RETALIATION, REMEDIES, AND TORTS

Sergio Mittlaender [*]

Fundação Getulio Vargas Law School in São Paulo (FGV Direito SP)

&

Max Planck Institute for Social Law and Social Policy

Abstract

Traditional tort models analyze the role of liability rules to induce socially optimal levels of activity and of care, focusing on the law's ability to deter accidents by imposing a cost on tortfeasors. Legal liability, however, further provides compensation for victims, and is apt to reduce victims' levels of aggrievement and thereby crowd out socially costly forms of private redress. This article presents a model of accidents and liability that endogeneizes aggrievement experienced by victims and its behavioral consequence in form of retaliation. It reveals how liability rules crowd-out retaliation and how strict liability performs better than negligence both to minimize retaliation and to avoid overinvestment in care that is predicted to arise under negligence. The predictions are tested in a laboratory experiment that provides evidence for these effects and for the underlying assumption of the model that experienced losses lead victims to retaliate. Compensation is justified not only because of its effects on potential tortfeasors, but also because of its effects on victims' behavior, and thereby on social welfare.

Keywords: Torts, Punishment, Accidents, Negligence, Strict Liability.

JEL codes: K13 (Tort Law), C91 (Laboratory, Individual Behavior).

[*] Professor of Law, Fundação Getulio Vargas Law School in São Paulo (FGV Direito SP), Rua Rocha 233, São Paulo 01330-000, SP, Brazil; Senior Research Fellow, Max Planck Institute for Social Law and Social Policy, Amalienstrasse 33, Munich 80799, Germany; email: mittlaender@mpisoc.mpg.de.

The author would like to thank Ulrich Becker and the Max Planck Institute for Social Law and Social Policy for financing of the experiment, Paola Bertoli, Lubomír Cingl, and Tomáš Miklánek for the support in implementing the experiment, and participants at the European Association of Law & Economics annual meeting in Tel-Aviv, 2019.

I. Introduction

Liability for accidents performs different economic functions. It provides incentives for potential tortfeasors to invest in care by imposing upon them, in case an accident occurs, the duty to compensate the victim for losses sustained (Calabresi 1965, 1968; Posner 1972; Brown 1973; Shavell 1980). Legal liability, however, does not prevent all accidents from occurring, for even if the potential tortfeasor invests optimally in care and precautions, there will still be a strictly positive probability of an accident. And once an accident occurs, the victim will feel aggrieved and entitled to be compensated for the loss she suffers because of the conduct of the tortfeasor. If she is not entitled to a legal remedy, aggrievement will often lead to retaliation and other acts to punish the tortfeasor. Legal liability performs a crucial function of crowding out aggrievement and retaliation by victims of wrongs in providing them compensation, and thereby avoiding a deadweight loss from decentralized forms of punishment.

While strict liability and negligence can both provide compensation for victims, the negligence rule often leaves victims without compensation, for whenever the injurer takes the required due level of care, then she is not considered negligent, and hence does not have to pay compensation for the victim if an accident occurs nevertheless. Victims might still understand that they should not bear those losses since they did not cause them, nor profited from the activity that led to them. A conflict emerges in which both the tortfeasor who was not negligent and the victim who suffered the loss will tend to understand that each should not bear the loss. Strict liability avoids the conflict and the resulting tendency of victims to punish injurers by inevitably imposing upon injurers the duty to fully compensate victims for the losses sustained.

The consequences of considering the role of punishment and private redress under the negligence rule and strict liability is not restricted to what happens after the accident. Potential injurers will anticipate that victims will punish them once an accident occurs and the victim, if unable to prove negligent behavior by the injurer, given the latter's absence of fault, must bear

the losses herself. In effect, under the negligence rule, injurers must consider not only the expected costs of legal liability, but also the expected losses from retaliation by the victim whenever they take the due level of care, but an accident nevertheless occurs. Potential tortfeasors are hence predicted to overinvest in care under the negligence rule, with two potential disadvantages of the negligence rule with respect to strict liability, none of which is considered in traditional tort models: social losses from retaliation and social losses from overinvestment in care.

This article provides, firstly, a model of accidents with aggrievement and retaliation based on Hart and Moore's model (2008) of contracts as reference points. Secondly, it presents results from an economic experiment that tests the model's predictions in the unilateral model of accidents (*see* Shavell 2007). The main predictions under empirical scrutiny are that (i) while both rules crowd out retaliation by victims, strictly liability performs better, and that (ii) while both rules induce potential tortfeasors to invest in care and precautions, the negligence rule will lead to overinvestment and hence performs worse.

In the experiment, participants took the role of either a potential injurer who could invest in precautions in order to reduce the expected costs of accidents or of victims who could punish injurers for their choices. In the three different treatments, there was either no liability, strict liability, or liability for negligence. Since the model's two main predictions both rely on the assumption that uncompensated losses suffered by victims lead them to punish tortfeasors even at a cost for themselves, the experiment further tests to what extent this is the case.

Results reveal that while both the negligence rule and strict liability substantially reduced observed rates of costly punishment by victims, strict liability did so by a larger extent, as victims still invested heavily in punishment when they were not, given the chosen due level of care by injurers, entitled to compensation under the negligence rule. Tortfeasors anticipated this effect and overinvested in care and precautions under the negligence rule, but not under

strict liability, when observed levels of care were the ones predicted by theory. Regression results further point to the fact that the experienced loss causes retaliation by victims whenever they are not compensated, providing evidence for the underlying assumption of the model.

There are few experiments in torts and legal liability. In an early one, Kornhauser and Schotter (1990) compared the performance of the negligence rule and of strict liability in the traditional model in which victims cannot react against tortfeasors except for claiming damages. More recently, Angelova and co-authors (2014) investigated both rules when third-parties are affected by accidents but cannot always be compensated by the tortfeasor. Deffains, Espinosa and Fluet (2017) study the interaction between liability rules and social norms when liability is perfectly or imperfectly enforced. Lastly, Guerra and Parisi (2018) investigate the symmetry of incentives created by strict liability and no liability on tortfeasors and victims, and how far they are reflected in their precautionary choices.

The present model and experiment investigate the effect of no liability, negligence, and strict liability on the behavior, firstly, of *victims*, who can react against the tortfeasor beyond just recovering damages. Secondly, it studies the impact of such behavioral patterns on tortfeasors' choice of levels of care and reveals a second social cost involved with the negligence rule in form of overinvestment in care. In doing so, it provides elements for the inclusion of the victim's behavior in the traditional economic analysis of tort law, as well as for its effect on overall social welfare.

The article is structured as follows. Section 2 develops the model and its predictions for the behavior of both parties. Section 3 explains the experimental design and provides details of the procedure. Section 4 presents the obtained results and the parametric analysis that attempts to disentangle the different elements leading victims to punish injurers in repeated interactions. The last section discusses further elements present in real interactions that are abstracted from in the model, but that can interact with its predictions, and concludes.

II. The model

In the traditional unilateral model of accidents (*see* Shavell 2007), let E_t and E_v denote the tortfeasor's and the victim's initial endowments, y the tortfeasor's gains from the activity, x her chosen level of care, c the cost of care, p the probability of accident, D the amount of damages to be paid by the tortfeasor to the victim if an accident happens, and $l(x)$ the losses suffered by the victim in case of accident, with $l'(x) < 0$ and $l''(x) > 0$.

The tortfeasor's and the victim's profits π_t and π_v are given by, respectively,

$$\pi_t = E_t + y - cx - pD \tag{1}$$

$$\pi_v = E_v - pl(x) + pD \tag{2}$$

Following Hart and Moore's (2008) model, let U_t and U_v denote the tortfeasor's and the victim's utility functions, which depend on their respective profits π_t and π_v, as well as on levels of aggrievement. Let r denote the amount of retaliation or punishment chosen by the victim, which imposes a loss m times higher on the tortfeasor. In Hart and Moore's model, aggrievement is equal to the maximum gross payoff that the individual could have achieved, taken over all feasible outcomes (Hart & Moore 2008: 8). In situations involving accidents, the victim's maximum gross payoff is her ex ante wealth, which could be achieved either by not suffering the loss at the first place, or by being fully compensated for the loss once it occurs. With respect to the tortfeasor's action, the victim's maximum gross payoff is not to bear the loss. Therefore, it is assumed that aggrievement is increasing in the loss, with $a_v'\big(l(x)\big) > 0$.

The tortfeasor's and the victim's utility U_t and U_v are given by, respectively,

$$U_t = \pi_t - pmr \tag{3}$$

$$U_v = \pi_v - p\ Max\{\theta_i a_v\big(l(x)\big) - mr, 0\} \tag{4}$$

where $0 \leq \theta_i \leq 1$ is an individual-specific parameter that gives the victim's desire to retaliate, capturing the strength of her reciprocal preferences and her propensity to reciprocate negatively against those who harm her (Levine 1998; Falk & Fischbacher 2006; Cox, Friedman, and Gjerstad 2007). The term $Max\{\theta_i a_v - mr, 0\}$ in (4) implies that the victim's aggrievement of \$1 causes a utility loss of $\$\theta_i$ to the victim, and that the victim can reduce this utility loss by retaliating against the injurer and imposing upon her a loss equal to mr. By assuming that retaliation is costless (something relaxed in the experiment), and that $m > 1$, the victim chooses $r = \theta_i a_v / m$, with $U_v = \pi_v$.

Moreover, and in line with Hart and Moore (2008), it is assumed that $a_v = l(x)$ since θ_i already captures an individual specific parameter that gives the victim's desire to punish. It implies that the victim will not always mechanically translate a loss of \$1 into punishment of \$1. Clearly, victims' aggrievement can depend on other factors, with inequality-aversion (Fehr & Schmidt 1999, Bolton & Ockenfels 2000) as the most prominent one. However, inequality is most often collinear with the loss suffered, as a loss of \$1 caused by the tortfeasor on the victim also creates inequality of \$1 between the payoff of the tortfeasor and of the victim. Therefore, the model is more in line with reciprocity models (Rabin 1993; Levine 1998; Dufwenberg & Kirchsteiger 2004; Falk & Fischbacher 2006; Cox, Friedman, and Gjerstad 2007).

The task of any liability rule is to maximize overall social welfare, which is, assuming that the accident does not impose any negative externality on third parties, and considering that, in equilibrium $r = \theta_i l(x)/m$, given by

$$SW(x, r) = U_t + U_v = E_t + E_v + y - cx - pl(x) - pmr \quad (5)$$

The first-best would be achieved with damages ensuring that $pl'(x) = -c$ and $r = 0$. Let $x^{**} \in arg\ max_{\ x}\ SW = E_t + E_v + y - cx - pl(x) - pmr$.

Under *no liability*, $D = 0$, and the victim chooses r in order to maximize her ex post utility, which depends on the loss she suffers, and is given by $U_v = \pi_v - Max\{\theta_i l(x) - mr, 0\} = E_v - pl(x) - Max\{\theta_i l(x) - mr, 0\}$. In contrast to strict rational choice models, which predict $r^* = 0$, the present model predicts $r^* = \theta_i l(x)/m$.

The injurer, on her turn, chooses x in order to maximize her expected utility given by $EU_t = (1-p)\pi_t + p(\pi_t - mr) = \pi_t - pmr = E_t + y - cx - p\theta_i l(x)$, with the first-order condition $p\theta_i l'(x) = -c$. In contrast to the traditional model of accidents without liability and without retaliation, the injurer chooses $x^* > 0$ for any $\theta_i > 0$. She will consider how increasing care reduces the losses suffered by the victim, and thereby the amount invested in retaliation, which will then reduce the injurer's own losses from retaliation. In other words, she will invest optimally to minimize losses from retaliation, but not losses from accidents. In fact, if $\theta_i = 1$, and all individuals would reciprocate by transferring back all losses they suffer to the tortfeasor, then the latter would invest optimally, with $pl'(x) = -c$. This, however, would still not ensure maximal welfare since whenever an accident occurs, the victim would still retaliate, the injurer would bear the losses, and social welfare would not be maximal.

Under the *negligence* rule, $D = l$ if $x^* < x^{**}$ and $D = 0$ if $x^* \geq x^{**}$, assuming that the law's required level of care corresponds to the socially optimal level x^{**}. The model predicts victims to choose $r^* = 0$ whenever $l = 0$, and hence whenever $x^* < x^{**}$, since in this case the injurer who did not take the minimum required due level of care must fully compensate the victim. However, when $x^* \geq x^{**}$, the victim must bear the losses, and she will retaliate by choosing $r^* = \theta_i l(x)/m$.

The injurer will anticipate this reaction and will choose x to maximize $EU_t = \pi_v - pmr = E_t + y - cx - pD(x) - pmr$. In the absence of the possibility for the victim to punish injurers, the first-order condition for the problem is given by $pl'(x) = -c$, and $x^* = x^{**}$. When

the victim can retaliate, the tortfeasor will choose $x^* \geq x^{**}$.[1] The first-order condition is then $pl'(x)(1+\theta_i) = -c$. In contrast to the traditional result in models that do not consider the victim's reaction, the potential tortfeasor will *overinvest in care* by considering how much she will lose from retaliation if the victim must bear the loss.

Under *strict liability*, $D = l$, and the victim will choose, in order to maximize her ex post utility $U_v = E_v - Max\{\theta_i l(x) - mr, 0\}$, simply $r^* = 0$. She never bears any loss, and hence never retaliates against the injurer.

The injurer maximizes $EU_t = E_t + y - cx - mr\ - pl(x)$. Since, in this case, $l = 0$ for the victim, and independent of the injurer's choice, $r^* = 0$ and the first-order condition is simply $pl'(x) = -c$, which corresponds to the traditional result.

Strict liability can therefore achieve both goals of inducing optimal levels of care and of minimizing retaliation by victims. Negligence, however, will lead to retaliation whenever the victim must bear the losses, but this happens less often than in the absence of any liability, and shall perform better than no liability for this goal. Although this possibility is very remote, negligence could perform worse than no liability when the social losses from overinvestment under negligence are larger than the social losses from underinvestment plus the social losses from higher levels of retaliation under no liability. When comparing liability rules, negligence is predicted to induce overinvestment in care and higher levels of retaliation than strict liability.

---

[1] If she chooses $x \geq x^{**}$, then $D = 0$ and $EU_t = E_t + y - cx - pmr$. Since $r^* = \theta_i l(x)/m$, then $EU_t(x \geq x^{**}) = E_t + y - cx - p\theta_i l(x)$.

If she chooses $x < x^{**}$, then $D = l$ and $EU_t = E_t + y - cx - pl(x) - pmr$. Since $r^* = 0$, then $EU_t(x < x^{**}) = E_t + y - cx - pl(x)$.

Since $0 \leq \theta_i \leq 1$, the assertion holds except for the case in which $\theta_i = 1$, when the injurer would choose any x.

III. The experiment

Subjects took the role, as in the unilateral model of accidents, of either an injurer who engages in an activity from which she profits, and that can cause losses upon the victim that the injurer can reduce by investing in precautions, or of a victim who could punish injurers for their choices (independent of whether an accident occurred or not, and of whether she received compensation or not).

The potential tortfeasor first chose the desired level of care and paid its corresponding cost. Subsequently, both parties observed if the accident occurred, and the victim observed how much the injurer had invested to reduce the loss. In the treatments *negligence* and *strict liability*, the injurer had to compensate the victim for her losses, which depended, as in the model before, on her investment in care. In the treatment *negligence*, she only had to compensate the victim if she chose a level of care below the predetermined due level of care (which was set equal to the socially optimal one). In the treatment *strict liability*, she always had to fully compensate the victim, independent of how much she had invested in care. The victim finally could spend points to reduce the number of points earned by the injurer.

Subjects played a one-shot game in one treatment and then played another 19 rounds of that same treatment with another participant necessarily different from the one they interacted before. They did not receive any feedback at the end of the one-shot game, and the decision of the victim was obtained with the strategy method, such that victims did not receive any information about how that particular injurer behaved, and injurers did not receive any information concerning how much punishment they suffered.[2]

[2] This means that the victim had to specify how many points she would invest in punishment for each possible level of care chosen by the injurer in the one-shot game. In the repeated game implemented afterwards, the strategy method was not used.

Hypotheses

Firstly, the experiment provides a test for the assumption that losses sustained by victims cause retaliation and punishment. It does so not only in the control treatment, and hence when victims never received compensation, but also in the negligence treatment, or when victims only received compensation if the injurer did not invest the socially optimal amount. In treatment strict liability, victims never lost any amount, and losses sustained could not have caused retaliation.

Secondly, the experiment tests whether liability for accidents crowds out retaliatory behavior, and hence whether either negligence or strict liability can reduce rates of retaliation observed in the control group. Moreover, it investigates to what extent strict liability is better apt to crowd out retaliation, given that, under the negligence rule, victims must often bear their losses.

Thirdly, the experiment tests whether the negligence rule leads to overinvestment in precautions and care, given that, if the underlying assumption of the model is correct, injurers must anticipate strictly positive rates of retaliation whenever an accident occurs, increasing the expected costs of accidents above the level which would prevail in the absence of retaliation.

Parameters

The parameters chosen for the experiment were as following. Each subject received an endowment of 120 points, and potential tortfeasors ("player A" in the experiment) engaged in a (fictive) activity that delivered them another 80 points. The marginal cost of care was 10 points, the probability that the activity would impose a loss on the victim ("player B") was of 0.5, and injurers could invest points to reduce the loss that the victim would suffer if the accident occurred as shown in the table 1 (reproduced from the instructions distributed to participants):

Table 1. Care levels of injurers and the resulting losses for victims

If person A invests 0 points, then person B loses 80 points if the loss happens.
If person A invests 10 points, then person B loses 48 points if the loss happens.
If person A invests 20 points, then person B loses 24 points if the loss happens.
If person A invests 30 points, then person B loses 8 points if the loss happens.
If person A invests 40 points, then person B loses 0 points if the loss happens.

The tortfeasor's and victim's payoff functions were

$$\pi_t = 120 + 80 - 10x - pD \quad (6)$$

$$\pi_v = 120 - pl(x) + pD \quad (7)$$

The victim could invest in punishment of the injurer by spending up to 30 points to impose a loss of four points on the injurer for each point she spent (i.e., costly punishment at the standard rate of 1:4, as in most economic experiments). The victim could punish the injurer in all occasions, i.e., when the accident occurred as well as when it did occur.

In the control treatment, the injurer invests x such that $p\theta_i l'(x) = -c$. For the chosen parameters, she will choose, in equilibrium, x such that $l'(x) = -20/\theta_i$. For $\theta_i = 1$, and hence considering a victim who would retaliate by transferring all losses back to the injurer, $x^* = 2$, and for $\theta_i \to 0$, or for pacifist victims with no desire nor tendency to retaliate, $x^* \to 0$.

In the negligence treatment, the injurer invests x such that $(1 + \theta_i)pl'(x) = -c$. For the chosen parameters, she will choose, in equilibrium, x such that $l'(x) = -20/(1 + \theta_i)$. Clearly, for $\theta_i = 0, x^* = 2$, corresponding to the result under strict rational choice assumptions.

In the strict liability treatment, victims never bear any loss and are predicted to choose $r^* = 0$. Injurers should choose the socially optimal level of care $x^* = 2$, or $cx^* = 20$ points.

They deliver the hypotheses:

$$x(negligence) \geq x(strict\ liability) \geq x(control) \quad (8)$$

$$r(control/loss) \geq r(negligence/loss) \geq r(strict\ liability) \quad (9)$$

$$\bar{r}(control) \geq \bar{r}(negligence) \geq \bar{r}(strict\ liability) \quad (10)$$

where $\bar{r}$ denotes average rates of retaliation, and the last inequalities in (10) follow from the fact that $x(negligence) \geq x(control)$, what implies that average losses suffered by victims will be larger in the control group than in the treatment negligence. In fact, unless θ_i takes on the extreme values of 0 and 1, all the weak inequalities in (8), (9), and (10) could be rewritten as strict inequalities.

In the experiment, subjects played one one-shot game in one treatment and then played a repeated game with the same partner, for 19 rounds, in the same treatment. Since the repeated game is finitely repeated, predictions remain the same for both cases, as executing costly punishment is never part of the subgame-perfect equilibrium. The implicit threat of punishment is not credible, and hence strict rational choice models would predict the same behavior in both cases.

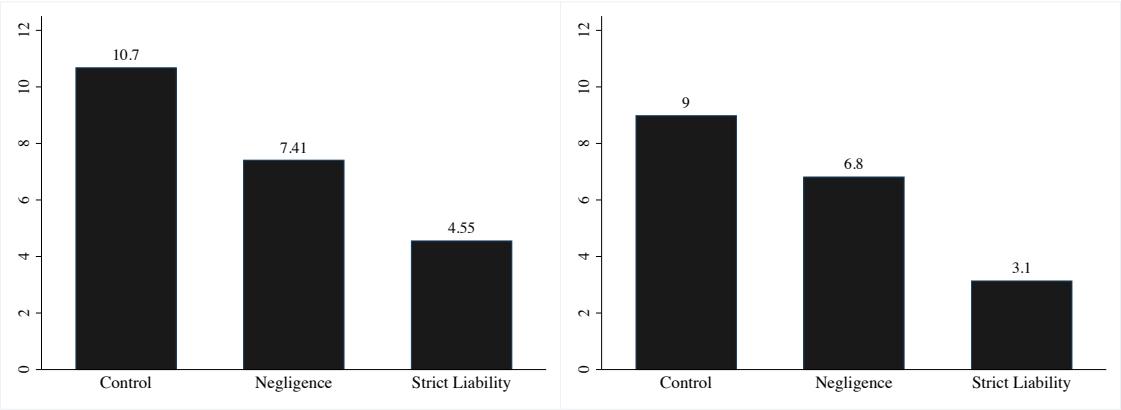
Punishment has no possibility of deterring future conduct by the injurer that can harm the victim in the one-shot games (in other words, potential victims could not use punishment to induce potential tortfeasors to invest in care in future interactions), and subjects were always informed that the subject they were playing with in the current game would never interact with them anymore throughout the whole experiment (even if they did not know the content of the subsequent part of the experiment). In the repeated game with partner matching, and dropping the assumption of common knowledge of rationality, punishment can have that function even if it is not credible. Multivariate regressions presented below disentangle the effect of aggrievement (and, ultimately, of the loss suffered) from the disciplinary effect of punishment (depending on the injurer's chosen level of care) and reveal how they both drive the victim's decision to punish in repeated games.

IV. Results

In total, 168 subjects took part in the nine sessions implemented with z-Tree (Fischbacher 2007) at the Laboratory of Experimental Economics at the *Vysoká škola ekonomická* in Prague in January 2019. A slight majority of the subjects (54%) were males, and the average age of participants was 23 years. They were recruited via ORSEE (Greiner 2015) and earned, on average, 13 Euros in each session that lasted about one hour.

Average rates of punishment and retaliation, in each treatment and type of game, are depicted in graph 1. In both types of games, victims punished injurers more often in the absence of compensation than in its presence. The difference between observed rates of punishment in control and in the treatment strict liability is highly significant in the one-shot game (Mann-Whitney test, N=28 *vs* N=29, one-sided, p = 0.006) and in the repeated game (*id.*, N=28 vs N=29, p = 0.007). While the difference between average punishment in the control treatment and in the treatment negligence is not statistically significant in the one shot game (Mann-Whitney test, N=28 vs. N=27, one-sided, p = 0.167) and in the repeated game (*id.*, N=28 vs. N=27, p = 0.170), regression results presented below, which control for the effect of punishment to induce future higher investments in care by the injurer report a statistically significant effect.

Graph 1. Average punishment



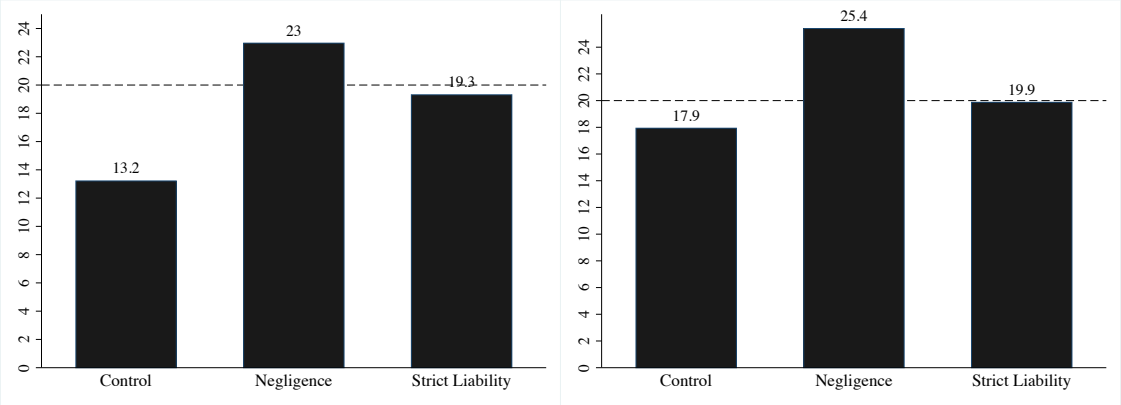
Moreover, strict liability was much better apt to crowd out punishment, and the difference between retaliation in the treatment negligence and in treatment strict liability is significant both in the one-shot game (Mann-Whitney, N=27 vs N=29, one-sided, $p = 0.018$) and in the repeated game (*id.*, N=27 vs N=29, $p = 0.009$).

It is worth noting that the slightly higher observed average rates of punishment in the one-shot games reflect the lower levels of care chosen by injurers in those games. In fact, without repetition, punishment has no function to induce injurers to invest more in care in future rounds and can only be motivated by the aggrievement experienced by the victim. Injurers did invest less in the absence of repetition, as shown below, what lead to higher experienced losses for victims, and therefore higher levels of punishment in the one-shot games.

Average investments in care, considering that the socially optimal amount, and which corresponds to the required level of care in treatment negligence, was 20 points, are depicted in graph 2. Both types of liability rules clearly induced injurers to invest in precautions, with substantially higher levels of care when comparing the negligence treatment with the control group (Mann-Whitney, N=28 vs N=27, one-sided, $p < 0.001$ in the one shot games; *id.*, $p = 0.007$ in the repeated game) as well as when comparing the strict liability treatment with the control group in the one shot games (Mann-Whitney, N=28 vs N=29, one-sided, $p = 0.018$), but not in the repeated games (*id.*, $p = 0.205$). The reason for this last result was the effective use of punishment to induce injurers to invest in the control group, in the repeated game, with quite high – albeit still not optimal (20) – levels of care.

Graph 2. Average care

Average level of care (one shot games) Average level of care (repeated games)



The experiment further provides evidence for the second hypothesis, namely that potential tortfeasors will overinvest in precautions under the negligence rule, but not under strict liability. In the one shot game, although point estimates point in this direction, the difference between retaliation in the negligence treatment and in the strict liability treatment is not significant (Mann-Whitney, N=27 vs N=29, one-sided, p = 0.177), but in the repeated game this is clearly the case (Mann-Whitney, N=27 vs N=28, p = 0.005). In fact, injurers invested optimally under strict liability in the repeated game (19.9 points instead of 20 points).

Regression results presented in table 2 reveal how the negligence rule also had the predicted effect of crowding out retaliation by the victim as soon as further variables are included in the model – i.e., in all models after (1). Still, as depicted in graph 1, the effect of strict liability is stronger. In all models, the results of the two-sided Wald test for the null hypothesis that the coefficients of *negligence* and *strict liability* are equal deliver p values lower than 0.05, except for model (1), where p = 0.054.

Table 2. Regression results on the victim's decision to retaliate

Dep. variable: retaliate	(1)	(2)	(3)	(5)	(6)
Negligence	-2.25	-4.23*	-4.07*	-4.04*	-3.87*
	(1.96)	(2.27)	(2.28)	(2.23)	(2.25)
Strict Liability	-6.06***	-8.79***	-8.71***	-8.52***	-8.43***
	(1.94)	(2.18)	(2.20)	(2.14)	(2.16)
Control # Loss		0.04***	0.04***	-0.02	-0.02
		(0.01)	(0.01)	(0.02)	(0.02)
Negligence # Loss		0.22***	0.22***	0.22**	0.22**
		(0.04)	(0.04)	(0.09)	(0.09)
Control # Loss # Negative				0.02***	0.02***
				(0.00)	(0.00)
Negligence # Loss # Negative				-0.00	-0.00
				(0.02)	(0.02)
Control # Care		-0.28***	-0.27***	-0.27***	-0.26***
		(0.03)	(0.03)	(0.03)	(0.03)
Negligence # Care		-0.12***	-0.12***	-0.12***	-0.12***
		(0.03)	(0.03)	(0.03)	(0.03)
Strict Liability # Care		-0.08***	-0.07**	-0.08***	-0.07**
		(0.03)	(0.03)	(0.03)	(0.03)
Lag(1) Loss			0.00		0.00
			(0.01)		(0.01)
Lag(1) Care			-0.02		-0.02
			(0.02)		(0.02)
Age	-0.01	0.01	0.00	-0.02	-0.02
	(0.25)	(0.24)	(0.24)	(0.24)	(0.24)
Gender	1.25	1.48	1.49	1.21	1.22
	(1.45)	(1.42)	(1.42)	(1.38)	(1.38)
Constant	6.56	9.73	9.97	10.47*	10.70*
	(6.61)	(6.51)	(6.54)	(6.33)	(6.37)
Observations	1,596	1,596	1,596	1,596	1,596
Wald χ^2	28.58	229.1	231.7	246.5	249.1
Prob > χ^2	0.157	0	0	0	0

LEGEND: Mixed effects models with clusters at the individual and session level; all models include period dummies; *** p<0.01, ** p<0.05, * p<0.1

Models (2) and (3) further consider how the loss suffered by the victim – before the decision to punish, i.e., after having suffered a loss, but before spending points to punish the injurer – affects retaliation, as well as how the chosen level of care by the injurer also affects

that decision. The effect of the loss in treatment strict liability is, of course, excluded, as victims never suffered any loss under that rule. Victims punished injurers more when the losses they suffered were higher, and punished injurers less when the injurer had taken higher levels of care. Model (3) further includes lagged variables, and results remain stable.

The last models (4) and (5) include a proxy for θ_i, obtained in the questionnaire implemented at the end of the experiment, namely the subject's answer (from 0 to 10) to *"If I am treated very unjustly, I will take revenge at the first occasion, even if there is a cost to do so."* (Falk et al. 2018).[3] At least in the control treatment, regression estimates point to the fact that retaliation is driven by the interaction between individuals' specific parameters and the loss experienced by victims. In other words, victims punished injurers more when their loss was higher and their individual specific parameter for negative reciprocity was higher.

V. Conclusion

Liability for accidents and the provision of compensation for victims' losses that arise from accidents not only provide incentives for potential tortfeasors to invest in precautions and care, thereby deterring accidents, but also provide incentives for victims not to take matters into their own hands, thereby avoiding a deadweight loss from decentralized forms of punishment. Punishment can surely also deter accidents, as revealed in treatment negligence, when injurers overinvested in care in order to avoid losses from punishment, and did so beyond what legal liability would require. Punishment, however, can only achieve this goal of deterrence at a very high costs, given by the losses incurred by its target and the costs incurred by victim to punish

[3] The questionnaire included all six questions of the "streamlined version of the preference module" in the appendix of Falk et al. (2018), including therefore a question on risk-taking, time discounting, trust, altruism, positive reciprocity, and negative reciprocity. While the focus concerns only negative reciprocity, all questions were asked in order to distract subjects from focusing on only one question, and perhaps inferring that this one would be compared by the experimenter with the subject's own previous behavior in the experiment.

whenever punishment is implemented. Compensation, in contrast, consists only in a monetary transfer from one party to the other, and is hence apt to achieve deterrence without those costs.

In a model with activity levels, retaliation might have another beneficial effect. It is apt to induce injurers to engage in lower than the predicted level of activity, which is, under the negligence rule, above the socially optimal one (Shavell 1980). If injurers anticipate that victims will punish them once accidents occur, then they might choose to engage less in the activity because the more they do so, the more accidents will occur, and the higher their losses from retaliation will be. This is a potential positive effect of retaliation under negligence that might balance its negative effect of inducing overinvestment in care. Still, for this beneficial effect, victims must spend resources to harm the tortfeasor, and the question then becomes what is more costly, whether the losses from punishment or the losses from too high levels of activity.

Another important factor in reality is the possibility for injurers to "buy off" vengeance and retaliation. The injurer can always, once an accident occurs, voluntarily attempt to reduce the victim's aggrievement by offering her some side-payment. This can, for sure, decrease the willingness to retaliate, avoiding social losses. It is, however, very often impossible to do so, such as in cases in which victims are many and the loss disseminated, or when victims differ in their feelings of entitlement and each of them would demand a different amount as compensation. Moreover, the behavioral reaction to retaliate is often taken at the heat of the moment, such as in traffic accidents, and it is an open question how far this state of mind will allow parties to successfully negotiate a settlement before the victim acts.

While apt to reduce punishment, side-payments do not solve the overinvestment problem. Injurers must anticipate that, in case of an accident, they will have to spend money in side-payments to avoid losing money from retaliation. They must still pay the victim, under the negligence rule, even if they took the due level of care, if they want to bribe her not to retaliate,

and this additional expected cost of accidents, above and beyond the legal one, will still lead them to overinvest in care.

References

Angelova, Vera, Olivier Armantier, Giuseppe Attanasi, and Yolande Hiriart (2014), "Relative performance of liability rules: experimental evidence", 77 THEORY AND DECISION 531.

Bolton, Gary & Axel Ockenfels (2000), "ERC: A Theory of Equity, Reciprocity, and Competition", 90 AMERICAN ECONOMIC REVIEW 166.

Brown, John (1973), "Toward an Economic Theory of Liability", 2 JOURNAL OF LEGAL STUDIES 323.

Calabresi, Guido (1965), "The Decision for Accidents: An Approach to Nonfault Allocation of Costs", 78 HARVARD LAW REVIEW 713.

___ . (1968), "Does the Fault System Optimally Control Primary Accident Costs?", 33 LAW AND CONTEMPORARY PROBLEMS 429.

Cox, James, Daniel Friedman and Steven Gjerstad (2007), "A Tractable Model of Reciprocity and Fairness", 59 GAMES AND ECONOMIC BEHAVIOR 17.

Deffains, Bruno, Romain Espinosa and Claude Fluet (2017), "Law and Norms: Experimental Evidence with Liability Rules", working paper.

Dufwenberg, Martin and Georg Kirchsteiger (2004), "A Theory of Sequential Reciprocity", 47 GAMES AND ECONOMIC BEHAVIOR 268.

Falk, Armin and Urs Fischbacher (2006), "A Theory of Reciprocity", 54 GAMES AND ECONOMIC BEHAVIOR 294.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman & Uwe Sunde (2018), "Global Evidence on Economic Preferences", 133 QUARTERLY JOURNAL OF ECONOMICS 1645.

Fehr, Ernst & Klaus Schmidt (1999), "A Theory of Fairness, Competition, and Cooperation", 114 QUARTERLY JOURNAL OF ECONOMICS 817.

Fischbacher, Urs (2007), "z-Tree: Zurich toolbox for ready-made economic experiments", 10 EXPERIMENTAL ECONOMICS 171.

Greiner, Ben (2015), "Subject pool recruitment procedures: organizing experiments with ORSEE", 1 JOURNAL OF THE ECONOMIC SCIENCE ASSOCIATION 114.

Guerra, Alice and Francesco Parisi (2018), "Accident Aversion: an Experiment", working paper.

Hart, Oliver and John Moore (2008), "Contracts as Reference Points", 123 QUARTERLY JOURNAL OF ECONOMICS 1.

Kornhauser, Lewis and Andrew Schotter (1990), "An Experimental Study of Single-Actor Accidents", 19 JOURNAL OF LEGAL STUDIES 203.

Levine, David (1998), "Modeling Altruism and Spitefulness in Experiments", 1 REVIEW OF ECONOMIC DYNAMICS 593.

Posner, Richard (1972), "A Theory of Negligence", 1 JOURNAL OF LEGAL STUDIES 29.

Rabin, Matthew (1993). "Incorporating Fairness into Game Theory and Economics", 83 AMERICAN ECONOMIC REVIEW 1281.

Shavell, Steven (1980), "Strict Liability Versus Negligence", 9 JOURNAL OF LEGAL STUDIES 1.

___ . (2007), "Liability for Accidents", *in* HANDBOOK OF LAW & ECONOMICS, vol. 1 (Mitchell Polinsky and Steven Shavell, eds.), 142.