

The Role of Social Image Concerns in the Design of Legal Regimes

Bruno Deffains* Claude Fluet^{†‡}

PRELIMINARY, January 2013

Abstract

We consider situations where liability yields insufficient incentives to control risks, e.g., parties are not always sued for harm done or are unable to pay fully for it. Some potential injurers nevertheless take appropriate precautions because of intrinsic prosocial concerns. Others have no such concerns but would like people to believe that they do. We show that the negligence rule is more effective than strict liability at transforming social image concerns into incentives to exert care. When evidence about care is imperfect, the rules of proof affect the inferences drawn from court decisions and therefore the stigma attached to a negligence ruling. If inadequate care is a rare event, plaintiffs should bear the burden of proving the defendant's negligence; otherwise there are cases where defendants should bear the burden of proving compliance with due care. Under either assignment of the burden of proof, incentives to comply are maximized by a standard of proof stronger than a mere preponderance of evidence.

KEYWORDS: Normative motivations, prosocial behavior, strict liability, negligence, burden of proof, standard of proof. (JEL: D8, K4, Z13)

*Université Paris 2 and European Business School at Wiesbaden. E-mail: bruno.deffains@u-paris2.fr

[†]Université du Québec à Montréal, Université Paris 2 and CIRPEE. E-mail: fluet.claude-denys@uqam.ca

[‡]We thank Cécile Bourreau-Dubois, Dominique Demougin, Yolande Hiriart, Giovanni Immordino, Éric Langlais, Urs Scheizer, Gerhard Wagner, and Ansgar Wohlschlegel for useful comments.

1 Introduction

In the *Theory of Moral Sentiments*, Adam Smith remarks that an individual found to have caused harm faces not only the possibility of a legal sanction — e.g., the damages he must pay — but also social disapproval or stigma. The Law and Economics literature has studied stigma mainly in relation to criminal activity (Rasmusen 1996, Harel and Clement 2007, Zasu 2007, among others). We inquire how a concern for social approval interacts with the incentives created by tort law and how this affects the relative performance of different legal regimes.

We consider the unilateral accident model when the strict liability and negligence rules yield insufficient incentives to control risks. Injurers are partially judgment-proof or are not always sued, e.g., it is not always feasible to prove harm or identify the injurer. Some potential injurers are nevertheless assumed to exert socially efficient care. They do so out of intrinsic moral or prosocial concerns. Other injurers have no such concerns but would like people to believe that they do; that is, they care about social approval. For instance, in a recent experimental study on liability rules (Angelova *et al.* 2012), half the subjects invested in safety measures even in the “No Liability” treatment and even though they could not be identified by the other subjects as having caused harm. Many experimental or field studies have also shown that social image concerns are important motivators of prosocial behavior (Dana *et al.* 2006, Ellingson and Johannesson 2008, Andreoni and Bernheim 2008, Ariely *et al.* 2010, Funk 2010, Lacetera and Macis 2010, among others).¹

In our framework, an individual’s actions are not directly observable by society at large. However, adverse court judgments provide public information from which inferences can be drawn about the individuals’ actions and

¹See Shavell (2004) for a general discussion of legal sanctions versus informal motivation as regulators of conduct.

therefore about their intrinsic predispositions. Under either strict liability or the negligence rule, social image concerns are shown to provide the non prosocial individuals with some incentives to mimic the virtuous. The issue is how this influences the optimal design of liability regimes, when the objective is to induce the greatest number of individuals to exert socially efficient care.

A basic result is that the negligence rule is more effective than strict liability in harnessing reputational concerns. The reason is simply that trial outcomes are then more informative. Under strict liability, an adverse ruling merely ascertains that the defendant caused harm, not that he took inadequate precautions. Under the negligence rule, a liability ruling also ascertains that the defendant exerted inadequate care, thereby providing more precise information about his intrinsic predispositions. Socially useful incentives are therefore derived from the signaling role of “fault” or “negligence”.

To further explore this possibility, we extend the analysis to imperfect evidence about a defendant’s precautions. A complete characterization of the negligence regime must now consider how courts deal with the risk of judicial error. The legal tools for this purpose are the assignment of the burden of proof — whether it is for the plaintiff to prove the defendant’s negligence or for the defendant to prove compliance with due care — and the standard of proof that needs to be satisfied by the party with the burden of proof. The burden of proof assignment and the standard of proof affect the inferences drawn from trial outcomes; that is, they bear on the “meaning” or “significance” of a ruling of negligence. We show that when injurers have social image concerns, and by contrast with the results in Demougin and Fluet (2006, 2008), compliance with due care is maximized by a standard of proof stronger than the common law preponderance of evidence standard. Roughly speaking, the assignment of the burden of proof depends on whether inadequate care is a frequent or infrequent event. In the usual case where

harm seldom occurs, incentives to comply with due care are maximized by assigning to the plaintiff the burden of proving the defendant's non-compliance.

A recent microeconomic literature has emphasized that one's actions may signal something about unobservable predispositions and that some predispositions are socially valued; see in particular Bernheim (1994), Bénabou and Tirole (2006, 2011), and Daughety and Reinganum (2010). Deffains and Fluet (2013) incorporate this approach in the unilateral accident model. The focus of that paper is the extent to which formal legal sanctions crowd-out or crowd-in informal motivations under different liability rules.² In the present paper, we also compare liability rules but the emphasis is in providing prescriptions about the law of evidence under evidentiary uncertainty.

Section 2 presents the basic setup. Section 3 compares the incentives to control risks under strict liability and the negligence rule with no judicial error. The next two sections introduce imperfect evidence about the defendant's precautions and derive the implications concerning the design of compliance maximizing legal regimes. Section 6 concludes. Proofs are in the Appendix.

2 The model

We start with a simple version of the unilateral accident model. Risk-neutral individuals are engaged in a socially valuable activity which may impose an accidental loss of amount L on third parties. The risk of causing harm depends on the potential injurer's level of care which is $e = 0$ for low (or no) care and $e = 1$ for high care respectively. The probability of accident is $p(0) = p_l$ and $p(1) = p_h$ where $p_l > p_h > 0$. The opportunity cost of low care is normalized to zero, that of high care is c distributed according to the

²How material penalties and rewards affect informal motivations has been explored in a vast experimental and empirical literature. See Frey and Jegen (2001) for a survey.

twice-differentiable cumulative function $G(c)$ with support $[0, \bar{c}]$. The interpretation is that an individual's cost of care depends on the circumstances he may be facing.

Social welfare is maximized if, depending on the circumstances, injurers choose the level of care minimizing the sum of precaution costs and of expected harm. When the cost of care is c , the socially efficient level of care minimizes $ce + p(e)L$. It is therefore

$$e^*(c) = \begin{cases} 1 & \text{if } c \leq c^* \equiv (p_l - p_h)L, \\ 0 & \text{otherwise,} \end{cases}$$

where c^* is the critical cost level below which high care should be exerted.

Liability. We consider situations where liability does not always ensure socially efficient precautions. First, injurers are sued for harm done only with some probability q . Secondly, injurers may not be able to pay fully for the harm they cause, i.e., the legal damages they will actually pay if found liable is $\min(w, L)$ where w denotes the injurers' assets.³ It is well known that, by contrast with strict liability, the negligence rule may then nevertheless induce efficient care. The following assumption rules out this possibility.

ASSUMPTION 1: $p_l q \min(w, L) < \min[\bar{c}, (p_l - p_h)L]$.

Any combination of q and w satisfying the assumption is sufficient for our purpose. Consider the strict liability rule. Injurers must then in principle pay for the harm they cause irrespective of circumstances. However, because he is not always sued or cannot pay fully for harm done, an individual with cost of care c exerts high care only if $c \leq q(p_l - p_h) \min(w, L)$. Given Assumption 1, there are therefore circumstances where individuals will take inadequate precautions.

Under the negligence rule, courts are assumed to observe *ex post* the cost c which the defendant faced and to set due care at the socially efficient

³A similar set-up is employed in Shavell (1984) to analyze the joint use of liability and regulation. Note that the cost of care is implicitly assumed to be non-pecuniary.

level. Accordingly, an individual faces a risk of liability only when he exerts low care and $c \leq c^* \equiv (p_l - p_h)L$. Because of the dilution of incentives due to the chance of not being sued or to partial judgment-proofness, the injurer complies with due care only if $c \leq p_l q \min(w, L)$. Again, given Assumption 1, there will be circumstances where precautions are inadequate, although inefficient care will be less frequent than under strict liability. See Figure 1. Note that for the time being we abstract from the possibility of judicial error.

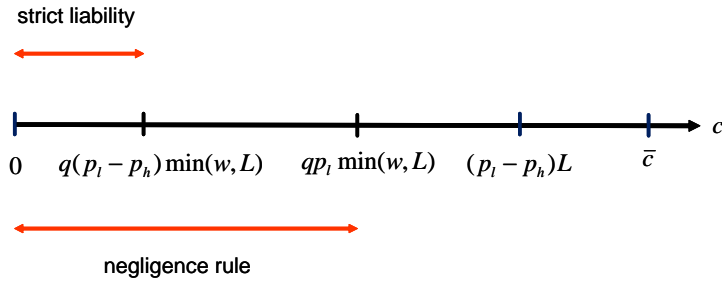


Figure 1. Inefficiencies

Social preferences. So far we have described the standard framework where the injurers’ behavior depends only on private costs and benefits as conventionally defined. We now consider informal motivations. We assume that there are two types of potential injurers. First, some potential injurers are “good citizens” with prosocial predispositions. They seek to behave in a socially or morally responsible manner by comparing their opportunity cost of care with the expected harm they impose on others.⁴ Such individuals, referred to as type $\theta = 1$, choose the socially efficient level of care irrespective of legal legal sanctions, i.e., they exert high care when $c \leq c^*$ and low care otherwise. There is a known proportion λ of such individuals.

Secondly, individuals who are thought to be intrinsically prosocial earn social esteem, a source of utility. For those individuals who are not prosocial,

⁴This is a simple version of Kant’s Categorical Imperative. See Brekke et al. (2003) for a formulation.

referred to as type $\theta = 0$, behavior is determined by the utility function $u = w_n + \beta\bar{\theta}_I$ where w_n is net final wealth, β is a positive parameter and $\bar{\theta}_I \equiv E(\theta | I)$ is the belief of society at large about the individual's type conditional on the information I . Given our definition of types, $\bar{\theta}_I$ is simply the posterior probability that the individual is intrinsically prosocial and $\beta\bar{\theta}_I$ is the utility derived by the individual from society's beliefs about his intrinsic predispositions.

An individual's type is private information. For society at large, so is the cost of care he faced, his chosen level of care and whether he caused harm, except insofar as these can be inferred from court judgments. Specifically, the only information “publicly” available about an individual — that is, in society at large — is either B for “bad news” or G for “good news”. Bad news refers to the case where the individual is known to have been found liable under the prevailing liability rule. Good news is the complementary event.

The meaning of good and bad news therefore depends on the liability regime. Under strict liability, B means that the individual caused harm, was subsequently sued (which arises only with probability q) and held liable. Under the negligence rule, B means that the individual caused harm, was sued and held liable, hence was found to have been negligent. In the case of a suit under the negligence rule, the court is assumed to observe the cost of care faced by the defendant and the level of care he exerted. Hence it can determine whether the defendant was negligent. However, society at large is only informed of adverse trial outcomes, i.e., the general public has no time for details.

3 Strict Liability versus Negligence

By assumption prosocial individuals always take adequate precautions, so we only need to examine the behavior of the non prosocial. Let $\alpha(e, c)$ denote

the probability that an individual is found liable given that he caused harm and is sued. The probability is determined by the liability rule and may depend on the defendant's level of care and the circumstances. Under strict liability, $\alpha(e, c) = 1$ irrespective of care and of circumstances. Under the negligence rule, $\alpha(e, c) = 1$ if $e = 0$ and $c \leq c^*$ and is otherwise zero.

Incentives. Given the cost of care c , the expected utility of a non prosocial as a function of his care level is

$$\begin{aligned} \bar{u} = & p(e)q\alpha(e, c)[w - \min(w, L) + \beta\bar{\theta}_B] \\ & + (1 - p(e)q\alpha(e, c))[w + \beta\bar{\theta}_G] - ce, \quad e \in \{0, 1\}, \end{aligned} \quad (1)$$

where $\bar{\theta}_B$ and $\bar{\theta}_G$ are society's beliefs about the individual's type conditional on "bad" and "good" news respectively. These beliefs are determined at equilibrium but are taken as given by the individual. It is easily seen that the individual exerts high care if and only if

$$c \leq q[p(0)\alpha(0, c) - p(1)\alpha(1, c)] [\min(w, L) + \beta\Delta], \quad (2)$$

where $\Delta \equiv \bar{\theta}_G - \bar{\theta}_B$ will be referred to as the reputational penalty from an adverse court judgment.

Under the strict liability rule, the condition (2) reduces to

$$c \leq q(p_l - p_h)[\min(w, L) + \beta\Delta] \equiv c_S(\Delta). \quad (3)$$

The right-hand side is the critical cost below which a non prosocial exerts high care under the strict liability rule (the subscript S refers to strict liability). The critical cost level is written as a function of the reputational penalty yet to be determined.

Under the negligence rule, $\alpha(1, c) = 0$ for all values of c and $\alpha(0, c) = 0$ if $c > c^*$. Therefore the condition (2) cannot be satisfied when $c > c^*$, i.e., the individual then necessarily exerts low care. He exerts high care if and only if

$$c \leq \min\{qp_l[\min(w, L) + \beta\Delta], c^*\} \equiv c_N(\Delta), \quad (4)$$

where the right-hand side is the critical cost below which a non prosocial exerts high care under the negligence rule (hence the subscript N). Observe that under the negligence rule the critical cost cannot be above the efficient threshold c^* .

The proportion of the non prosocial exerting high care is $G(c_r)$ where c_r is the critical cost level under the liability rule $r = S, N$. It will be useful to focus instead on the ratio, denoted y , of the population of non prosocial exerting high care over those who *should* be exerting high care. This ratio will be referred to as the compliance rate and is defined by $y \equiv G(c_r)/G(c^*)$. From the above discussion, the compliance rate is a function of the reputational penalty and satisfies

$$y = \psi_r(\Delta) \equiv \frac{G(c_r(\Delta))}{G(c^*)}, \quad r = S, N. \quad (5)$$

When the reputational penalty is nil, the compliance rate is the same as would be observed in the standard model without social preferences. Recall that Assumption 1 then ensures that $\psi_S(0) < \psi_N(0) < 1$. Under strict liability, the compliance rate is increasing in the reputational penalty (as long as $c_N(\Delta)$ is less than the upper bound \bar{c}). For a sufficiently large penalty, overcompliance (i.e., $y > 1$) is consistent with (3) and (5), although this will never arise at equilibrium as shown below. Under the negligence rule, the compliance rate is also increasing in the reputational penalty, but only up to the socially efficient compliance rate $y = 1$.

Beliefs. We now turn to the determination of the reputational penalty. Using Bayes' rule, the beliefs $\bar{\theta}_G$ and $\bar{\theta}_B$ under a given liability rule — and therefore the reputational penalty — can be expressed as a function of the compliance rate.

Lemma 1 *Let $p^* \equiv G(c^*)p_h + (1 - G(c^*))p_l$. Given the compliance rate y ,*

the reputational penalty satisfies $\Delta = \varphi_r(y)$, $r = S, N$. Under strict liability,

$$\varphi_S(y) \equiv \frac{\lambda(1 - qp^*)}{1 - q[p^* + (1 - \lambda)(1 - y)G(c^*)(p_l - p_h)]} - \frac{\lambda p^*}{p^* + (1 - \lambda)(1 - y)G(c^*)(p_l - p_h)}.$$

Under the negligence rule

$$\varphi_N(y) \equiv \frac{\lambda}{1 - (1 - \lambda)(1 - y)G(c^*)p_l q}.$$

Both $\varphi_S(y)$ and $\varphi_N(y)$ are strictly decreasing functions, with $\varphi_S(y) < \varphi_N(y)$ for $y \leq 1$ and $\varphi_S(1) = 0$, $\varphi_N(1) = \lambda$.

Note that p^* is the average probability of accident when all the non prosocial exert the socially appropriate level of care. For a given compliance rate and restricting attention to $y \leq 1$, the reputational penalty is smaller under strict liability because an adverse judgment provides noisier information about the individual's intrinsic predispositions. Under the negligence rule, bad news reveals perfectly that the injurer is non prosocial, hence $\bar{\theta}_B = 0$. Under strict liability, this only becomes more likely so.

When all injurers exert efficient care (i.e., when $y = 1$), bad and good news provide no information at all under strict liability. The prosocial and non prosocial then behave the same. Because they face the same probability of an adverse judgment, the updated beliefs do not differ from the prior, i.e., $\bar{\theta}_B = \bar{\theta}_G = \lambda$. By contrast, when all individuals comply with due care under the negligence rule, good news is uninformative ($\bar{\theta}_G = \lambda$) because it occurs with certainty, but bad news would remain perfectly revealing ($\bar{\theta}_B = 0$).⁵

Finally, the fact that $\varphi_S(y)$ and $\varphi_N(y)$ are strictly decreasing functions may be interpreted in terms of strategic substitutability. Everything else

⁵Bad news is then an out-of-equilibrium event with zero probability, hence $\bar{\theta}_B$ cannot be computed using Bayes' rule. The reputational penalty is then obtained from $\varphi_N(1) \equiv \lim_{y \rightarrow 1} \varphi_N(y) = \lambda$. The belief $\bar{\theta}_B = 0$ can also be rationalized in terms of Cho and Kreps' (1987) D1 criterion.

equal, when more individuals exert efficient care the reputational penalty decreases, hence the incentive to avoid an adverse judgment becomes smaller.

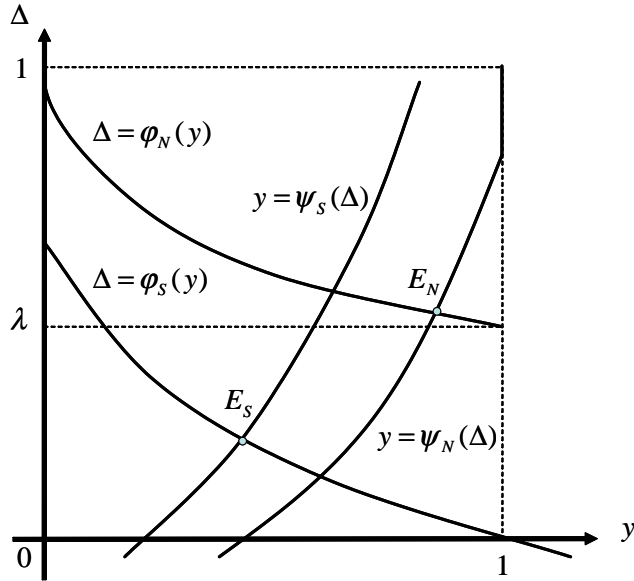


Figure 2. Equilibria

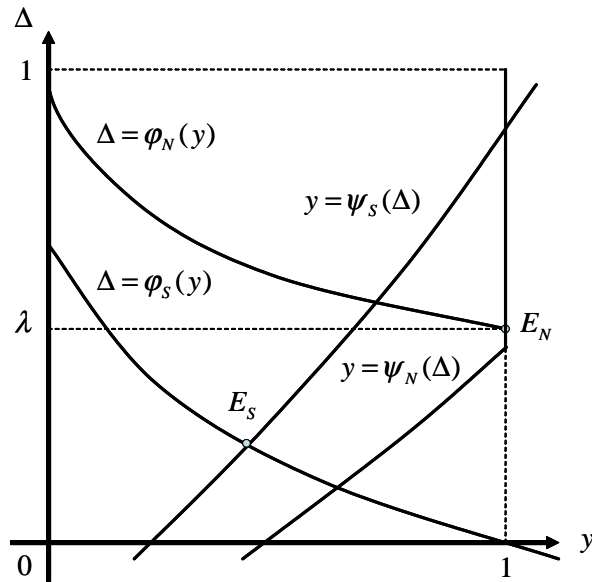


Figure 3. First best under the negligence rule

Equilibrium. An equilibrium is a compliance rate and a reputational penalty that simultaneously solve $y = \psi_r(\Delta)$ and $\Delta = \varphi_r(y)$. Denote an equilibrium by $E_r = (y_r, \Delta_r)$. Figure 2 provides an example for both the strict liability and negligence rules. Figure 3 provides yet another example with a larger β , i.e., the non prosocial care more about social approval. The figure illustrates the case where β is large enough for everyone to comply with due care under the negligence rule.⁶ The next proposition summarizes the main properties of the equilibria.

Proposition 1 *Under strict liability or the negligence rule there is a unique equilibrium (y_S, Δ_S) or (y_N, Δ_N) respectively. Under Assumption 1,*

- (i) $y_S < y_N \leq 1$ and $\Delta_S, \Delta_N > 0$ with $\Delta_N = \lambda$ when $y_N = 1$;
- (ii) y_S and y_N are increasing in w, q or β ; Δ_S and Δ_N are decreasing in w and β ;
- (iii) Δ_N and y_N are increasing in λ ; Δ_S and y_S are increasing (decreasing) in λ when λ is small (large).

Part (i) states that, in terms of providing incentives to exert efficient care, the negligence rule always does better than strict liability. In particular, the negligence rule deters optimally if esteem concerns are strong enough, while strict liability always underdeters. Part (ii) states that relaxing the sources of inefficiency, i.e., increasing the probability of suit or reducing the extent of judgment-proofness, increases compliance although there may be some “motivational crowding-out” (i.e., the reputational penalty may decrease). Similarly, a greater concern for social image improves compliance. Due to esteem concerns, there is an externality between prosocial individuals and the behavior of the non prosocial. Part (iii) considers whether the externality has greater force the greater the number of prosocial individuals. Under the negligence rule, a greater proportion of virtuous individuals shifts the φ_N curve upwards and therefore induces a greater number of the non prosocial

⁶The condition is $p_l q [\min(w, L) + \beta \lambda] \geq c^*$.

to exert efficient care. Under strict liability, the effect is ambiguous: if the proportion of virtuous individuals is sufficiently small, more of them shifts the φ_S curve upwards, thereby increasing deterrence; the opposite obtains when the proportion of virtuous individuals is sufficiently large.

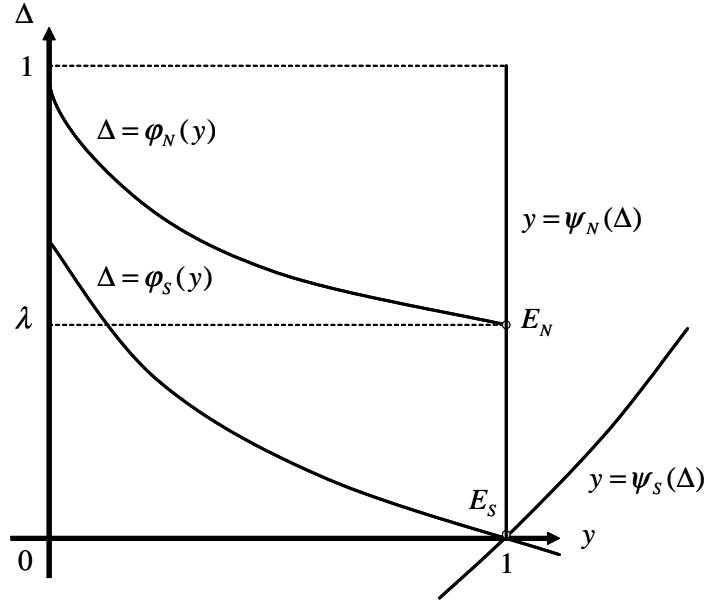


Figure 4. Relaxing Assumption 1

To conclude this section, we briefly describe the equilibria when Assumption 1 is relaxed. Suppose injurers can always be sued and have sufficient assets to pay in full for the harm done. In the standard model without social image concerns, the first best then obtains under either strict liability or the negligence rule. With respect to the functions defined in (5), we now have $\psi_N(0) = \psi_S(0) = 1$. The equilibria are represented in Figure 4. Under strict liability, the equilibrium is $(y_S, \Delta_S) = (1, 0)$; that is, an adverse court judgment imposes no reputational penalty, implying that the non prosocial are now motivated solely by formal legal incentives. Under the negligence rule, the equilibrium is $(y_N, \Delta_N) = (1, \lambda)$. Although having been found negligent would impose a reputational penalty, the incentives

provided by reputational concerns are now superfluous with respect to inducing compliance with due care, by contrast with the case illustrated in Figure 3.

4 Judicial Error under the Negligence Rule

The above analysis shows that the negligence rule is more efficient at transforming the externality due to social image concerns into incentives to exert socially appropriate care. As noted, the reason is that trial outcomes are more informative about a defendant's behavior and therefore about his intrinsic predispositions. It is not clear, however, whether the negligence rule would remain more efficient if its greater informational requirements cannot be met perfectly. To address this issue, we extend our analysis of the negligence rule to the case where trial outcomes are noisy signals. Specifically, we assume that the evidence about the injurers' behavior is not perfectly informative, so that can courts make mistakes in assessing whether the defendant complied with due care. They can erroneously rule against the defendant (referred to as a "false positive" or type I error) or erroneously rule against the plaintiff (a "false negative" or type II error).

The possibility of judicial error implies that the negligence rule will have elements of strict liability. Moreover, a complete description of the negligence rule must now take into account how the judicial system trades-off type I and type II errors. Different trade-offs define different legal regimes. In practice, the trade-off is determined by the prevailing rules of proof, by which we mean the burden of proof assignment and the standard of proof. The party with the burden of proof needs to persuade the court that he is entitled to a judgment in his favor, otherwise the default decision is that he loses the case. The standard of proof refers to the weight of evidence needed to discharge the burden.

Burden and standard of proof. Because accidents occur more often

under low care, the mere occurrence of harm provides some information about the injurer’s behavior. Any additional information that might be used to assess behavior is summarized by the random variable x with cumulative distribution functions $F_h(x)$ and $F_l(x)$ that depend on the defendant’s level of care. The distributions have continuously differentiable density functions, denoted $f_h(x)$ and $f_l(x)$, and the same support $[\underline{x}, \bar{x}]$ where the bounds need not be finite. The “invariant support” condition means that no realization x perfectly reveals the defendant’s care level.

ASSUMPTION 2: $f_l(x)/f_h(x)$ is strictly decreasing with $f_l(\underline{x})/f_h(\underline{x}) = \infty$ and $f_l(\bar{x})/f_h(\bar{x}) = 0$.

The distributions satisfy the monotone likelihood ratio property (MLRP) with the convention that a small value of x is more indicative of low care. The condition that f_l/f_h goes from infinity to zero means that extreme values are tantamount to perfectly informative evidence.⁷

The plaintiff has the burden of proving the occurrence of harm and the injurer’s identity. As before, this can either be done without ambiguity (with probability q) or not at all. When this requirement is satisfied, a suit is feasible. Both the plaintiff and the defendant are then assumed to have access to the additional evidence x about the injurer’s behavior, as well as to perfect evidence concerning the circumstances c faced by the defendant. The complete evidence eventually submitted to the court therefore comprises the occurrence of harm and the injurer’s identity, the cost of care c faced by the injurer and the realization x . When $c > c^*$ the injurer would not be found negligent, so the victim does not sue. When $c \leq c^*$ there are two possibilities:

- (i) If, as is usually the case, the plaintiff has the burden of proving the

⁷The assumption is not essential but it simplifies the exposition by eliminating the possibility of corner solutions in what follows. The invariant support condition is not essential either. If the supports of f_h and f_l overlap only partly, the evidence will sometimes (but not always) reveal care perfectly.

defendant's negligence, he succeeds only if he can submit x such that

$$\frac{p_l f_l(x)}{p_h f_h(x)} > k \tag{6}$$

where k is the standard of proof that the plaintiff must satisfy to discharge the burden. The expression on the left-hand side is the likelihood ratio of low care versus high care on the part of the defendant, given the evidence "occurrence of harm and x ". Condition (6) states that the plaintiff must adduce evidence showing that inadequate care is k times more likely than due care.

(ii) If the injurer bears the burden of proving that he complied with due care, he avoids liability only if he can submit x such that

$$\frac{p_h f_h(x)}{p_l f_l(x)} > k.$$

The interpretation is similar except that the left-hand side is now the relative likelihood of high versus low care. To escape liability, the defendant must show that compliance with due care is k times more likely than non compliance.

Note that court rulings are assumed to be based purely on the evidence pertaining to the particular case before the court. The rationale is that rulings are independent of the "priors" the court may hold about the general prevalence of high care among the population of injurers with cost of care c . The contested issue is the particular defendant's actions. Both in common law and in civil law, priors in the form of a "known" (e.g., at equilibrium) proportion of injurers exerting high care would not be considered as relevant or admissible evidence.⁸

We consider standards of proof satisfying $k \geq 1$. The case $k = 1$ is the common law *preponderance of evidence* standard. For the party with the burden of proof, it then suffices to show that the relevant evidence

⁸Or would amount to statistical discrimination. For a discussion, see Demougis and Fluet (2006) and the references therein.

gives greater weight to his contention, however slightly; that is, the party with the burden of proof need only prove his claim on a “more likely than not” basis. A threshold $k > 1$ means a stronger standard. For instance, it is sometimes said that $k = 3$ roughly conveys the standard of *clear and convincing evidence* (see Schauer and Zeckhauser 1996).

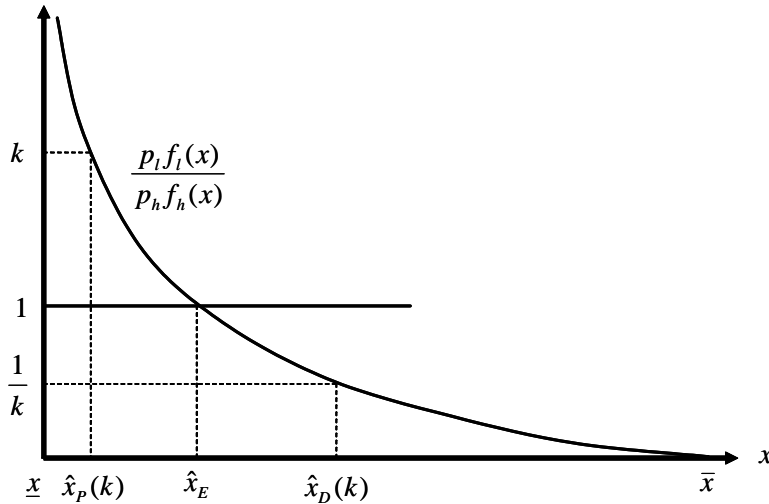


Fig. 5. Likelihood ratios

For a given standard of proof, the assignment of the burden of proof yields different *evidentiary thresholds* for the court to rule in favor of the plaintiff or the defendant. In figure 5, $\hat{x}_P(k)$ is the evidentiary threshold when the plaintiff bears the burden of proof. The defendant is found negligent when $x < \hat{x}_P(k)$. When the defendant bears the burden of proving compliance with due care, the evidentiary threshold is $\hat{x}_D(k)$. The defendant then escapes liability only if $x > \hat{x}_D(k)$. For the preponderance of evidence standard, the evidentiary threshold does not depend on the assignment of the burden of proof; it is denoted \hat{x}_E in the figure.

The liability risk differential. Let \hat{x} be the evidentiary threshold for some assignment of the burden and some standard of proof. Conditional on the occurrence of harm and a suit being filed, the probability that the

injurer will be found liable is $\alpha_j \equiv F_j(\hat{x})$ depending on his care level $j = h, l$. When the injurer complied with due care, the probability of a type I error is α_h ; when he exerted inadequate care, the probability of a type II error is $1 - \alpha_l$. For any evidentiary threshold, the monotone likelihood ratio property implies that $\alpha_h < \alpha_l$ except when the threshold is at the bounds of the support, in which case the equality holds. Ex ante, given the possibility of court error, the liability risk differential between low and high care is $\delta = q(p_l \alpha_l - p_h \alpha_h)$.

It is useful to express α_l as a function of the type I error $\alpha_h = F_h(\hat{x})$, i.e., $\alpha_l(\alpha_h) \equiv F_l(F_h^{-1}(\alpha_h))$.

Lemma 2 $\alpha_l(\alpha_h)$ is strictly concave with $\alpha_l(0) = 0$, $\alpha_l(1) = 1$ and

$$\alpha_l'(\alpha_h) = \frac{f_l(F_h^{-1}(\alpha_h))}{f_h(F_h^{-1}(\alpha_h))}, \alpha_h \in [0, 1]. \quad (7)$$

Written as a function of the type I error, the liability risk differential is $\delta(\alpha_h) = q(p_l \alpha_l(\alpha_h) - p_h \alpha_h)$ and is therefore concave in α_h . Figure 6 provides an illustration. Observe that the differential is zero when $\alpha_h = 0$ and that it is equivalent to the one under strict liability when $\alpha_h = 1$, except for the fact that injurers now escape liability when c is above c^* .

Lemma 3 Under Assumption 2, $\delta(\alpha_h)$ has a strict interior maximum at $\alpha_h = \alpha_h^E \equiv F_h(\hat{x}_E)$ where \hat{x}_E solves

$$\frac{p_l f_l(\hat{x}_E)}{p_h f_h(\hat{x}_E)} = 1.$$

The lemma states that, irrespective of the burden of proof assignment, the liability risk differential is maximized if courts decide the case on the basis of the preponderance of evidence standard of proof. Assigning the burden of proof to the plaintiff (resp. the defendant) and using a standard of proof stronger than preponderance would yield a type I error smaller (resp. larger) than α_h^E .

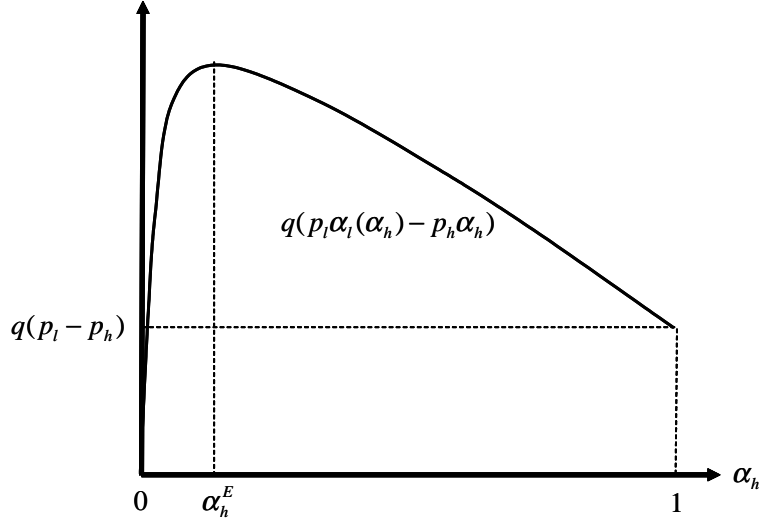


Fig. 6. Liability Risk Differential

5 Compliance Maximizing Negligence Rules

We now consider how to design the rules of proof in order to maximize incentives to comply with due care. From the preceding section, a liability regime can be summarized by the type I error α_h in rulings of negligence. Bad news and good news are defined as before.

Replicating the approach in Section 3, it is easily seen that a non prosocial individual with cost of care c now exerts high care if and only if

$$c \leq \min\{q(p_l \alpha_l(\alpha_h) - p_h \alpha_h)[\min(w, L) + \beta \Delta], c^*\} \equiv c_N(\Delta, \alpha_h) \quad (8)$$

The interpretation is the same as for the condition (4) of Section 3. The critical cost below which an individual exerts high care is now written as a function of the rules of proof.⁹ Similarly, the compliance rate now satisfies

$$y = \psi_N(\Delta, \alpha_h) \equiv \frac{G(c_N(\Delta, \alpha_h))}{G(c^*)}, \quad (9)$$

⁹When the evidence about the defendant's precautions is perfectly informative, $\alpha_h = 0$ and $\alpha_l = 1$. The condition (8) then reduces to (4).

As before, the function is strictly increasing in the reputational penalty (as long as $y < 1$).

Clearly, when the non prosocial have no social image concerns (i.e., $\beta = 0$), $\psi_N(\Delta, \alpha_h)$ does not depend on the reputational penalty and therefore compliance is maximized by maximizing the liability risk differential, thus by setting $\alpha_h = \alpha_h^E$.¹⁰ When social image matters, however, there is an additional consideration because the legal regime will also affect the reputational penalty.

Lemma 4 *Under the negligence regime with type I error α_h , the reputational penalty satisfies $\Delta = \varphi_N(y, \alpha_h)$ where*

$$\varphi_N(y, \alpha_h) \equiv \frac{\lambda[1 - G(c^*)qp_h\alpha_h]}{1 - G(c^*)q[(\lambda + (1 - \lambda)y)p_h\alpha_h + (1 - \lambda)(1 - y)p_l\alpha_l(\alpha_h)]} - \frac{\lambda p_h \alpha_h}{(\lambda + (1 - \lambda)y)p_h\alpha_h + (1 - \lambda)(1 - y)p_l\alpha_l(\alpha_h)}.$$

The function is decreasing in y with $\varphi_N(1, \alpha_h) = 0$.

For a given negligence regime α_h , an equilibrium is a solution to $y = \psi_N(\Delta, \alpha_h)$ and $\Delta = \varphi_N(y, \alpha_h)$. As before, the solution is unique. We denote the equilibrium by $E_N(\alpha_h) = (y_N(\alpha_h), \Delta_N(\alpha_h))$. Choosing the best liability regime from a deterrence point of view requires maximizing $y_N(\alpha_h)$ with respect to its argument. Let us first define

$$\pi(\alpha_h, y) \equiv G(c^*)q[(\lambda + (1 - \lambda)y)p_h\alpha_h + (1 - \lambda)(1 - y)p_l\alpha_l(\alpha_h)]$$

The expression is the proportion of individuals found liable under the negligence regime α_h when the rate of compliance is y . At equilibrium under the regime α_h , the proportion of individuals found negligent is $\pi(\alpha_h, y_N(\alpha_h))$.

¹⁰The rate of compliance would be smaller than under perfect evidence because $q[p_l\alpha_l(\alpha_h^E) - p_h\alpha_h^E] < qp_l$, where the right-hand side is the liability risk differential without judicial error. As is well known, judicial error reduces incentives (see Polinsky and Shavell 1989).

We will say that finding negligence is a rare event if the proportion of individuals found negligent is less than one half; conversely, it is a frequent event if the proportion is greater than one half. We can now state the following.

Proposition 2 *Suppose the negligence regime maximizes compliance with due care. Then the plaintiff bears the burden of proving negligence (resp. the defendant bears the burden of proving compliance with due care) if finding negligence is a rare (resp. frequent) event. In either case the standard of proof is stronger than preponderance of evidence.*

The intuition is a simple one. Suppose α_h^* is compliance maximizing. Consider a marginal increase in the type I error; that is, it now becomes easier for the plaintiff to prove the defendant's negligence or it becomes more difficult for the defendant to prove that he complied with due care. Suppose this shifts the ψ_N curve to the right in the neighborhood of the equilibrium. Observe that this can arise only when α_h^* is below α_h^E , the evidentiary threshold under the preponderance of evidence standard. In other words, the compliance maximizing regime is characterized by the plaintiff bearing the burden of proof and by a standard of proof stronger than preponderance of evidence. If at the same time the φ_N curve shifts upwards in a neighborhood of the equilibrium, then α_h^* cannot have been compliance maximizing. Thus, it must be that the φ_N curve shifts downwards, i.e., the optimal regime trades-off the effects on the liability risk differential and on the reputational penalty. More generally, at a compliance maximizing regime, a small change in the type I error must have "opposite" effects on the ψ_N and φ_N curves. As shown in the Appendix, whether the φ_N curve shifts upwards or downwards depends on the frequency of negligence rulings.

Corollary 1 *If $qG(c^*)(\lambda p_h + (1 - \lambda)p_l) \leq 1/2$, maximizing compliance requires the plaintiff to bear the burden of proving the defendant's negligence and to do so to a standard greater than preponderance of evidence.*

The corollary provides a straightforward sufficient condition. The expression in the corollary is an upper bound for the proportion of individuals found negligent under any negligence regime. Hence the condition follows trivially from Proposition 2. As a particular case, the condition is satisfied when $p_l \leq 1/2$. Thus, when the occurrence of harm seldom arises even under low care, the plaintiff should be assigned the burden of proof.

Sufficient conditions for the injurer to bear the burden of proof are not as straightforward to characterize. We describe one possible case. Suppose $q = 1$ so that underdeterrence is solely due to the injurers' inability to pay fully for the harm done. Suppose further that $c^* \geq \bar{c}$ so that $G(c^*) = 1$, i.e., high care is the due care standard in all circumstances. Let \hat{y}_N be the equilibrium compliance rate when the evidence is perfectly informative as in Section 3. The corresponding proportion of individuals found negligent is

$$\hat{\pi}_N = (1 - \lambda)(1 - \hat{y}_N)p_l.$$

There will be cases where $\hat{\pi}_N > 1/2$. In such cases the finding of negligence is a frequent event because most injurers are not prosocial, few of them are induced to comply with due care even with perfect evidence, and accidents occur often when inadequate care is taken. Consider next the negligence rule with imperfectly informative evidence. Under the compliance maximizing regime α_h^* , the proportion of individuals found negligent is

$$\pi_N(\alpha_h^*, y_N(\alpha_h^*)) = (\lambda + (1 - \lambda)y_N(\alpha_h^*))p_h\alpha_h^* + (1 - \lambda)(1 - y_N(\alpha_h^*))p_l\alpha_l(\alpha_h^*).$$

It is easy to produce examples where this probability is larger than $\hat{\pi}_N$.

6 Concluding Remarks

Liability rulings do not have the same “social meaning” under strict liability as under the negligence rule. Under either rule, the meaning also differs depending on the proportion of virtuous individuals in the population of

potential injurers and the extent to which formal legal sanctions underdeter. When assessing a defendant's level of care is subject to error, the meaning of a finding of negligence will also depend on the risk of type I and type II judicial errors and therefore on the rules of proof.

In most situations, accidental harm and legal suits will be rare events. A regime that seeks to maximize compliance with due care should then make it relatively difficult to find negligence. This is achieved by assigning to the plaintiff the burden of proving the defendant's negligence and imposing a standard of proof stronger than preponderance of evidence. The intuition is that, when suits are rare events, not finding negligence is banal, i.e., posterior beliefs then do not differ much from the prior. By contrast, a finding of negligence yields substantial disesteem. Making it harder still to find negligence increases the stigma and therefore the incentives to exert care. On the other hand, when the occurrence of harm and suits are frequent events, not being found negligent may provide significant prestige. The reputational gain — hence the incentives to comply with the due care — can be increased by making it relatively difficult to escape liability. The best regime is then one that imposes on the defendant the burden of proving that he complied with due care and to do so to a standard stronger than preponderance of evidence.

Our results are reminiscent of Bénabou and Tirole's (2011) discussion of how acceptable behavior arises from the interplay of "honor" and "stigma". High stigma is attached to a behavior that "is just not done", i.e., only the worst type will do it. Alternatively, when "everyone does it", the same behavior carries little stigma. But then "not doing it" yields prestige. In the case of trial outcomes under the negligence rule, whether the finding of negligence imposes significant "stigma" or whether not finding negligence confers significant "honor" depends on the underlying situation, but to some extent can also be influenced by the liability regime for the purpose of increasing incentives to comply with due care.

Appendix

To be completed.

References

- [1] Andreoni, J. and B.D. Bernheim (2009), “Social Image and the 50-50 Norm,” *Econometrica* 77, 1607-1636.
- [2] Ariely, D., A. Bracha and S. Meier (2009), “Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially,” *American Economic Review* 99, 544-555.
- [3] Bénabou, R. and J. Tirole (2006), “Incentives and Prosocial Behavior,” *American Economic Review* 96, 1652-1678.
- [4] Bénabou, R. and J. Tirole (2011), “Laws and Norms,” NBER wp 17579.
- [5] Bernheim, B.D. (1994), “A Theory of Conformity,” *Journal of Political Economy*, 102, 905-953.
- [6] Brekke, K.A., Kverndokk, S. and K. Nyborg (2003), “An Economic Model of Moral Motivation,” *Journal of Public Economics* 87, 1967-1983.
- [7] Cho, I. K. and D. Kreps (1987), “Signaling games and stable equilibria,” *Quarterly Journal of Economics* 102, 179-221.
- [8] Dana, J., D.M. Cain and R. Dawes (2006), “What You Don’t Know Won’t Hurt Me: Costly (but Quiet) Exit in Dictator Games,” *Organizational Behavior and Human Decision Processes* 100, 193-201.
- [9] Daughety, A. and J. Reinganum (2010), “Public Goods, Social Pressure, and the Choice Between Privacy and Publicity,” *American Economic Journal: Microeconomics* 2, 191-222.

- [10] Deffains, B. and C. Fluet (2013), “Legal Liability when Individuals Have Moral Concerns”, *Journal of Law, Economics, and Organization*, forthcoming.
- [11] Demougin, D. and C. Fluet (2005), “Deterrence versus judicial error: a comparative view of standards of proof”, *Journal of Institutional and Theoretical Economics* 161(2), 193-206.
- [12] Demougin, D. and C. Fluet (2006), “Preponderance of evidence”, *European Economic Review* 50, 963-976.
- [13] Demougin, D. and C. Fluet (2008), “Rules of proof, courts, and incentives”, *RAND Journal of Economics* 39, 20-40.
- [14] Ellingsen, T. and M. Johannesson (2008), “Pride and Prejudice: The Human Side of Incentive Theory,” *American Economic Review* 98, 990-1008.
- [15] Frey, B. and R. Jegen (2001), “Motivation Crowding Out Theory,” *Journal of Economic Surveys* 15, 589-611.
- [16] Funk, P. (2010), “Social Incentives and Voter Turnout: Theory and Evidence,” *Journal of the European Economic Association* 8, 1077-1103.
- [17] Harel, A. and A. Klement (2007), “The Economics of Stigma: Why More Detection of Crime May Result in Less Stigmatization,” *Journal of Legal Studies* 36, 355-378.
- [18] Hay, B.L. and K.E. Spier (1997), “Burdens of proof in civil litigation: An economic perspective,” *Journal of Legal Studies* 26, 413–431.
- [19] Lacetera, N. and M. Macis (2010), “Social Image Concerns and Prosocial Behavior: Field Evidence from a Nonlinear Incentive Scheme,” *Journal of Law, Economics, and Organization* 76, 225-237.

- [20] McAdams, R.H. and E. B. Rasmusen (2007), “Norms in Law and Economics”, in *Handbook of Law and Economics*, M. Polinsky and S. Shavell eds., Elsevier, ch. 20, 1573-1618.
- [21] Polinsky, M.A. and S. Shavell, (1989), “Legal error, litigation, and the incentive to obey the law,” *Journal of Law, Economics and Organization* 5, 99–108.
- [22] Rasmusen, E. (1996), “Stigma and Self-Fulfilling Expectations of Criminality,” *Journal of Law and Economics* 39, 519-544.
- [23] Schauer, F. and Zeckhauser, R. (1996), “On the degree of confidence for adverse decisions,” *Journal of Legal Studies* 25, 27–52.
- [24] Shavell, S. (1984), “A Model of the Optimal Use of Liability and Safety Regulation,” *RAND Journal of Economics* 15, 271-280.
- [25] Shavell, S. (2002), “Law versus Morality as Regulators of Conduct,” *American Law and Economics Review* 4, 227-257.
- [26] Zasu, Y. (2007), “Sanctions by Social Norms and the Law: Substitutes or Complements?” *Journal of Legal Studies* 36, 379-396.