

The Signal-Tuning Function of Liability Regimes

Claude Fluet*
Université Laval
CRREP, CRED

Murat C. Mungan†
George Mason University,
Antonin Scalia Law School

August 2017

Abstract

Fault-based liability regimes require an inquiry into the nature of the defendant's conduct, whereas this type of inquiry is absent in strict liability regimes. Therefore, verdicts reached through fault-based liability regimes can convey superior information compared to verdicts reached through strict liability regimes. Further reflection reveals that this advantage is enjoyed by fault-based liability regimes only if the evidence related to the nature of defendants' actions is sufficiently informative. Otherwise, admitting such evidence can add noise to the information conveyed through verdicts. Therefore, liability regimes have a function of tuning signals conveyed on to third parties, which, in turn, causes deterrence effects by affecting the informal sanctions imposed on defendants who are found liable. We construct a model wherein this function is formalized, and we identify the optimal liability regime and burden of proof as a function of various factors (e.g. the commonality of the harmful act, and the informativeness of the evidence).

Keywords: Informal sanctions, reputational sanctions, fault-based liability, strict liability, burden of proof.

*Email: claudio.fluet@fsa.ulaval.ca.

†Email: mmungan@gmu.edu

1 Introduction

Many legal doctrines call for an inquiry into the particular circumstances in which a person's acts gave rise to harm. To do this, courts often draw distinctions between wrongful and justifiable acts. Examples abound. Criminal law inquires into the mental state of offenders through its *mens rea* requirements; tort law often asks whether a person's conduct was negligent; and agency law makes distinctions between good faith and bad faith conduct. For most people, this distinction makes a lot of intuitive sense: society should punish people only when they are blameworthy. However, for many economists, this conclusion is not always warranted from an instrumentalist perspective. In particular, if courts are not well-equipped (or do not have enough information) to accurately draw such distinctions, they may frequently make judgement errors, which, in turn, may negatively distort people's incentives. Moreover, there are important exceptions to this legal approach, including the use of strict liability.

In this article, we highlight the *signal-tuning function* of courts, which provides a rationale for why it may be useful to use different liability rules in different contexts. The primary idea is that the choice of liability rule affects the type of information available to third parties, who may use this information to adjust their beliefs about people who have been found liable. Whether inquiring into the nature of the act that led to harm leads to superior information for third parties depends on many factors, including, the court's accuracy in making such determinations; how much third parties care about whether the actor was in fact blameworthy; and the dangerousness of the act. To explain the dynamics that drive these results, we focus on the frequently studied distinction between fault-based liability regimes and strict liability regimes.

The comparative advantages of strict liability versus fault-based liability has been extensively studied in the law and economics literature.¹ In this literature, strict liability refers to instances in which an actor is found liable whenever his action causes harm. On the other hand, a rule is fault-based, if "a party who has been found to have caused harm is sanctioned only if he failed to obey some standard of behavior or regulatory requirement" (Polinsky and Shavell (2007 p. 407)). Examples of both types of liability

¹Polinsky and Shavell (2007), for instance, reviews the comparative advantages of these regimes in the public law enforcement context. In the torts context, there is a very large literature, and a relatively recent review of the literature can be found in Mueller-Langer and Schäfer (2009).

are abundant in the public law enforcement context (which includes the enforcement of criminal laws, civil infractions, and governmental regulations) as well as in the torts context where simple negligence regimes resemble fault-based liability. Another, seemingly independent, literature investigates the reputational costs² one incurs as a result of being held liable.³ These costs can take the form of a reduction in stock prices for a corporation⁴ as well as reduced job prospects for individuals who obtain criminal records.⁵

Studying the intersection of these two strands of the literature reveals how liability regimes impact the magnitude of reputational costs. This allows the identification of meaningful factors that affect which liability regime maximizes reputational costs, such as the dangerousness and the commonality of the act, as well as how accurately courts can ascertain whether a person exercised care while engaging in the activity. The analysis reveals, consistent with conventional wisdom and prior work, that the case for using fault-based liability is stronger when courts perform well in deducing whether defendants took care, and when the act leads to harm with low probability even absent care.⁶ On the other hand, as we show in the present paper and perhaps counter-intuitively, the case for using strict liability is strengthened when the act is common rather than uncommon.

This last result has interesting implications. First, as a general matter, that the social desirability of strict versus fault-based liability depends on the commonality of the regulated act is a deviation from some of the important findings in the prior literature.⁷ Second, and relatedly, the same

²These costs are analyzed in both the public and private law enforcement contexts. The similarities between these costs in the two contexts are explained in Mungan (2016), which also provides a brief review of this literature.

³Although most economic analyses focus on either reputational sanctions or the comparison of negligence and strict liability regimes, an exception, which focuses on this intersection is Deffains and Fluet (2013). However, Deffains and Fluet (2013) excludes the essential ingredients of our analysis, e.g. erroneous determinations of liability, which allows us to identify key factors like the commonality of the regulated act that mediate the comparative advantages of negligence versus strict liability regimes.

⁴See, e.g., Helland (2006) and Karpoff et al. (2008).

⁵Many empirical studies provide support for this statement. Pager (2003) and Pager et al. (2009), for instance, compare the job application outcomes of people who have criminal records and who do not, and Lott (1992a) and (1992b) provide estimates of the size of informal sanctions. On the other hand, Murray (2016) and the sources cited therein are examples of legal scholarship which argue that a conviction has a negative impact on one's success in the labor market.

⁶See, e.g., Demougin and Fluet (2005) and (2006).

⁷Demougin and Fluet (2006), for instance, considers a model where there are no reputational sanctions, and concludes that the deterrence maximizing decision rule is independent

dynamics that drive this result also imply that the optimal burden of proof in determining liability depends on the commonality of the act. Third, the result suggests that a prevailing approach in United States tort law, which requires an act to be *uncommon* in addition to being dangerous for the imposition of strict liability, may be in conflict with the objective of optimal deterrence.⁸

The above results pertain to the question of which liability regime maximizes deterrence. Additionally, the analysis reveals important insights about which party ought to carry the burden of proof in establishing whether or not the defendant took care. This question naturally arises only under a regime of fault-based liability, because under strict liability, the person is found liable regardless of whether he took care. Our analysis reveals that, it is optimal to allocate the burden of proof on to the plaintiff, unless liability is frequent in equilibrium, and the court is relatively skilled in determining whether a defendant took care.

A point worth highlighting is that the above discussion implicitly assumes that increasing reputational costs is a desirable goal. One may naturally question this assumption, which we maintain in our formal analysis. This assumption is defensible in a variety of circumstances where formal sanctions lead to under-deterrence. As the existing literature illustrates, under-deterrence often becomes a problem in both the public and private enforcement contexts. In public enforcement, a well known result is the Beckerian maximum fine result, which, in turn, generates under-deterrence via low probabilities of detection (Becker 1968). In the standard torts context, under-deterrence occurs when wrong-doing does not lead to liability with certainty, since damages are capped at the compensatory level. A similar result is obtained in both the public enforcement as well as the torts context when defendants are judgement proof. Due to these, and other reasons explored in the literature, it becomes socially desirable to increase the level of deterrence, and large reputational sanctions can achieve this result.

Moreover, even though our formal analysis is geared towards identifying which liability regime maximizes deterrence, the broader implications of our observations are relevant even when the objective might be different. In particular, our analysis highlights that the liability regime has a direct impact on deterrence through its effect on the magnitude of reputational sanctions. Ignoring these effects can lead to incorrect conjectures regarding

of the frequency of the act.

⁸ See, e.g., Restatement (Third) of Torts: Phys. & Emot. Harm § 20 (2010).

the normative desirability of a particular liability regime. The importance of this omitted consideration depends, of course, on the ratio of reputational sanctions to formal sanctions. One need only look at criminal law to observe that this ratio can be quite high in many contexts. Thus, interactions between liability regimes and reputational sanctions ought to be considered in thinking about the normative desirability of various liability regimes.

In the next section we intuitively, and with minimal resort to mathematical concepts, explain the dynamics behind our analysis. Later, in section 3, we introduce a model which we use to conduct our formal analysis. In section 4, we discuss the implications of our analysis, and conclude in section 5. An appendix, in the end, contains proofs of all lemmas and propositions.

2 Intuitive Explanation

That individuals are stigmatized upon being convicted of a crime is a well documented phenomenon.⁹ These individuals have a harder time finding jobs in addition to facing difficulties in forming social relationships.¹⁰ Similarly, in the commercial context, firms often suffer great reputational losses when they are found liable, in addition to the formal sanctions they incur.¹¹ A plausible explanation for the emergence of these reputational sanctions is that third parties use a person's or a firm's liability as a (noisy) signal of its underlying characteristics. Moreover, as noted in the literature, third parties can make inferences based not only on whether the actor was found liable, but, also based on the severity of the applicable sanctions.¹² If a person commits a crime despite the existence of an extreme sanction, this signals to third parties that the person has nothing to lose, or that he has very high criminal tendencies. Thus, higher sanctions can signal lower 'quality'. A similar inference can be drawn in commercial contexts: if a firm fails to comply with regulations despite the existence of large sanctions, it presumably signals a type of internal organizational failure; high compliance costs due to technological inferiority; or a disregard for long term costs. In

⁹ See, note 4, above.

¹⁰ The references cited in note 4, above, provide support for this proposition. These effects are also described in Rasmusen (1996), which formalizes stigma in a law enforcement model.

¹¹ See, note 3, above, and the references cited therein.

¹² Iacobucci (2014) explains how this type of stigmatization can occur. Mungan (2016) generalizes this idea and provides a model where the expected sanction provides information regarding individuals' quality.

all cases, a consumer (or potential investors) may draw negative inferences regarding the firm. To abbreviate descriptions, in what follows, we call the unobservable characteristic about which liability provides information ‘productivity’ regardless of the context, and we continue to use this phrase in the next section.

When productivity is negatively related to an entity’s propensity to commit illegal acts, intuitively, fault-based liability provides superior information compared to strict liability to third parties. This happens when, despite compliance with the law, a person may nevertheless accidentally cause the type of harm that the law seeks to mitigate. In such circumstances, under strict liability, third parties cannot infer whether a person caused harm because he violated the law or, whether, despite taking care his actions resulted in harm. On the other hand, under fault-based liability, lack of care is a pre-requisite to finding the defendant liable. Thus, strict liability only informs people about whether a defendant caused harm, whereas fault-based liability provides information about whether the defendant ‘meant to’ avoid harm. This piece of information is important, because it relates to the person’s intention, which, in turn, reveals information about his character.

Although correct, the above reasoning implicitly assumes that the court makes no errors in determining whether the defendant took care. When this assumption is violated, the informational superiority of fault-based liability can vanish. The extreme case where the court’s determination is only as good as a random guess helps in demonstrating why. In this extreme case, the best the court can do is to choose a probability, say β , with which it randomly decides whether the defendant took care. In addition, suppose p_0 and p_1 denote the probabilities with which harm occurs when the person does not take care, and when he takes care, respectively. Thus, a person is found liable with probability $p_i\beta$ where $i \in \{0, 1\}$ depending on whether the person took care. Therefore, under fault-based liability, the ratio of accurate findings of liability to false findings of liability is $\frac{\theta p_0 \beta}{(1-\theta)p_1 \beta}$, where θ is the proportion of individuals who have actually not taken care. It is easy to note that this proportion is the same under strict liability, namely, $\frac{\theta p_0}{(1-\theta)p_1}$, since fault is not a prerequisite and liability follows whenever there is harm. Thus, across the two regimes, a finding of liability implies the same probability that the defendant actually took care.¹³ However, the

¹³This simplified explanation holds the degree of deterrence, i.e. $1 - \theta$, constant across regimes. This is for expositional purposes only and deterrence is endogenously determined in section 3.

same cannot be said about a finding of no liability. In particular, under fault-based liability, the ratio of accurate findings of no liability to false findings of no liability is: $\frac{(1-\theta)[(1-p_1)+p_1(1-\beta)]}{\theta[(1-p_0)+p_0(1-\beta)]}$. This ratio is smaller than the analogous ratio under strict liability, which is given by $\frac{(1-\theta)(1-p_1)}{\theta(1-p_0)}$.¹⁴ This is essentially because fault-based liability adds an element of randomness in determining whether a person who has caused harm is nevertheless not found liable. Under strict liability, however, this determination is completely based on whether harm has occurred. Since the lack of harm is a better signal of whether a person took care than a random signal, a finding of no-liability ends up being a better proxy of whether a person in fact took care under strict liability.

The above observations demonstrate that across the two regimes, the accuracy of the information provided to third parties, and, therefore, the size of reputational sanctions, depend on how well the court performs in determining whether the defendant actually took care. In particular, when the court makes no mistakes, fault-based liability performs better, and, when the court's determination is only as reliable as a random guess, strict liability performs better. These are obviously extreme and unrealistic cases, but they demonstrate that there is no clear ranking between the two regimes. In intermediate (and more realistic) cases the relative ranking of the two regimes is often *a priori* ambiguous, and depends on meaningful factors, such as the frequency of liability. The dynamics that are responsible for this result are less apparent, and describing them requires some explanation of how the frequency of liability, holding all else constant, affects reputational costs.

Reputational costs are given by the difference between third parties' estimates of the productivity of people who are found liable and people who are not found liable. Intuitively, when a very large proportion of individuals are not liable in equilibrium, liability signals a large deviation from the average person in society, and, thus, being found liable generates a large magnitude of reputational costs. Similarly, when a very small proportion of individuals are liable in equilibrium, no liability signals exceptionally high productivity. Thus, having no liability carries a large reward. However, when the proportion of individuals who are found liable and the proportion of individuals who are not found liable are roughly equal, being in either category suggests only a moderate deviation from the population's average quality.

¹⁴This relies on the implicit assumption that $p_0 > p_1$, i.e. care reduces the likelihood of harm.

Therefore, if a move from strict liability to fault-based liability brings down the proportion of liability closer to 50% it reduces reputational costs. Indeed, when the proportion of liability is 50%, the binary ‘liable-not liable’ signal has maximum variance, and is therefore very noisy and has little impact.¹⁵ Holding individuals’ behavior constant, strict liability always increases the proportion of individuals found liable, simply by reducing the requirements that must be met for a finding of liability. Thus, if the act is uncommon, and, therefore, the maximum proportion of liability is less than 50%, switching from strict liability to fault-based liability makes liability an even rarer event, and, thereby, increases reputational costs. The opposite conclusion holds due to similar reasons when the maximum proportion of liability is more than 50%. Thus, counter-intuitively, the range of circumstances under which fault-based liability enhances reputational costs is broader when the act in question is uncommon.¹⁶

The interactions between the frequency of liability and reputational costs also reveals an important insight about the optimal burden of proof. When liability is infrequent, as explained, the magnitude of reputational sanctions can be increased by making liability a rarer event. Allocating the burden on the plaintiff is then efficient, because it maximizes reputational sanctions by reducing the frequency of liability. Essentially the opposite conclusion holds when liability is frequent, and it is optimal to place the burden on the defendant whenever possible.¹⁷

3 Model

We consider a continuum of individuals who can take care to reduce the likelihood with which their acts cause harm to others. These probabilities are, respectively, p_0 if the individual does not take care, and $p_1 < p_0$ if the individual exercises care. The cost of taking care varies among individuals and equals $g \in [0, \infty)$, with corresponding density function $k(g)$ and cumulative

¹⁵In the parlance of information theory, a binary variable achieves maximum entropy when it takes either value with 50% probability. *See, e.g.*, MacKay (2003), p.2. In this context, entropy can be thought of as a measure of uncertainty.

¹⁶This is reminiscent of Bénabou and Tirole’s (2006) discussion of the interplay of honor and shame in terms of the commonality of behavior, e.g., whether an act “is just not done” versus “everyone does it”.

¹⁷As will be made apparent in the next section, it is possible for evidence regarding care to be so irrelevant that it can never be used to over-turn the presumption caused by the occurrence of harm that the defendant did not take care. In these cases, allocating the burden on the defendant is equivalent to holding him strictly liable.

distribution function $K(g)$.

The act is regulated through either a strict liability regime or a fault-based liability regime. In both regimes, harm is a necessary condition for liability, but under fault-based liability a person is held liable only if the court rules that the person did not take care, and is not liable if the court rules that he took care. To simplify the analysis, we assume that the probability of enforcement is unity and the sanction is s .¹⁸

To capture the signalling effect of legal determinations of liability, we assume that an agent's productivity is inversely related to his cost of taking care. Specifically, the productivity is given by $\varphi(g)$ with $\varphi' < 0$. Interactions between third parties and agents generate surplus equal to the productivity of the agent. Therefore, third parties who face perfect competition, offer agents transfers that equal their expected productivity.¹⁹ To simplify the exposition of results, we call third parties employers, and we call transfers from employers to agents 'wages'. Employers form expectations regarding agents' productivities based on all available information, including whether they were found liable. This leads to a gap between the wages offered to people who have and have not been found liable. We refer to this gap as the stigma associated with being found liable. As will become clear, the magnitude of stigma depends on the liability regime.

3.1 Evidence and court decision rules

Under strict liability, a person is held liable whenever his action causes harm. Whether the court believes the person took care is irrelevant. Under fault-based liability, the court needs to assess behavior. It is unable to perfectly observe whether or not an actor took care and must rely on imperfect evidence to make a decision. We represent the potential evidence when harm occurs as a random variable x with support $[\underline{x}, \bar{x}]$ whose distribution is affected by whether or not the person took care.²⁰ The density functions associated with x are f_1 and f_0 when the person has and has not taken care respectively. The densities satisfy the monotone likelihood ratio property with f_0/f_1 monotonically decreasing, i.e., a large x suggests that the

¹⁸The analysis is identical when the probability of enforcement is fixed and smaller than unity. Our assumption eases the notation slightly by allowing us to carry around one variable fewer in derivations.

¹⁹We are following prior work, e.g. Rasmusen (1996), in making this assumption. However, results easily extend to cases where the transfer is not equal, but only proportional to the expected productivity of the agent.

²⁰The bounds of the support need not be finite.

person is more likely to have taken care. F_0 and F_1 are the corresponding cumulative distribution functions and it naturally follows that F_1 first-order stochastically dominates F_0 .

In this setting, an efficient decision rule for the court is to find liability only if x is below a critical value, denoted x^s , and which we refer to as the evidence threshold. By doing so, the court implicitly fixes the probability of erroneously finding a person who has taken care. We refer to this as the probability of type-1 error (or “false positive”), which equals

$$\alpha \equiv F_1(x^s) \tag{1}$$

The probability of correctly finding a person who has not taken care is

$$\beta \equiv F_0(x^s) \tag{2}$$

so that the probability of type-2 error (“false negative”) is $1 - \beta$. A threshold decision rule is efficient in the sense that the probability of type-2 error is minimized for any level of type-1 error. It is worth noting that these type-1 and type-2 error probabilities are conditional on adjudication. The marginal probabilities of such errors are $p_1\alpha$ and $1 - p_0\beta$, respectively.

Although fault is assessed on the basis of a critical evidence threshold, it is convenient to conduct the analysis as if the court directly chooses the type-1 error $\alpha \in [0, 1]$, implying the evidence threshold $x^s = F_1^{-1}(\alpha)$. Choosing the type-1 error yields a β satisfying

$$\beta(\alpha) = F_0(F_1^{-1}(\alpha)) \tag{3}$$

It is easily seen that $\beta(0) = 0$, $\beta(1) = 1$, and that the derivatives satisfy

$$\beta'(\alpha) = \frac{f_0(F_1^{-1}(\alpha))}{f_1(F_1^{-1}(\alpha))} \tag{4}$$

and $\beta''(\alpha) < 0$. The concavity follows from the monotone likelihood ratio property. Note that the preceding implies $\beta(\alpha) > \alpha$ for all $\alpha \in (0, 1)$, i.e., conditional on the occurrence of harm, for any positive type-1 error less than unity, an individual who did not exercise care is more likely to be found liable than one who did. Note also that $\alpha = 1$ is equivalent to strict liability. Accordingly, we will say that a regime is fault-based if it is characterized by a type-1 error $\alpha < 1$.

There is a natural and intuitive relationship between the court’s decision rule and familiar legal notions of burdens and standards of proof. We explain these relationships, next.

Burden and standard of proof. The plaintiff (or the public prosecutor as the case may be) always bears the burden of proving the occurrence of harm and causation. By assumption, this can always be established without error. When the liability rule is fault-based, the remaining issue is whether the defendant exercised care. The overall evidence for a ruling on this issue is the occurrence of harm and the realization x of the additional evidence about the defendant’s care.

When the plaintiff bears the burden of persuasion, as is usually the case in practice, the court will rule in his favor only if ‘no care’ is sufficiently likely relatively to ‘care’. The critical likelihood ratio to discharge the burden is the standard of proof, i.e., the required weight of evidence. Formally, the court finds for the plaintiff when x satisfies

$$\frac{p_0 f_0(x)}{p_1 f_1(x)} > t_P \tag{5}$$

where the left-hand side is the relative likelihood of ‘no care’ versus ‘care’ on the basis of all the available evidence (occurrence of harm and x) and where t_P is the standard of proof that the plaintiff must meet.²¹ In particular, $t_P = 1$ is akin to the preponderance standard of proof in common law, i.e., for the plaintiff to succeed it suffices that ‘no care’ is merely more likely than not.²² When $t_P > 1$, the standard of proof is stronger than preponderance.

In some situations, it turns out that the optimal legal regime will require a threshold likelihood ratio smaller than unity in condition (5). Consistent with the notion that standards of proof are at least as strong as preponderance, the optimal rule can then be interpreted in terms of a reverse burden of proof assignment. The defendant now has the burden of proving ‘care’ and succeeds when the inverse likelihood ratio satisfies

$$\frac{p_1 f_1(x)}{p_0 f_0(x)} > t_D \tag{6}$$

where t_D is the standard of proof that the defendant must meet.²³

²¹For a given evidence generating process as defined by the densities f_0 and f_1 , the chosen t_P yields a critical evidence threshold x^s . However, the concept of standard of proof is more general because the same standard can be applied in different categories of cases and under very different evidence generating processes.

²²‘More likely’ is solely in terms of the likelihood ratio on the basis of the particular evidence for the case at hand, irrespective of the court’s priors derived from the proportion of agents taking or not taking care in the overall population. The latter would generally not be considered relevant evidence for assessing the defendant’s behavior. See Demougin and Fluet (2006).

²³In this interpretation, when the standard equals unity (the preponderance standard),

Optimal legal design, as we shall illustrate, depends much on the precision of the evidence about care. Therefore, we proceed by describing how the properties of $\beta(\alpha)$ capture the informativeness of the evidence generating process (henceforth ‘EGP’).

Accuracy of the evidence. Any pair of densities $\{f_0, f_1\}$ defines an information system. According to a well known criterion, a system is more informative than another if it yields a smaller type-2 error for any given level of type-1 error.²⁴ For our purpose, it will be useful to allow for all possibilities ranging from completely uninformative to nearly perfectly informative evidence.

Possible information systems are represented by the set of continuously differentiable and strictly increasing concave functions $\beta(\alpha)$, $\alpha \in [0, 1]$, satisfying $\beta(0) = 0$ and $\beta(1) = 1$. From this set, one can extract families whose elements are ordered in terms of the relation ‘more informative than’.

DEFINITION 1: $\{\beta_\gamma\}_{\gamma \in \mathbb{R}_+}$ is an ordered family with system γ' more informative than system γ if $\gamma' > \gamma$ implies $\beta_{\gamma'}(\alpha) > \beta_\gamma(\alpha)$ for all $\alpha \in (0, 1)$, with $\beta_0(\alpha) \equiv \alpha$ and $\lim_{\gamma \rightarrow \infty} \beta_\gamma(\alpha) = 1$ for $\alpha > 0$.

At the lower bound $\gamma = 0$, the EGP is completely uninformative and $f_0/f_1 \equiv 1$. At the other end, for γ very large, f_0/f_1 ranges from very large values to nearly zero. From (4), this implies that the slope $\beta'_\gamma(\alpha)$ is very large for small values of the type-1 error and approaches zero as α approaches unity. We will repeatedly use the latter property.

Lemma 1 *In an ordered family of information systems, $\lim_{\gamma \rightarrow \infty} \beta'_\gamma(1) = 0$.*

When the EGP is very informative, a small reduction in the type-1 error from its value $\alpha = 1$ in the strict liability regime has negligible consequences in terms of the resulting increase in type-2 error.

Another important consideration is whether the EGP provides additional useful information compared with the mere knowledge that harm occurred.

it does not matter whether the plaintiff or the defendant bears the burden of proof, as the decision rule is essentially symmetric. When the standard is greater than unity (i.e., stronger than preponderance), it does matter.

²⁴This criterion is discussed in Blackwell and Girschick (1954) among other equivalent conditions. It is an instance, for the case of dichotomies, of Lehmann’s (1988) probability-probability plots. The criterion is equivalent to the likelihood ratio f_0/f_1 having more dispersion (in the sense of a mean preserving spread) in more informative systems; see Demougis and Fluet (2001) for a simple proof and Jewitt (2007) for a comprehensive analysis.

Because $p_0 > p_1$, the occurrence of harm is information that is more consistent with low care. If the evidence generating system is so poorly informative that it cannot produce any piece of evidence that would be sufficient for the defendant to over-turn this presumption, then evidence pertaining to care would be irrelevant for establishing preponderance of the evidence. We call EGPs that are informative enough to potentially switch the preponderance of the evidence in favor of the defendant *preponderance-relevant*. Conversely, we call EGPs that cannot fulfill this function, *preponderance-irrelevant*. This concept can formally be defined as follows.²⁵

DEFINITION 2: An evidence generating process is preponderance-relevant if

$$\frac{p_0 f_0(x)}{p_1 f_1(x)} < 1 \text{ for some realizations } x. \quad (7)$$

It is worth noting that when the EGP is *preponderance-irrelevant*, a defendant bearing the burden of proof never prevails under the preponderance standard, i.e., he can never show that due care is more likely than not. Therefore, in these cases, allocating the burden of proof on the defendant amounts to using strict liability. Next, we identify conditions under which EGPs possess this property.

Lemma 2 *The evidence generating process is preponderance-relevant if and only if $p_0\beta(\alpha) - p_1\alpha$ has an interior maximum.*

The function $p_0\beta(\alpha) - p_1\alpha$ is the differential liability risk between not taking care and taking care under a liability regime with type-1 error equal to α . The differential is depicted in Figure 1 for three information systems from an ordered family with $\gamma''' > \gamma'' > \gamma'$. The positively sloped dotted line is the differential liability risk in the limiting case of completely uninformative evidence (i.e., $\gamma = 0$). The negatively sloped dotted line is the upper bound, never reached, of completely informative evidence ($\gamma = \infty$). Under system γ' , the EGP is preponderance-irrelevant. The differential liability risk then reaches a corner maximum at $\alpha = 1$, i.e., strict liability. The two other systems are preponderance-relevant, so the curves have an interior maximum. Obviously, (7) holds if the evidence is very informative (i.e., a large γ) and it does not hold if the evidence is sufficiently poor (i.e., a small γ).

In the sequel, we take the EGP as given for the category of cases considered. However, we also make comparative statics statements relating the

²⁵See Demougin and Fluet (2006) for a similar discussion.

quality of the potential evidence to the optimal legal regime. *Sufficiently informative* evidence means that there is a sufficiently large γ in some ordered family of information systems such that the statement is true. *Sufficiently uninformative* means that there is a sufficiently small γ in some such family such that the statement holds.

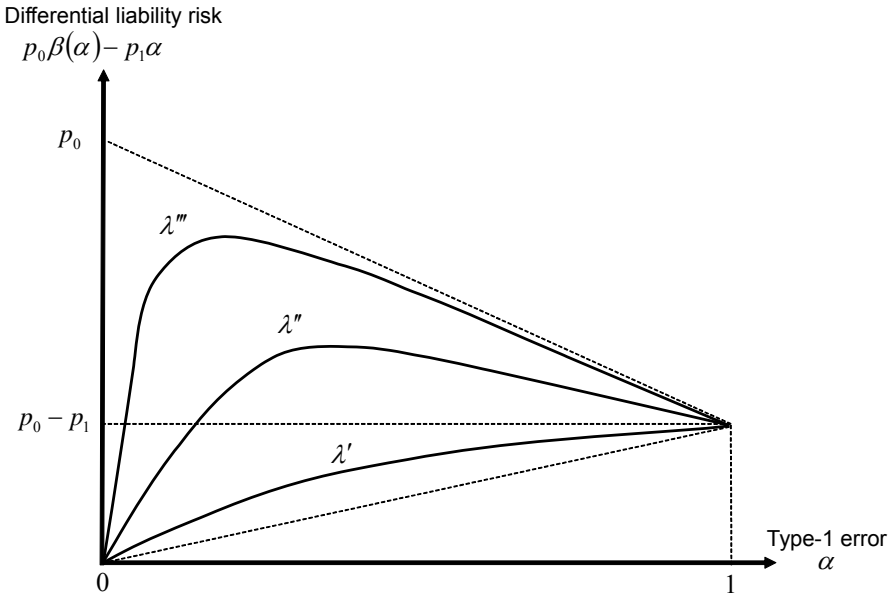


Fig. 1. Information systems

3.2 Individuals' Decisions to take Care

As a preliminary observation, note that the pools of individuals who are and are not found liable will have different average productivities, denoted as w_L and w_N respectively. The difference between these average productivities, $\sigma = w_N - w_L$, is the stigma imposed on an individual who is found liable. As there is a continuum of individuals, each individual takes the (endogenously determined) equilibrium level of stigma as given while making decisions because his behavior has no effect on this value. Thus, given the stigma generated by all other individuals' behavior, denoted σ^e where the superscript refers to an equilibrium value, a person's expected net benefit from taking care is

$$\begin{aligned} (1 - p_1\alpha)w_N + p_1\alpha(w_L - s) - g = \\ w_N - p_1\alpha(\sigma^e + s) - g \end{aligned} \quad (8)$$

whereas the net expected benefit from not taking care is

$$\begin{aligned} (1 - p_0\beta)w_N + p_0\beta(w_L - s) = \\ w_N - p_0\beta(\sigma^e + s) \end{aligned} \quad (9)$$

Thus, a person takes care if

$$g < \hat{g}(\sigma^e, \alpha) \equiv (p_0\beta - p_1\alpha)(s + \sigma^e) \quad (10)$$

The function $\hat{g}(\sigma^e, \alpha)$ characterizes the individuals' best-responses in terms of taking care versus not taking care as a function of their anticipated level of stigma σ^e and given the liability regime α . On the other hand, the magnitude of stigma that emerges in the market is itself a function of individuals' decisions with respect to taking care. Specifically, it depends on the equilibrium threshold g^e defining the proportion of individuals who take care. We first derive the magnitude of stigma that emerges as a function of g^e , which we later use to characterize the equilibrium.

3.3 Stigma as a Function of g^e and α

Given g^e , the measure of type g people who are found liable is given by:

$$\begin{aligned} p_0\beta k(g) & \text{ if } g \geq g^e \\ p_1\alpha k(g) & \text{ if } g < g^e \end{aligned} \quad (11)$$

Similarly, the measure of people who are not found liable is given by:

$$\begin{aligned} & (1 - p_0\beta)k(g) & \text{if } g \geq g^e \\ & (1 - p_1\alpha)k(g) & \text{if } g < g^e \end{aligned} \quad (12)$$

Thus, the average productivity of people who are found liable is given by

$$w_L \equiv \frac{p_0\beta \int_{g^e}^{\infty} \varphi(g)k(g)dg + p_1\alpha \int_0^{g^e} \varphi(g)k(g)dg}{\psi} \quad (13)$$

where the function

$$\psi(g^e, \alpha) \equiv p_0\beta(\alpha)(1 - K(g^e)) + p_1\alpha K(g^e) \quad (14)$$

denotes the measure of individuals who are found liable. The average productivity of individuals without liability is given by

$$w_N \equiv \frac{(1 - p_0\beta) \int_{g^e}^{\infty} \varphi(g)k(g)dg + (1 - p_1\alpha) \int_0^{g^e} \varphi(g)k(g)dg}{1 - \psi} \quad (15)$$

To simplify notation, let the average quality of individuals who do and do not take care be η and λ respectively:

$$\lambda(g^e) \equiv \frac{\int_{g^e}^{\infty} \varphi(g)k(g)dg}{1 - K(g^e)} \quad \text{and} \quad \eta(g^e) \equiv \frac{\int_0^{g^e} \varphi(g)k(g)dg}{K(g^e)} \quad (16)$$

Given g^e and the legal regime, the stigma attached to being liable is therefore:

$$\hat{\sigma} = \frac{(1 - p_0\beta)\lambda(1 - K) + (1 - p_1\alpha)\eta K}{1 - \psi} - \frac{p_0\beta\lambda(1 - K) + p_1\alpha\eta K}{\psi} \quad (17)$$

Simplifying this expression yields the stigma function:

$$\hat{\sigma}(g^e, \alpha) = \frac{[p_0\beta(\alpha) - p_1\alpha][K(g^e)(1 - K(g^e))][\eta(g^e) - \lambda(g^e)]}{\psi(g^e, \alpha)(1 - \psi(g^e, \alpha))} \quad (18)$$

All components that affect the magnitude of stigma have intuitive interpretations. The first factor in the numerator, $p_0\beta(\alpha) - p_1\alpha$, is the liability risk differential discussed in the preceding section. This depends only on the

legal regime and reflects its *discriminating power* because it is the difference in the frequency with which a person is correctly versus incorrectly found liable. The second factor, $K(1-K)$, depends only on the equilibrium g^e and is the *variance in the distribution of individuals who do or do not take care*. The third factor, $(\eta - \lambda)$, also depends only on g^e and is *the difference in the average quality of individuals who do and who do not take care*. Finally, the denominator, $\psi(1 - \psi)$, depends on both the legal regime and g^e , and is the *variance in the distribution of individuals who are or are not found liable*. As (18) immediately reveals, stigma is increasing in the first three factors and decreasing in the fourth.

The next lemma makes a few observations regarding the relationship between stigma and type-1 error, taking g^e as given.

Lemma 3 *Liability generates positive stigma, i.e., $\hat{\sigma}(g^e, \alpha) > 0$ for all $\alpha > 0$: (i) $\hat{\sigma}_\alpha(g^e, \alpha) > 0$ if the evidence generating process is sufficiently uninformative, or if it is preponderance-irrelevant and $\psi(g^e, \alpha) \geq 1/2$; (ii) $\hat{\sigma}_\alpha(g^e, 1) < 0$ if the evidence generating process is sufficiently informative, or if it is preponderance-relevant and $\psi(g^e, 1) \leq 1/2$.*

The lemma implies that stigma is maximized by a strict liability regime if the EGP is sufficiently uninformative and by a fault-based regime if it is sufficiently informative. When the condition is merely whether the EGP is preponderance relevant or irrelevant, one can only sign the effect of a local marginal increase in the type-1 error, taking g^e as given. The sign then depends on the proportion $\psi(g^e, \alpha)$ of agents who are found liable. Observe, however, that the condition $\psi(g^e, 1) \leq 1/2$ in the second part of the lemma is trivially satisfied when $p_0 \leq 1/2$. The occurrence of harm is then “infrequent” even under low care. Therefore, so is the finding of liability.

Having determined how stigma and type-1 error affect individuals’ behavior, and how stigma is affected by individuals’ behavior and the type-1 error, we can now investigate the equilibrium properties of stigma and deterrence as a function of the legal regime.

3.4 Equilibrium and Optimal Liability Regimes

At equilibrium, given the behavior of all other individuals, no individual can make himself better off by deviating from his behavior. Thus, an equilibrium is a pair (g^e, σ^e) that solves:

$$g^e = \hat{g}(\sigma^e, \alpha) \text{ and } \sigma^e = \hat{\sigma}(g^e, \alpha) \tag{19}$$

where the functions are as defined in (10) and (18) respectively. Substituting from the second equation, the equilibrium condition can be expressed as:

$$g^e - [p_0\beta(\alpha) - p_1\alpha](s + \hat{\sigma}(g^e, \alpha)) = 0 \quad (20)$$

Stable equilibria satisfy the condition:

$$1 - [p_0\beta(\alpha) - p_1\alpha]\hat{\sigma}_{g^e}(g^e, \alpha) > 0 \quad (21)$$

At a stable equilibrium, the effect on deterrence of a marginal change in the legal regime is given by

$$\frac{dg^e}{d\alpha} = \frac{[p_0\beta'(\alpha) - p_1](s + \hat{\sigma}(g^e, \alpha)) + [p_0\beta(\alpha) - p_1\alpha]\hat{\sigma}_\alpha(g^e, \alpha)}{1 - [p_0\beta(\alpha) - p_1\alpha]\hat{\sigma}_{g^e}(g^e, \alpha)} \quad (22)$$

A small increase in the type-1 error increases deterrence if the numerator in (22) is positive. The sign depends on two effects. The first is the direct effect of a change in the differential liability risk due to a change in the type-1 error. The second is the indirect effect through the change in the stigma.

Strict versus fault-based liability. The preceding observation, combined with Lemma 3, allows the identification of conditions under which strict liability or fault-based liability maximizes deterrence.

Proposition 1

A) *Fault-based [resp. strict] liability is optimal if the evidence generating process is sufficiently informative [resp. uninformative].*

B) *In an optimal legal regime: (i) if the evidence generating process is preponderance-relevant and liability is infrequently found (i.e., $\psi \leq 1/2$), then liability is fault-based; (ii) if the process is preponderance-irrelevant and liability is frequently found (i.e., $\psi > 1/2$), then liability is strict.*

Proposition 1 shows that one regime does not unambiguously lead to more deterrence than the other. Table 1 summarizes the results.

Table 1 *Optimal Liability Regimes*

| [Liability is] → | Infrequent | Frequent |
|----------------------------|--------------------|--------------------|
| [Evidence gen. proc. is] ↓ | | |
| Very uninformative | Strict Liability | Strict Liability |
| Preponderance-irrelevant | Strict or Fault-B. | Strict Liability |
| Preponderance-relevant | Fault-Based | Strict or Fault-B. |
| Very informative | Fault-Based | Fault-Based |

Burden and standard of proof in a fault regime. Next, we investigate how the burden of proof ought to be allocated when fault-based liability is optimal and discuss the required standard of proof. Let $\alpha^* < 1$ refer to the optimal type-1 error. When $p_0\beta'(\alpha^*) \geq p_1$, recalling the discussion in Section 2.1, the court's decision rule can be interpreted as putting the burden on the plaintiff with the requirement that the plaintiff succeeds only if he submits evidence satisfying

$$\frac{p_0 f_0(x)}{p_1 f_1(x)} > t_P \equiv \frac{p_0 \beta'(\alpha^*)}{p_1} \geq 1 \quad (23)$$

Conversely, when $p_0\beta'(\alpha^*) < p_1$, the court's decision rule can be interpreted as putting the burden on the defendant with the requirement that he submits evidence satisfying

$$\frac{p_1 f_1(x)}{p_0 f_0(x)} > t_D \equiv \frac{p_1}{p_0 \beta'(\alpha^*)} > 1 \quad (24)$$

Proposition 2 *When the optimal regime is fault-based, (i) if the evidence generating process is preponderance-irrelevant, then liability is infrequent and the burden of proof is on the plaintiff; (ii) if the evidence generating process is preponderance-relevant, the burden of proof is on the plaintiff [resp. the defendant] if liability is infrequent [resp. frequent].*

As Proposition 2 demonstrates, there is a strict relationship between the optimal assignment of the burden of proof and the frequency of liability. Table 2 summarizes these results.

Table 2 *Burden of Proof under Fault-Based Liability*

| [Liability is] \longrightarrow | Infrequent | Frequent |
|---------------------------------------|------------|-----------|
| [Evidence gen. proc. is] \downarrow | | |
| Preponderance-irrelevant | Plaintiff | Plaintiff |
| Preponderance-relevant | Plaintiff | Defendant |

Note that, when $p_0 \leq 1/2$, liability is always infrequent. An optimal fault-based regime then puts the burden on the plaintiff.

An interesting result worth emphasizing is that the preponderance of the evidence standard is optimal only in a very peculiar case. The next corollary highlights this result.

Corollary 1 *In an optimal regime, the standard of proof is stronger than preponderance of evidence, except when liability is found with a frequency of 1/2.*

The proof follows immediately from part *ii* of the proof of proposition 2. The case where the frequency of finding fault equals one half is clearly non-generic. A stronger formulation would therefore be that the preponderance of evidence standard is ‘never’ optimal. This result clearly presents a deviation from findings in the prior literature in which preponderance of the evidence emerges as an optimal decision rule. We discuss this, and related implications of our results, in the next section.

4 Discussion

Our analysis highlights the very general point that the choice of liability regime affects incentives through its impact on expected reputational sanctions, in addition to affecting the incremental expected formal sanctions one faces by violating laws. The latter corresponds to the effect of the liability regime on what we termed the *differential liability risk* (see, e.g. figure 1). One important aspect of this risk is that it is completely determined by how informative the EGP is (i.e. the shape of $\beta(\alpha)$) and how much exercising care reduces the probability of harm (i.e. p_0/p_1). On the other hand, reputational sanctions, and how they are affected by different liability rules, depend very much on the potential level of deterrence, as well as the population’s characteristics (e.g. the shape of φ), in addition to the factors that affect the differential liability risk. This distinction has implications pertaining to the optimality of exclusionary rules as well as the uncommonality requirement for using strict liability. Next, we discuss these issues, and how they relate to prior work.

4.1 Exclusionary Rules, Revisited

That the differential liability risk depends only on the informativeness of the EGP and the reduction in the probability of harm caused by taking care implies that, in a setting where there are no reputational costs, these two factors completely determine which liability regime is optimal. As explained in Demougin and Fluet (2006), this observation serves as a strong rationale for using rules that exclude evidence from the consideration of decision makers in returning verdicts. For instance, in determining whether a person has

more likely than not acted without due care, the jury ought not to consider the proportion of individuals who have taken care. This is because doing so has the potential of distorting the decision rule away from the rule that maximizes the differential liability risk.²⁶

However, as our analysis demonstrates, when reputational sanctions are incorporated into the analysis, the optimal liability regime, as well as the standard of proof that ought to be used in implementing it, depends on the population's characteristics and the level of deterrence. Thus, optimal results can only be obtained by using different liability regimes in different contexts, and by requiring courts to adjust the standard of proof applicable in different cases. However, in reality, making small adjustments to the standard of proof may be impracticable. In fact, there are only a handful of such standards that are frequently used by courts. If the use of a discrete number of standards of proof is imposed as a constraint, it may be desirable to allow evidence that relates to the underlying population's characteristics to mitigate losses in deterrence (caused by the impossibility of using a continuum of standards of proof). Further research focusing on this possibility may generate fruitful results, and may have important implications regarding the optimal exclusion of evidence that relates to population characteristics.

4.2 The Uncommonality Requirement

In the United States, courts generally hold an injurer strictly liable in tort law if two requirements are met: "(1) the injurer's activity must generate a highly significant danger even when undertaken with reasonable care; and (2) the injurer's activity must be uncommon" (Shavell 2017 p. 1 citing Restatement (Third) of Torts). The wisdom behind the second of these requirements has been questioned, recently, in Shavell (2017), which argues that all acts that meet the first requirement ought to be regulated through strict liability.²⁷ Shavell's primary claim is that, by making uncommonity a requirement, the legal system forgoes the opportunity to increase the deterrence of acts that are dangerous but common, and that this causes reductions in welfare.

In our model, the uncommonality requirement corresponds to the equilibrium level of deterrence (i.e. $K(g^e)$) being high. As table 1 illustrates,

²⁶Further discussion of this point can be found in Demougin and Fluet (2006).

²⁷Shavell (2017) also notes that the American approach appears to be an outlier, because no other countries seem to adopt a similar requirement for the imposition of strict liability.

strict liability is optimal under a broader set of conditions when liability is frequent compared to when it is infrequent. Liability is more frequent, in turn, when the level of deterrence is low. Thus, our analysis suggests that, contrary to the legal doctrine in the United States, fault-based liability enjoys a comparative advantage over strict liability when the act is uncommon. The rationale behind this result is explained section 2: when liability is infrequent, switching from strict liability to fault based liability causes a further reduction in the frequency of liability. The end result is a more extreme separation between the groups of individuals who take care and who do not take care, and, thus a finding of liability has a greater signal value. Therefore, fault-based liability generates additional deterrence effects. Hence, although our analysis complements Shavell (2017) in suggesting that the commonality of an act should not be seen as a reason to not use strict liability, it also suggests that there is a rationale to use fault-based liability more often when the act is uncommon.

5 Conclusion

The comparative advantages of strict liability regimes versus fault-based liability regimes has attracted the interest of many law and economics scholars. An overwhelming majority of scholarship in this field focuses on how liability regimes affect various parties' incentives through their effect on formal sanctions. We have demonstrated here, that the choice of liability regime alters the quality of the information conveyed through findings of (no-)liability. Therefore, liability regimes can affect incentives also through their impact on informal sanctions. Although our analysis centered around a specific issue, our approach can be used to study the more general issue of defining wrongful acts. For instance, the number and nature of the elements included in the definition of a crime or tort can affect the stigma generated from being found liable. Thus, future research focusing on the informational aspects of legal design are likely to enhance our understanding of the incentive effects generated by various doctrines and institutions.

6 Appendix

Proof of Lemma 1: Define $\beta'_\gamma(1)$ as the left-derivative at $\alpha = 1$. Because β_γ is increasing and concave, for $\alpha > 0$,

$$0 \leq \beta'_\gamma(1) \leq \frac{1 - \beta_\gamma(\alpha)}{1 - \alpha}$$

The result then follows from $\lim_{\gamma \rightarrow \infty} \beta_\gamma(\alpha) = 1$. ■

Proof of Lemma 2: Condition (7) cannot hold for all x because $p_0 > p_1$ and

$$\int_{\underline{x}}^{\bar{x}} (f_0(x)/f_1(x))f_1(x) dx = 1$$

Hence, (7) is true if and only if, for some $x^s \in (\underline{x}, \bar{x})$,

$$\frac{p_0 f_0(x^s)}{p_1 f_1(x^s)} = 1$$

Then (7) holds for $x > x^s$ and the reverse inequality holds for $x < x^s$. Defining $\alpha_P \equiv F_1(x^s)$, it follows that $\alpha_P \in (0, 1)$ and from equation (4)

$$\frac{p_0 \beta'(\alpha_P)}{p_1} = \frac{p_0 f_0(x^s)}{p_1 f_1(x^s)} = 1$$

implying that $p_0 \beta(\alpha) - p_1 \alpha$, a concave function, is maximized at α_P . ■

Proof of Lemma 3: From (16), $\eta(g^e) > \lambda(g^e)$ because $\varphi' < 0$. For all $\alpha > 0$, $p_0 \beta(\alpha) - p_1 \alpha > 0$. Hence $\hat{\sigma} > 0$. To prove the remaining, observe that $\hat{\sigma}_\alpha$ has the same sign as

$$\frac{\partial}{\partial \alpha} \left(\frac{p_0 \beta(\alpha) - p_1 \alpha}{\psi(g^e, \alpha)(1 - \psi(g^e, \alpha))} \right),$$

which in turn has the same sign as

$$\begin{aligned} Q(g^e, \alpha) &\equiv [p_0 \beta'(\alpha) - p_1] \psi(g^e, \alpha)(1 - \psi(g^e, \alpha)) \\ &\quad - [p_0 \beta(\alpha) - p_1 \alpha] \psi_\alpha(g^e, \alpha)(1 - 2\psi(g^e, \alpha)). \end{aligned} \quad (25)$$

We now introduce the index γ to identify information systems in an ordered family and omit the argument g^e to simplify notation. Thus,

$$\begin{aligned} Q_\gamma(\alpha) &= [p_0 \beta'_\gamma(\alpha) - p_1] \psi_\gamma(\alpha)(1 - \psi_\gamma(\alpha)) \\ &\quad - [p_0 \beta_\gamma(\alpha) - p_1 \alpha] \psi'_\gamma(\alpha)(1 - 2\psi_\gamma(\alpha)) \end{aligned} \quad (26)$$

where $\psi_\gamma(\alpha) = p_0\beta_\gamma(\alpha)(1 - K) + p_1\alpha K$ and therefore $\psi'_\gamma(\alpha) > 0$.

Claim (i). If the EGP is preponderance-irrelevant, then $p_0\beta'_\gamma(\alpha) > p_1$ for all $\alpha < 1$, hence $Q_\gamma(\alpha) > 0$ if $\psi_\gamma(\alpha) > 1/2$. Next we show that $Q_\gamma(\alpha) > 0$ for all $\alpha > 0$, irrespective of $\psi_\gamma(\alpha)$, if the EGP is sufficiently uninformative. It suffices to show that this is true for $\gamma = 0$, in which case $\beta_0(\alpha) = \alpha$ and therefore $\psi_0(\alpha) = \alpha[p_0(1 - K) + p_1K]$. Substituting in (26) then yields

$$Q_0(\alpha) = (p_0 - p_1)[p_0(1 - K) + p_1K]^2\alpha^2 > 0 \text{ for all } \alpha > 0$$

Claim (ii). By Lemma 2, if the EGP is preponderance-relevant, then $p_0\beta'_\gamma(1) < p_1$, hence $Q_\gamma(1) < 0$ if $\psi_\gamma(1) \leq 1/2$. Next we show that $Q_\gamma(1) < 0$, irrespective of $\psi_\gamma(1)$, if the EGP is sufficiently informative. By Lemma 1, $\lim_{\gamma \rightarrow \infty} \beta'_\gamma(1) = 0$ and therefore $\lim_{\gamma \rightarrow \infty} \psi'_\gamma(1) = p_1K$. Substituting in (26) evaluated at $\alpha = 1$ then yields

$$\begin{aligned} \lim_{\gamma \rightarrow \infty} Q_\gamma(1) &= -p_0p_1[1 - p_0(1 - K) - p_1K] \\ &\quad - p_1(1 - p_0)[p_0(1 - K) + p_1K]K \\ &< 0. \end{aligned}$$

which completes the proof. ■

Proof of Proposition 1: A) From lemmas 2 and 3, $p_0\beta'(\alpha) > p_1$ and $\hat{\sigma}_\alpha > 0$ for all α when the evidence is sufficiently uninformative. The numerator of (22) is then always positive, so the optimal regime is strict liability. From the same lemmas, $p_0\beta'(1) < p_1$ and $\hat{\sigma}_\alpha(g^e, 1) < 0$ when the EGP is sufficiently informative. The numerator of (22) is then negative at $\alpha = 1$, which therefore cannot be optimal.

B) The proof is by contradiction. Denote the optimal regime by α^* and let g^e be the equilibrium at α^* . Suppose the EGP is preponderance-relevant and $\psi(g^e, \alpha^*) \leq 1/2$. Then $\alpha^* = 1$ yields a contradiction because, from the lemmas 2 and 3, $p_0\beta'(1) < p_1$ and $\hat{\sigma}_\alpha(g^e, 1) < 0$, hence the numerator of (22) is negative. Similarly, suppose the evidence is preponderance-irrelevant and $\psi(g^e, \alpha^*) > 1/2$. Then $\alpha^* < 1$ yields a contradiction because, from the same lemmas, $p_0\beta'(\alpha^*) > p_1$ and $\hat{\sigma}_\alpha(g^e, \alpha^*) > 0$, hence the numerator of (22) is positive. ■

Proof of Proposition 2: Let α^* refer to the optimal type-1 error.

Claim (i). In this case, $p_0\beta'(\alpha) > p_1$ for all α . By definition, the burden is therefore on the plaintiff. By Proposition 1 B-ii), $\alpha^* < 1$ can be optimal only if liability is infrequent.

Claim (ii). In this case, $p_0\beta(\alpha) - p_1\alpha$ has a strict maximum at some $\alpha_P < 1$, hence $p_0\beta'(\alpha_P) = p_1$. At the optimum, $dg^e/d\alpha = 0$. Therefore, from (22),

$$N \equiv [p_0\beta'(\alpha^*) - p_1](s + \hat{\sigma}(g^e, \alpha^*)) + [p_0\beta(\alpha^*) - p_1\alpha^*]\hat{\sigma}_\alpha(g^e, \alpha^*) \quad (27)$$

must be equal to zero, where g^e is the equilibrium at $\alpha = \alpha^*$. Borrowing from the proof of Lemma 1, the sign of $\hat{\sigma}_\alpha(g^e, \alpha^*)$ is the same as the sign of

$$Q = [p_0\beta'(\alpha^*) - p_1]\psi(g^e, \alpha^*)(1 - \psi(g^e, \alpha^*)) \\ - [p_0\beta(\alpha^*) - p_1\alpha^*]\psi_\alpha(g^e, \alpha^*)(1 - 2\psi(g^e, \alpha^*))$$

The proof is by contradiction. If $\alpha^* > \alpha_P$ and $\psi(g^e, \alpha^*) \leq 1/2$, then $Q < 0$ and therefore $N < 0$, a contradiction. Similarly, if $\alpha^* \leq \alpha_P$ and $\psi(g^e, \alpha^*) > 1/2$, then $Q > 0$ and therefore $N > 0$, again a contradiction. The only possibilities are therefore $\alpha^* > \alpha_P$ and $\psi(g^e, \alpha^*) > 1/2$, or $\alpha^* < \alpha_P$ and $\psi(g^e, \alpha^*) < 1/2$, or $\alpha^* = \alpha_P$ and $\psi(g^e, \alpha^*) = 1/2$. ■

References

- [1] Blackwell, D. and M.A. Girschick (1954) *Theory of Games and Statistical Decisions*, Wiley.
- [2] Bénabou, R. and J. Tirole (2006) "Incentives and Prosocial Behavior" 96 *American Economic Review* 1652-1678.
- [3] Deffains, D. and C. Fluet (2013) "Legal Liability when Individuals Have Moral Concerns" 29 *Journal of Law, Economics, and Organization* 930-955.
- [4] Demougin, D. and C. Fluet (2001) "Ranking of Information Systems in Agency Models: An Integral Condition" 17 *Economic Theory* 489-496.
- [5] Demougin, D. and C. Fluet (2005) "Deterrence versus Judicial Error: A Comparative View of Standards of Proof" 161 *Journal of Institutional and Theoretical Economics* 193-206.
- [6] Demougin, D. and C. Fluet (2006) "Preponderance of Evidence" 50 *European Economic Review* 963-976.
- [7] Iacobucci, E. (2014) "On the Interactions between Legal and Reputational Sanctions" 43 *Journal of Legal Studies* 189-207.

- [8] Helland, E. (2006) "Reputational Penalties and the Merits of Class-Action Securities Litigation" *Journal of Law and Economics* 49, 365-395.
- [9] Jewitt, I. (2007) "Information Orders in Decision and Agency Problems" Unpublished.
- [10] Karpoff, J., D. Lee, and G. Martin. (2008) "The cost to firms of cooking the books." *43 Journal of Financial and Quantitative Analysis* 581-611.
- [11] Lott, J. (1992a) "An Attempt at Measuring the Total Monetary Penalty from Drug Convictions: The Importance of an Individual's Reputation" *21 Journal of Legal Studies* 159-187.
- [12] Lott, J. (1992b) "Do We Punish High Income Criminals Too Heavily?" *30 Economic Inquiry* 583-608.
- [13] MacKay, David. *Information Theory, Inference, and Learning Algorithms* Cambridge: Cambridge University Press, 2003.
- [14] Mueller-Langer, Frank; Schäfer, Hans-Bernd (2009). Strict liability versus negligence, in: Michael Faure (Hg.), *Tort Law and Economics*, Chapter 1 (*Encyclopedia of Law and Economics*), 3-45. Edward Elgar: Cheltenham; Northampton, Mass. 2009.
- [15] Mungan, M. (2016) "A Generalized Model for Reputational Sanctions and the (Ir)relevance of the Interactions between Legal and Reputational Sanctions" *46 International Review of Law and Economics* 86-92.
- [16] Murray, B. (2016) "A New Era for Expungement Law Reform? Recent Developments at the State and Federal Levels" *10 Harvard Law and Policy Review* 361-383.
- [17] Pager, D. (2003) "The Mark of a Criminal Record" *108 American Journal of Sociology* 937-975.
- [18] Pager, D., B. Western, and N. Sugie. (2009) "Sequencing Disadvantage: Barriers to Employment Facing Young Black and White Men with Criminal Records." *623 The ANNALS of the American Academy of Political and Social Science* 195-213.
- [19] Polinsky, M. and S. Shavell "Public Enforcement of Law" in *Handbook of Law and Economics* 403-454 (A. M. Polinsky & S. Shavell, eds. 2007)

- [20] Rasmusen, E. (1996) "Stigma and Self-Fulfilling Expectations of Criminality" 39 *Journal of Law and Economics* 519-543.
- [21] Shavell, S. (2017) "Why Strict Liability Should Apply To All Dangerous Activities—Both Common and Uncommon" Unpublished Manuscript.