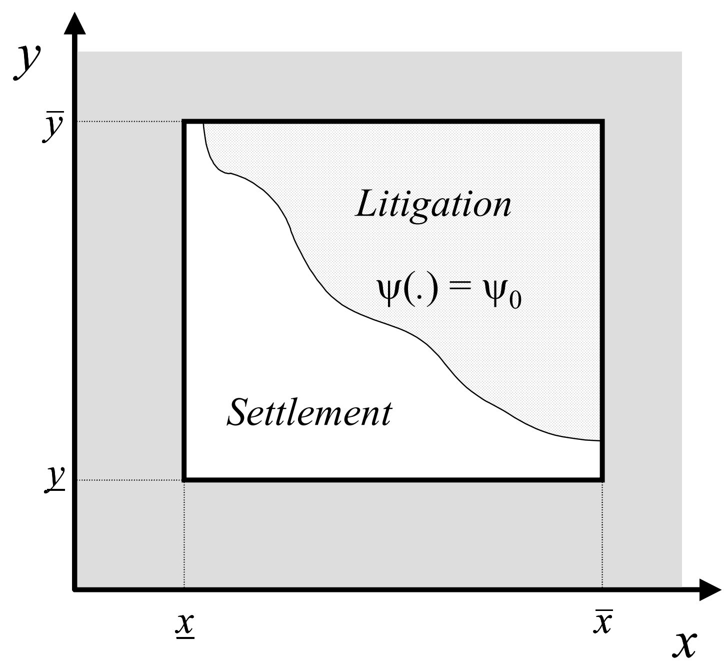
Moreover, this solution satisfies monotonicity conditions (a) and (b) from Lemma 2, and thus solves program (*) as well.

[13] Remember, the solution of this problem does not automatically satisfy the rent monotonicity conditions in Lemma 1 (conditions (*a*) and (*b*) in Lemma 2). Thus we have to re-check the relaxed problem to confirm it is the solution to the general problem.

Proposition 2 states that settlement will occur only in those situations where litigation costs are sufficiently high to offset the combined effects of strategic rent extraction by the parties and the effect of bias on the parties behavior. In particular, settlement fails (and suit occurs) in this model when and only when the following condition holds:

$$
\begin{aligned}
0 &< \underbrace{\frac{(\gamma+\lambda)}{(1+\lambda)}\left(\varepsilon_{\pi}(x,y)+\varepsilon_{\Delta}(x,y)\right)}_{\text{Bias Rents}} && (12)\\
&\quad +\underbrace{\frac{\lambda}{(1+\lambda)}\left(\left(J_1(x,y)+\varepsilon_{\pi_1}(x,y)\right)\left(\frac{F(x)}{f(x)}\right)+\left(-J_2(x,y)+\varepsilon_{\Delta_2}(x,y)\right)\left(\frac{G(y)}{g(y)}\right)\right)}_{\text{Information Rents}}\\
&\quad -\left(c_{\pi}+c_{\Delta}\right) && (13)\\
&= \psi\left(x,y;\gamma,\lambda,c_{\pi},c_{\Delta}\right)
\end{aligned}
$$

A stylized, conceptual representation of this suit/settlement frontier appears in the figure below. Notice that in the figure, litigation occurs for (x, y) pairs that are the strongest – a feature that is typical in these sorts of models,[14] since stronger parties find it optimal to signal their strength by manifesting a willingness to litigate. We will utilize this geometric relationship in testing the models against one another in the empirical section below.



[14] The downward sloping shape in frontier the figure generally obtains so long as the cross partial $J_{12}(x, y)$ does not take on values that are too extreme.

Within this general setting, it is possible to glean some insight about how shocks to the exogenous parameters $(c_\pi, c_\Delta, \gamma)$ parameters might affect the selection of cases for judgment. Inspection of expression (12) and the implicit function theorem generate the following comparative statics on litigation rates– that is $E_{x,y}\{1-p(x,y)\}$:

- As total litigation costs increase, litigation rates decrease. Equivalently, $\frac{\partial E_{x,y}\{1-p(x,y)\}}{\partial c_\pi} < 0$ and $\frac{\partial E_{x,y}\{1-p(x,y)\}}{\partial c_\Delta} < 0$. This is an artifact of the fact that $\frac{\partial \psi}{\partial c_\pi} < 0$ and $\frac{\partial \psi}{\partial c_\Delta} < 0$. This should not be surprising, since this comparative static holds for both the optimism and the information model. Consequently, in the information model it means that progressively only the stronger cases (on both sides) are selected for litigation as litigation costs increase. In the optimism model it means that only litigants possessing more extreme types of cognitive bias are selected for trial.
- As lawyers increasingly mimic their client's biases, litigation rates increase. Equivalently, $\frac{\partial E_{x,y}\{1-p(x,y)\}}{\partial \gamma} > 0$. This is an artifact of the fact that $\frac{\partial \psi}{\partial \gamma} > 0$. Consequently, under an information model, as lawyers come to dampen their clients' preferences, only the stronger cases are selected for trial. Note that under the 'pure' information based model (the $\varepsilon_i(.)$'s are both zero), however, shocks to γ should have no affect on litigation rates rates (see our analysis below).

While comparative statics such as these are good for understanding the intuitions behind the model (and assessing its plausibility), they are not particularly helpful for current purposes, as the jury verdicts data set we are interested in testing contains only those cases that are actually litigated. Given this fact, it is necessary to generate comparative statics not on settlement rates, but rather on some observable aspects of final judgments Equivalently, we are interested in generating comparative statics on:

$$E_{x,y}\{h(x,y)\,|p(x,y)=0\} = \frac{\int\int_{\psi(x,y)>0} h(x,y)\,dF(x)\,dG(y)}{E_{x,y}\{1-p^*(x,y)\}}$$

where $h(x,y)$ is some function of interest (like plaintiff win outcomes, damages amounts, etc.) which is observable in the data. To generate such comparative statics, it will be necessary to add additional structure to the model by assuming greater structure to the model. It is to this task that we turn in the next section.

3 Testable Parameterization

As noted above, in order to test the information model against the optimism model in our data set, we must add additional structure to the substantive and distributional parameters posited above. In this section, therefore, we consider a specification of the general framework above in order to generate testable comparative statics. In so doing, it is clearly important for us to exercise some caution: for any parameterization must (if it is to produce meaningful predictions) accord with one's intuitions and existing experimental evidence about how both information and self serving biases are likely to enter into settlement negotiations. The analysis that follows reflects our effort to provide such a check while simultaneously adding some structure to our framework. It can, therefore, be conceptualized at least as a starting point from which to launch into our empirical investigation.

Consider a judgment function takes the following functional form:

$$J(x, y) = \alpha \cdot D(x) \cdot \rho(y) + \eta \tag{14}$$

where α is a nonnegative parameter (discussed below), $D(x)$ represents expected damages conditional on liability, $\rho(y)$ represents the probability that the defendant will be found liable, and η represents an error term with mean zero that is distributed independent of either x or y. We assume that $D'(x) > 0$ and $\rho'(y) < 0$. This specification captures a generalization of what is perhaps the most commonly modeled litigation environment in the bargaining literature, where defendants are in a better position to know of their previously negligent acts, while plaintiffs are in a better position to know the harm they have suffered (e.g., Reinganum & Wilde (1986); Bebchuk (1984); Spier (1992); Farmer & Pecorino (2003)).[15]

In addition to this possible information asymmetry, suppose that each party is subject to additive optimism in the form $\varepsilon_\pi(x, y) = \varepsilon_\pi$ for the plaintiff and $\varepsilon_\Delta(x, y) = \varepsilon_\Delta$ for the defendant, where ε_π and ε_Δ are both nonnegative. Suppose further that ε_π and ε_Δ are distributed independently of one another according to distribution functions $H_\pi(e_\pi)$ and $H_\Delta(e_\Delta)$, with respective means of μ_π and μ_Δ.

Taking expectations of this function and applying Proposition 2, the probability of settlement under an optimal bargaining mechanism is:

$$p^*(x, y, \varepsilon_\pi, \varepsilon_\Delta) = \begin{cases} 1 & \text{if } (c_\pi + c_\Delta) \geq \frac{(\gamma+\lambda)}{(1+\lambda)}(\varepsilon_\pi + \varepsilon_\Delta) \\ & + \frac{\lambda\alpha}{(1+\lambda)}\left(D'(x)\rho(y)\left(\frac{F(x)}{f(x)}\right) + -D(x)\rho'(y)\left(\frac{G(y)}{g(y)}\right)\right) \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

Given this optimal settlement rule, we now compare the "pure information" model to the "pure optimism" model in terms of their relative comparative statics, focusing particularly on expected win rates of plaintiffs.

[15] As above, we continue to suppose that $\alpha \cdot D(\underline{x}) \cdot \rho(\overline{y}) > c_\pi$, so as to rule out the possibility of non-credible suits in expected value.

3.1 Pure information model

Consider first the pure information version of the model, where ε_π and ε_Δ are constrained to be zero by hypothesis. Here, suit occurs whenever

$$\frac{(c_\pi + c_\Delta)}{\alpha} < \frac{\lambda}{(1+\lambda)} \left(D'(x)\,\rho(y) \left(\frac{F(x)}{f(x)} \right) + -D(x)\,\rho'(y) \left(\frac{G(y)}{g(y)} \right) \right) \quad (16)$$

Referring to Figure 1 above, under a regularity condition,[16] the suit-settlement frontier will be downward sloping in $x - y$ space. Analysis of this frontier permits us to generate the following predictions:

Prediction 3a: Under the pure information model, the plaintiff's expected win rate, $E_{x,y}\{\rho(y)\,|p^* = 0\}$, is strictly decreasing in c_π and c_Δ, strictly increasing in α, and invariant in γ. The plaintiff's expected damage level conditional on liability, $E_{x,y}\{D(x)\,|p^* = 0\}$, is strictly increasing in c_π, c_Δ, strictly decreasing in α, and invariant in γ.

The intuition behind Prediction 3 is relatively straightforward. An increase in the litigation costs of either party (i.e., c_π or c_Δ) makes bargaining more attractive, and in so doing slackens the incentive compatibility constraints. This causes the settlement frontier to move up and to the right, so the plaintiff and defendant types that continue to litigate are stronger on average than those who would have been willing to litigate prior to the cost shock. Consequently, for litigated cases the expected strength of both the defendant's and the plaintiff's case increases as well. This translates into a decreased likelihood of liability and an increased level of expected damages.

3.2 Pure optimism model

Under the pure optimism model, information asymmetries are constrained to be zero, and thus the distributions of x and y collapse to mass points at their true value. Consequently, it must be that $\lambda = 0$, since the complete information case means that the incentive compatibility constraint is no longer binding. We therefore need only ensure that settlement is individually rational. From 15), a necessary condition (and, with appropriate transfers, a sufficient one) for litigation to occur in the pure optimism model for a given (x, y) is:

$$(c_\pi + c_\Delta) < \gamma\,(\varepsilon_\pi + \varepsilon_\Delta) \quad (17)$$

Note that this condition does not turn on the relative values of either x or y. Consequently, the expected win rate and damages under the pure optimism model should remain constant after a change in underlying parameters. This intuition is formalized in the following Prediction:

Prediction 3b: Under the pure optimism model, both the plaintiff's expected win rate, $E_{x,y}\{\rho(y)\,|p^* = 0\}$, and the plaintiff's expected damages, $E_{x,y}\{D(x)\,|p^* = 0\}$, are invariant to changes in c_π c_Δ, α, and γ.

[16] That is that $|D'(x)\,\rho'(y)|$ is sufficiently small.

3.3 Synthesis

A comparison of Predictions (3a) and (3b) yields the following tables of comparative statics on how shocks to various parameters are related to changes in expected plaintiff win rates:

Table 1: Liability (Comparative Statics on $E(\rho(y))$)			
	$\frac{\partial}{\partial\alpha}$	$\frac{\partial}{\partial c_\pi}$ or $\frac{\partial}{\partial c_\Delta}$	$\frac{\partial}{\partial\gamma}$
Pure Information Model	$(+)$	$(-)$	0
Pure Optimism Model	0	0	0

Table 2: Damages (Comparative Statics on $E(D(x))$)			
	$\frac{\partial}{\partial\alpha}$	$\frac{\partial}{\partial c_\pi}$ or $\frac{\partial}{\partial c_\Delta}$	$\frac{\partial}{\partial\gamma}$
Pure Information Model	$(-)$	$(+)$	0
Pure Optimism Model	0	0	0

Note that the pure information and the pure optimism models do not always predict the comparative statics. In particular, the two models make opposite predictions about how the expected values of $\rho(y)$ and $D(x)$ will react to a shock in parameters α, c_π and c_Δ. At the same time, the "pure" models both predict that win rates and damages will be invariant to changes in γ.[17] In the following section, we will exploit these comparative statics to test between the pure information and the pure optimism model as an explanatory model of settlement.

4 Testing the Theory

In this section, we turn to a strategy for testing the above parameterization of the model: using an exogenous shock to the legal environment to test the comparative statics predictions given above. In particular, we consider the introduction of the Medical Injury Compensation Reform Act (MICRA), a well-known 1975 California statute,[18] to test how plaintiff win rates responded to a shock in expected damage awards. The comparative statics presented in Table 1 above illustrate that the pure information model and the pure optimism model provide distinct predictions about how the pool of cases that go before a jury should change in response to certain parameters. Because, as we shall discuss, MICRA only affected medical malpractice cases in California, this statutory shock provides a natural experiment worth exploring, with an eye toward comparing trends in plaintiff win rates in medical malpractice cases in San Francisco versus other cases in San Francisco, and versus similar cases in Cook County.

[17] Interestingly, a combined version of this parameterization in which both information and optimism are at play generates non-zero comparative statics for shocks to γ. We do not exploit this fact in the current draft, however.

[18] Medical Injury Compensation Reform Act (1975) (codified in scattered sections of Cal. Civ. Code and Cal. Bus. & Prof. Code (2002)).

Using these comparisons, it is possible to test whether the observed changes are more consistent with the pure optimism model or the pure information model.[19]

MICRA was introduced in California in 1975. There were many components to the legislation,[20] but the main provision for our purposes — and the one that received most attention at its passage — was a ceiling on noneconomic damage awards (commonly referred to pain and suffering awards) in medical malpractice verdicts equal to $250,000. This provision was codified in section 3333.2 of the California Civil Code, which reads (in relevant part):

> **§ 3333.2. Negligence of health care provider; noneconomic losses; limitation**
>
> (a) In any action for injury against a health care provider based on professional negligence, the injured plaintiff shall be entitled to recover noneconomic losses to compensate for pain, suffering, inconvenience, physical impairment, disfigurement and other nonpecuniary damage.
>
> (b) In no action shall the amount of damages for noneconomic losses exceed two hundred fifty thousand dollars ($250,000).

According to the bill's legislative history, the cap implemented by § 3333.2 was intended to help alleviate what was perceived at the time to be a crisis in the medical industry in California due to escalating costs from medical malpractice claims (see, e.g., Zeiler (2002)). Whether such a crisis indeed existed at the time or not, the MICRA approach to capping pain and suffering awards is currently being used as a model for the Bush Administration's proposed federal ceiling on noneconomic damage awards.

Relating this statutory change to the parameterization of our model developed in Section 3, we can capture the effect of § 3333.2 by considering changes in the judgment parameter α. In particular, suppose that the probability of liability is given by $\rho(y)$, as before, and that damages are given by

$$Damages = \alpha \cdot D(x) \tag{18}$$

so that the judgment parameter α captures the extent to which the underlying legal rule dampens or multiplies the plaintiff's "true" damages of $D(x)$. Viewed in this sense, the introduction of § 3333.2 was tantamount to a reduction in the value of α.[21] While our data contains insufficient information to pinpoint which

[19] In this version of the paper, we do not ask whether a 'combined' model outperforms either of the pure models. In later drafts, however, we hope to be able to generate some comparative statics that might facilitate such a test.

[20] For example, MICRA allowed doctors and patients to contract for binding arbitration, allowed for admissibility of collateral source evidence (Cal. Civ. Code § 3333.1), limited contingency fees (in a fashion that is strikingly at odds with conventional incentive theory) (Cal. Bus. & Prof. Code § 6146(a)), and permitted the payment of malpractice damages as part of a periodic payment scheme (Cal. Code of Civ. Proc. § 667.7).

[21] To be sure, this is but one interpretation of the effect of MICRA. Another might be that § 3333.2 only had the effect of dampening large damages, leaving smaller damages unaffected. The comparative statics predictions would not be affected by this alternative interpretation.

verdicts would have been affected by the MICRA cap, recent work by Pace (2003) suggests that approximately 40% of malpractice awards in California from 1995-1999 were ultimately affected by the cap, as measured ex post.[22] To the extent, however, that any case has at least some chance of collecting extremely large non-economic damages, the effect of MICRA seems likely to be one of reducing expected damages across the board (though perhaps especially so for large damages cases). *This effect, then, would lead us to predict (as per Table 1) that under a pure information model, expected win rates should decrease (all else held constant) after the introduction of MICRA. Under the pure optimism model, however, we would predict no change in expected win rates after MICRA.*[23]

Given these predictions, the optimism model essentially serves as a null hypothesis predicting no effect on win rates against the alternative hypothesis of a negative effect under the information model. We use data on jury verdicts from San Francisco and Cook county from 1960-1999 to test these hypotheses using a difference-in-differences estimator in a natural experiment framework. Medical malpractice verdicts in San Francisco after the introduction of MICRA provide us with a treatment group, while our control groups consist of (1) other tort cases in San Francisco and (2) medical malpractice cases in Cook County, IL; and (3) other tort cases in Cook County. In what follows, we describe the data in more detail, outline our empirical strategy, present our results and briefly discuss the power of our estimates given our sample size.

4.1 Data

Our data come from the RAND Jury Verdicts Database, which has collected information on nearly all civil jury verdicts in Cook County, IL and San Francisco County, CA reported in the *Cook County Jury Verdict Reporter* of Chicago, IL and *Jury Verdicts Weekly* of Santa Rosa, CA, respectively from 1960-1999. The jury verdict reporters are regular publications that provide jurisdiction level information on jury verdicts as reported either by in-court witnesses or summaries sent in by litigants, or both. The summaries in these reporters, intended primarily to be a source of information on the value of cases to potential litigants, vary somewhat in the information provided but almost universally provide data on key variables of interest. For our purposes the information the information of primary importance are the case type (i.e. the issues litigated over in court) and whether the jury ruled in favor of the plaintiff or the defendant.

The data include information on 20,326 verdicts in San Francisco County and Cook County from 1960-1999. However, we kept only verdicts involving

[22] The cap has not been adjusted for inflation since its introduction, so we expect that it affects a larger fraction of cases today than it would have upon its induction.

[23] Note that given this specification, the predicted effect on observed damages would be indeterminate. Explcitly, the predicted effect would be:

$$\frac{\partial E(Damages)}{\partial \alpha} = \alpha \frac{\partial E_x \{D(x)\}}{\partial \alpha} + E_x \{D(x)\}$$

whose sign is clearly ambiguous. We therefore concentrate our efforts on plaintiff win rates.

tort cases and dropped 1,653 verdicts that involved issues of contract disputes and financial injury, because the reporting of these cases appear to have changed over time, particularly in Cook County. The types of tort cases included as "other cases" in our data include automobile, common carrier, products liability, premises liability, intentional tort, other professional malpractice and other tort liability cases. We also restrict our sample to only those verdicts in which a single case type is reported, eliminating an additional 2,240 observations. Some verdicts might report more than one case type, and it is possible that in some of these cases medical malpractice might be an issue, but is such a minor issue that MICRA would not be applicable (at least in a practical sense). The final data set we use, therefore, contains information on 16,433 verdicts.

Table 2 reports summary statistics for the data we use in our analysis. We provide mean win rates for medical malpractice and other cases in San Francisco and before and after the introduction of MICRA. Note that the MICRA damage cap took effect for injuries that occurred on or after December 12, 1975, so a case is coded as being affected by MICRA only if the incident being litigated over is listed as being on or after that date.[24]

Table 3
Summary Statistics

	San Francisco		Cook County	
	Mean	St. Dev.	Mean	St. Dev.
Pre-MICRA				
Plaintiff win rate: Malpractice	31.48%	46.53	33.96%	47.43
Plaintiff win rate: Other	54.20%	49.83	46.62%	49.89
Malpractice Cases (%)	8.10%	27.29	4.46%	20.64
Number of Verdicts	3,333		7,131	
Post-MICRA				
Plaintiff win rate: Malpractice	32.74%	47.07	35.81%	47.96
Plaintiff win rate: Other	60.18%	48.98	56.84%	49.54
Malpractice Cases (%)	12.95%	33.58	29.12%	45.43
Number of Verdicts	1,298		4,671	

Note from the table that plaintiff win rates in medical malpractice are much lower than that for other cases on average, as has been noted previously in Sieg (2000). If one were attempting to test between the information and optimism models by concentrating on win rates alone (e.g. Waldfogel (1998)), the table might suggest that the relatively unbalanced win rate for malpractice cases suggests that the information model was at play in those circumstances. As noted above, however, analysis of aggregate win levels alone is not particularly

[24] There were provisions of MICRA other than the damage cap, most notably the introduction of a sliding scale on attorney fees. The sliding scale affected attorney-plaintiff contracts that were agreed upon before or after December 12, 1975. We have no information on the attorney contracts, so it is possible some cases that involve injuries that occured before this date but where a lawyer was not hired until after might be affected by MICRA, at least to the extent that attorneys influence the settlement decision.

useful for testing the two models of settlement against one another, since under either model, virtually any probability of liability is possible with appropriate stakes and distributional assumptions.

Also worth noting is that the total number of verdicts in SF fell off sharply after the passage of MICRA, although the number of medical malpractice verdicts fell by far less than the number of other tort verdicts. On the other hand, the number of medical malpractice verdicts in CC grew substantially, while the number of other verdicts fell somewhat. The possibility that other, non-observable factors might have been affecting the SF cases (or the CC cases, for that matter) over this time is a factor that deserves attention in formulating one's empirical strategy. It is this question to which we now turn.

4.2 Empirical Strategy

As mentioned above, in our data we have one treatment group, medical malpractice verdicts in San Francisco after the passage of MICRA, and three potential control groups, other verdicts in San Francisco (which were not affected by the passage of MICRA), medical malpractice verdicts in Cook County and other verdicts in cook County. Using different combinations of these controls we estimate three separate parameters that test for the presence of the information model. Using different control groups allows us to control for the possibility of confounding trends that might have affected other cases in San Francisco or cases in Cook County.

Reviving the parametrization from Section 3, let ρ represent the probability of a plaintiff victory. In individual cases, we observe observe a realization of 1 if a jury finds for the plaintiff and 0 otherwise, and we test for differences in the predicted probability of a plaintiff win using a linear probability model.[25] The first parameter is the average difference in plaintiff win rates between medical malpractice and other verdicts in San Francisco before and after the implementation of MICRA. Formally we can define this parameter as

$$\partial_1 = \left(\rho_{m,sf,1} - \rho_{o,sf,1}\right) - \left(\rho_{m,sf,0} - \rho_{o,sf,0}\right) \tag{19}$$

where $\rho_{i,j,t}$ represents the average plaintiff win rate for verdict type i in jurisdiction j in period t. Note that hereafter i equals m if a verdict is a medical malpractice verdict and o if it is another tort verdict, j equals sf if the verdict occurred in San Francisco and cc if it occurred in Cook County and t equals 1 if a verdict was subsequent to (and subject to) MICRA and 0 otherwise.

The second parameter we estimate is the difference in average plaintiff win rates for medical malpractice cases in San Francisco and Cook County before and after the passage of MICRA, which we define as

$$\partial_2 = \left(\rho_{m,sf,1} - \rho_{m,cc,1}\right) - \left(\rho_{m,sf,0} - \rho_{o,cc,0}\right). \tag{20}$$

[25] Our qualitative results do not change if we use a probit or a logit model. All standard errors reported in the following section for the estimation results are corrected for heteroskedasticity.

The third parameter that we estimate is a "difference-in-difference-in-differences" where we take the difference between medical malpractice and other verdicts in San Francisco before and after MICRA and subtract the difference between medical malpractice and other verdicts in Cook County before and after MICRA. This parameter is defined as

$$\begin{aligned}\partial_3 \quad = \quad & \left[\left(\rho_{m,sf,1} - \rho_{o,sf,1}\right) - \left(\rho_{m,sf,0} - \rho_{o,sf,0}\right)\right] \\ & - \left[\left(\rho_{m,cc,1} - \rho_{o,cc,1}\right) - \left(\rho_{m,cc,0} - \rho_{o,cc,0}\right)\right]. \qquad (21)\end{aligned}$$

Under the parameterization of the pure information model studied above, all three parameters ∂_1, ∂_2 and ∂_3 should all be negative. MICRA only capped awards for medical malpractice cases in San Francisco, so we would expect that plaintiff win rates for medical malpractice verdicts should fall relative to win rates in other types of verdicts. Likewise, Cook County verdicts were unaffected, so win rates in San Francisco should fall relative to both types of verdicts there relative to Cook County. If the pure optimism model holds, then the estimated parameters should be zero, and MICRA should have no effect on win rates.[26]

4.3 Results

Table 4 presents our estimates of the effect of MICRA on plaintiff win rates. The top section of the table presents the estimated coefficients of ∂_1, ∂_2 and ∂_3 for the full sample of verdicts. The column labeled Control Group 1 presents the estimated difference in difference between win rates in medical malpractice and other verdicts in San Francisco before and after MICRA. The estimated value of this parameter is -0.046, suggesting that — consistent with the information model — plaintiff win rates in medical malpractice verdicts fell relative to other tort cases after the passage of MICRA. However, the standard error associated with this parameter is sufficiently large (with a value of 0.051) to render the estimate statistically insignificant.

The parameter for the second control group, comparing medical malpractice verdicts in San Francisco to those in Cook County, is also negative. This suggests that plaintiff win rates in medical malpractice cases in San Francisco fell relative to those in Cook County after the passage of MICRA; but once again, the difference is not statistically significant.

The estimate using the third control group, the difference-in-difference-in-differences estimate using differences in win rates between medical malpractice and other verdicts in San Francisco and Cook County, is actually positive with a value of 0.052, but as with the other estimates it is not statistically significant.

[26] Again, note that in the very special case where $\kappa \in (0, 1)$ MICRA will have a negative effect on win rates, and we would expect that ∂_1, ∂_2 and ∂_3 would be negative. However, we only explicitly test the one-tailed hypothesis that the coefficients are positive.

Table 4 **The Estimated Effect of MICRA on Plaintiff Win Rates** [27]			
	Control Group 1: ∂_1	Control Group 2: ∂_2	Control Group 3: ∂_3
1960-1999			
	-0.046 (0.051)	-0.008 (0.055)	0.052 (0.060)
Number of Obs	4,631	2,116	16,433
1965-1985			
	-0.024 (0.076)	0.015 (0.096)	0.036 (0.101)
Number of Obs	2,991	615	9,888
1970-1980			
	-0.035 (0.103)	0.103 (0.157)	0.051 (0.165)
Number of Obs	1,675	371	5,243

We now turn our attention to the bottom two parts of Table 4, which provide estimates on restricted subsamples of verdicts. The availability of almost twenty years of data before and after the passage of MICRA is convenient because it provides relatively large sample sizes. However, the time frame is long enough that it is worthwhile to estimate the parameters using data on verdicts closer to the actual passage of the act. This allows us to check whether or not MICRA might have had a stronger effect closer to its original enactment, and helps to avoid the possibility of contamination in the treatment or control groups by other, unobserved factors that might have changed over time. For example, in 1987 the sliding scale on attorney fees was loosened somewhat.[28] If attorneys have the ability to influence the settlement process then this would mitigate the effects of MICRA and bias us towards finding no effect in verdicts after the reform.

The center portion of Table 4 presents the estimated coefficients when the sample is limited to those verdicts that occurred from 1965-1985, ten years before and ten years after the passage of MICRA. The third and final portion of the table limits the sample to five years before and five years after the passage of MICRA. Inspection of the estimated coefficient values on both of these populations show that limiting the data to these subsamples has very little effect on the results. In no cases are the estimated parameters statistically significant. In the estimates using the first and third control groups, corresponding to parameters ∂_1 and ∂_3, the parameters have the same sign as and magnitudes similar to their estimated values using the entire sample of verdicts. Only the value

[27] Heteroskedsticity-consistent standard errors in parentheses. A regressions are weighted using to reflect that some cases have a different probability of being selected. This sampling reflects the fact that for automobile cases, which are a large fraction of tort cases, only one out of every four (or one out of every five in some years) of these verdicts was included in the data.

[28] Cal. Bus. & Prof. Code § 6146(a).

of ∂_2 changes noticeably, becoming positive in both of the restricted samples and with a (relatively large) point estimate of 0.103 in the 1970-1980 sample. However, the sample sizes using the second control group are quite small, so the resulting standard errors on these estimates are high.

The results presented in Table 4 demonstrate that the passage of MICRA had no perceptible effect on plaintiff win rates in medical malpractice jury verdicts in San Francisco. This result is robust to several different control groups and is consistent as we look at different spans before and after the act's passage.[29] In general, then, our results fail to provide any support for the information model of pre-trial settlement behavior over the optimism model. Explicitly, we fail to find any significant effect of MICRA on win rates, and thus we fail to reject the null hypothesis that the optimism model best explains pre-trial settlement behavior.

4.4 Robustness

At this juncture, it is worthwhile considering a few possible ways in which either our approach may be subject to criticism questioning whether our approach genuinely provides support for the optimism model over the information model. We consider three such criticisms, each in turn: (1) Insufficient power in empirical tests; (2) Misspecification of the parameterized theoretical model; and (3) Failure to account for ex ante behavioral adjustments in comparative statics.

4.4.1 Empirical Power

Whenever interpreting a failure to reject the null hypothesis it is important to consider sample size and the power of one's estimates. In order to address these concerns, we briefly explore the extent to which our failure to detect any statistically significant effect of MICRA in the previous section can be explained by insufficient sample sizes. To see the importance of sample size, consider the case of ∂_1, where we estimate the difference-in-differences using the regression:

$$p_i = a_0 + a_1 Medmal_i + a_2 MICRA_i + \partial_1 Medmal\_MICRA_i + e_i \qquad (22)$$

where p_i is an indicator variable equalling 1 if the jury found for the plaintiff in verdict i, $Medmal$ is an indicator variable for whether or not a verdict involved medical malpractice, $MICRA$ is an indicator variable for whether or not incident leading to the litigation occurred after the effective date of MICRA, and $Medmal\_MICRA$ is the "treatment" variable that indicates whether or

[29] An obvious strategy that is not discussed here is the inclusion of additional covariates in the estimating equations reported in Table 3. In regressions not reported here, we included various additional explanatory variables including indicators for case type, measures of reported economic losses, and indicators for litigant types (i.e. whether the plaintiff or defendant is an individual or a business or government agency), but none of these have any affect on the qualitative results of this section.

not a verdict involved medical malpractice and an injury that occurred after MICRA's effective date.[30]

In this example the t-statistic for the OLS estimate of ∂_1 is given by

$$t = \frac{\hat{\partial}_1}{\frac{\hat{\sigma}_e}{\sqrt{N\hat{\sigma}^2_\partial(1-R^2_\partial)}}} \tag{23}$$

where $\hat{\sigma}_e$ is the root mean square error of the difference-in-differences regression, N is the number of observations, $\hat{\sigma}^2_\partial$ is the sample variance of the variable $Medmal\_MICRA$ and R^2_∂ is the r-square of the regression of $Medmal\_MICRA$ on $Medmal$ and $MICRA$. Recognizing the fact that our estimate $\hat{\partial}_1$ will only be statistically significant at the 5% level if t is greater than 1.96, we can rearrange Equation 23 into the following expression for the minimum sample size needed to obtain statistical significance

$$N > \frac{(1.96)^2\hat{\sigma}^2_e}{\hat{\partial}^2_1\hat{\sigma}^2_\partial(1-R^2_\partial)}. \tag{24}$$

Equation 24 illustrates that the sample size we need is increasing in the mean square error of the regression, but decreasing in the absolute value of the parameter estimate, the sample variance of the treatment, and the portion of variation in the treatment that is uncorrelated with the other explanatory variables. Note that this equation can be easily adapted to apply to any of the regressions used to produce the estimates in Table 4 above, so by filling in values for the parameters we can estimate the sample size needed to find support for the information model.

Table 5 provides estimates of the sample sizes necessary to find a significant difference coefficient of at least 0.05 using each of the three control groups estimated on the whole sample. Estimates of the parameters $\hat{\sigma}^2_\partial$ and R^2_∂ are easily derived from the sample. The mean square error in the regressions used to create Table 4 range from 0.227 0.246 so to be conservative we use 0.246 as our estimate of $\hat{\sigma}^2_e$. The table provides the sample sizes necessary to obtain significance at either the 5% or 10% level.[31]

Table 5
Minimum Sample Sizes Necessary for Statistical Significance

Significance Level	Control Group 1:	Control Group 2:	Control Group 3:
	∂_1	∂_2	∂_3
5%	7,731	10,877	46,598
10%	5,445	7,662	32,824
Actual #	4,631	2,116	16,433

[30] This is the exact form of the regression that was used to provide the estimate of ∂_1 presented in Table 3.

[31] The 10% critical value for the t-statistic used here is 1.645.

The results of Table 5 suggest that a lack of power may explain at least part of our failure to reject the optimism model. For none of the parameters do we have enough observations to observe a statistically significant coefficient value of 0.05 at even the 10% level. Our estimates of ∂_1 appear to have the most power, we could observe a statistically significant coefficient at the 10% level if the sample size included an additional 814 observations, an increase of about 17%. To observe statistical significance at the 5% level, however, we would require a much larger increase of 3,100, or about 67%. The power of our estimates for ∂_2 and ∂_3 is worse, requiring at least a doubling of the number observations simply to observe significance at the 10% level. Because the restricted samples in Table 4 have even fewer observations, those estimates have less power and would require much greater increases in the number of observations to obtain significance.

These calculations suggest that our failure to reject the optimism model may, in part, be due to the relatively small size of our data set rather than a failure of the information model. It is worth noting, however, that the only parameter estimated in the full sample that has a numerically significant sign consistent with the information model, parameter ∂_1, is also the one in which the actual sample size is relatively close to the required sample size in the power simulations. It nevertheless seems possible that an increase in our sample size might provide more support for the information model than we are currently finding.

4.4.2 Theoretical Misspecification

One possible concern with our findings is that we were estimating a special case of our general model, and that other parameterizations would have given rise to different results. In particular, our parameterization presupposed that the underlying information asymmetry was such that the defendant had private information about liability, and the plaintiff had private information about damages. Under different parameterization, one might argue, possibly different (and even contradictory) comparative statics might obtain, thereby leading one to question what our empirical tests were attempting to falsify.

While the misspecification concern is a legitimate one, we believe that there are a number of responses to it that are relatively convincing. First of all, even if the specification we posited did not capture all possible litigated cases, it is at the very least a plausible one that probably captured at least some sorts of cases. Given that the generality of the model requires some sort of parameterization to conduct a meaningful econometric test, perhaps plausibility is as good a criterion as any.

Second, the parameterization we considered could be changed without affecting the qualitative predictions. Indeed, a number of variations on our parameterization (including multiplicative biases and one-sided information on liability favoring the defendant) all lead to precisely the same comparative statics predictions as those we tested.

Finally, even under an alternative parameterization where our comparative statics would change, the empirical result may persist. In particular, virtually every parameterization of the judgment function would result in some type of selection effect in litigated cases. Only in the case of the private information model would this effect be equal to zero. Thus, even if an alternative parameterization of the information model were to produce the opposite comparative static prediction, the results reported in Table 4 are not significantly different from zero. As such, only in the special case where the information model predicts no effect whatsoever would our empirical inquiry lose its ability to discern between the two models.

4.4.3 Ex Ante Behavior

Finally, our parameterization has not considered the effect that MICRA may have had on ex ante behavior by defendants. If responded to the dampening in expected exposure introduced by the damages cap, then the distribution of defendant 'types' would likely not be held constant in the pre- and post-MICRA worlds. Failure to account for this endogeneity, one might argue, might skew our predictions. (See, e.g., Bernardo, Talley & Welch (2000); Zeiler (2002)).

If the endogeneity of ex ante behavior were sufficiently strong, it might cause us to reject erroneously the information model. Indeed, it is conceivable that potential medical malpractice defendants began to act more carelessly after the passage of MICRA, anticipating a smaller liability exposure. Consequently, of the universe of cases filed after MICRA, there are more defendants with objectively 'weak' cases than before passage of the Act, pushing expected win rates up. This effect works against the direct effect of reducing win rates that our pure information model generated. Consequently, we cannot unambiguously reject the possibility that ex ante behavioral adjustments might be working in conjunction with an information model. We conjecture, however, that it would be rather surprising if these effects cancelled one another out exactly.

5 Conclusion

This paper has presented a model of pre-trial settlement behavior that integrates two competing strands of the literature, private information and self-serving biases (or "optimism"). Focusing on the bargaining outcomes that are supported by direct revelation mechanisms characterized the selection of jury trials allowing for a two-sided asymmetric information and general forms of cognitive biases. Our model not only provides a more general framework for studying the bargaining process in the civil justice system, it can also be used to generate testable hypotheses about how characteristics of the cases that go to trial respond to exogenous shocks in the underlying parameters of the model.

With a relatively straightforward parameterization, we are able to construct a direct test of the two competing models of settlement behavior. Under this parameterization, the information model predicts that plaintiff win rates should rise in response to a decrease in expected damage awards. The optimism model, however, predicts that win rates should be unchanged in response to a shock to expected damages. We used data on civil jury verdicts in California and Illinois before and after the passage of a noneconomic damage cap in California to test whether or not win rates rose, as would be predicted by the information model. We found no significant effect of the damage cap on win rates, providing implicit support for the optimism model of settlement behavior. While the sample sizes we use may be too small to estimate the effect of MICRA precisely enough to rule out the information model completely, our results ultimately offer it little support.

In many respects, this paper is but the tip of the iceberg in suggesting possible avenues for future research. A great deal of empirical work needs to be done before we can definitely answer which model of settlement behavior best fits existing data. For example, in addition to malpractice reforms, many jurisdictions within our data set implemented mandatory, non-binding arbitration for various lower stakes cases. Such reforms may be interpreted as an attempt to de-bias the potential litigants and encourage settlement. Our parameterized model would be capable of predicting differential effects of such reforms in comparing the pure version of either model against a hybrid model of behavior. We expect to incorporate this analysis into the existing framework in a future draft. Nonetheless, at the very least the approach presented here provides both a fruitful framework for formulating predictions, and a helpful empirical approach for uncovering the dynamics of settlement — by focusing not at cases that actually settle, but rather at cases that fail to do so.

6 References

1. Babcock, Linda & Loewenstein, George, Explaining Bargaining Impasse: The Role of Self-Serving Biases, Journal of Econ. Perspectives 1: 109-126 (1997).

2. Babcock L. and G. Loewenstein (1997), "Explaining Bargaining Impasse: The Role of Self-Serving dispositions," *Journal of Economic Perspectives,* 11: 109-126.

3. Babcock, Linda & Loewenstein, George, Issacharoff Samueal & Camerer, Colin (1995), "Biased Judgmenets of Fairness in Bargaining," *American Economic Review* 85: 1337-43.

4. Bebchuk, L., (1984), "Litigation and Settlement under Imperfect Information," Rand J. Econ. 15: 404.

5. Bernardo, A., E. Talley & I. Welch, "A Theory of Legal Presumptions," J. Law Econ. & Org. 16: 1-50.

6. Bester H. and W. Güth (1998), "Is Altruism Evolutionary Stable?", *Journal of Economic Behavior and Organization* **34(2),** pp. 211-221.

7. Bolle F. (2000), "Is Altruism Evolutionarily Stable? And Envy and Malevolence? - Remarks on Bester and Güth," *Journal of Economic Behavior and Organization* **42(1):**. 131-133.

8. Chatterjee K. and W. Samuelson (1983), "Bargaining under Incomplete Information", *Operations Research* **31**: 835-851.

9. Cramton P. (1992), "Strategic Delay in Bargaining with Two Sided Uncertainty", *Review of Economic Studies* **59**: 205-225.

10. Eisenberg, T & Farber, H., (1997), "The Litigious Plaintiff Hypothesis," Rand J. Econ. 28: S92.

11. Farmer, A. & P. Peccorino (2003) "Pretrial Bargaining with Self-Serving Bias and Asymmetric Information," J. Econ. Behavior & Org. (forthcoming).

12. Farmer, A. & P. Peccorino (2002) "The Causes of Bargaining Failure: Evidence from Major League Baseball"

13. Guesnerie, R. & J.J. Laffont (1984), "A Complete Solution to a Class of Principal-Agent Problems with Application to the Control of a Self-Managed Firm," *Journal of Public Economics,* **25:** 329-69.

14. Kakalik, J., M. Selvin, & N. Pace (1990), "Averting Gridlock: Strategies for Reducing Civil Delay in the Los Angeles Superior Court," RAND Institute for Civil Justice.

15. Kessler, D., Meites, T., & G. Miller, (1996), "Explaining Deviations from the Fifty Percent Rule: A mUlitmodal Approach to the Selection of Caswes for Litigation," J. Legal Studies 25: 233.

16. Klerman, D. (2001) "Settlement and the Decline of Private Prosecution in Thirteenth Century England." Law and History Review 19: 1.

17. Pace, N. (2003) "Capping Pain and Suffering: The California MICRA Experience." RAND Institute for Civil Justice, *forthcoming.*

18. Shavell, S. (1996) "Any Frequency of Plaintiff Victory at Trial is Possible," Journal of Legal Studies 25: 493-502.

19. Sieg, H. (2000) "Estimating a Bargaining Model with Asymmetric Information: Evidence from Meidcal Malpractice Disputes." Journal of Political Economy 108: 1006-1021.

20. Siegelman, P. & Waldfogel, J., (1999), "Toward a Taxonomy of Disputed: Nw Evidence Through the Prism of the Priest/Klein Model" J. Legal Studies 27: 101-30.

21. Spier, K. (1994), "Pretrial Baragining and the Design of Fee-Shifting Rules," Rand J. Econ 25: 197-214.

22. Stole, L. "The Economics of Liquidated Damages Clauses in Contractual Environments with Private Information," J. Law Econ. & Org. 8: 582-606 (1992).

23. Talley, E. (1995), "Liability-Based Fee-Shifting Rules and Settlement Mechanisms Under Incomplete Information," Chi-Kent L. Rev. 71: 461-503.

24. Waldfogel, J. (1998), "Reconciling Asymmetric Information and Divergent Expectations Theories of Litigation," J. Law & Economics ___.

25. Zeiler, K. (2002), Medical Malpractice and Contract Disclosure: An Equilibrium Model of the Effects of Legal Rules on Behavior in Health Care Markets (unpublished draft).

7 Appendix

This Appendix presents a number of the derivations and proofs omitted in Sections 2 and 3 of the paper.

7.1 Derivation of the Parties' Net Payoff Functions:

Consider first the derivation of the parties' net payoff functions given in expressions (3) and (4) in the text. If one thinks of settlement as a type of dummy indicator variable, k, which equals one when and only when settlement succeeds, then k is distributed Bernoulli with success/fail parameters $(p, 1-p)$ .When settlement fails (i.e., when $k = 0$), the parties will end up in court. In such a circumstance, the parties will Bayesian update their expectations of the expected judgment. In particular, the plaintiff whose type is x and who has reported ξ updates her assessment of the expected litigation payoff to be

$$
\begin{aligned}
& E_y \left\{ J(x,y) + \varepsilon^{\pi}(x,y) - c_{\pi} | x, \xi, k=0 \right\} \qquad (25) \\
= & \int_{\underline{y}}^{\overline{y}} \left(J(x,y) + \varepsilon^{\pi}(x,y) - c_{\pi} \right) \left(\frac{\Pr\{k=0|x,\xi,y\} \cdot g(y|x,\xi)}{\Pr(k=0|x,\xi)} \right) dy \\
= & \frac{\int_{\underline{y}}^{\overline{y}} \left(J(x,y) + \varepsilon^{\pi}(x,y) - c_{\pi} \right) (1 - p(\xi,y)) \cdot g(y)\, dy}{E_y\{1 - p(\xi,y)\}} \\
= & \frac{E_y \left\{ J(x,y) + \varepsilon^{\pi}(x,y) - c_{\pi} (1 - p(\xi,y)) \right\}}{E_y\{1 - p(\xi,y)\}}
\end{aligned}
$$

For the plaintiff,then, the expected *net* gain from participation is therefore:

$$
\begin{aligned}
r_{\pi}(\xi|x) = & E_y \left\{ p(\xi,y) \cdot t(\xi,y) \right\} \qquad (26) \\
& + E_y \left\{ (1 - p(\xi,y)) | x, \xi \right\} \cdot E_y \left\{ J(x,y) + \varepsilon^{\pi}(x,y) - c_{\pi} | x, \xi, k=0 \right\} \\
& - E_y \left\{ J(x,y) + \varepsilon^{\pi}(x,y) - c_{\pi} \right\} \\
= & E_y \left\{ p(\xi,y)\, t(\xi,y) + (1 - p(\xi,y)) \left(J(x,y) + \varepsilon^{\pi}(x,y) - c_{\pi} \right) - \left(J(x,y) + \varepsilon^{\pi}(x,y) \right) - c_{\pi} \right\}
\end{aligned}
$$

This is the expression given in the text. An identical approach applies to the defendant.■

7.2 Proof of Lemma 1

From Guesnerie and Laffont (1984), for players π and Δ who report (respectively) ξ and θ, and whose expected payoffs are (respectively) $r_{\pi}(\xi|x)$ and $r_{\Delta}(\theta|y)$,incentive compatability implies that.

$$
\begin{aligned}
r'_{\pi}(x|x) &= \left. \frac{\partial r_{\pi}(\xi|x)}{\partial x} \right|_{\xi=x} & r'_{\Delta}(y|y) &= \left. \frac{\partial r_{\Delta}(\theta|y)}{\partial y} \right|_{\theta=y} \\
r''_{\pi}(x|x) &= \left. \frac{\partial^2 r_{\pi}(\xi|x)}{\partial \xi \partial x} \right|_{\xi=x} \geq 0 & r''_{\Delta}(y|y) &= \left. \frac{\partial^2 r_{\Delta}(\theta|y)}{\partial \theta \partial y} \right|_{\theta=y} \geq 0
\end{aligned}
$$

Imposing the definitions of $r_{\pi}(\xi|x)$ and $r_{\Delta}(\theta|y)$ from the text produces the conditions that appear in the Lemma.■

7.3 Proof of Lemma 2

Taking expectations of the expression in (9) over x yields the plaintiff's ex ante expected gain:

$$\begin{aligned}
E_x\{r_\pi(x)\} &= r_\pi(\overline{x}) + \int_{\underline{x}}^{\overline{x}}\int_{x}^{\overline{x}} E_y\{p(\xi,y)\cdot(J_1(\xi,y)+\varepsilon_{\pi_1}(x,y))\}\,d\xi dF(x) \qquad (27)\\
&= r_\pi(\overline{x}) + \int_{\underline{x}}^{\overline{x}}\int_{\underline{x}}^{\xi} E_y\{p(\xi,y)\cdot(J_1(\xi,y)+\varepsilon_{\pi_1}(x,y))\}\,dF(x)\,d\xi\\
&= r_\pi(\overline{x}) + \int_{\underline{x}}^{\overline{x}} E_y\{p(\xi,y)\cdot(J_1(\xi,y)+\varepsilon_{\pi_1}(x,y))\}\cdot F(\xi)\cdot d\xi\\
&= r_\pi(\overline{x}) + E_{x,y}\left\{p(x,y)\cdot(J_1(\xi,y)+\varepsilon_{\pi_1}(x,y))\cdot\left(\frac{F(x)}{f(x)}\right)\right\}
\end{aligned}$$

Similarly, Taking expectations of (9) over y yields the plaintiff's ex ante expected payoff:

$$\begin{aligned}
E_y\{r_\Delta(y)\} &= r_\Delta(\overline{y}) + \int_{\underline{y}}^{\overline{y}}\int_{y}^{\overline{y}} E_x\{p(x,\theta)\cdot(-J_2(x,\theta)+\varepsilon_{\Delta_2}(x,y))\}\,d\theta dG(y) \qquad (28)\\
&= r_\Delta(\overline{y}) + E_{xy}\left\{p(x,\theta)\cdot(-J_2(x,\theta)+\varepsilon_{\Delta_2}(x,y))\cdot\frac{G(\theta)}{g(\theta)}\right\}
\end{aligned}$$

Summing these expected payoffs yeilds:

$$\begin{aligned}
E_{xy}\{r_\pi+r_\Delta\} &= r_\pi(\overline{x})+r_\Delta(\overline{y}) \qquad (29)\\
&\quad +E_{x,y}\left\{p(x,y)\left[(J_1(\xi,y)+\varepsilon_{\pi_1}(x,y))\left(\frac{F(x)}{f(x)}\right)+(-J_2(x,y)+\varepsilon_{\Delta_2}(x,y))\left(\frac{G(y)}{g(y)}\right)\right]\right\}
\end{aligned}$$

But from (6), we know also that

$$E_x\{r_\pi(x)\}+E_y\{r_\Delta(y)\} = E_{x,y}\{p(x,y)((c_\pi+c_\Delta)-(\varepsilon_\pi(x,y)+\varepsilon_\Delta(x,y)))\}$$

Combining the two above expressions therefore yields:

$$\begin{aligned}
& r_\pi(\overline{x})+r_\Delta(\overline{y}) \qquad (30)\\
=\;& E_{x,y}\left\{p(x,y)\left[\begin{array}{c}(c_\pi+c_\Delta)-(\varepsilon_\pi(x,y)+\varepsilon_\Delta(x,y))\\ -(J_1(\xi,y)+\varepsilon_{\pi_1}(x,y))\cdot\left(\frac{F(x)}{f(x)}\right)+(-J_2(x,y)+\varepsilon_{\Delta_2}(x,y))\cdot\left(\frac{G(y)}{g(y)}\right)\end{array}\right]\right\}\\
\equiv\;& E_{x,y}\{p(x,y)\cdot\phi(x,y,c_\pi,c_\Delta,\varepsilon_\pi,\varepsilon_\Delta)\}
\end{aligned}$$

Noting that a necessary (and with appropriate transfers, sufficient) condition for an IC/IR mechanism is that $r_\pi(\overline{x})+r_\Delta(\overline{y})\geq 0$, we arrive at the expression given in the lemma.■

[To Be Completed]