ISSN 1045-6333

STRICT LIABILITY VERSUS NEGLIGENCE UNDER INCOMPLETE INFORMATION

Omri Ben-Shahar*

and

Philipp von Randow**

Discussion Paper No. 127

4/93

Harvard Law School
Cambridge, MA 02138

The Program in Law and Economics is supported by a grant from the John M. Olin Foundation.

*S.J.D. Candidate, Harvard Law School, and Ph.D. Candidate, Harvard Department of Economics.

**Assistant Professor of Law, University of Osnabrűck, Germany.

STRICT LIABILITY VERSUS NEGLIGENCE UNDER INCOMPLETE INFORMATION*

Omri Ben-Shahar**

and

Philipp von Randow***

ABSTRACT

Accident victims vary according to the level of harm suffered. If courts cannot observe victims' harm perfectly, low-harm victims may pretend to be high-harm victims, to induce courts to award them high compensatory damages. Can the level of care that a victim takes reveal his true harm? This question is addressed in model in which courts observe the victims' level of care and adjust the damages award accordingly. The magnitude of liability can depend only on the agents' levels of care, not on the victim's harm. The main task is to check whether the ordinary liability rules, strict liability and negligence, can implement the optimal levels of care. It is demonstrated that a negligence rule can induce victims to reveal their true harm through their choice of care, while the strict liability regime fails to do so. However, in contrast to the complete information case, first-best levels of care cannot be implemented.

* Research support from the Olin Foundation at Harvard Law School is gratefully acknowledged. We thank Lucian Bebchuk, Louis Kaplow, Kathy Spier and participants at the Harvard Research Seminar in Law and Economics for helpful comments.

** S.J.D. Candidate, Harvard Law School, and Ph.D. Candidate, Harvard Department of Economics.

*** Assistant Professor of Law, University of Osnabrück, Germany.

INTRODUCTION

One of the basic principles of Tort and Accident Law is the idea that the liability threatened to be imposed on an injurer should be related to the actual harm suffered by the victim (the "compensation principle"). Economic analysis has emphasized the value of accurate compensation and illuminated the conditions under which the magnitude of injurer's liability ought to be precisely the victim's harm.

In order to carry out this policy, courts must observe the victim's actual harm accurately. Often, the precise magnitude of loss is not readily observable. In many situations, a victim may claim to have preferences other than his real ones and compensation for losses which exceed his true harm, exploiting the courts' difficulty in observing and determining the actual harm that occurred.

One way of dealing with this phenomenon, an "accounting approach", is to establish lists of standardized compensation awards, based on actuarial data (Fleming, 1987). Here, the compensation scheme redresses what regulators determine as reasonable or average levels of loss. While this is a simple and convenient mechanism, it usually fails to satisfy the compensation principle and generates sub-optimal precautionary behavior. If an injurer expects a liability payment different from the actual loss, his incentives to reduce risks are either too high or too low.

This paper offers a different solution to the problem of unobservable harm, an "economic approach". The approach builds on the idea that the victim's behavior **prior** to the injury, in particular his precautionary activity, may signal

information about the magnitude of harm he expects to suffer, if injury occurs. Thus, if the court can observe the victim's behavior that preceded the accident, it may be able to learn more about the actual harm that the accident caused, and set liability of magnitude that varies according to the victim's level of care.[1]

Previous literature has analyzed extensively how the choice of liability rules determines the strategies of the victim and injurer in choosing their levels of care. Most of this literature assumes explicitly that harm is observable by courts.[2] In a recent paper, Kaplow and Shavell (1992a) argued that when harm is difficult to assess, it may be socially desirable to resort to the accounting approach and award average sums, rather than invest resources in assessing the actual harm. In particular, if injurers cannot anticipate the actual harm, ex-post accuracy will have no ex-ante incentive effects and will not induce better precautions. Our analysis makes several points that relate to Kaplow and Shavell's observation. First, better accuracy may be achieved simply by observing the victim's care. Second, accuracy always has a desirable effect on the victim's incentives to take care. Third, if an injurer does not know the victim's actual harm but observes the victim's level of care, accuracy can still improve the injurer's incentives to take care. Thus, by relating the victim's harm to the level of care that he has taken,

---

[1]To our knowledge, the idea of relating harm to the observation of the agents' care, has appeared in only one article, Emons (1990). His focus is on deriving the optimal mechanism, while our focus is comparing the performance of the ordinary mechanisms, the strict liability and negligence rules.

[2]There is a growing literature that addresses problems arising from incomplete information about other issues. It includes Craswell and Calfee (1986), on errors in assessing negligence; Png (1986), Polinsky and Shavell (1989) and Kaplow and Shavell (1992b), on errors in determining the identity of an offender.

some of the distortions arising from unobservable harm can be rectified.

The main results of the paper are as follows. A liability rule that makes the victim's reward dependent on his choice of care may be susceptible to manipulative care behavior by victims. We show that only a negligence regime can induce victims with different potential harm level to signal their types by engaging in different levels of care. A strict liability regime is always susceptible to pooling equilibria, i.e., different types of victims choosing the same level of care, hence cannot be screened. We show that first-best levels of care (that can be implemented if there is complete information about all parameters, including harm), cannot be implemented when harm is unobservable. If the injurer observes only the victim's level of care before choosing his own level of care, the victim cannot be induced to reveal his type through his choice of care. There may always exist victim types that pool with other types. However, the second-best, defined as the optimal levels of care given that the injurer does not vary his care with victim's types, can be implemented through a negligence rule.

This last result lends a new efficiency rationale to the comparative negligence doctrine. Under this doctrine, a negligent victim may receive partial compensation, which declines with the victim's degree of carelessness. We argue that the comparative negligence rule does not deny careless a victim his full compensation, but rather fully compensates him for the lower harm which he signalled through his degree of carelessness.

The paper is organized as follows. Section I presents the basic model, in which both parties choose care simultaneously, and demonstrates that the

negligence regime implements the second-best. Section II extends the model to allow for sequential choices of care. It demonstrates that although the injurer can receive a signal about the victim's type that will induce him to take optimal care, in general the victim cannot be driven to reveal his type truthfully. Section III discussed the new interpretation that this model offers to the comparative negligence doctrine. Section IV concludes by offering some extensions and remarks on the analysis.

I. A Model with Care as a Signal for Harm

A. The Framework of the Analysis

Agents are either injurers or victims. Injurers are all identical. victims may be either one of two types, ***L*** or ***H***, according to the level of harm they suffer at the event of an accident. Denote by h the actual level of victim's harm, where $h \in \{h_L, h_H\}$, $h_L < h_H$. The frequency of type L in the population of victims is π, $\pi \in (0,1)$. Only the victim knows his type. The injurer knows π.

The victim and injurer may engage in precautionary activity. Let x,y denote the injurer's and the victim's costs of care, respectively ($x,y \in \mathbb{R}_+$). It is assumed that both parties choose their precautions simultaneously.[3] The parties' precautions affect the probability of harm, but not its magnitude.[4] Specifically, write the probability of harm as $p(x,y)$, where $p(.)$ is continuous, twice

---

[3]Section II of the paper extends the analysis to the case in which parties choose precautions sequentially.

[4]This type of loss reduction is referred to in the literature as "self protection" (see Becker and Ehrlich (1972)). The other type of loss reduction, in which the victim's care can affect the magnitude of harm but not its probability is known as "self insurance". Later, we will extend the analysis to capture self-insurance as well.

differentiable in both arguments, with $p_x<0$, $p_y<0$ and $p_{xx}>0$, $p_{yy}>0$, $p_{xy}>0$, and these are assumed to be common knowledge. All parties are assumed to be risk-neutral.

If harm occurs, a liability rule comes into effect. The court, while it cannot observe the actual harm, does observe the parties' choices of care, (x,y), with accuracy. Based on this observation it determines a compensation award according to the function $f(x,y)$, which is the liability rule.

Before turning to analyze behavior under specific liability rules, it is useful to make precise the notion what the optimal levels of care are. We distinguish between two efficiency criteria that apply to the situation at hand. The "ex-post efficient", or the "first-best", versus the "interim efficient" or the "second-best" levels of care.[5] The first-best levels of care are the solution to the following:

$$Min_{x_i,y_i}\{x_i + y_i + p(x_i,y_i)h_i\}, \quad i = L,H. \tag{1}$$

and are denoted by $\{(x_L^1, y_L^1),(x_H^1, y_H^1)\}$. Notice that the first-best involves both injurer and victim adjusting their levels of care according to the victim's type.

When the injurer cannot observe the victim's type, nor the victim's care level, he is constrained to choose one level of care, and cannot vary it according to the victim's type. In this case, the "second-best" levels of care are the solution to:

$$Min_{x,y_L,y_H}[x + \pi(y_L + p(x,y_L)h_L) + (1-\pi)(y_H + p(x,y_H)h_H)], \tag{2}$$

[5]These two concepts of efficiency differ with respect to the information allocation at the time of welfare evaluation. Ex-post efficiency is Pareto-efficiency when all relevant information is public. Interim efficiency is Pareto efficiency when some individuals have private information. For details, see Holmström and Myerson (1985).

and are denoted by (x^0, y_L^0, y_H^0). Given our assumptions about $p(.)$, the solution satisfies $y_L^0 < y_H^0$.

For every given liability rule, the injurer's and victim's choices of care involve a strategic interaction. The victim, through his choice of care, has to decide whether to reveal his type to the court, or to pose as the other type. He chooses the strategy that minimizes his total costs:

$$U_i = y_i + p(x,y_i)[h_i - f(x,y_i)], \quad i = L,H. \tag{3}$$

The injurer's optimal strategy, when he cannot observe the victim's level of care, is the one that minimizes his cost function:

$$V = x + \pi p(x,y_L)f(x,y_L) + (1 - \pi)p(x,y_H)f(x,y_H). \tag{4}$$

We wish to investigate whether there exists a liability regime that (Nash) implements the second-best levels of care. I.e., can the strict liability or negligence liability rules be constructed in a way that induces the parties to take second-best levels of care. Once we establish the result that only a negligence rule can achieve the second-best, we will turn to consider the implementability of the first-best.

B. Negligence Rule with Defense of Contributory Negligence

Under this rule, the injurer will not be liable if he takes due care or if the victim fails to take due care. It is well established that in the full information case, if the due levels of care are set at the socially optimal levels, this rule will induce actual levels of care that are equal to the optimal ones.[6]

The problem we encounter when adding unobservability of harm is that the

---

[6]Shavell (1987), pp. 14-15.

court cannot set the due levels of care optimally, because it does not observe the particular types. Suppose, therefore, that the court views the victim's choice of care as a signal for his type, i.e., for his harm level. Specifically, suppose that the court sets two "threshold" levels of due care for the victim, a low one, y_L^0, and a high one, y_H^0. If the victim takes at least y_L^0 but less than y_H^0 he signals that he is a low harm type and is entitled to be compensated only h_L. If the victim has engaged in y_H^0 level of care or more, he signals himself as a high harm type and, at the event of the injurer being negligent, is awarded h_H. If he takes anything less than y_L^0 he is contributorily negligent and gets 0. Formally, the rule can be written in the following form:

$$f(x,y) = \begin{cases} h_H & \text{if } y \geq y_H^0 \text{ and } x < x^0 \\ h_L & \text{if } y_L^0 \leq y < y_H^0 \text{ and } x < x^0 \\ 0 & \text{if } y < y_L^0 \end{cases}$$

Consider the parties' optimization problems. The injurer will never take more than the due care level x^0, because x^0 suffices to escape liability. Further, if there exists a separating equilibrium in which victim types self select, the injurer necessarily chooses x^0 in this equilibrium. Given this choice by the injurer, for a separating equilibrium to arise, it must be that the following incentive constraints are satisfied:

$$y_i^0 = argmin_{y_i} [y_i + p(x^0,y_i)h_i], \quad \text{for } i = L,H \tag{5}$$

where the expression in brackets is the cost borne by type i, given that the injurer takes due care. Yet from the definition of y_i^0, we know that it minimizes the

expression in brackets, which guarantees that expression (5) is satisfied.[7] This establishes the following proposition:

Proposition 1: Under the negligence rule with defense of contributory negligence, the second-best separating equilibrium always arises.

Remarks: (i) *Why negligence works?* The reason that the negligence rule implements the second-best levels of care is that under the negligence regime, the victim has no one to cheat but himself. Given a level of care by the injurer which satisfies the due care standard, the victim bears the entire cost of harm. He can reduce it by incurring his own cost of care, till he balances the incremental cost of care with expected reduction in harm. Since each victim type faces a different expected harm, each strikes a different balance. We will see that this logic fails under the strict liability regime.

(ii) *Errors by the court.* This result does not in any way depend on the court knowing π. Even if the court is mistaken in its estimate of π and thus sets sub-optimal standards of due care, as long as the injurer knows these standards and is induced to exercise his due level of care, the victims will self-select. Thus, a court's failure to estimate π correctly can have adverse effects on the injurer's behavior (by setting sub-optimal "x^0"), but not on the victim's behavior.

(iii) *Uniqueness.* The separating equilibrium identified is the unique equilibrium that can arise under the negligence rule. To verify this, note that for any level of care greater or equal to x^0 that the injurer chooses, the victims self-select. Thus, no level of care greater than x^0 can be optimal for the injurer and be part of an

---

[7] If $(x^0,y_i^0) = argmin_{x, y_i}[x + y_i + p(x,y_i)h_i]$, then necessarily $y_i^0 = argmin_{y_i}[y_i + p(x^0,y_i)h_i]$.

equilibrium. For any level of care less than x^0 that the injurer chooses, at least type H takes $y_H{}^0$ (and perhaps type L as well). Thus, being negligent exposes the injurer to high liability, which is not optimal for him. Hence, the injurer takes x^0, and the victims self-select.

(iv) *Negligence rule without a defense.* While the defense of contributory negligence as defined above incorporated into the analysis the idea that victim's care signal his harm, this defense is not necessary for there to be adequate incentives to take second-best care. Incentive-wise, a negligence rule without any defense can work just as well. As long as the injurer is induced to take the due level of care x^0, the victim has to bear his losses and takes the optimal level of care, given his type. Thus, a negligence regime needs only to set the injurer's due level of care correctly, and need not set a standard for the victim's behavior. Put differently, the negligence regime displays the property that in equilibrium, no compensation is actually awarded, thus, in particular, there is no opportunity to effectuate the idea of higher compensation to more careful victims.

If, for a reason outside the scope of the model, injurers are occasionally found negligent and liability is actually imposed, then, to overcome the problem of unobservable harm, it is useful to have the defense of contributory negligence as suggested above. If the magnitude of compensation rises with the victim's level of care, a victim will perfectly reveal his harm type through his choice of care. This a situation in which the defense of contributory negligence guarantees both accurate compensation and optimal care by victims.

C. Strict Liability with Defense of Contributory Negligence

Under this regime, the injurer has to compensate the victim for his full harm, unless the victim's level of care has fallen short of his due level of care. It is well established that if type i victim's due level of care is set at the socially optimal level y_i^0, then this rule generates optimal care by both parties.[8] In the incomplete information case, it is impossible to set the due levels of victim's care optimally, if h_i is unobservable. Suppose, again, that the court sets two threshold levels of due care for the victim. If the victim has taken at least y_L^0 but less then y_H^0, he signals himself as low-harm type and is entitled to be compensated only h_L. If, however, he has engaged in y_H^0 level of care or more, he signals himself as high-harm type and is awarded h_H. This adjustment incorporates the idea that higher level of care is associated with a higher magnitude of loss.

Formally, the strict liability with defense of contributory negligence rule can be written as:

$$f(x,y) = \begin{cases} h_H & \text{if } y \geq y_H^0 \\ h_L & \text{if } y_L^0 \leq y < y_H^0 \\ 0 & \text{if } y < y_L^0 \end{cases}$$

Consider the victim's optimization problem. Regardless of his type, it is clear that given this liability rule the victim will choose either y_L^0 or y_H^0. Anything less than y_L^0 is sub-optimal, since he loses the right for compensation, which, by the efficiency of y_L^0, is more valuable. Anything strictly between y_L^0 and y_H^0 yields

[8]See Shavell (1987), pp. 12-13.

the same coverage as y_L^0, but costs more. Any care level higher than y_H^0 does not add to the damage award but is costlier than y_H^0. However, it is not clear that each type would be driven to choose his correct level of care. For a separating equilibrium to arise, it must be the case that the incentive compatibility constraints of both parties are satisfied:

$$(IC_L) \qquad y_L^0 + p(x^0,y_L^0)(h_L - h_L) \leq y_H^0 + p(x^0,y_H^0)(h_L - h_H) \tag{6}$$

$$(IC_H) \qquad y_H^0 + p(x^0,y_H^0)(h_H - h_H) \leq y_L^0 + p(x^0,y_L^0)(h_H - h_L) \tag{7}$$

The first condition, type L's incentive constraint, implies that if the injurer chooses x^0, type L finds it beneficial to choose y_L^0 rather than pretend to be type H (by taking y_H^0). He does so when the incremental increase in expected compensation associated with pretending to be type H is less than the added cost of care it inflicts on him. Type L may, in a sense, purchase a "lottery": pay a certain amount $(y_H^0 - y_L^0)$ for an uncertain prize $(h_H - h_L)$. As the probability $p(x^0,y_H^0)$ of this prize decreases, so does the willingness to participate in the gamble. Analogously, the second condition is type H's incentive constraint.

Conditions (6) and (7) have to hold only at $x=x^0$. The reason is that if a separating equilibrium indeed arises, then the injurer finds it optimal to choose x^0.[9]

Conditions (6) and (7) can be simplified and combined to be written as:

$$p(x^0,y_H^0) \leq p^0 \leq p(x^0,y_L^0) \tag{8}$$

where the p^0 denotes a ratio which is defined by:

---

[9]This can be seen simply by comparing the F.O.C of the injurer's problem, given choice of $(y_L^0 y_H^0)$ by the victim and the resulting compensation awards, to the F.O.C of the second-best welfare problem. They are the same.

$$p^0 \equiv \frac{y_H^0 - y_L^0}{h_H - h_L}.$$

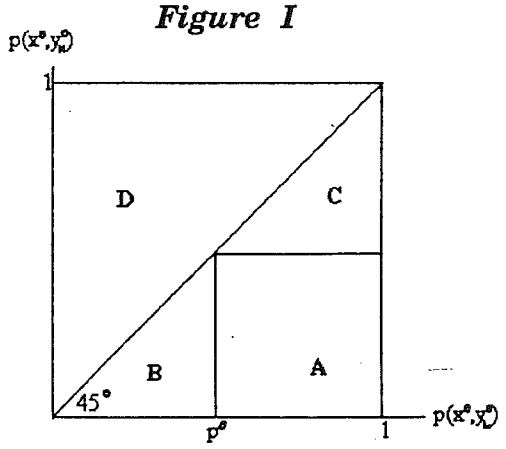
We can state this result in the following proposition:

Proposition 2: In the strict liability with defense of contributory negligence regime the second-best separating equilibrium arises if and only if condition (8) holds. Otherwise there is a pooling equilibrium that is worse than the second-best.

Remarks: (i) *Separating v. pooling equilibria.* This separating equilibrium is the unique separating equilibrium that could possibly arise. In particular, there cannot be a "bad" separating equilibrium in which each type pretends to be the other. (To verify, note that the inequality signs in expression (8) need to be reversed for the other separating equilibrium to arise, but this cannot occur.)

However, we may have two pooling equilibria: one in which both types take y_L^0 (if condition (6) but not (7) holds): and another in which both take y_H^0 (if (7) but not (6) holds). Thus the parameters that characterize each situation may lead to either a separating or a pooling equilibrium, and an equilibrium always exists. This can be depicted graphically:

Figure I



The exogenous parameters h_L, h_H and the specific function $p(.)$ determine the values of $p(x^0,y_L^0)$, $p(x^0,y_H^0)$. The different possible parametric values are measured on the axes. The parameter space may be partitioned into four areas, as in the figure. Area "D" is not feasible, since we can never have $p(x^0,y_L^0) < p(x^0,y_H^0)$. Area "A" includes all cases in which a separating equilibrium arises. In area "B" are the cases in which a pooling equilibrium of (x^0,y_L^0) arises. Area "C" includes the case in which a pooling equilibrium of (x^0,y_H^0) arises.

The emergence of pooling equilibria is unfortunate, from a policy perspective. When courts observe a particular care level by the victim, they generally cannot infer whether it is part of a pooling or a separating equilibrium, being unable to observe the particular parameters that distinguish the two. Therefore, they do not know if the care level is indeed a useful signal for the level of harm. Thus, in a strict liability regime, the problem of the victim's unobservable harm cannot be solved solely by observing his level of care.

(ii) *Decoupling liability and compensation.* Can the victim's incentive to "cheat" be overcome by decoupling liability and compensation? Suppose the victim receives a fixed compensation sum, k, regardless of his type, while the injurer pays to the court an additional fine, c (may be negative), which reflects the information the court has inferred about the victim's type. In particular,the court can set a higher fine if it observes a higher level of care by the victim. This regime will always implement a separating equilibrium: each type of victim will take a different level of care and self-select. The reason is that being compensated a fixed sum, the victim bears his residual loss, $h_i - k$. Because each type faces a

different residual loss, each takes a different level of care. Therefore, the court can infer the victim's type from his level of care and adjust the injurer's fine, such that the injurer's total payment, $k + c$, sums up to the actual harm the victim revealed.

Notice, however, that the separating equilibrium that is achieved is not second-best. While the injurer has optimal incentives to take care, the victim's incentives are distorted. Because k does not vary with victim types, each victim type is either overcompensated or undercompensated. If the victim is overcompensated, he takes too little care, in order to increase the likelihood of profitable accident. If the victim is undercompensated, he takes excessive care, to reduce the likelihood of an uncompensated accident. While the distortion of too little care can be solved by setting a standard of due care for the victim, the problem of excessive care cannot be addressed through contributory negligence standards. In sum, a decoupled strict liability regime does guarantee a separating equilibrium, but not the second-best one.

(iii) The analysis in this section has demonstrated that only a negligence regime can implement the second-best levels of care unconditionally. On a positive note, it points out that even when harm is unobservable, the victim can be driven to take optimal care. However, the result is unsatisfactory in the sense that both regimes fail to induce the injurer to take the first-best level of care. If the focus of the tort system is incentives for care, then the negligence regime is better due to its stronger incentive effects on the victim. If, in contrast, the objective of the tort system is compensating victims for their losses, then this goal is not achieved

by a negligence regime, and is better served by strict liability. In the next section of the paper, we turn to examine a family of situations in which relaing the victim's harm to his care may impove even the injurer's incentives.

II. Extension: A Model with Sequential Actions

The analysis thus far has examined a model in which both parties choose their care simultaneously. Being the ordinary model analyzed in the literature, it captures a variety of situations in which each party's choice of care is not observable to the other party when she has to make her own care decision, but is verifiable by the court (perhaps at some verification cost). We demonstrated that if the court does not observe the victim's harm, observing care as a signal for harm can at most lead to second-best levels of care. Next, the analysis is extended to include the possibility of sequential behavior, i.e., that one of the parties may observe the other's care before choosing his own. This extension enriches the information structure of the model, as it allows the injurer to observe a signal about the victim's private information by observing his choice of care.[10,11] Can the court implement the first-best levels of care by inferring the victim's harm from his level of care and compensating him accordingly?

---

[10]The first analysis of a sequential choice accidents model has appeared in Shavell (1983). His was a complete information model, which did not address the information dissemination effects of the sequential structure.

[11]We consider situations in which the sequence of actions is given exogenously and is outside the parties strategic interaction. In a more general framework, the order of actions may also result from strategic choices: each party may try to claim a first or last mover position. An analysis of the parties' incentives to choose a place in the order of actions under each liability rule is omitted, mainly because it yields ambiguous results. The analysis is pursued assuming that the order in any particular case is an exogenous characteristic of it.

A. Injurer Moves First

This is the simpler of the two extensions — the one in which the injurer chooses his care level before the victim. Consider, for example, a recreation facility (swimming pool, amusement park) in which the owner employs various fixed safeguards, all of which are observable by the users (e.g., warning signs, information brochures, guards and inspectors). Each participant, becoming fully aware of the hazards involved, can engage on her part in additional care. This would be a particular case where the injurer moves first.

The interaction that takes place in this setting is strategically equivalent to the one in the simultaneous moves model. By moving first, the injurer, who possesses no private information, is unable to observe a signal from the victim, and cannot condition his level of care on the victim's true or revealed type. Thus, in equilibrium, his move is determined by the same program as in the simultaneous moves model. The victim, knowing the injurer has no private information, can fully predict the injurer's move. Thus, the victim's action is dependent solely on his private information, and there is no additional information embodied in his observation of the injurer's move. In sum, this model generates the same reduced strategic-form game and predicts the same patterns of equilibrium behavior as the simultaneous moves model. Hence, all the previous results carry over.

B. Victim Moves First

Consider a situation in which the injurer can observe the choice of care by each potential victim, and has the opportunity to adjust his own care according to

this observation. An example is the famous flax and locomotive sparks story. A locomotive that passes through farmers' fields may cause different magnitudes of harm, depending on specific features of the crop, the season, the irrigation system, the market price, etc. These attributes cannot be observed by the railroad company or its driver, before or as the locomotive passes through the farms. However, the locomotive can observe the care actions taken by the farmers, such as the height of fences, warning signs and distance from the track. He can then adjust his level of care accordingly, by reducing speed or applying spark arresters. Here, the locomotive (injurer) receives a signal from the victims about their types, a verifiable signal, and can react accordingly.

In this setting, can the court implement the first-best levels of care, by imposing on the injurer a duty to observe the victim's care and to adjust his care level to fit the harm that the victim signals? In other words, the court can require the injurer to be more careful when he interacts with highly careful victims than otherwise, reasoning that such victims signal themselves to the injurer as more vulnerable, and their higher expected harm justifies higher precautions. When this liability approach is followed, the injurer receives valuable information by actually observing the victims's care. He in fact can learn the magnitude of liability that he will be exposed to, if the (unobservable) harm realizes. Therefore, the injurer can adjust his level of care to correspond to the observed care by the victim.

The victim, on his part, may exploit this feature to manipulate the injurer's care decision. He may find it optimal to engage in a care level that signals a harm

different than his true one. In particular, his signal generates two effects: first, the court determines the perceived harm according to it; and second, the injurer chooses his level of care according to it, thus it affects the likelihood of harm. In a sense, the victim is a "Stakelberg" leader, and chooses a point on the injurer's reaction curve, that is best for him. What follows is a formal analysis of this interaction.

To analyze this situation, let the chronological sequence of events in the model be: the victim's type is determined, only the victim knows her type. The victim chooses y, her choice is observable. The injurer observes y, and his choice of x can be dependent on this observation, thus write $x=x(y)$. Harm realizes with probability $p(x,y)$. The magnitude of harm is $h \in \{h_L,h_H\}$, known only to victim; court or injurer cannot observe it. The court selects a damages award according to the liability rule $f(x,y)$.

After the victim selects y, this action generates two resulting choices: $x(y)$ by the injurer, and $f(x(y),y)$ by the court. Thus, the victim's expected cost can be written as:

$$U_i = y + p(x(y),y)[h_i - f(x(y),y)] \tag{9}$$

The injurer, who can observe the victim's signal and anticipate the precise magnitude of liability he may face, minimizes his cost:

$$V = x(y) + p(x(y),y)f(x(y),y) \tag{10}$$

The next two sub-sections examine whether a strict liability or negligence regimes can implement the first-best levels of care, $\{(x_L^1,y_L^1),(x_H^1,y_H^1)\}$.

1. Negligence rule with defense of contributory negligence.

Unlike the situation in the simultaneous moves model, here the negligence regime does not guarantee a separating equilibrium with optimal levels of care. The feature that arises in this regime is the potential incentive for a victim of low-harm type to engage in an excessive level of care. By investing more in care and posing as a high-harm type, the victim induces the injurer to do the same, thus reducing the likelihood of an (uncompensated) accident. Formally, the rule can be written in the following form:

$$f(x,y) = \begin{cases} h_H & \text{if } y \geq y_H^1 \text{ and } x < x_H^1 \\ h_L & \text{if } y_L^1 \leq y < y_H^1 \text{ and } x < x_L^1 \\ 0 & \text{if } y < y_L^1 \end{cases}$$

Notice that the injurer will always "match" the victim's level of care. If the victim takes y_L^1 (respectively, y_H^1), the injurer — to escape liability in magnitude of h_L (respectively, h_H) — takes x_L^1 (respectively, x_H^1). The injurer's reaction function is:

$$x(y) = \begin{cases} x_H^1 & \text{if } y \geq y_H^1 \\ x_L^1 & \text{if } y_L^1 \leq y < y_H^1 \\ 0 & \text{if } y < y_L^1 \end{cases}$$

Therefore, it is again straightforward that the victim will choose either one of the two "threshold" levels of due care, y_L^1 or y_H^1. The victim's incentive constraints are:

$$(IC_L) \qquad y_L^1 + p(x_L^1, y_L^1)h_L \leq y_H^1 + p(x_H^1, y_H^1)h_L \qquad (11)$$

$$(IC_H) \qquad y_H^1 + p(x_H^1, y_H^1)h_H \leq y_L^1 + p(x_L^1, y_L^1)h_H \qquad (12)$$

Yet from the definition of x_H^1, y_H^1 we know that:

$$x_H^1 + y_H^1 + p(x_H^1, y_H^1)h_H \leq x_L^1 + y_L^1 + p(x_L^1, y_L^1)h_H \quad (13)$$

With $x_L^1 \leq x_H^1$, condition (12) follows immediately from (13), thus type H always reveals herself. Rearranging the remaining condition (11), we establish the following claim:

Proposition 3: Under the negligence rule with defense of contributory negligence, a separating equilibrium arises if and only if:

$$y_H^1 - y_L^1 \geq [p(x_L^1, y_L^1) - p(x_H^1, y_H^1)]h_L \quad (14)$$

Otherwise, there is a pooling equilibrium in which both types of victim choose y_H^1 and injurer chooses x_H^1.

Remarks: (i) *The manipulative motive.* The reason that the negligence rule does not necessarily drive the victims to self-select and take optimal care, whereas it did so in the simultaneous moves model, is that in this model the victim "has someone to cheat". Indeed, the injurer will never be negligent and the victim will have to bear his own loss. But by pretending to be of a different type, the victim manages to manipulate the injurer's care level, and alter the probability of bearing his loss. Regardless of his type,the victim remains uncompensated; yet he may find it beneficial to incur an additional cost of care to reduce the likelihood of harm.

(ii) *The value of accuracy.* Kaplow and Shavell (1992a) have argued that courts do not have to bother with assessing the victim's harm accurately, if this determination has no ex-ante incentive effects. The analysis above highlights a situation in which spending resources in order to accurately ascertain harm does

have desirable incentive effects. Accurate determination of harm implies that the victim cannot mislead the court about his true harm, not even by engaging in manipulative care. And if the victim takes the first-best level of care, so does the injurer, as long as he is required to react to what he observes from the victim. Thus, even if the injurer does not know the magnitude of harm (which will only be ascertained ex-post), as long as he observes a signal by the victim and is held to a standard of care that depends on the victim's signal, he can anticipate the harm, and accuracy in assessing harm has valuable incentive effects.

2. Strict Liability with Defense of Contributory Negligence

As in the basic model, the strict liability rule incorporates two threshold levels of contributory negligence, one for each type of victim. Thus, the rule is written as:

$$f(x,y) = \begin{cases} 0 & \text{if } y < y_L^1 \\ h_L & \text{if } y_L^1 \le y < y_H^1 \\ h_H & \text{if } y \ge y_H^1 \end{cases}$$

In contrast to the results of a complete information model (see Shavell (1983)), even if (y_L^1, y_H^1) are set at the first-best levels, optimal care may not actually arise. The low-harm victim type, for example, may prefer to engage in excessive level of care. This way he may convince the court that he is a high-harm type and receive the higher damages award. He would do so if the increase in the expected damages award is greater than the incremental (certain) increase in cost of care. To see this, note that for similar reasoning as was discussed in the previous model, it is again true that the victim will always choose either one of the

due levels of care, y_L^1 or y_H^1. To solve explicitly for the victim's strategy, we first have to derive the injurer's reaction function. If the victim chooses y_L^1, the injurer, expecting a strict liability of h_L, responds with the socially optimal care level, namely $x(y_L^1) = x_L^1$. Likewise, if the victim chooses y_H^1, the injurer's best response is x_H^1. Whether each type of victim will choose the first-best levels of care depends on the satisfaction of the incentive constraints:

$$(IC_L) \qquad y_L^1 + p(x_L^1, y_L^1)(h_L - h_L) \leq y_H^1 + p(x_H^1, y_H^1)(h_L - h_H) \tag{15}$$

$$(IC_H) \qquad y_H^1 + p(x_H^1, y_H^1)(h_H - h_H) \leq y_L^1 + p(x_L^1, y_L^1)(h_H - h_L) \tag{16}$$

Or:

$$p(x_H^1, y_H^1) \leq p^1 \leq p(x_L^1, y_L^1), \tag{17}$$

where p^1 is defined as:

$$p^1 \equiv \frac{y_H^1 - y_L^1}{h_H - h_L}.$$

To summarize,

Proposition 4: In the strict liability with defense of contributory negligence regime with the victim moving first, the first-best separating equilibrium arises if and only if condition (17) holds. Otherwise, there is a pooling equilibrium that is socially worse that even the second-best.

Proof: The first part, the existence of a separating equilibrium that is first-best, is straightforward. Condition (17) guarantees that victim reveals his true type, thus the problem of unobservable harm is indirectly solved, and no welfare loss is required to generate this separation. As for the second part of the proposition, pooling equilibria emerge whenever one of the two inequalities of condition (17) is violated. Any pooling equilibrium is necessarily worse that the second-best.

The reason, stated simply, is that in a pooling equilibrium, the injurer does not vary his level of care with the victim's type. Depending on the equilibrium, the injurer takes either x_L^1 or x_H^1. Yet by the definition of the second-best, the optimal levels of care for the injurer under the constraint that he dos not vary his level of care with victim types, is x^0, not x_L^1 nor x_H^1.

Remark: *Decoupling:* As argued in section I above, a strict liability regime in which injurer's liability and victim's compensation are decoupled, can eliminate the pooling equilibria. The victim, who gets a fixed amount regardless of his type, has no incentive to cheat. The injurer, who pays an additional fine that depends on the victim's revealed type, can be induced to by the appropriate fine to exercise optimal care. However, such a regime cannot implement the first-best separating equilibrium. For the same reasons as were argued in section I, the victim may have incentives to take either excessive or too little care, to influence the probability of an undercompensated or a profitable accident.

III. Legal Application: A Theory of Comparative Negligence

The Doctrine of comparative negligence states that when both parties, the victim and the injurer, are negligent, damages are apportioned according to the relative degrees of fault.[12] While the ordinary justification for this principle is a fairness argument, the "relative degrees of fault" can also be given an economic content. In the economic literature, the doctrine is interpreted as allocating the loss according to the proportion by which each party departed from his due level

[12]See Li v. Yellow Cab Co. of California, 13 Cal.3d 804, 532 P.2d 1226, 119 Cal. Rptr. 858 (1975).

of care. This literature has argued that while there may be circumstances where comparative negligence generates optimal incentives, it may also be dominated by other liability rules. In general, there is no characterization of the instances in which comparative negligence is optimal.[13]

Our analysis leads to a new interpretation of the comparative negligence doctrine. In our model, the comparative negligence rule does not address issues of relative fault in the deviation from standards of due care, because rational parties never deviate from the standards that are set optimally. Instead, the comparative negligence rule addresses the problem of unobservable harm. Unable to determine the true magnitude of harm suffered by the victim, courts deduce information about it from what they do observe – the care the victim had taken. Courts presume that as the magnitude of harm rises, the victim – who knows the true harm – will be taking more care. Thus, care is a signal, not for relative fault, but for the magnitude of subjectively anticipated harm. Lower care makes the victim comparatively negligent and reduces his compensation, yet in fact he had taken the optimal care given her individual harm, and indeed, he is fully compensated. However, in order to guarantee that this rationale works, truthful

[13]See Shavell (1987), pp. 39-40, 103-4 (establishing some cases where comparative negligence works efficiently, and arguing that in the existence of uncertainty about the victims ability to take care, comparative negligence works best); Haddock and Curran (1985) (arguing that comparative negligence may be superior an all-or-nothing contributory negligence defense when courts make errors in determining the actual care taken, or when parties err in predicting the due level of care); Cooter and Ulen (1986) (arguing that comparative negligence generates smaller excess levels of care when the parties expect courts to set non-optimal due levels of care); Rubinfeld (1987) (arguing that comparative negligence induces parties with lower cost of care to take higher precautions); Rea (1987) (demonstrating that comparative negligence is the optimal rule when some agents do not respond legal incentives and cannot be deterred from being negligent). See also Schwartz (1978).

revelation must be implemented, namely that victims do not send false, manipulative signals. We have proved that only under a negligence regime the incentive constraints hold, thus the mechanism works.

If our interpretation is true and the comparative negligence doctrine is part of a mechanism that addresses the problem of unobservable harm, and if only under a negligence liability regime it implements the second-best separating equilibrium, then we are led to hypothesize that the comparative negligence defense is more often applied under negligence regimes than under strict liability regimes. As a positive theory, out hypothesis could be tested by checking whether the defense of comparative negligence is more often used in areas where negligence is the reigning rule. And indeed, in areas of tort law where the cause of action is strict liability, courts have often held that the comparative negligence defense does not apply.[14]

IV. Extensions and Discussion

(a) *Summary of results.* The main idea analyzed in the paper is whether courts can overcome the problem of unobservable harm by observing the victim's level of care, and attributing higher harm to a victim that took more care. It was demonstrated that in most cases, this method of inference is not strategy-proof, and may lead some victim types to manipulate their level of care. In one case -- under a negligence regime when the victim and the injurer choose care simultaneously, care can perfectly reveal harm. But when the injurer chooses care after observing the victim's choice of care, neither strict liability (with or without

[14]See Seary v. Chrysler Corp., 609 P.2d 1382 (Wash. 1980).

decoupling), nor negligence can implement the first-best levels of care. Unobservability of harm inflicts a social cost.

Kaplow and Shavell (1992) argued that when the injurer cannot observe the victim's harm, there is no sense in courts investing resources in order to determine the victim's harm ex-post, as this assessment has no incentive effects. Our analysis qualifies this observation. For accuracy in the assessment of harm to generate ex-ante incentive effects, it is sufficient that the injurer can observe the victim's level of care before the injurer acts. The injurer, who is then required to take care that "matches" the victim's care, and the victim, who cannot mislead the court, are both led to behave optimally.

The following remarks extend the analysis and generalize it.

(b) *Self-insurance.* Another class of precautionary behavior, not discussed thus far, is one which reduces the magnitude of harm, not its probability. It is usually referred to as "self-insurance".[15] We can extend the model to include the possibility of self-insurance. In such a framework, victims take two kinds of care, one that reduces the probability of harm and another that reduces its magnitude. Victims vary according to their particular ability to reduce the magnitude of harm.[16] All the results derived above carry over to a model that incorporates both kinds of care activity. As long as the standards of due care can be defined with respect to both care variables, and both can be observed by the court, the

---

[15]See Ehrlich and Becker (1972) for the original distinction between "self-protection" and "self-insurance".

[16]If z denotes the level of self-insurance, harm is now written as $h(z)$, and different types of victim have different such function. Specifically, type L's function is $h_L(z)$, and type H's function is $h_H(z)$, with $h_L(z) < h_H(z)$ for all z.

situation is strategically identical to the one analyzed thus far. In particular, a liability rule can require the victim to meet due levels of care in both self-protection and self-insurance, to qualify for compensation. It follows, again, that only a negligence regime can implement second-best levels of care.

(c) *More than two types.* How robust are the results of the two-type model? All the results discussed above continue to hold if the model has n types or if victim's types are drawn from a continuous set. While a technical demonstration of this claim is omitted, the reason that it is true is the following. Take a negligence regime that implements the second-best separating equilibrium. It does so because the victim, who bears his loss, has nobody to cheat and is best-off taking optimal care. But this is true regardless of how many and how different the other victims are. Next, take any other regime. It was shown that there may always be a victim type that wants to cheat. But if this victim is better-off posing as the other type when there is only one additional type, he is also better-off doing so when there are many types. Hence, the model's results are robust.

(d) *Application to criminal law.* The idea that the victim's behavior reveals valuable information when harm is not readily observable, can be extended beyond the tort law area. It applies to criminal law, in which the victim's behavior is often relevant for both the definition of a crime and the magnitude of the penalty imposed on the offender. Different kinds of precautions by victims give rise to different criminal classifications of the same act by the offender.[17] Thus, the

[17]Harel (1992) studies the efficiency and the equity of "blaming the victim".

level of deterrence depends to some extent on the victim's care, and consequently, different types of victims take different levels of care.

REFERENCES

Cooter, Robert D., and T.S. Ulen (1986), *An Economic Case for Comparative Negligence*, N.Y.U. Law Review 61: 1067-1100.

Craswell, Richard, and J.E. Calfee (1986), *Deterrence and Uncertain Legal Standards*, Journal of Law, Economics and Organization 2: 279-303.

Ehrlich, Issac and G. Becker (1972), *Market Insurance, Self-Insurance and Self-Protection*, Journal of Political Economy 80: 623-648.

Emons, Winand (1990), *Some Recent Developments in the Economic Analysis of Liability Law: An Introduction*, Journal of Institutional and Theoretical Economics 146: 237-248.

Fleming, John (1987), *The Law of Torts*, 7th ed., Sydney.

Haddock, David, and C. Curran (1985), *An Economic Theory of Comparative Negligence*, Journal of Legal Studies 14: 49-72.

Harel, Alon (1992), *The Efficiency and Equity of Blaming the Victim*, Mimeo. Hebrew University Faculty of Law.

Holmström, Bengt and R. Myerson (1983), *Efficient and Durable Decision Rules with Incomplete Information*, Econometrica 51: 1799-1819.

Kaplow, Louis, and S. Shavell (1992a), *Accuracy in the Assessment of Damages*, Mimeo., Harvard Law School.

_____ (1992b), *Accuracy in the Determination of Liability*, Mimeo., Harvard Law School.

Png, Ivan P.L. (1986), *Optimal Subsidies and Damages in the Presence of Judicial Error*, International Review of Law and Economics 6: 101-105.

Polinsky, A. Mitchell and S. Shavell (1989), *Legal Error, Litigation, and the Incentive to Obey the Law*, Journal of Law, Economics and Organization 5: 99-108.

Rea, Samuel A. (1987), *The Economics of Comparative Negligence*, International Review of Law and Economics 7: 147 -162.

Rubinfeld, Daniel L. (1987), *The Efficiency of Comparative Negligence*, Journal of Legal Studies 16: 375-394.

Schwartz, Gary T. (1978), *Contributory and Comparative Negligence: A Reappraisal*, Yale Law Journal 87: 697-727.

Shavell, Steven (1983), *Torts in Which Victim and Injurer Act Sequentially*, Journal of Law and Economics 26: 589-612.

_____ (1987), *Economic Analysis of Accident Law*, Harvard University Press.