CENTRE DE RECHERCHE
SUR LES STRATÉGIES ÉCONOMIQUES

Relative Performance of Liability Rules: Experimental Evidence

Vera Angelova, Olivier Armantier, Giuseppe Attanasi, Yolande Hiriart

March 2013

Working paper No. 2013 – 03

CRESE
30, avenue de l'Observatoire
25009 Besançon
France
http://crese.univ-fcomte.fr/

UNIVERSITÉ DE FRANCHE-COMTÉ

Relative Performance of Liability Rules: Experimental Evidence*

Vera Angelova,** Olivier Armantier,† Giuseppe Attanasi,‡ Yolande Hiriart††

March 12, 2013

Abstract

We compare the performance of liability rules for managing environmental disasters when third parties are harmed and cannot always be compensated. A firm can invest in safety to reduce the likelihood of accidents. The firm's investment is unobservable to authorities. The presence of externalities and asymmetric information call for public intervention in order to define rules aimed at increasing prevention. We determine the investments in safety under *No Liability*, *Strict Liability* and *Negligence* rules, and compare these to the first best. Additionally, we investigate how the (dis)ability of the firm to fully cover potential damage affects the firm's behavior. An experiment tests the theoretical predictions. In line with theory, *Strict Liability* and *Negligence* are equally effective; both perform better than *No Liability*; investment in safety is not sensitive to the ability of the firm to compensate potential victims. In contrast with theory, prevention rates absent liability are much higher, and liability is much less effective, than predicted.

JEL classification: D82, K13, K32, Q58.

Keywords: Risk Regulation, Liability Rules, Incentives, Insolvency, Experiment.

*We thank David Alary, Marie Obidzinski, participants at the *15th Conference on the Foundations and Applications of Utility, Risk and Decision Theory* (FUR) in Atlanta, the *10th Annual Conference of the German Law and Economics Association* in Magdeburg, *19th Annual Conference of the European Association of Environmental and Resource Economists* (EAERE) in Prague, the *French Experimental Economics Association Annual Meeting* (ASFEE) in Montpellier, the *LAMETA Seminar* (University of Montpellier), the *CREM Seminar* (University of Rennes), the *CES-ifo Conference on Law and Economics* in Munich, the *ESREL Annual Conference* in Troyes, the *20th SRA-Europe Meeting* in Stuttgart and the *International Workshop on Economic and Financial Risks* in Niort, as well as two anonymous referees for their useful comments and suggestions. The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) Grant Agreement no. 230589. Financial support from the *Foundation for an Industrial Safety Culture* (*FONCSI*) and from the French *Agence Nationale pour la Recherche (ANR)* for the project *Environmental Regulation and Market Imperfections* is also acknowledged, together with CESifo sponsorship.

**Technical University Berlin, Germany: vera.angelova@tu-berlin.de.

†Federal Reserve Bank of New York, US: olivier.armantier@ny.frb.org.

‡University of Strasbourg and Toulouse School of Economics: attanasi@unistra.fr.

††Université de Franche-Comté and Institut Universitaire de France: yolande.hiriart@univ-fcomte.fr.

1 Introduction

We study the design of a suitable public policy for managing environmental disasters. Beginning in the 1970s, there has been a major wave of health, safety and environmental regulation. Pioneered by the United States, this has led not only to the establishment of new regulatory agencies with broad responsibilities for risk and environmental policy, but also to an increase in the importance of courts of law (see Viscusi, 2007). The rationale for the later is twofold. First, liability is often viewed as a successful legal response to finance the remediation of hazardous sites or to indemnify victims (compensation role). Second, it may also create incentives for prevention by inducing private actors to internalize environmental damage (incentive role). Both dimensions are valuable, in particular if one does not want to use public funds for site restoration, a common practice until now in Europe.

There are a number of ways of attributing liability.[1] The field of environmental risk does not depart from more general contexts of accident law in its use of *Strict Liability* and *Negligence* as the two main means of holding (or not holding) the responsible party liable for damage. A quick look at the main North American or European laws reveals that these two liability rules prevail. For instance, the *Comprehensive Environmental Response, Compensation, and Liability Act* (*CERCLA*), enacted in 1980 in the U.S.A., is a *Strict Liability* rule forcing any responsible party to pay for the cleanup of contaminated sites.[2] The U.S. Oil Pollution Act of 1990 is another example.[3] The Euro-

[1]See Posner (1992) or Shavell (2004) for a textbook description of liability rules.

[2]Under *CERCLA*, liability is strict and (at least presumptively), it is joint and several. "A potentially responsible party (*PRP*) cannot simply say that it was not negligent or that it was operating according to industry standards. If a *PRP* sent some amount of the hazardous waste found at the site, that party is liable" (http://www.epa.gov/compliance/cleanup/superfund/liability.html). Such liability is said to be strict because exercise of due care is not a defense. Although *CERCLA* does not explicitly provide for Joint and Several Liability, Congress confirmed when passing the Superfund Amendments and Reauthorization Act (*SARA*) amendments in 1986 that it was its intention. Accordingly, all of the responsible parties at a cleanup site are said to be jointly and severally liable for the response costs at the site. "Any one potentially responsible party may be held liable for the entire cleanup of the site (when the harm caused by multiple parties cannot be separated)". In practical terms, if five out of twenty responsible parties are meaningfully solvent at the time that response costs are sought under Section 107 of *CERCLA*, these five may be liable for the entire cost of the cleanup (see Ashford and Caldart 2008, page 761). *CERCLA* is a good illustration for our analysis, excluding the case of harm with multiple responsible parties where the contribution of each cannot be distinguished.

[3]When oil is discharged from a vessel or facility into the navigable waters of the United States,

pean Community's 2004 Directive for contaminated sites is a *Negligence* rule.[4] Despite this international tendency toward liability regimes for environmental damage, there is no general agreement on the rationale for relying on them (see Faure and Skogh, 2003). In particular, it is widely accepted that the insolvency of potential injurers is a serious impediment to the effectiveness of any liability rule.[5] It is thus essential to understand the behavior of judgment-proof firms (i.e. firms whose assets cannot fully cover potential damage) when subjected to these policies.

In this paper we investigate which liability rule is most effective in reducing the probability of an accident. Furthermore, we study the role of insolvency, i.e. whether the firm's willingness to invest in safety depends on the ability of the firm to compensate third parties.[6] More specifically, we compare the performance of *No Liability*, *Strict Liability*, and *Negligence* rules enforced against firms that can potentially harm third parties (the environment, or human beings in their health or property). In our analysis, we assume that the firm does not directly suffer damage when an accident occurs: only third parties, who do not have any contractual or market relations with the firm, suffer harm. Employees of the firm and consumers of the firm's products are thus excluded from our analysis. Notice that we restrict attention to unilateral accidents. While firms (potential injurers) have influence on the probability of occurrence of the harm, third parties (potential victims) play a passive role: they have no means to affect the probability and the size of the damage.[7]

---

adjacent shorelines, or the exclusive economic zone, the Oil Pollution Act makes each "responsible party" liable for "removal costs" and "damages". The Oil Pollution Act explicitly adopts the standard of liability of section 311 of the Clean Water Act, knowing that courts have consistently construed section 311 to establish Strict Liability (see Murchison 2011).

[4] *Directive 2004/35/CE on environmental liability regarding the prevention and compensation of environmental damage*, adopted by the European Parliament and Board of Ministers on April 21, 2004.

[5] Other limits include the low probability of a suit, the difficulty of proving causality between the decisions of injurers and harm, the time required, scientific ignorance, and uncertainty about court judgments stemming from mistakes or judges' subjectivity (Shavell, 1984b). Furthermore, liability may change the contractual or market relationships in risky sectors, which may lead to under-investment (Hiriart and Martimort, 2006b). On the other hand, liability is a very natural way to align private and public interests; there is therefore a strong tendency to introduce it as a part of traditional regulation all over the world.

[6] Since we are talking about the firm and third parties, the reader may ask who the second party is: the second party is the public authority.

[7] See Shavell (2004) for *unilateral* accidents in the Law and Economics literature.

The focus of our decision model is a potential disaster due to the firm's moral hazard when investing in prevention. With a small probability, the firm causes a great deal of damage to third parties; however, the firm can reduce the likelihood of such an accident by investing in safety measures. The firm's safety investments could include, for example, buying new equipment, educating and training employees, or increasing watchfulness. Safety measures taken by the firm are not directly observable by the authorities. Since prevention is both costly to the firm and unobserved by the rest of the world, we model safety care as a moral hazard variable. The potential externality caused by a disaster, together with asymmetric information, require authorities to provide incentives for the firm to reduce risk. Public intervention takes the form of liability rules, which induce the firm to reduce risk and/or to compensate the victim in case damage occurs. We assume that if an accident occurs, victims lose their entire wealth. The injurer's assets, however, may not suffice to fully compensate victims. Since the firm is protected by limited liability, the firm can only be held liable for damage up to the value of her wealth.[8]

In our theoretical model, we determine the conditions leading the firm to invest in prevention under *No Liability*, *Strict Liability* and *Negligence* and compare these conditions to the first-best. An experiment is implemented to test the main theoretical insights. Our experimental results show that, in line with the theory, both *Strict Liability* and *Negligence* perform better than *No Liability*: the firm increases her level of safety care under these rules. Furthermore, there is no significant difference in the effectiveness of *Strict Liability* and *Negligence*, confirming another theoretical result. Finally, investment in safety does not change when the firm is unable to cover the losses of third parties. We also find two results which are not in line with the theoretical predictions: prevention rates absent liability are much higher, and liability is much less effective, than predicted.

This paper belongs to the Law and Economics literature devoted to tort law; specif-

[8]Although grammatically incorrect, when talking about the firm we will use the pronouns "her" and "she" in order to be consistent with the literature.

ically, to the public control of agents that can potentially and unintentionally harm third parties. We investigate the incentives provided by public authorities to foster prevention when safety care is unobservable. In this sense, this paper also relates to the Principal-Agent literature.[9] At the same time, our paper differs from this literature because the set of instruments available to public authorities (the Principal) is restricted to fall within the definition of liability rules. More in line with the Law and Economics literature, in this setting there is no direct and personalized regulation (such as a regulatory contract). Liability rules are common knowledge and apply equally to all agents. The theoretical model of this paper is related both to Shavell (1980), and Shavell (1986). Shavell (1980) analyzes thoroughly *Strict Liability* versus *Negligence* rules, while Shavell (1986) provides insights on the judgment-proof problem.

The original part of our work remains the empirical one. The empirical literature on environmental liability rules is small. The few econometric studies that exist cover topics such as the adoption of *Strict Liability* within the U.S.A. (Alberini and Austin, 1999a), the effectiveness of *Strict Liability* when handling toxic spills (Alberini and Austin, 1999b), and how firms escape *Strict Liability* (Alberini and Austin, 2001, 2002). To the best of our knowledge, there is no study that compares the relative performance of liability rules. We aim to fill this gap. Since the variable of interest, in this case the investment in safety, is not observable in the field, we make use of the experimental method.

To date, there are not many experiments on liability rules. King and Schwartz (1999, 2000) and Dopuch and King (1992) study the special case of liability rules for auditors. Dopuch, Ingberman and King (1997) explore liability rules applied to the multi-defendant case, namely *proportionate* versus *joint and several* liability rules. Wittman et al. (1997) investigate the learning of liability rules. The experimental study which is most similar to ours is by Kornhauser and Schotter (1990) (KS thereafter). The main

[9] See Pitchford (1995), Newman and Wright (1990), or Hiriart and Martimort (2006a), although all these papers study essentially *extended liability*.

difference between our framework and theirs is that we consider accidents of substantial size compared to the injurer's level of assets. Disasters pose a particular problem for public authorities due to the frequent insolvency of the responsible parties. Our main contribution is to shed light on the level of prevention when potential injurers cannot fully compensate victims. This is a question that, to our knowledge, has not yet been the object of experimentation. Three important design issues set our study apart from KS. First, in our experiment there is a real third party, i.e. a subject who is sitting in the laboratory and can potentially suffer losses resulting from the behavior of the injurer. In KS, if an accident occurs, the injurer is called upon to pay for the damage but no one gets "hurt". Second, subjects in our experiment get their endowment through a real effort task. They are thereby induced to perceive the money at risk as their own, which makes the decision situation more realistic and may lead to a different behavior, as compared to the case where the money at risk is provided as a windfall by the experimenter. Third, KS lack a *No liability* treatment, for which we find rather intriguing results. It should also be noted that we have adopted a binary level of care, whereas KS modeled care as a "continuous" variable. Given these differences in the experimental design, it is not surprising that our results differ. While we find evidence in favor of equivalence between *Strict Liability* and *Negligence*, KS do not find such equivalence, even when the standard of due care is set at the optimal level.

The paper is organized as follows. Section 2 presents the model. The experimental design, procedures, and behavioral predictions are described in Section 3. Results are given in Section 4. Section 5 briefly concludes by pointing out alleys for further work.

2 The Model

A firm can cause damage of a given size h to third parties (human beings and/or the environment). The firm can exercise some safety care $e \in \{0, 1\}$, i.e. invest in safety, in order to reduce the probability of an accident from p_0 to p_1 (both being in $[0, 1]$), with

$\Delta p = p_0 - p_1 > 0$. The firm has initial assets w_0. Let us denote by w_t the assets at time t. It costs an amount $c > 0$ to the firm to invest in safety ($e = 1$), whereas not investing in safety ($e = 0$) costs nothing. The level of investment in safety is privately known by the firm: it is neither observable by public authorities, nor by third parties.[10]

In each period t the firm decides whether to invest in safety or not. Each unit invested in safety decreases the firm's remaining wealth by the same amount. The firm's wealth is large compared to the total amount she may invest in safety, even if she would invest in each period: the intertemporal dimension is eliminated as there is no binding constraint in the firm's period-by-period problem. The investment decision in each period is thus independent from the investment decision in the other periods. The framework reduces to a static decision model, which is repeated a finite number of times.

When an accident takes place, the firm is held liable or not to pay for the harm caused, depending on the liability rule.

In the absence of liability, i.e. a *No Liability* (henceforth *NoL*) rule, the firm does not pay anything, and the third party bears the losses.

Under a *Strict Liability* (henceforth *SL*) rule, the firm responsible for the harm caused to a third party must compensate the third party, regardless of the firm's behavior in the conduct of the operations that have led to the damage. In other words, even if the firm has been cautious and the damage could not have been avoided by the exercise of due care, the third party must be compensated.[11]

Under a *Negligence* (henceforth *Ne*) rule, the firm is not held liable for the harm caused, unless she is found to be negligent. That is, the firm is liable only if she has not satisfied a standard of due care in the conduct of the operations that have led to the damage.[12] Here the standard of due care is set at $e = 1$.

[10] This model is an adaptation of Shavell (1984a).

[11] See Shavell (2004).

[12] An injurer firm is held liable for losses if her level of care is less than a level called "*due care*" specified by the courts. See Posner (1992).

■ *Social Optimum*. Investing in safety is socially optimal if the expected social cost of investing is smaller than the expected social cost of not investing:

$$p_1 * h + c \leq p_0 * h.$$

This condition can be rewritten as

$$c \leq \Delta p * h. \tag{1}$$

Investing in safety care is thus socially optimal when the prevention cost borne by the firm is smaller than the incremental expected harm affecting third parties. Let us assume that this inequality holds in what follows: the objective of public authorities is then to implement this high level of care. Since investment in safety is unobservable, the best authorities can do is to impose policies so as to induce the firm to exercise $e = 1$. This is the role assigned to liability rules. We will first assess the efficiency of these rules from a theoretical viewpoint before testing the theory with an experiment. To this end, we characterize the circumstances under which the firm invests in safety care: if a liability rule induces the firm to exercise $e = 1$ in any circumstances, then the rule is socially efficient.

We now characterize the firm's cost-minimizing choice under each liability rule. Whatever the rule, the firm invests in safety if the expected private cost of investing is smaller than the expected private cost of not investing.

■ Under an *NoL* rule, the firm invests in safety care if

$$p_1 * 0 + (1 - p_1) * 0 + c \leq p_0 * 0 + (1 - p_0) * 0,$$

which never holds true since $c > 0$. Hence, the firm is never induced to exercise $e = 1$: an *NoL* rule is always inefficient.

■ Under an *SL* rule, a responsible firm with assets w_t at time t has to pay an amount equal to $\min\{h, w_t\}$ since the firm is protected by limited liability. The firm thus invests at time t if

$$p_1 * \min\{h, w_t\} + (1 - p_1) * 0 + c \leq p_0 * \min\{h, w_t\} + (1 - p_0) * 0,$$

which can be rewritten as

$$c \leq \Delta p * \min\{h, w_t\}. \tag{2}$$

Comparing (2) to (1), it is clear that an *SL* regime will induce the proper investment in safety (or choice of care) when the firm is rich enough. The firm will always invest in safety when her assets are sufficient to cover the external harm, i.e. when $w_t \geq h$. In this favorable case, the firm exercises $e = 1$ and, if an accident occurs, the firm is able to (and will) fully compensate the third party for losses. The firm will also invest when her assets fall in a medium range, i.e. when $w_t \in \left[\frac{c}{\Delta p}, h\right)$: in this range *SL* provides sufficient incentives for investment, although the firm will compensate the third party only partially. Conversely, when the firm's assets fall below the threshold $\frac{c}{\Delta p}$, the firm will not invest in safety. In this case an *SL* regime is inefficient and the compensation provided by the firm to the harmed third party will be lower than in the medium range case.

■ Under an *Ne* rule, the firm is held liable only if she did not exercise $e = 1$. Hence, the firm invests in safety if

$$p_1 * 0 + (1 - p_1) * 0 + c \leq p_0 * \min\{h, w_t\} + (1 - p_0) * 0,$$

which can be rewritten as

$$c \leq p_0 * \min\{h, w_t\}. \tag{3}$$

Comparing (3) to (2), it is clear that an *Ne* rule will induce the firm to invest in safety more often than an *SL* rule: (3) is less demanding than (2), since $p_0 \geq \Delta p$.[13] There are

[13]This means that the firm's choice $e = 1$ will be induced by an *Ne* rule for a larger set of parameters, i.e. also for $w_t \in \left[\frac{c}{p_0}, \frac{c}{\Delta p}\right)$.

now two relevant intervals of the firm's wealth that determine the rule's efficiency and the extent of the compensation. When $w_t \geq \frac{c}{p_0}$, the firm takes care and, since she has satisfied the standard of due care, she does not have to compensate the third party. When the firm's wealth is below the threshold $\frac{c}{p_0}$, the firm does not take care and has to compensate the third party for the harm caused, although she will do it only partially.

The results above hold for risk-neutral firms. The analysis for risk-averse firms using CARA or CRRA utility functions is provided in Appendix A. Under risk aversion, the conditions for which the firm chooses to invest in safety under each liability regime are slightly different but the qualitative results hold. Namely, the firm should not invest in safety under *NoL* but should do so under *SL* or *Ne*, provided that both the prevention cost and the – relative or absolute – degree of risk aversion are not too large. *Ne* induces prevention for a larger set of parameters than *SL*.

3 Experiment

3.1 Design

We implemented six distinct treatments (see Table 1).[14] They differed in the type of liability rule (*NoL*, *SL*, or *Ne*) and in whether the harm potentially caused could be fully compensated by the injurer. In the *low* (*high*) damage treatments, the potential injurer would always (never) be able to fully compensate the potential victim.

	low damage	*high* damage
No Liability (*NoL*)	*NoL-low*	*NoL-high*
Strict Liability (*SL*)	*SL-low*	*SL-high*
Negligence (*Ne*)	*Ne-low*	*Ne-high*

Table 1: Treatments

[14] All data and instructions are available upon request.

The experiment consisted of two phases with 5 periods each. Phase 2 was merely a repetition of phase 1. One of the phases was randomly selected for payment at the end of the experiment.

The 30 participants were randomly assigned to roles (*A* and *B*) and *A*-*B* pairs at the beginning of each phase.[15] One can think of subject *A* as the firm or potential injurer and of subject *B* as the third party, or potential victim.[16] Just like in the model, the third party was passive in all treatments of the experiment: *B* was going to be affected by the decisions of *A*, but could not do anything to influence them. Since there was no interaction between *A* and *B*, subjects in the experiment faced an individual decision-making situation in a non-strategic set-up. In addition, the investment decision remained private information to *A*: no other participants could observe it.

Subjects were not informed until the end of the phase about their role, so that first everyone was asked to decide as if they had been assigned the role of *A*. This way we collected 30 individual decision paths as *A* for each phase and each treatment. At the end of each phase, subjects were informed about their actual role. The decision path of the subject who had been assigned the role of *A* became relevant for the payoff outcome of the *A*-*B* pair. In contrast, the decision path of *B* was irrelevant for the payoff outcome of the same *A*-*B* pair.

In each of the five periods, with a probability of 5%, an accident[17] would occur and lead to a loss of endowment by *A*, *B*, or both, depending on the treatment. In each period, subject *A* was asked to decide whether she wanted to reduce this probability to 1% by investing 1 ECU (Experimental Currency Unit). An endowment of 40 ECU was given to subject *A* at the beginning of each phase. Subject *B*'s endowment depended on the treatment: it was 30 ECU in the *low* damage treatments and 50 ECU in the *high* damage treatments. When an accident occurred, the victim lost her entire wealth. In the *low* damage treatments, *A* was able to fully compensate *B* in every period t of the

[15] Phase 1 was independent from phase 2 in both the random draw of roles, and the random assignment to pairs.

[16] To keep a neutral frame, we never used the terms "injurer" and "victim" in the instructions.

[17] Instead of an "accident" we referred to an "event" in the instructions, to maintain the neutral frame.

phase, independent of *A*'s investment in prevention. Conversely, in the *high* damage treatments, *A* was never able to fully compensate *B*.

In each period, each subject first made her decision and then learned whether she was hit by an accident. Subjects also learned how many other subjects were hit. This information was restricted to the subject's "feedback group", i.e. to 14 other subjects with whom the subject in question had not been paired.[18]

Whenever a subject was hit by an accident, that subject was not allowed to make decisions in the remaining periods of the current phase.[19]

At the end of the phase, subjects were informed about their actual role, whether their pair was hit by an accident and the resulting payoffs. If the pair was not hit by an accident, *A*'s payoff was her initial endowment less her total prevention cost (the ECU paid for investment in safety in the five periods of the phase). *B* simply kept her initial endowment. In the case of accident, payoffs depended on the treatment. First, *B* lost her initial endowment (30 ECU in the *low* damage treatments and 50 ECU in the *high* damage treatments). Then,

- in the *NoL treatments*, *A* was not required to compensate *B*. Therefore, *A*'s payoff amounted to her initial endowment less her total prevention cost, independent of *B*'s initial endowment. *B* was left with nothing.
- in the *SL treatments*, *A* was required to compensate *B* up to the level of *A*'s remaining wealth, independent of whether *A* had invested in safety in that period or not.[20] In treatment *SL-low*, *B* received 30 ECU from *A*. *A* was left with a positive amount of money: 40 ECU less the 30 ECU compensation less the total prevention cost (at most 5 ECU). In the *SL-high* treatment, *B* received from *A* 40 ECU less *A*'s total prevention cost (i.e., she received at most 40 ECU, which is

[18]Recall that until the end of the phase subjects did not know their role in the pair. Therefore, receiving information about an accident occurring to a subject potentially paired with oneself would unnaturally influence one's decision path.

[19]An accident introduces an asymmetry in the decision situation. Before the accident, *A* holds her endowment less the cost for investment in safety. After the accident, depending on the liability rule, *A* may not have any more resources to invest in safety, even if she wants to.

[20]Since *A* was protected by *limited liability*, *A* was not asked to give to *B* more than what *A* owned.

10 less than what *B* initially possessed). *A* was left with nothing since her entire wealth was transferred to *B*.

- in the *Ne treatments*, payoffs were conditional on *A*'s decision in the period of the accident. If *A* had invested in safety in the period of the accident, *A* was not required to compensate *B*: by investing 1 ECU, *A* had complied with the standard of due care and was not liable for the harm caused. Hence, payoffs were exactly as in the *NoL* treatments. If *A* had not invested in safety in the period of the accident, *A* was required to compensate *B* up to the level of *A*'s wealth. The resulting payoffs were, therefore, the same as in the *SL* treatments.

The last task was a post-experimental questionnaire consisting of three hypothetical questions. They aimed at collecting information about the risk attitudes of subjects and their perceptions of others' and own selfishness.[21]

At the end of the experiment, payoffs were converted from ECU into euros at the exchange rate of 6 ECU = 1 euro. A show-up fee of 2.5 euros was added to that amount and paid to subjects in cash.

3.2 Procedures

We performed one session per treatment, for a total of six sessions. A total of 192 undergraduate students from the University of Jena (32 per session) participated in this experiment. They were recruited with the online recruitment system for economic experiments ORSEE (Greiner, 2004). Another 64 subjects took part in the two pilot sessions. On average, participants earned 8.65 euros and spent 60 minutes (including 15 minutes of written and oral instructions) in the laboratory of the Max Planck Institute of Economics in Jena (Germany).

[21] The translation from German read: 1. "How do you judge yourself: are you generally a risk loving person, or do you try to avoid risks?"; 2. Would you say that most of the time people try to help others or only follow their own interests? 3. Would you say that most of the time you try to help others or only follow your own interests?. The answer to 1. was on a scale from 0 (very risk-averse) to 10 (very risk-loving). Answers to 2. and 3. were on a scale from 0 (help others) to 6 (follow own interests).

Upon arrival in the laboratory, subjects were randomly assigned to a cubicle with a computer. Initially, subjects received instructions only for Part I of the experiment, which consisted of a real effort task they had to perform in order to "earn" their endowment for Part II. Earning the initial endowment was equivalent to earning the right to participate in Part II, which was the main experiment. Subjects were given five minutes in which to solve the maximum number of mathematical tasks: in this case, summing up five two-digit numbers (as in Niederle and Vesterlund, 2007). At the end of this real effort task subjects were ranked according to the number of math tasks they had solved correctly. The 30 best performers received their initial endowment and instructions for Part II.[22] The two worst performers did not earn an endowment and had to leave the laboratory. They were compensated 3 euros each.

For both parts of the experiment, after subjects read the instructions individually, instructions were also read aloud by the experimenter. The experimenter clarified the instructions in private, when necessary. Additionally, before Part II of the experiment began, subjects answered a list of questions to verify their understanding of the instructions. Part II did not start until everyone had answered all the questions correctly. The experiment was programmed in z-Tree (Fischbacher, 2007).

In Part II of the experiment, the realization sequences for accident/no accident for each subject, under both the 1% and 5% probability regimes, were drawn in advance. Within a treatment, each subject faced an independent sequence of realizations. In order to ensure the comparability of treatments, subjects with the same identification number in different treatments (e.g. subjects with number 1 in all treatments) were confronted with the same sequence of realizations. To help subjects calculate the objective probability of being hit by an accident in one of the remaining periods of the current phase, we supplied them with an on-screen calculator.[23] Subjects could enter their planned decisions until the end of the phase into the calculator, which would return the probability of an accident occurring and the complementary probability

[22]See Appendix B for an English translation of the instructions.
[23]See Instructions in Appendix B for a picture of the on-screen calculator.

of no accident occurring in that time. Displaying both probabilities insured that we did not influence subjects in an optimistic or a pessimistic way. The calculator was accessible to subjects at all times.

During the experiment, eye contact was not possible. Although participants saw each other at the entrance to the lab, there was no way for them to guess with which person(s) from the crowd of 30 students they would be matched in the two phases of Part II of the experiment. Most subjects were experienced – only 10 out of 256 had never participated in an experiment before.

3.3 Discussion of Design and Procedures

The following is a critical discussion of some key features of our experimental design.

The purpose of the real effort task in Part I was to make subjects "earn" their endowment and hence perceive the money at risk as their own or as that of the potential victim. It is true that the chance of "earning" this endowment – and with it the right to participate in Part II of the experiment – is very high ($30/32 = 0.9375$). However, participation is not guaranteed, and does ultimately depend on subjects' abilities and efforts in the specific task.

About Part II, the predetermination of the random sequence of accidents has advantages and disadvantages. As described above, these random draws were made prior to the experiment, with the explicit goal of having the same set of random draws in each treatment. This strategy reduces noise when comparing behavior across treatments. On the other hand, there is a risk that, as subjects do not visually observe the random draws, this may lead them to doubt the true randomness of the events. The feedback provided about accidents that occurred to others in the previous period is one way to counteract this. The primary reason we opted for general feedback about accidents is to add more realism to the experiment.[24] Nonetheless, this may

[24] In reality information about major environmental disasters is usually provided in the form of statistics in the news and may lead to a change in the behavior of potential injurers.

have helped subjects to check, period by period, that other subjects were indeed hit by accidents within their "feedback group" and that the probability of an accident occurring was actually between 1% and 5%.

In designing the decision setting, we opted for a repeated set-up with a restart (i.e. phase 2) to check whether decisions of experienced subjects differed from those of inexperienced ones.[25] This offers a control for learning. Notice that Bayesian updating may happen due to the information released between periods about the number of other subjects hit within one's "feedback group". Our null hypothesis is that both the number of accidents that occurred to others in the previous periods and whether a subject was hit by an accident in the previous phase should not influence investment behavior, since accidents occur independently. However, it is well-known that people fall prey to fallacies when faced with a random sequence of events.[26] This hypothesis is explicitly tested in section 4.2.

As a final note on the decision setting recall that, within each phase, subjects who were assigned the role of the victim and would otherwise have been passive were asked to make decisions. The procedure of randomly assigning participants to one of two roles and informing them about their actual role only at the end of the phase is known as *random dictatorship*. This was implemented in order to collect data on twice as many subjects. However, it introduces a further random draw in the payment protocol, making it more elaborate.

The payment protocol itself deserves further comment. The specific protocol we used may be classified as a mixture of a "pay all decisions sequentially" (henceforth *PAS*) and a "pay one decision randomly" (henceforth *POR*) mechanism.[27] In fact, within each of the two phases of the experiment, there is sequential payoff accumulation

[25]The two phases were chosen to be not too long (only 5 periods), in order to avoid noise in the decisions caused by boredom or fatigue.

[26]For example, two fallacies that may apply here are the *gambler's fallacy* and the *hot hand fallacy*. Given a fair coin, after a sequence of heads, people suffering from the former would expect tails while people suffering from the latter would expect heads (see, e.g., Slovic, 2000). For our experiment this would mean, respectively, that a person who was hit by an accident in phase 1 would not expect to be hit in phase 2, or would expect to be hit once again in phase 2.

[27]These terms are borrowed from Cox, Sadiraj and Schmidt (2012).

over the five periods constituting the phase. However, this accumulated payoff is finally assigned according to two independent random draws: one at the beginning of each phase (random dictatorship) and one at the end of the experiment (random selection of one phase). On the one hand, the main problem of *PAS* protocols is the possibility that earnings on previous periods may influence behavior in later periods within the same phase. However, several studies that have tested for wealth effects when using *PAS* in experiments with decisions under risk reported that these effects were insignificant.[28] On the other hand, being a random-lottery incentive mechanism, *POR* protocols may generate negative distortions to the proportion of risky choices in the population. However, the sequential payoff accumulation within each phase of our experiment may have counterbalanced this effect.[29]

A final remark concerns the role of a subject's risk attitude in our decision setting. Even if firms are often assumed to be risk-neutral, individual subjects are not.[30] However, the theoretical analysis in Appendix A shows that, for the specific predictions we test in our experiment, risk aversion is not an issue. Indeed, for the values of parameters adopted in our experimental design a subject with CRRA preferences should use the same optimal safety investment rule as a risk-neutral one, for any positive degree of relative risk-aversion: she should not invest in safety under *NoL* and invest under both *SL* and *Ne*. A similar result holds for a CARA specification.[31] We do not have

[28]See, e.g., Cox and Epstein (1989) and Cox and Grether (1996).

[29]In an individual choice experiment with five lottery pairs, Cox, Sadiraj and Schmidt (2012) have compared the performance of several payment mechanisms, among those *POR*, *PAS* and a one-task (henceforth *OT*) design. They find that *POR* data are biased towards significantly more risk aversion compared to *OT* data, while *PAS* data show less risk aversion than *OT* data. This result seems to support the insensitivity to risk attitudes of our payment protocol, if we interpret it as a mixture of *PAS* and *POR*.

[30]We acknowledge that the risk attitude of an individual cannot directly be translated to the risk attitude of a company. It is debatable whether firms are risk-averse, like individuals, or not. The attitude towards risk of companies is certainly related to their size and to their financial constraints. The framework of our experiment is by far too simple to take into account such parameters. The argument that firms are not necessarily risk-neutral and that, as a result, their decisions can look like the decisions of individuals, has been stressed by Leland and Pyle (1977). These authors show that the assumption of risk aversion has traction for small companies that suffer from restricted access to financial markets. However, in order to convince investors that their project is worthwhile, these risk-averse small firms accept to bear some risk and, finally, seem to behave like risk-neutral big companies.

[31]To be precise, the only difference in predictions with respect to the CRRA case is that a risk-averse subject with a (constant absolute) degree of risk aversion $r > 9.65$ should not invest under *SL*. However, such extremely risk-averse subjects are quite rare in the real world (see, e.g., Binswanger, 1980).

specific theoretical predictions for risk-loving subjects under *SL* and *Ne*.[32] However, experimental studies of choice under risk comparable to ours in terms of both subject pool and subjects' expected earnings report of only a few risk-loving subjects.[33] For all these reasons, and provided that the behavioral predictions we want to test only involve comparisons of aggregate behavior across treatments, we do not consider risk attitude as an explanatory variable in our data analysis.[34]

3.4 Behavioral predictions

In this section, we derive hypotheses based on the theoretical model from Section 2, its extension with risk aversion (in Appendix A), and the following parameter values: $c = 1$, $p_0 = 0.05$, $p_1 = 0.01$ (hence $\Delta p = 0.04$), $h = \{30, 50\}$, and $w_t \in [35, 40]$.

It is straightforward to see that (1) is satisfied for both harm sizes. Hence, investment in safety (i.e. $e = 1$) is socially optimal. From an individual standpoint, subjects should not invest in safety under an *NoL* rule,[35] but should do so both under the *SL* rule[36] and the *Ne* rule.[37]

For instance, such subjects would prefer a sure amount of 1 euro to a lottery giving them 1000 euros with probability 0.99 and 0 euro otherwise.

[32]Obviously, any expected utility maximizing subject should not invest under *NoL*, whatever her attitude toward risk.

[33]For example, both Holt and Laury (2002) and Harrison et al. (2005) find that only 8% of subjects in their pool of undergraduate students are risk-loving in the same "low real payoff" individual decision under risk. Both studies provide evidence that this percentage decreases as long as payoffs are scaled up. Notice that also participants in our study are undergraduate students and that the expected payoff in our safety investment game is around three times as much as in Holt and Laury's (2002) "low real payoff" task.

[34]Our experimental design includes a question aimed at eliciting risk attitude (see footnote 21). However, this risk-elicitation instrument – based on a hypothetical question – is hardly adequate for our analysis. First of all, the measure that it provides is self reported. Moreover, it can hardly be incorporated into the CARA and CRRA analysis in Appendix A.

[35]In both treatments *NoL-high* and *NoL-low*, the subject pays 1 ECU if she invests in safety and nothing if she does not invest. The occurrence of an accident does not lead to any cost for this subject. Hence, the subject should not invest, regardless of her attitude towards risk.

[36]In treatment *SL-low*, $\min\{h, w_t\} = h = 30$; $c = 1$ is smaller than $0.04*30 = 1.2$. In treatment *SL-high*, $\min\{h, w_t\} = w_t$; $c = 1$ is smaller than $(0.04 * w_t) \in [1.4; 1.6]$. Condition (2) is thus satisfied in all *SL* treatments, meaning that a risk-neutral subject should invest in safety in each period. From Appendix A, it is easy to see that the same result holds for CRRA subjects, and for CARA subjects when their degree of risk aversion is not extreme.

[37]In treatment *Ne-low*, $\min\{h, w_t\} = h = 30$; $c = 1$ is smaller than $p_0 * 30 = 1.5$. In treatment *Ne-high*, $\min\{h, w_t\} = w_t$; $c = 1$ is smaller than $(p_0 * w_t) \in [1.75; 2]$. Condition (3) is thus satisfied in all *Ne* treatments, meaning that a risk-neutral subject should invest in safety in each period. From Appendix

In accordance with the model and the chosen parameter values, the potential injurer should invest in safety in the presence of liability rules; the corollary is that the potential injurer should not invest in safety in the absence of any liability rule. If this result holds at the individual level, it should also hold when considering a group of individuals faced with the same decision task, without any strategic interaction among them. This leads us to the following hypotheses:

H1. On average, investment in safety under SL and Ne will be above investment in safety under NoL.

H2. On average, investment in safety under Ne should not differ from investment in safety under SL.

In her desire to reduce the probability of an accident, the potential injurer should be driven by her own loss (the amount the injurer will be asked to pay as a compensation for the harm caused to third parties), but not directly by the loss borne by third parties. In particular, in the model the potential injurer is protected by limited liability. Hence, in her investment decision, she should never take into account the losses that exceed her own liabilities, i.e. the losses that she is unable to compensate. Therefore:

H3. For a given liability regime, whether the potential damage can be fully compensated or not will not influence investment in safety.

4 Results

In analyzing the results, we proceed as follows. In the subsection "Main results", we first compare investment in safety under the three liability rules and then investigate whether or not the ability to fully compensate the victim influences investment in safety. In the subsection "Controls", we look at how additional factors – such as individual perception of own and general selfishness, and learning from own and others'

A, it is easy to see that the same result holds for both CARA and CRRA subjects, for any positive degree of risk aversion.

experience – influence the decision to invest in safety.

4.1 Main results

We define a subject's "investment ratio" as the number of times the subject decided to invest in safety divided by the number of times the subject had to make such a decision.[38] We report in Table 2 descriptive statistics for the investment ratios of the 60 subjects in each main treatment. Remember that for each liability rule we ran two sessions (60 subjects in total) – one in which the victim could be fully compensated for potential damage (*low* damage) and one in which she could not (*high* damage). In Table 2, we pool the data within each liability rule. Thus, for the time being, we neglect the ability to compensate the victim, focussing instead on the influence of the liability rule on investment behavior.

Treatment	*NoL*	*SL*	*Ne*
Average Investment Ratio	0.50	0.66	0.70
(Standard Deviation)	(0.38)	(0.35)	(0.36)

Table 2: Investment ratio by treatment, pooled.

Investment in safety is highest under *Ne*, lower under *SL*, and the lowest under *NoL*. To assess whether these differences are significant, we compare pairwise the three distributions using a Mann-Whitney test. The distribution of investment ratios is significantly lower under *NoL* than under either *Ne* ($p = 0.012$) or *SL* ($p = 0.002$). The relatively small difference in investment ratios under *SL* and *Ne* is not statistically significant ($p = 0.473$).

The results in Table 2 suggest noisy behavior, whereby subjects alternate between investing in safety and not investing in safety. These descriptive statistics, however,

[38] Recall that the number of times a subject had to decide whether or not to invest in safety is lower than 10 when a subject was hit by an accident.

are slightly misleading. To better appreciate differences in investment patterns across treatments, we plot in Figure 1 the histogram of investment ratios in each treatment. While Figure 1 confirms that some subjects do not act systematically (i.e. always invest or never invest), it also shows that the investment ratios are not uniformly distributed. Instead, the mode of the investment ratios is around 0 (i.e. never invest) in the *NoL* treatment and around 1 (i.e. always invest) in the *SL* and *Ne* treatments. Consistent with our prediction, we can see that the proportion of subjects who (almost) never invest is more than double in the *NoL* treatment (27%) compared to the *SL* (8%) and *Ne* (10%) treatments. Conversely, the proportion of subjects who (almost) always invest is more than double in the *Ne* than in the *NoL* treatment (43% versus 20%), and slightly higher in the *Ne* than in the *SL* treatment (43% versus 37%).

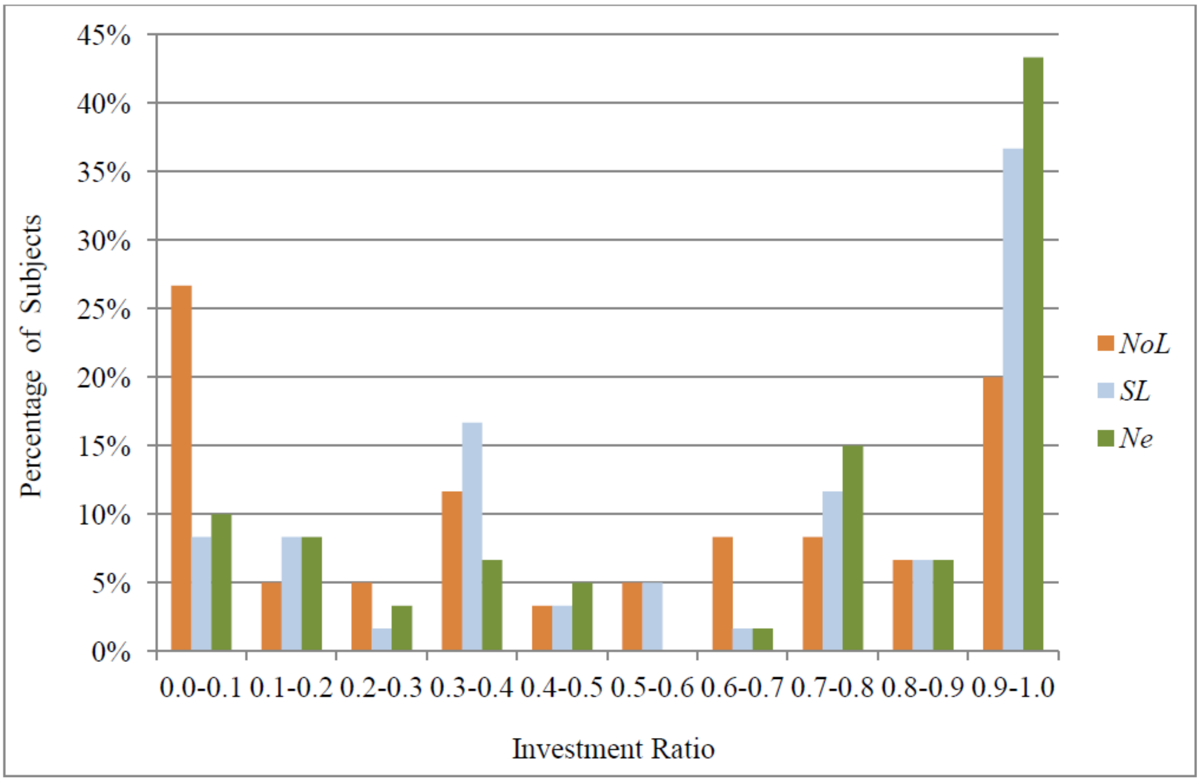

Figure 1: Distribution of investment ratios by treatment, pooled.

To test for treatment effects while controlling for the randomness in subjects' decisions, we adopt the error-rate model proposed by Harless and Camerer (1994). The basic idea behind the model is that, in each treatment, there is a proportion P of subjects who prefer to invest in safety and a proportion $(1 - P)$ who prefer not to invest in safety. The model, however, allows for the possibility that a subject deviates randomly from his/her underlying preferences. More precisely, there is a probability ε that a subject who prefers to invest in safety may erroneously decide not to invest (and vice versa). To test for treatment effects, we define the proportion of subjects who prefer to invest in safety as $P_{NoL} = 0.5 + \Delta_{NoL}$ in the *NoL* treatment, $P_{SL} = P_{NoL} + \Delta_{SL}$ in the *SL* treatment, and $P_{Ne} = P_{SL} + \Delta_{Ne}$ in the *Ne* treatment. The parameter Δ_{NoL} therefore indicates whether more or less than half of the subjects in the *NoL* treatment preferred to invest in safety. The parameter Δ_{SL} (respectively Δ_{Ne}) captures any differences in investment proportions between the *NoL* and *SL* treatments (respectively the *SL* and *Ne* treatments). Following Harless and Camerer (1994), the error rate ε is assumed to be independent across periods and subjects.[39] The parameters, estimated by maximum likelihood, are reported in Table 3.

Parameters	Δ_{NoL}	Δ_{SL}	Δ_{Ne}	ε
Estimates	-0.088^{**}	0.217^{**}	0.064	0.147^{***}
(Standard Deviation)	(0.029)	(0.078)	(0.081)	(0.016)

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3: Error rate model

We find that Δ_{NoL} is negative and significant, thereby suggesting that most subjects in the *NoL* treatment (59%) preferred not to invest in safety. The parameter Δ_{SL} is positive and significant, confirming that the proportion of subjects who prefer to invest in safety is larger under the *SL* treatment than under the *NoL* treatment. Finally, we find no evidence of differences in investment proportion between the *SL* and *Ne* treatments, as Δ_{Ne} is positive but not significantly different from 0. As we shall see below, these treatment effects are confirmed when we estimate panel logit regressions.

[39] We estimate similar treatment effects if we allow the error rate ε to differ across treatments.

Figure 2 shows the proportion of investment in safety decisions over time. Notice that learning (within a phase and between the two phases) neither causes an increase nor a decrease in investment in safety. This visual observation will be confirmed statistically below when we estimate various regressions models.

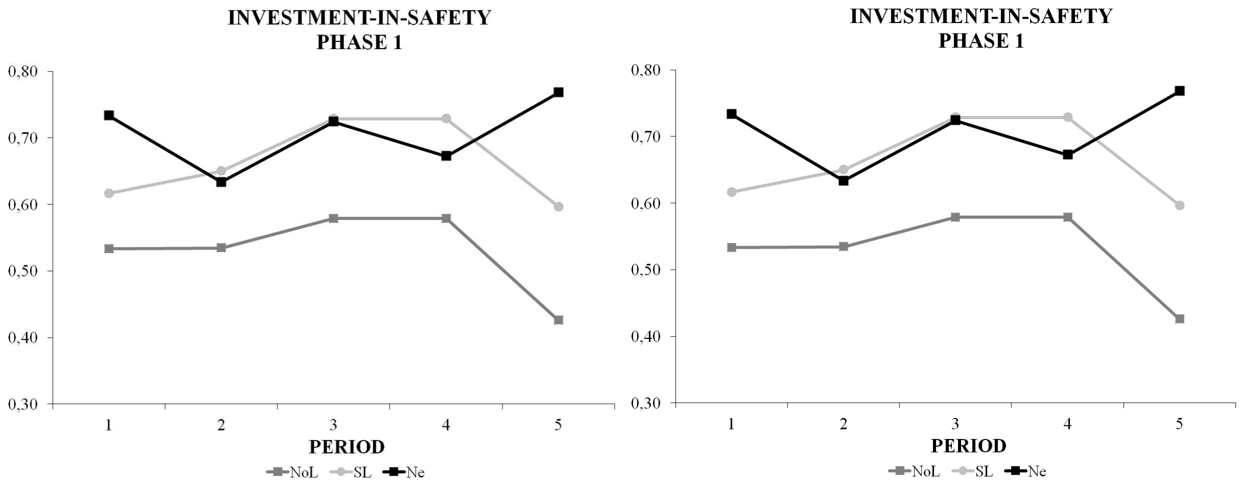

Figure 2: Investment in safety decisions over time (in %), by liability rule (pooled for ability to compensate the victim).

We summarize our results so far as follows:

Result 1 : *Subjects invest more in safety in the presence of liability rules.*

Result 2 : *Investment in safety under the Ne rule does not differ significantly from investment in safety under the SL rule.*

Our results differ from Kornhauser and Schotter (1990), who find that *SL* and *Ne* are not equivalent, even when the standard of due care is set at the socially optimal level.[40] Under *Ne*, they find compliance to the standard of due care (when set at its optimal level or not too far above), with remarkably stable investment in safety

[40]The standard of due care is also set at its optimal level $e = 1$ in our experiment, so the difference in results cannot come from this specification.

behavior over 35 periods. Under *SL* however, behavior in their experiment is quite volatile, showing over-investment in the first periods and under-investment in the final periods. We observe rule equivalence and stable investment in safety under both rules.

Next, we shed light on how the ability to compensate victims' losses affects investment in safety. Table 4 presents summary statistics by treatment.

Treatment	*NoL-low*	*NoL-high*	*SL-low*	*SL-high*	*Ne-low*	*Ne-high*
Average Investment Ratio	0.47	0.53	0.68	0.64	0.69	0.72
(Standard Deviation)	(0.38)	(0.39)	(0.36)	(0.34)	(0.39)	(0.33)

Table 4: Investment ratio by treatment.

To assess whether being able to fully compensate the victim influences investment in safety, we compare treatments where the victim can be compensated with treatments where she cannot, maintaining a constant liability rule. Using a Mann-Whitney test, we check whether the distribution of investment ratios is different in the *SL-low* and *SL-high* treatments, and in the *Ne-low* and *Ne-high* treatments. These pairwise comparisons of treatments do not yield any significant differences. This is further confirmed in the regressions below.

Result 3 : *Investment in safety in the presence of liability rules is not sensitive to the size of the potential damage.*

Holding constant the size of the damage while pairwise comparing liability rules confirms Results 1 and 2: liability significantly increases prevention, and *SL* and *Ne* are equally effective.[41]

[41] The distribution of investment ratios under *NoL-low* is significantly below the distribution of investment ratios under both *SL-low* (Mann-Whitney test $p = 0.028$) and *Ne-low* (Mann-Whitney test $p = 0.021$). The same is true when comparing *NoL-high* to *SL-high*, and *NoL-high* to *Ne-high* , with $p = 0.081$ and $p = 0.048$ respectively from a Mann-Whitney test.

Figure 3 depicts the proportions of investment decisions by treatment over time.

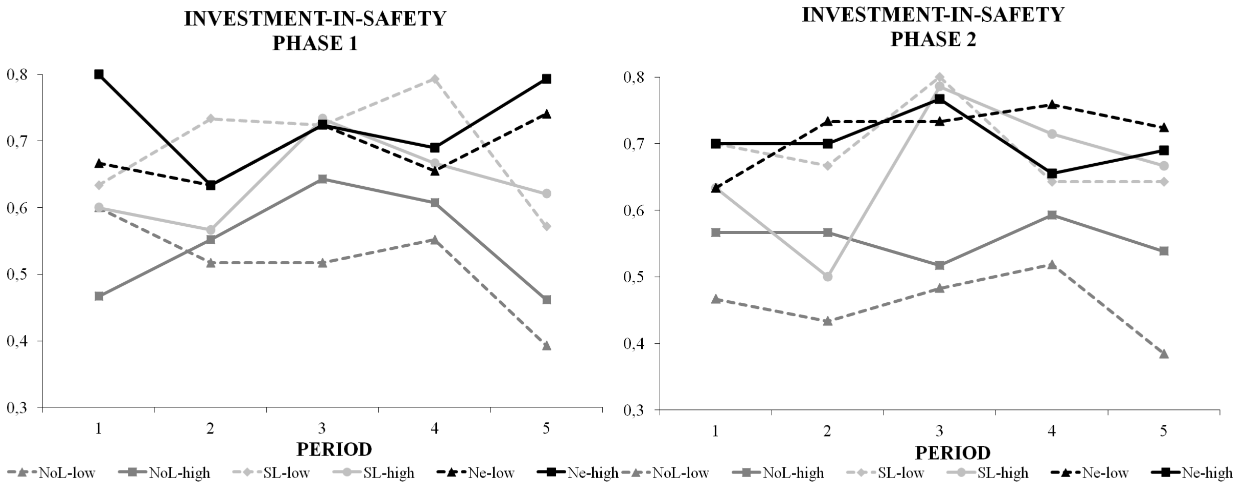

Figure 3: Investment in safety decisions over time (in %), by treatment.

Comparing the effect of the size of the damage phase by phase under a constant liability rule yields significant results only for the *NoL* treatments. In phase 2 of the experiment, investment in safety under *NoL-high* is significantly above investment in safety under *NoL-low* (Mann-Whitney test $p = 0.042$). Only in the absence of liability rules do we observe that subjects, once they have gained experience in phase 1, become sensitive to the potential damage caused to third parties. It seems that this (second-phase) sensitivity to damage is crowded out by liability.

4.2 Controls for individual characteristics and learning

In this section, we present evidence on whether and how additional factors, such as individual characteristics and learning, affect investment in safety. In particular, subjects may have a preference for helping others that can interfere with liability rules. Additionally, subjects may learn from their own experiences (if, for instance, they are hit by an accident in phase 1) and from the feedback about the experiences of others (the number of accidents in the "feedback group"). We run panel logit regressions in which the dependent variable is a dummy variable equal to 1 when a subject in a given period invests in safety. To account for the possible correlation in a subject's decisions across periods, we include individual-specific random effects in the regressions.

Dep.var.: INVEST	Ia	Ib	Ic	IIa	IIb	IIc
(*SL*+*Ne*) dummy	$.31^{**}$(.13)	$.32^{**}$(.13)	–.01(.22)	$.32^{**}$(.14)	$.31^{**}$(.14)	–.02(.24)
Ne dummy	.07(.10)	.08(.12)	–.08(.25)	.09(.11)	.11(.10)	–.16(.28)
High damage dummy	.03(.09)	.02(.09)	.05(.09)	.02(.10)	.02(.09)	.05(.10)
Others–selfish	—	$–.07^{*}$(.04)	$–.07^{*}$(.04)	—	$–.09^{**}$(.04)	$–.09^{**}$(.04)
Me–selfish	—	–.05(.04)	—	—	–.06(.04)	—
Me–selfish * *NoL* dummy	—	—	$–.16^{**}$(.07)	—	—	$–.18^{**}$(.08)
Me–selfish * *SL* dummy	—	—	–.03(.06)	—	—	–.03(.07)
Me–selfish * *Ne* dummy	—	—	.03(.06)	—	—	.05(.07)
Phase 2 dummy	—	–.04(.13)	–.04(.13)	—	—	—
Phase 1 dummy * Period	—	–.005(.02)	–.005(.02)	—	—	—
Phase 2 dummy * Period	—	.002(.01)	.002(.01)	—	.001(.01)	.001(.01)
N. of accidents in *t*–1	—	–.04(.03)	–.04(.03)	—	–.04(.04)	–.04(.04)
Accident in phase 1 dummy	—	—	—	—	–.08(.23)	–.10(.25)
Log Likelihood	–773.6	–678.1	–675.9	–410.4	–402.0	–399.3
N. of observations	1549	1549	1549	856	856	856

Note: Standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Regressions Ia, Ib, Ic use the whole data set. Regressions IIa, IIb, IIc use data from phase 2 only.

Table 5: Marginal effects from logit regressions with individual random effects explaining investment in safety.

The explanatory variables in Table 5 may be partitioned into three groups: treatment variables (the first three), individual characteristics (the next five), and learning (the last five). The treatment variables include a dummy, (*SL+Ne*) *dummy*, equal to 1 when the data was collected in the *SL* or the *Ne* treatment (and therefore equal to 0 when the data was collected in the *NoL* treatment). Because the logit models include a constant term, the parameter associated with (*SL+Ne*) *dummy* picks up any differences in investment probability in the *NoL* treatment compared to the other two treatments. Likewise, the dummy variable *Ne dummy* captures differences in investment probability in the *Ne* treatment compared to the *SL* treatment. By construction, these two dummy variables ignore the size of the damage. Therefore, we introduced a *High damage dummy* to measure the effect of the size of the damage on investment behavior, independent of the liability rule.

The individual characteristics were elicited in the post-experimental questionnaire. *Others-selfish* is a subject's response to the question "Would you say that most of the time people try to help others or only follow their own interests?", while *Me-selfish* is a subject's response to the question "Would you say that most of the time you try to help others or only follow your own interests?". Both variables are measured on a scale from 0 (help others) to 6 (follow own interests).

The variables aimed at testing for learning effects include *Phase 1 dummy* (respectively *Phase 2 dummy*) which takes the value of 1 for data collected in phase 1 (respectively phase 2). The interaction between the phase dummies and the period (from 1 to 5) accounts for any gradual learning within each phase.[42] *N. of accidents in* $t-1$ is the number of accidents that occurred in the subject's "feedback group" in the previous period. Here, we assume that accidents from at most the previous period may affect decisions in the current period. *Accident in phase 1 dummy* takes the value 1 if a subject was hit by an accident in phase 1.

The marginal effects produced by the panel logit models are reported in Table 5. We start in column Ia with a simple specification that includes only the treatment variables. The panel logit results confirm the treatment effects identified in the previous

[42]E.g. *Phase 1 dummy * period* takes the value of 1 if we are in phase 1, period 1. The same variable takes the value of 2 if we are in phase 1, period 2, and so on until phase 1, period 5. *Phase 1 dummy * period* takes the value 0 if we are in phase 2.

section. Indeed, the parameter associated with (*SL+Ne*) *dummy* indicates that investment in safety was 31% more likely in the *SL* and *Ne* treatments than in the *NoL* treatment. In contrast, we find no statistical evidence that the propensity to invest differs in the *SL* compared to the *Ne* treatment, or when the size of the damage caused to a third party increases.

In column Ib of Table 5, we augment the specification by accounting for individual characteristics and learning variables. We find that the probability of investing in safety decreases with the individual's perception of others being selfish. In contrast, whether or not an individual considers herself selfish does not influence her behavior. Furthermore, all the learning variables are found to be insignificant, providing no evidence that a subject's investment decisions evolved over the course of the experiment.

In column Ic, we interact the variable *Me–selfish* with the treatment variables. We find that *Me–selfish * NoL dummy* is negative and significant, meaning that the more selfish an individual rated herself, the less likely she was to invest in safety in the *NoL* treatment. Note also in column Ic that the parameter associated with the (*SL+Ne*) *dummy* is no longer significant. Thus, when we control for selfishness in treatment *NoL*, investment behavior in the absence of liability rules is not statistically different from investment behavior in the presence of liability rules. In other words, it appears that liability rules induce selfish subjects to invest in prevention: in the end, the pool of prosocial and selfish subjects under liability invest in safety as much as prosocial subjects under no liability.

To test the robustness of our results, and to confirm the absence of learning, we re-estimate the models in columns Ia, Ib and Ic with the data collected in phase 2 only. The results, presented in column IIa, IIb and IIc, indicate that neither the signs nor the magnitude of the estimated parameters are significantly different. In other words, we identify the same effects with the entire sample as with the sub-sample of data from the second half of the experiment.

In summary, the panel logit regressions confirm our previous conclusions: *SL* and *Ne* rules induce more investment in safety than *NoL*, and insolvency does not change investment behavior. Learning (from own and others' experience) does not change investment behavior, and believing others to be selfish decreases investment in safety. Finally, the difference between treatments without liability and with liability appears to be driven by the increased investment in safety of the more selfish subjects in the liability treatments.

5 Conclusions

In this paper, we compare the performance of three liability rules, *No Liability*, *Strict Liability* and *Negligence*, enforced against a firm that can potentially cause a disaster and thereby harm third parties. We model the firm's investment in safety as a moral hazard variable. The predictions of our theoretical model are tested in an experiment. In line with the theory, *Strict Liability* and *Negligence* perform better than *No Liability*: agents increase their level of care when they can be held liable for the harm caused. Furthermore, there is no significant difference in the effectiveness of *Strict Liability* and *Negligence* rules. Finally, for a given size of own wealth, agents do not invest more when losses to third parties increase (i.e. when the insolvency problem is more stringent). In contrast with the theory, which predicts zero prevention under *No Liability* and 100% prevention under liability (for risk-neutral and risk-averse subjects), prevention rates are as high as 50% in the former and significantly below 100% in the latter case.

Our work can be extended in several directions. Most of the theoretical predictions were confirmed by a subject pool of German undergraduates; however, the substantial level of investment that appeared under *No liability* requires further exploration. Other-regarding preferences, arising from subjects caring for the well-being of third parties, may be responsible for this outcome. This conjecture would be in line with Brennan et al. (2008), who show that once own outcome is not at risk, subjects care for the risk borne by others. More research will also be needed to provide explanations for the relatively low investment in prevention in the presence of liability rules.

Further, in our setting the size of the harm is given, and the only way of reducing expected losses is to reduce the probability of an accident. However, one could consider a more general model where both the probability of an accident and the size of the harm can be influenced by prevention. In such a setting, the size of the harm can be linked to the firm's scale of activity, and the probability of an accident can be linked to the intensity of safety effort. From the Law and Economics literature[43] we know that *Strict Liability* is effective in providing incentives for both activity and probability reduction, since the responsible firm is held liable for the entire loss, regardless her conduct in the operations that had led to damage. The firm thus has incentives to use all the available means to reduce expected losses. *Negligence* rule, on the contrary, is only effective for probability reduction. Since the injurer is not held liable if she complied with a standard of due care, only her level of prevention matters: her level of activity has no influence on the court's liability decision. It would be worth developing an experiment to test such differences in firm's incentives when managing potential harm to third parties.

A careful adaptation of the present experiment could also provide empirical arguments for a number of long-standing theoretical debates in the Law and Economics and Incentive Regulation literatures. For instance, one could test the effectiveness of extended liability,[44] and also whether the risk of an accident is better controlled with ex-ante (standard regulation implemented by agencies) or with ex-post (liability rules, enforced by courts of law) instruments.[45]

[43]See Segerson (2002) for informal arguments and Shavell (1980) for formal ones.

[44]See Pitchford (1995) or Hiriart and Martimort (2006a) and the references therein.

[45]See Shavell (1984a), Kolstad, Ulen and Johnson (1990), or Hiriart, Martimort and Pouyet (2008, 2010).

6 References

Alberini, A., and Austin, D.H. (1999a), "On and off the liability bandwagon: Explaining state adoptions of strict liability in hazardous waste programs", *Journal of Regulatory Economics*, 15, 41–63.

Alberini, A., and Austin, D.H. (1999b), "Strict liability as a deterrent in toxic waste management: Empirical evidence from accident and spill data", *Journal of Environmental Economics and Management*, 38, 20–48.

Alberini, A., and Austin, D.H. (2001), "An analysis of the preventive effect of environmental liability," Study commissioned by DG ENV of the European Commission.

Alberini, A., and Austin, D.H. (2002), "Accidents waiting to happen: Liability policy and toxic pollution releases", *Review of Economics and Statistics*, 84, 729–741.

Ashford, N., and Caldart, C. (2008), "*Environmental law, policy, and economics*", Cambridge, MA: MIT Press.

Binswanger, H. (1980), "Attitudes toward risk: Experimental measurement in rural India", *American Journal of Agricultural Economics*, 62, 395–407.

Brennan, G., Gonzales, L.G., Güth, W., and Levati, M.V. (2008), "Attitudes toward private and collective risk in individual and strategic choice situations", *Journal of Economic Behavior and Organization*, 67: 253–262.

Cox, J.C., and Epstein, S. (1989), "Preference reversals without the independence axiom", *American Economic Review*, 79, 408–426.

Cox, J.C., and Grether, D. A. (1996), "The preference reversal phenomenon: Response mode, markets, and incentives", *Economic Theory*, 7, 381–405.

Cox J.C., Sadiraj V., and Schmidt U. (2012), "Paradoxes and mechanisms for choice under risk", Experimental Economics Center Working Paper Series 2012-08, Experimental Economics Center, Andrew Young School of Policy Studies, Georgia State

University.

Dopuch, N., and King, R.R. (1992), “Negligence versus strict liability regimes in auditing: An experimental investigation”, *The Accounting Review*, 67, 97–120.

Dopuch, N., Ingberman, D., and King, R.R. (1997), “An experimental investigation of multi-defendant bargaining in ‘joint and several’ and proportionate liability regimes”, *Journal of Accounting and Economics*, 23, 189–221.

Faure, M., and Skogh, G. (2003), “*The Economic analysis of environmental policy and law: An introduction*”, Cheltenham, UK/Northampton, MA: Edward Elgar.

Fischbacher, U. (2007), “Z-Tree: Zurich toolbox for ready-made economic experiments”, *Experimental Economics*, 10, 171–178.

Greiner, B. (2004), “An online recruitment system for economic experiments”, in K. Kremer and V. Macho (eds.), *Forschung und wissenschaftliches Rechnen 2003. GWDG Bericht 63*, Göttingen: Ges. für Wiss. Datenverarbeitung, 79–93.

Harless, D., and Camerer. C.F. (1994), The predictive utility of generalized expected utility theories, *Econometrica*, 62, 1251–1290.

Harrison G.W., Johnson E., McInnes M.M., and Rutström, E.E. (2005). “Risk Aversion and Incentive Effects: Comment”, *American Economic Review*, 95, 897–901.

Hiriart, Y., and Martimort, D. (2006a), “The benefits of extended liability”, *RAND Journal of Economics*, 37, 562–582.

Hiriart, Y., and Martimort, D. (2006b), “Liability rules and regulation for environmentally risky ventures”, in M. Boyer, Y. Hiriart and D. Martimort (eds.), *Frontiers in the Economics of Environmental Regulation and Liability*, Williston VT: Ashgate.

Hiriart, Y., Martimort, D., and Pouyet, J. (2008), “The regulator and the judge: The optimal mix in the control of environmental risk”, *Revue d’Economie Politique*, 119, 941–967.

Hiriart, Y., Martimort, D., and Pouyet, J. (2010), “The public management of environmental risk: Separating ex-ante and ex-post monitors", *Journal of Public Economics*, 94, 1008–1019.

Holt C.A., and Laury, S.K. (2002), "Risk aversion and incentive effects", *American Economic Review,* 92, 1644–1655.

King, R., and Schwartz, R. (1999), "Legal penalties and audit quality: An experimental investigation", *Contemporary Accounting Research*, 16, 685–710.

King, R., and Schwartz, R. (2000), "An experimental investigation of auditors' liability: implications for social welfare and exploration of deviations from theoretical predictions", *The Accounting Review*, 75, 429–451.

Kolstad, C., Ulen, T., and Johnson, G. (1990), "Ex-post liability for harm vs. ex-ante safety regulation: Substitutes or complements", *American Economic Review*, 80, 888–901.

Kornhauser, L., and Schotter, A. (1990), "An experimental study of single-actor accidents", *Journal of Legal Studies*, 19, 203–233.

Leland, H., and Pyle, D.J. (1977), "Informational asymmetries, financial structure and financial intermediation", *Journal of Finance*, 32, 381–387.

Murchison, K.M. (2011), "Liability under the oil pollution act: Current law and needed revisions, *Louisiana Law Review*, 71, 917–956.

Newman, H., and Wright, D. (1990), "Strict liability in a principal-agent model", *International Review of Law and Economics*, 10, 219–231.

Niederle, M., and Vesterlund, L. (2007), "Do women shy away from competition? Do men compete too much?", *Quaterly Journal of Economics*, 122, 1067–1101.

Pitchford, R. (1995), "How liable should a lender be? The case of judgment-proof firms and environmental risks", *American Economic Review*, 85, 1171–1186.

Posner, R.A. (1992), "*Economic analysis of law*", Boston: Little, Brown and Company, Fourth edition.

Segerson, K. (2002), "*Economics and liability for environmental problems*", Aldershot, UK/Burlington, VT: Ashgate.

Shavell, S. (1980), "Strict liability versus negligence", *Journal of Legal Studies*, 9, 1–25.

Shavell, S. (1984a), "A model of the optimal use of liability and safety regulation", *RAND Journal of Economics*, 15, 271–280.

Shavell, S. (1984b), "Liability for harm versus regulation of safety", *Journal of Legal Studies*, 13, 357–374.

Shavell, S. (1986), "The judgment proof problem", *International Review of Law and Economics*, 6, 45–58.

Shavell, S. (2004), "*Foundations of economic analysis of law*", Cambridge, Mass.: Harvard University Press.

Slovic, P. (2000), "*The perception of risk*", London, UK: Earthscan.

Viscusi, K. (2007), "Regulation of health, safety, and environmental risks", in A.M. Polinsky and S. Shavell (eds.), *Handbook of Law and Economics*, vol. I, Amsterdam: Elsevier, 591–645.

Wittman, D., Friedman, D., Crevier, S., and Braskin, A. (1997), "Learning liability rules", *Journal of Legal Studies*, 26, 145–164.

Appendix A

■ **Risk-aversion** with a **CARA** utility function. Assume that the firm is risk-averse and her preferences are captured by a CARA utility function: $u(x) = \frac{1-e^{-rx}}{r}$, where the parameter $r > 0$ measures the absolute risk aversion and x is a monetary payoff.

Social Optimum. Prevention is socially optimal as long as:

$$p_1 u(w_t - h - c) + (1 - p_1)u(w_t - c) \geq p_0 u(w_t - h) + (1 - p_0)u(w_t), \tag{4}$$

which can be rewritten:

$$c \leq \frac{1}{r} * \ln\left(\frac{1 - p_0 + p_0 e^{rh}}{1 - p_1 + p_1 e^{rh}}\right). \tag{5}$$

No Liability. The firm chooses to invest in prevention as long as:

$$p_1 u(w_t - c) + (1 - p_1)u(w_t - c) \geq p_0 u(w_t) + (1 - p_0)u(w_t), \tag{6}$$

a condition that reduces to $u(w_t - c) \geq u(w_t)$ and that, obviously, never holds. Hence, the firm never invests in safety in the absence of liability.

Strict Liability. The firm chooses to invest in prevention as long as:

$$p_1 u(w_t - \min\{h, w_t\} - c) + (1 - p_1)u(w_t - c) \geq p_0 u(w_t - \min\{h, w_t\}) + (1 - p_0)u(w_t), \tag{7}$$

which can be rewritten:

$$c \leq \frac{1}{r} * \ln\left(\frac{1 - p_0 + p_0 e^{r \min\{h, w_t\}}}{1 - p_1 + p_1 e^{r \min\{h, w_t\}}}\right). \tag{8}$$

Comparing (5) and (8), it is straight forward to see that the firm will make the socially optimal decision if she is wealthy enough, i.e. if her wealth w_t is sufficient to cover harm h.

Negligence. The firm chooses to invest in prevention as long as:

$$p_1 u(w_t - c) + (1 - p_1)u(w_t - c) \geq p_0 u(w_t - \min\{h, w_t\}) + (1 - p_0)u(w_t), \qquad (9)$$

a condition that can be rewritten:

$$c \leq \frac{1}{r} * \ln\left(1 - p_0 + p_0 e^{r \min\{h, w_t\}}\right). \qquad (10)$$

Comparing (8) and (10), we can easily show that the former is more demanding than the latter: the firm is induced to exercise care for a larger set of parameters when submitted to *Negligence* rather than *Strict Liability*.

The qualitative theoretical results obtained with a risk-neutral firm therefore do not change when moving to a CARA case. In particular, for the set of parameters (c, w_t, h, p_0, p_1) that characterize our experimental setting, a risk-averse firm should behave as a risk-neutral one both under *No Liability* and under *Negligence*: for every $r > 0$, she should not invest in safety in the former regime and invest in the latter. Under *Strict Liability*, a risk-averse firm should invest in safety for each $r \in \{0, 9.65\}$, i.e. so long as she is not extremely risk-averse.

■ **Risk-aversion** with a **CRRA** utility function. Assume that the firm is risk-averse and her preferences are captured by a CRRA utility function: $u(x) = \frac{x^{1-\gamma}}{1-\gamma}$, where the parameter $\gamma > 0$ ($\gamma \neq 1$) measures the relative risk aversion and x is a monetary payoff.[46]

Social Optimum. Prevention is socially optimal when condition (4) is satisfied. Using the fact that $f(x+y) = f(x) + yf'(x)$ when y is small,[47] this condition can be rewritten as:

$$c \leq \frac{\Delta p\left[u(w_t) - u(w_t - h)\right]}{p_1 u'(w_t - h) + (1 - p_1)u'(w_t)}. \qquad (11)$$

[46] In the case where $\gamma = 1$, $u(x) = \ln x$.

[47] Hence, $u(w_t - h - c) \simeq u(w_t - h) - cu'(w_t - h)$ and $u(w_t - h) \simeq u(w_t) - cu'(w_t)$ for c small enough.

No Liability. The firm chooses to invest in prevention when condition (6) is satisfied, a condition that, again, reduces to $u(w_t - c) \geq u(w_t)$, which never holds. Hence, the firm never invests in safety in the absence of liability.

Strict Liability. The firm chooses to invest in prevention when condition (7) is satisfied. This condition can be rewritten as:

$$c \leq \frac{\Delta p \left[u(w_t) - u(w_t - \min\{h, w_t\})\right]}{p_1 u'(w_t - \min\{h, w_t\}) + (1 - p_1)u'(w_t)}. \tag{12}$$

Hence, (12) coincides with (11) when the firm is wealthy enough, i.e. she takes the socially optimal decision if her wealth w_t is sufficient to cover harm h.

Negligence. The firm chooses to invest in prevention when (9) is satisfied. This condition can be rewritten as:

$$c \leq \frac{p_0 \left[u(w_t) - u(w_t - \min\{h, w_t\})\right]}{u'(w_t)}. \tag{13}$$

Comparing (13) and (12), we see that the former is less demanding than the latter:[48] regardless her wealth, the firm is induced to exercise care for a larger set of parameters when submitted to *Negligence* rather than *Strict Liability*. Using the same arguments, in the case where $\min\{h, w_t\} = h$, we also see that (13) is less demanding than (11). The set of parameters for which wealthy firms with CRRA preferences invest in prevention is larger under *Negligence* than at the social optimum.

For the values of parameters (c, w_t, h, p_0, p_1) adopted in our experiment, a firm presenting CRRA preferences should behave as a risk-neutral one under any liability regime and for any positive degree of relative risk-aversion γ: she should not invest in safety under *No Liability* and invest under *Strict Liability* and *Negligence*. Indeed, the term on the right-hand side in (12) is equal to 1.206 when $w_t = 40$ and $\gamma = 0.01$, and strictly increases with a lower wealth level w_t or a higher degree of relative risk-aversion. Since $c = 1$, condition (12) is satisfied for all possible values taken by the

[48]This is because $u'(w_t) < p_1 u'(w_t - \min\{h, w_t\}) + (1 - p_1)u'(w_t)$ and $p_0 \geq p_0 - p_1$.

parameters in our experiment: the subjects should always invest under *Strict Liability*. The same happens under *Negligence* since the lowest value obtained on the right-hand side of (13) is also 1.206.