The Role of Social Interactions in the Design of Legal Regimes

Bruno Deffains* Claude Fluet†

September 2012

Abstract

We consider situations where, due to imperfect enforcement, tort law yields insufficient incentives to prevent harm. Some injurers nevertheless exert socially efficient care because of intrinsic prosocial concerns. Others have no such concerns but would like people to believe that they do. We show that negligence is more effective than strict liability at transforming esteem concerns into incentives to exert care. This extends to the case where courts may err in assessing the defendants' behavior. The regime of proof then affects the inferences drawn from the finding of negligence and can therefore be chosen to make the best use of reputational incentives. When suits are rare, the plaintiff should bear the burden of proving negligence; when they are frequent, the defendant should bear the burden of proving compliance with due care. In either case, incentives to comply are maximized by a standard of proof stronger than 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

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 the victim — 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 what role it may play in a civil law context and how it impacts the effectiveness of tort regimes in the unilateral accident model.

We consider situations where the strict liability and negligence rules yield insufficient incentives to take precautions. Injurers are judgment-proof or it is not always feasible for victims to file a valid claim, e.g., they cannot always prove harm or identify the injurer. We assume that some injurers nevertheless exert socially efficient care. They do so out of intrinsic moral or prosocial concerns about causing harm. Other injurers have no such concerns. However, they would like people to believe that they do, i.e., they care about social esteem.[1] In our framework, an individual's actions are not observable by society at large. However, an adverse court judgment represents information publicly available from which inferences can be drawn about the individual's actions and intrinsic predispositions. Under either strict liability or the negligence rule, there will be purely reputational incentives on the part of non prosocial individuals to mimic the virtuous. Our focus is how such incentives differ between tort regimes and how they affect the optimal design of the tort regime when the purpose is to induce the greatest number of individuals to exert socially efficient care.

A basic result its that the negligence rule is more effective than strict liability in harnessing the externality due to esteem concerns. For instance, under the negligence rule, an increase in the number of intrinsically prosocial individuals induces a greater proportion of the non prosocial to exert efficient

[1] See Shavell (2004) for a general discussion of legal sanctions versus informal motivation as regulators of conduct and McAdams and Rasmusen (2005) for a survey on the Law and Economics of norms.

care. This is not necessarily so under strict liability because trial outcomes are then less informative, i.e., an adverse court judgment merely conveys that the defendant caused harm, not that he took inadequate precautions. In order to better characterize the informational role of fault or negligence, we therefore extend the analysis to the case of imperfect evidence about a defendant's precautions.

A complete characterization of the negligence regime must then consider how the risk of judicial error is dealt with. 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. We show that when injurers have reputational incentives, and by contrast with the results in Demougin and Fluet (2005, 2006), compliance with due care is maximized by a standard of proof stronger than the common law preponderance of evidence standard. Roughly speaking, whether the plaintiff or the defendant should bear the burden of proof depends on whether legal suits are frequent or infrequent events.

A recent literature emphasizes the idea that one's actions may signal something about unobservable predispositions and that some predispositions are socially esteemed (Bernheim 1994, Bénabou and Tirole 2006, 2011, and Daughety and Reinganum 2010, among others). In Deffains and Fluet (2011), we apply this approach to a tort situation where potential injurers have both intrinsic moral concerns about causing harm and esteem concerns. One focus of that paper is the extent to which formal legal sanctions crowd-out or crowd-in informal motivations. Whether material penalties or rewards affect informal motivations has been explored in a vast experimental and empirical literature (see Frey and Jegen 2001 for a survey). In the present paper, the emphasis is different. We use a much simpler model to analyze how different legal regimes are effective at using stigma to provide incentives to exert care. In particular, we provide prescriptions about the regime of proof under evidentiary uncertainty.

Section 2 presents the basic setup. Section 3 compares the reputational outcomes and incentives to exert care under strict liability and the negligence

rule with no judicial error. In section 4, we introduce imperfectly informative evidence about the injurers' precautions. In section 5, we derive the implications concerning the design of compliance maximizing legal regimes. Section 6 concludes. Most of the argument is obvious from the main text; details are in the Appendix.

2 The model

We start with a simplified version of the unilateral accident model. Individuals are engaged in a socially valuable activity which may impose an accidental loss L on third parties. The risk depends on the injurer's precautions which are h or l for high and low care respectively, with probabilities of accident $p_l > p_h > 0$. The opportunity cost of high care is c distributed according to the cumulative function $G(c)$, with support $[0, \overline{c}]$ and a continuously differentiable density. The interpretation is that injurers are heterogeneous or an injurer's cost of care depends on the circumstances.

We consider situations where the threat of a legal suit does not always ensure socially efficient precautions. First, victims may not always be able to submit a valid claim. They can prove the occurrence of harm or identify the injurer only with some probability q. Secondly, injurers may be judgment-proof or the legal damages awarded may be capped below the actual amount of harm, as is often the case for non-pecuniary or environmental losses. We denote by D the damages actually paid by an injurer found liable. The injurers' initial wealth is W, hence $D \leq W$. When $W < L$, injurers are partially judgment-proof; when $D < L < W$, legal damages are capped.

Under the strict liability rule, an injurer with cost of care c faces the expected cost $p_l q D$ if he exerts low care and $c + p_h q D$ if he exerts high care. He is therefore better off exerting high care when $c \leq q(p_l - p_h)D$. We will refer to $q(p_l - p_h)$ as the liability risk differential under strict liability. Under the negligence rule with high care as the legal due care standard, the potential injurer faces the cost $p_l q D$ if he exerts low care and incurs only c if he complies with due care. He is better off complying when $c \leq p_l q D$, where $p_l q$ is the liability risk differential under the negligence rule. As is well

known, the negligence rule may induce efficient care even though victims do not always sue or damages paid are less than the amount of harm. We introduce conditions ruling out perfect compliance.

ASSUMPTION 1: $p_l qW < \overline{c} < (p_l - p_h)L$.

The second inequality means that high care is always the socially efficient action. The first inequality implies that some potential injurers will take inadequate precautions even under a negligence rule with damages paid up to the injurers' wealth, i.e., $D = W$. Any combination of q and W satisfying Assumption 1 is sufficient for our purpose. For instance, inefficient care will arise even when injurers are not judgment-proof and punitive damages $D > L$ are awarded.

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 values. They seek to behave in a socially (morally) responsible manner by comparing their opportunity cost of care with the expected harm they impose on others.[2] Alternatively, they could be described as complying with an informal social norm prescribing high care. Such individuals, referred to as type $\theta = 1$, choose the socially efficient level of care irrespective of legal legal sanctions. There is a 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, referred to as type $\theta = 0$, behavior is determined by the utility function $u = w + \beta\overline{\theta}_I$ where w is net final wealth, β is a positive parameter and $\overline{\theta}_I \equiv E(\theta \mid I)$ is society's belief about the individual's type conditional on the information I. Given our definition of types, $\overline{\theta}_I$ is simply the posterior probability that the individual is intrinsically prosocial and $\beta\overline{\theta}_I$ is the utility derived from society's beliefs.

An individual's type is private information and so is his cost of care, his

[2]This is a simple version of Kant's Categorical Imperative. See Brekke et al. (2003) for a discussion.

chosen level of care and whether he caused harm, except insofar as these can be inferred from court judgments. The 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 has been found liable under the prevailing liability rule. Good news is the complementary event. The meaning 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.

3 Strict Liability versus Negligence

Consider first the strict liability rule. As a function of his level of care, the expected utility of a non prosocial individual is

$$u_i = p_i q(W - D + \beta\overline{\theta}_B) + (1 - p_i q)(W - D + \beta\overline{\theta}_G),\ i = h, l,$$

where $\overline{\theta}_B$ and $\overline{\theta}_G$ are the beliefs about his type under bad and good news respectively. These beliefs will be determined at equilibrium and will satisfy $\overline{\theta}_B < \overline{\theta}_G$, i.e., $\overline{\theta}_B$ is the stigma imposed by an adverse court judgment and $\overline{\theta}_G$ is the esteem otherwise earned. The potential injurer exerts high care if $c \leq q(p_l - p_h)(D + \beta\Delta)$ where $\Delta \equiv \overline{\theta}_G - \overline{\theta}_B$ is the reputational penalty from a liability ruling.

Given the penalty, the proportion of non prosocial individuals exerting high care is $y = y_S(\Delta)$, where the subscript S refers to strict liability and

$$y_S(\Delta) \equiv G\left[q(p_l - p_h)(D + \beta\Delta)\right]. \tag{1}$$

For a given y and applying Bayes' rule, the reputational penalty satisfies $\Delta = \Delta_S(y)$ where

$$\Delta_S(y) \equiv \frac{\lambda(1 - p_h q)}{(\lambda + (1 - \lambda)y)(1 - p_h q)) + (1 - \lambda)(1 - y)(1 - p_l q)} - \frac{\lambda p_h q}{(\lambda + (1 - \lambda)y)p_h q + (1 - \lambda)(1 - y)p_l q}. \tag{2}$$

The first term on the right-hand side is the posterior belief following good news, the second term the belief following bad news. An equilibrium under strict liability is a pair (y, Δ) solving (1) and (2).

Under the negligence rule, an individual is not held liable if he complies with due care. Expected utility is then

$$u_i = \begin{cases} W - c + \beta\overline{\theta}_G \text{ if } i = h, \\ p_l q(W - D + \beta\overline{\theta}_B) + (1 - p_l q)(W - D + \beta\overline{\theta}_G) \text{ if } i = l. \end{cases}$$

The individual complies with due care if $c \leq qp_l(D + \beta\Delta)$ where Δ is the reputational penalty from an adverse court judgment.

The proportion of complying non prosocial individuals is $y = y_N(\Delta)$, where N refers to the negligence rule and

$$y_N(\Delta) \equiv G\left[qp_l(D + \beta\Delta)\right]. \tag{3}$$

The penalty satisfies $\Delta = \Delta_N(y)$ where

$$\Delta_N(y) = \frac{\lambda}{(\lambda + (1 - \lambda)y) + (1 - \lambda)(1 - y)(1 - p_l q)} \tag{4}$$

The right-hand side is the posterior belief about an individual who has not been found liable. An individual found liable must have been negligent, revealing that the individual is not the prosocial type, i.e., the belief following bad news is now $\overline{\theta}_B = 0$. An equilibrium under the negligence rule solves 3 and (4).

Proposition 1 *The negligence rule provides greater deterrence than strict liability. Negligence deters optimally if esteem concerns are strong enough; strict liability always underdeters.*

The argument is straightforward. From (1) and (3), the proportion of non prosocial injurers exerting efficient care is increasing in the reputational penalty, with $y_S(\Delta) < y_N(\Delta)$ because the liability risk differential is greater under the negligence rule. From (2) and (4), it is easily seen that under either rule the reputational penalty is decreasing in the proportion of injurers exerting efficient care. Moreover $\Delta_S(y) < \Delta_N(y)$ with $\Delta_S(1) = 0$ and

$\Delta_N(1) = \lambda$. Other things equal, the reputational penalty is smaller under strict liability because a liability ruling provides noisier information than under the negligence rule. Under the latter, bad news reveals that the injurer is non prosocial; under strict liability, this is only more likely so. In particular, when all injurers exert efficient care, bad and good news provide no information at all under strict liability because in either case the updated belief is identical to the prior; under the negligence rule, good news is uninformative ($\overline{\theta}_G = \lambda$) but bad news is perfectly revealing ($\overline{\theta}_B = 0$).[3]

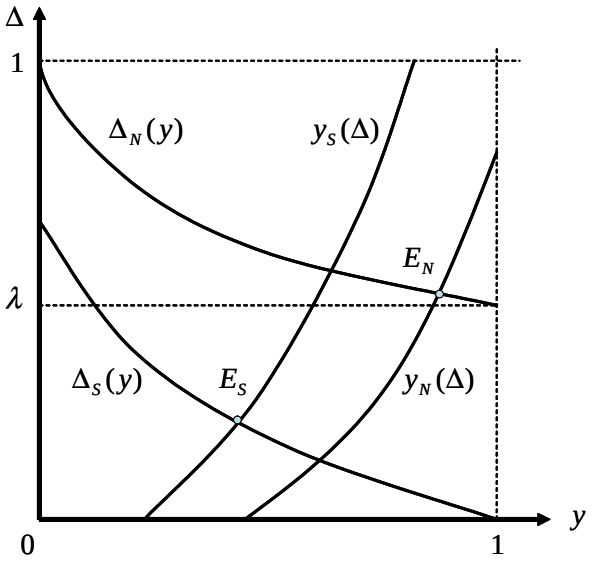

Fig. 1. Equilibria

The functions are as represented in Figure 1. At equilibrium (the intersections E_S and E_N in the figure) there are more injurers exerting efficient care under the negligence rule. Note that the equilibrium reputational penalty need not always be larger. However, if it is smaller, it must be because deterrence is sufficiently worse under strict liability. Figure 2 illustrates a situation where the negligence rule optimally deters. When all injurers comply with due care, the reputational penalty from a ruling of negligence equals λ. All injurers comply if $p_l q(D + \beta\lambda) \geq \overline{c}$, i.e., when β is large enough.

[3]When $y = 1$, bad news is off the equilibrium path under the negligence rule. We obtain the stigma from $\Delta_N(1) \equiv \lim_{y \to 1} \Delta_N(y)$.

Assumption 1 implies that strict liability always underdeters. Suppose this is due not to judgment-proofness but to to the fact that suits are seldom filed. When $W > D = L$, deterrence is increased by the addition of punitive damages. In the Figures 1 and 2 this yields a rightward shift in the Δ_S and Δ_N curves. Under either liability rule, there is then some crowding-out of informal motivation because the equilibrium reputational penalty decreases. This tends to be less important under the negligence rule because the Δ_N curve is bounded below by λ. As a result, punitive damages will tend to have a greater deterrence effect both because the liability risk differential is larger and because the crowding-out effect is smaller.

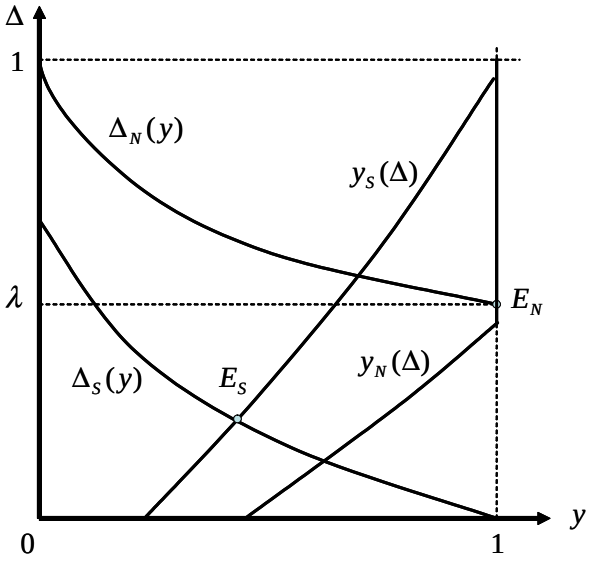

Fig. 2. First Best under the Negligence Rule

Due to esteem concerns, there is an externality between prosocial individuals and the behavior of non prosocial injurers. It is interesting to inquire whether the externality has greater force the greater the number of prosocial individuals. Here again the negligence rule fares better.

Proposition 2 *Under the negligence rule, a greater proportion of prosocial individuals induces more non prosocial individuals to exert efficient care. Under strict liability, the effect is ambiguous: if the proportion of prosocial individuals is sufficiently small (resp. large), more of them increases (resp. reduces) deterrence.*

Under strict liability, a larger λ increases both $\overline{\theta}_B$ and $\overline{\theta}_G$, other things equal. When λ is small, the first effect dominates and the Δ_S curve shifts upwards; when λ is large, the second effect dominates and the Δ_S curve shifts downwards (see the Appendix). Under the negligence rule, $\overline{\theta}_B$ is unaffected (it equals zero) while $\overline{\theta}_G$ increases, hence the Δ_N curve unambiguously shifts upwards. The negligence rule fares better because it induces strategic complementarity: the greater the number of virtuous individuals, the greater the incentives to mimic virtue on the part of the non virtuous. Under strict liability, because a liability ruling is a noisy signal about one's intrinsic predisposition, there may be strategic substitutability. When the number of virtuous is already large, more of them merely reduces the reputational penalty from court rulings, hence the non virtuous have less incentives to avoid causing harm.

4 Judicial Error

The above shows that the negligence rule is more efficient at transforming the externality due to esteem concerns into incentives to exert care. Obviously, the rule has greater informational requirements. We now extend the analysis to the case where the evidence about the injurers' behavior is not perfectly informative. Courts can now make mistakes. 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 determine different legal regimes. We characterize a regime in terms of the prevailing rules of proof, by which me 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 looses the case. The standard of proof refers to the weight of evidence needed to

discharge the burden.[4]

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 level of care. The distributions have continuously differentiable density functions, denoted $f_h(x)$ and $f_l(x)$, and the same support $[\underline{x}, \overline{x}]$, where the bounds need not be finite. The "invariant support" condition means that no realization x perfectly reveals the injurer'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 $f_l(\overline{x})/f_h(\overline{x}) = 0$.

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.[5]

A victim 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, both the victim and the injurer are assumed to have access to the additional evidence x about the injurer's behavior; the complete evidence eventually submitted to the court therefore comprises both the occurrence of harm and the realization x. There are two possibilities:

(i) If, as is usually the case, the plaintiff has the burden of proving the injurer's negligence, he succeeds only if he can submit x such that

$$\frac{p_l f_l(x)}{p_h f_h(x)} > k. \tag{5}$$

[4]This section draws on Demougin and Fluet (2005, 2006).

[5]The 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.

The expression on the left-hand side is the likelihood ratio of low care versus high care given the evidence "occurrence of harm and x". The parameter on the right-hand side is the standard of proof that the plaintiff must satisfy to discharge the burden. Condition (5) states that the plaintiff must adduce evidence showing that inadequate care is k times more likely than due care.[6]

(ii) If the injurer bears the burden of proving that he complied with due care, he will avoid 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.

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 evidence gives greater weight to his contention, however slightly. 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*.[7]

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 3, $\widehat{x}_P(k)$ is the evidentiary threshold when the plaintiff bears the burden of proof. The defendant is found negligent when $x < \widehat{x}_P(k)$. When the defendant bears the burden of proving compliance with due care, the evidentiary threshold is $\widehat{x}_D(k)$. The defendant then escapes liability only if $x > \widehat{x}_D(k)$. For the preponderance of evidence standard, the evidentiary threshold does not depend on the assignment of the burden of proof; it is denoted $\widehat{x}_E$ in the figure.

[6] Observe that court rulings are then purely "evidence based", i.e., independent of the priors the court may hold. In the common law rules of evidence, priors in the form of a "known" proportion of injurers exerting low care would not be considered as admissible evidence. See Demougin and Fluet (2005, 2006) for a discussion.

[7] See Schauer and Zeckhauser (1996).

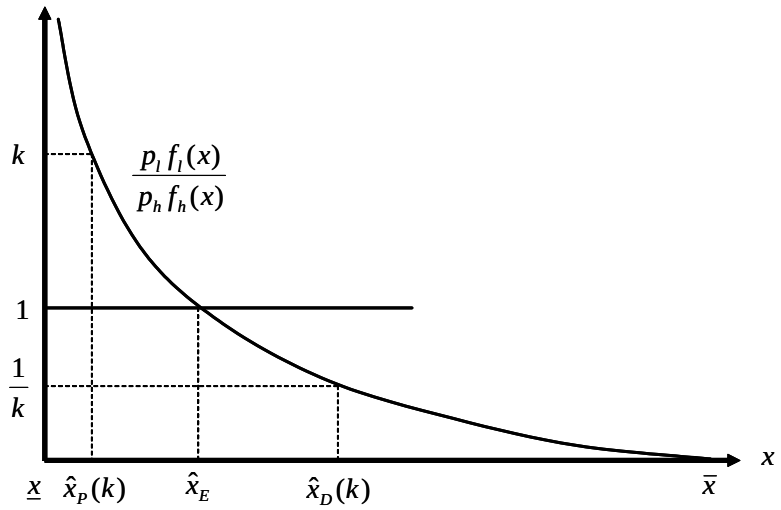

Fig. 3. Likelihood ratios

The liability risk differential. Let $\widehat{x}$ 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 $\alpha_j \equiv F_j(\widehat{x})$ 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 $1 - a_l$. 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. Ex ante, given the possibility of court error, the liability risk differential under the negligence rule is therefore $q(p_l\alpha_l - p_h\alpha_h)$.

It is useful to express α_l as a function of the type I error α_h. The function is $\alpha_l(\alpha_h) \equiv F_l\left(F_l^{-1}(\alpha_h)\right)$ and is increasing and concave:

$$\alpha_l'(\alpha_h) = \frac{f_l(\widehat{x})}{f_h(\widehat{x})} \text{ where } \alpha_h \equiv F_h(\widehat{x}), \tag{6}$$

$$\alpha_l''(\alpha_h) = \frac{d\left(f_l(\widehat{x})/f_h(\widehat{x})\right)}{dx}\left(\frac{1}{f_h(\widehat{x})}\right) < 0,$$

where the sign follows from MLRP. Note that $\alpha_l(0) = 0$ and $\alpha_l(1) = 1$.

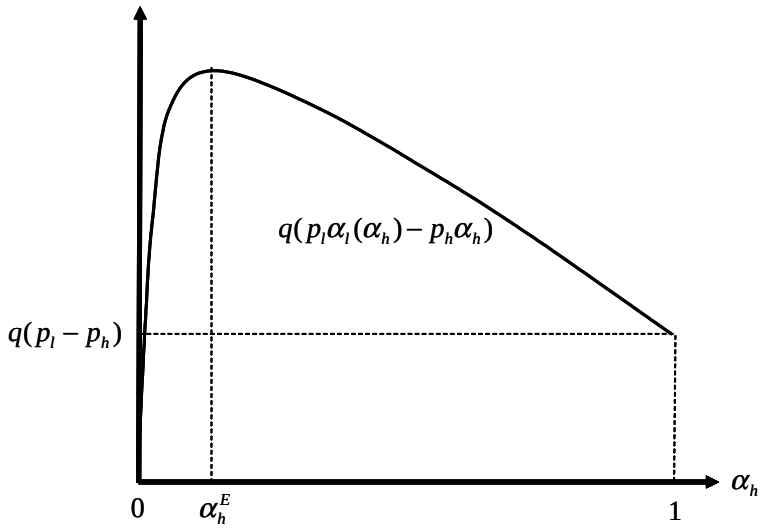

Fig. 4. Liability Risk Differential

The liability risk differential is therefore concave in α_h. Figure 4 provides an illustration. We make three observations. First, Assumption 2 implies that there always exists levels of the type I error such that the differential under the negligence rule is larger than under strict liability.[8] Secondly, the type I error that maximizes the liability risk differential is the one obtained under the preponderance of evidence standard. The first-order condition is $p_l\alpha_l'(\alpha_h) - p_h = 0$. From (6), this is easily seen to require $\widehat{x}_E$ as defined in Figure 3. We denote by α_h^E the corresponding type I error. Thirdly, assigning the burden of proof to the plaintiff (resp. the defendant) and using a standard of proof stronger than preponderance yields a type I error smaller (resp. larger) than α_h^E.

5 Compliance Maximizing Regimes

We now consider the design of the liability regime for the purpose of maximizing incentives to comply with due care. We take D and q as given. When

[8] The slope of the liability risk differential is $p_l\alpha_l'(\alpha_h) - p_h$. At $\alpha_h = 1$, the negligence rule is equivalent to strict liability and the slope of the differential is then negative because $\alpha_l'(1) = f_l(\overline{x})/f_h(\overline{x}) = 0$.

injurers have no image concerns, deterrence is then maximized by maximizing the liability risk differential. This is achieved by the preponderance of evidence standard, the assignment of the burden of proof being irrelevant.[9] When image concerns matter, there is an additional consideration because maximizing the liability risk differential generally does not maximize the liability stigma.

From the results in the preceding section, a liability regime is summarized by the type I error in rulings of negligence. Bad and good news are defined as before. In particular, the general public is only informed of adverse trial outcomes, not of the detailed evidence submitted at trial. In other words, the general public has no time for details. Replicating the analysis of Section 3, we now have

$$y_N(\Delta, \alpha_h) \equiv G\left[q(p_l\alpha_l(\alpha_h) - p_h\alpha_h)(D + \beta\Delta)\right] \tag{7}$$

and

$$\Delta_N(y, \alpha_h) \equiv \frac{\lambda(1 - qp_h\alpha_h)}{1 - \pi(y, \alpha_h)} - \frac{\lambda q p_h \alpha_h}{\pi(y, \alpha_h)}. \tag{8}$$

where

$$\pi(y, \alpha_h) \equiv (\lambda + (1 - \lambda)y)qp_h\alpha_h + (1 - \lambda)(1 - y)qp_l\alpha_l(\alpha_h) \tag{9}$$

is the probability of finding negligence.

As in section 3, $y_N(\Delta, \alpha_h)$ is increasing in the reputational penalty while $\Delta_N(y, \alpha_h)$ is decreasing in the rate of compliance with due care. When $\alpha_h = 1$, the functions are the same as under strict liability. When $\alpha_h = 0$, we have $y_N(\Delta, 0) = 0$ for all values of Δ because the liability risk differential then vanishes; moreover, $\Delta_N(y, 0) = \lambda$ for all y. When the probability of a "false positive" is arbitrarily small, escaping liability reveals nothing, i.e., the first term in (8) is equal to the prior λ; by contrast, being found negligent is perfectly revealing, i.e., the second term equals zero.[10]

Solving the above equations yields the equilibrium under the given the liability regime α_h. Denote the equilibrium values by $\Delta_0(\alpha_h)$ and $y_0(\alpha_h)$, and

[9] The latter no longer holds when the parties do not always have access to the same evidence ex post. See Demougin and Fluet (2008).

[10] The result follows from Assumption 2 which implies $\alpha_l(\alpha_h)/\alpha_h \to \infty$ as $\alpha_h \to 0$.

let $\pi_0(\alpha_h)$ be the probability of finding negligence at equilibrium. Choosing the best liability regime from a deterrence point of view requires maximizing $y(\alpha_h)$ with respect to its argument. We will say that, under the given liability regime, finding negligence is a rare event if $\pi_0(\alpha_h) < 1/2$; conversely, it is a frequent event if $\pi_0(\alpha_h) > 1/2$ We can now state the following.

Proposition 3 *Suppose the liability regime maximizes deterrence. 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 deterrence maximizing. Consider a marginal increase in α_h. Say it shifts the y_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 legal regime α_h^* is then 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 Δ_N curve shifts upwards following a marginal increase in α_h, then α_h^* cannot have been deterrence maximizing. Thus, it must be that the Δ_N curve shifts downwards, i.e., the optimal regime trades-off the effects on the liability risk differential and the reputational penalty. More generally, at a deterrence maximizing regime, the effect of a marginal change in the type I error on the y_N and Δ_N curves must be of opposite signs. Now, as shown in the Appendix, whether the Δ_N curve shifts upwards or downwards depends on the frequency of negligence rulings.

Corollary 1 *If $q(\lambda p_h + (1-\lambda)p_l) < 1/2$, maximizing deterrence 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. When the occurrence of harm together with the feasibility of proving its occurrence is

relatively infrequent, the plaintiff should be assigned the burden of proof. The claim follows trivially from Proposition 3 noting that

$$\pi_0(\alpha_h) < q(\lambda p_h + (1-\lambda)p_l)$$

under any negligence regime.

Sufficient conditions for the injurer to bear the burden of proof are not as straightforward to characterize. We provide one possible illustration. Suppose $q = 1$ so that underdeterrence is due to the injurers' judgment-proofness. Let $\widehat{y}_N$ be the proportion of non prosocial injurers complying with due care when the evidence is perfectly informative, as in Section 3. Denote the corresponding probability of finding negligence by $\widehat{\pi}_N$ and suppose

$$\widehat{\pi}_N = (1-\lambda)(1-\widehat{y}_N)p_l > \tfrac{1}{2}. \tag{10}$$

In this situation, the finding of negligence is a frequent event because most injurers are not prosocial, few of them are induced to comply with due care, and accidents occur often when inadequate care is taken. Consider next the negligence rule with imperfectly informative evidence. The probability of finding negligence is

$$\pi_0(\alpha_h) = (\lambda + (1-\lambda)y_0(\alpha_h))p_h\alpha_h + (1-\lambda)(1-y_0(\alpha_h))p_l\alpha_l(\alpha_h). \tag{11}$$

Obviously, $y_0(\alpha_h) < \widehat{y}_N$ for any type I error.

Now suppose that p_h is quite small compared to p_l. It is easily verified that the preponderance of evidence threshold α_h^E will then be relatively large[11]; $\alpha_l(\alpha_h^E)$ will be larger still, perhaps not too far from unity if the evidence is reasonably informative. Given (10) and the fact that $y_0(\alpha_h^E) < \widehat{y}_N$, there will therefore be situations where $\pi_0(\alpha_h^E)$ is greater than one half. It follows that $\pi_0(\alpha_h)$ will be greater than one half for all $\alpha_h > \alpha_h^E$; by continuity, this will also be true for some $\alpha_h < \alpha_h^E$ in a neighborhood of the preponderance of evidence threshold. Thus, there will be situations where, at the optimal threshold, $\pi_0(\alpha_h^*) > 1/2$. Observe that, with p_h very small, a relatively large α_h^* is not inconsistent with $p_h\alpha_h^*$ being very small, i.e., a compliant

[11] The threshold satisfies $\alpha_l'(\alpha_h^E) = p_h/p_l$. Given Assumption 2, $\alpha_h^E \to 1$ when $p_h \to 0$.

injurer is rarely found negligent. In the situation just described, the mere occurrence of harm is by itself relatively strong evidence that inadequate care was taken. Making it difficult to escape a negligence ruling then increases the reputational penalty.

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 differ 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 error and on the regime 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 esteem. 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 manipulated by the liability regime for the purpose of increasing incentives to comply with due care.

Appendix

Proof of Proposition 2. We complete the argument in the text for the case of strict liability. (2) can be rewritten as

$$\Delta_S(y) = \frac{\lambda(1-\lambda)(1-y)q(p_l - p_h)}{\pi(y)(1-\pi(y))}$$

where

$$\pi(y) \equiv (\lambda + (1-\lambda)y)p_h q + (1-\lambda)(1-y)p_l q$$

is the probability that an injurer will be held liable. It follows that

$$\frac{\partial \Delta_S(y)}{\partial \lambda} = \frac{\pi(y)(1-\pi(y))(1-2\lambda) + \lambda(1-\lambda)(1-\pi(y))(d\pi(y)/d\lambda)}{\pi(y)^2(1-\pi(y))^2}.$$

When λ is close to zero, the second term in the numerator is negligible and the first term is positive. When λ is close to unity, the second term in the numerator is again negligible but the first term is negative. QED

Proof of Proposition 3. Equation (8) can be rewritten as

$$\Delta_N(y, \alpha_h) = \frac{\lambda(1-\lambda)(1-y)q(p_l\alpha_l(\alpha_h) - p_h\alpha_h)}{\pi(y,\alpha_h)(1-\pi(y,\alpha_h))}.$$

Hence

$$\begin{aligned}\frac{\partial \Delta_N(y,\alpha_h)}{\partial \alpha_h} &= \frac{\lambda(1-\lambda)(1-y)q(p_l\alpha_l'(\alpha_h) - p_h)}{\pi(y,\alpha_h)(1-\pi(y,\alpha_h))} \\ &\quad - \frac{\lambda(1-\lambda)(1-y)q(p_l\alpha_l(\alpha_h) - p_h\alpha_h)}{\pi(y,\alpha_h)^2(1-\pi(y,\alpha_h))^2}(1-2\pi(y,\alpha_h))\pi_{\alpha_h}(y,\alpha_h).\end{aligned}$$

Moreover

$$\frac{\partial y_N(\Delta, \alpha_h)}{\partial \alpha_h} = G' \left[q(p_l \alpha_l(\alpha_h) - p_h \alpha_h)(D + \beta\Delta) \right] q(D + \beta\Delta)(p_l \alpha_l'(\alpha_h) - p_h).$$

From the argument in the text,

$$\left(\left. \frac{\partial \Delta_N(y_0(\alpha_h^*), \alpha_h)}{\partial \alpha_h} \right|_{\alpha_h = \alpha_h^*} \right) \left(\left. \frac{\partial y_N(\Delta_0(\alpha_h^*), \alpha_h)}{\partial \alpha_h} \right|_{\alpha_h = \alpha_h^*} \right) \leq 0.$$

Therefore

$$\left. \frac{\partial y_N(\Delta_0(\alpha_h^*), \alpha_h)}{\partial \alpha_h} \right|_{\alpha_h = \alpha_h^*} \gtreqless 0 \text{ if } \pi_0(y_0(\alpha_h^*), \alpha_h^*) \lesseqgtr \tfrac{1}{2}.$$

Recalling that

$$\frac{\partial y_N(\Delta, \alpha_h)}{\partial \alpha_h} \gtreqless 0 \text{ if } \alpha_h \lesseqgtr \alpha_h^E$$

then completes the proof. QED

References

[1] Bénabou, R. and J. Tirole (2006), "Incentives and Prosocial Behavior," *American Economic Review* 96, 1652-1678.

[2] Bénabou, R. and J. Tirole (2011), "Laws and Norms," NBER wp 17579.

[3] Bernheim, B.D. (1994), "A Theory of Conformity," *Journal of Political Economy*, 102, 905-953.

[4] Brekke, K.A., Kverndokk, S. and K. Nyborg (2003), "An Economic Model of Moral Motivation," *Journal of Public Economics* 87, 1967-1983.

[5] Daughety, A. and J. Reinganum (2010), "Public Goods, Social Pressure, and the Choice Between Privacy and Publicity," *American Economic Journal: Microeconomics* 2, 191-222.

[6] Deffains, B. and C. Fluet (2012), "Legal Liability when Individuals Have Moral Concerns", *Journal of Law, Economics, and Organization*, forthcoming (doi: 10.1093/jleo/ews002).

[7] 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.

[8] Demougin, D. and C. Fluet (2006), "Preponderance of evidence", *European Economic Review* 50, 963-976.

[9] Demougin, D. and C. Fluet (2008), "Rules of proof, courts, and incentives", *RAND Journal of Economics* 39, 20-40.

[10] Frey, B. and R. Jegen (2001), "Motivation Crowding Out Theory," *Journal of Economic Surveys* 15, 589-611.

[11] 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.

[12] 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.

[13] Rasmusen, E. (1996), "Stigma and Self-Fulfilling Expectations of Criminality," *Journal of Law and Economics* 39, 519-544.

[14] Schauer, F. and Zeckhauser, R. (1996), "On the degree of confidence for adverse decisions," *Journal of Legal Studies* 25, 27–52.

[15] Shavell, S. (2002), "Law versus Morality as Regulators of Conduct," *American Law and Economics Review* 4, 227-257.

[16] Zasu, Y. (2007), "Sanctions by Social Norms and the Law: Substitutes or Complements?" *Journal of Legal Studies* 36, 379-396.