Laws and Norms: Experimental Evidence with Liability Rules

B. Deffains[*], R. Espinosa[†], C. Fluet[‡]

March 14, 2017

PRELIMINARY DRAFT, PLEASE DO NOT CIRCULATE

Abstract

We study an experiment where participants choose between actions which provide private benefits but may also impose losses on strangers. Depending on circumstances, private benefits are greater or smaller than the negative externalities. We compare three legal environments: no law, strict liability for the harm caused to others, and an efficiently designed negligence rule whereby damages are paid only if the harm exceeds the private benefit. The law may be perfectly enforced (Strong Law) or only weakly so (Mild Law). Strong Law efficiently regulates behavior and does much better than no law. Mild Law also regulates behavior better than no law even though legal sanctions are non-deterrent. Moreover, strict liability does better than the negligence rule when self and group interests conflict, although monetary incentives are the same. We investigate how legal sanctions and social preferences interact to yield this pattern. Individuals trade-off self-interest and social efficiency concerns. Legal obligations reinforce the latter and do so to a greater extent under the strict liability rule.

[1]Université Paris 2, CRED, University of Liverpool, IUF.
[2]Université Paris 2,CRED.
[3]Université Laval, CRREP, CRED.

1 Introduction

A main purpose of liability rules is to deter behavior generating negative externalities. Under the so-called strict liability rule, individuals must compensate the harm they impose on others, i.e., the harm is internalized irrespective of precautions or circumstances. Under a negligence or fault-based rule, individuals must compensate if their behavior fell short of some legal standard of conduct. In either case, and provided the legal standard of conduct in the negligence rule is appropriately set, perfectly enforced liability rules yield socially efficient incentives to avoid causing harm. Conversely, when legal liability is nonexistent or is only imperfectly enforced, e.g., injurers are seldom detected or victims seldom file suit, individuals are under-deterred. The usual prediction is then that behavior will be socially inefficient and negative externalities will arise too often.

Casual observation, however, and quite a few empirical studies suggest that there are many situations where people avoid privately profitable but socially harmful behavior irrespective of legal sanctions. There is also a voluminous experimental literature on social dilemma situations, e.g., experiments on voluntary contributions to a public good, showing that individuals are not solely motivated by their own material payoff. One strand of this literature examines how contributions to the public good might be influenced by the "legal obligation" to contribute, even if such an obligation is weakly enforced. We study a variant of the public good game where individuals must trade-off their private gain against the harm caused to others and may or may not face legal sanctions in the form of liability rules.

In our experiment, subjects repeatedly interact with strangers in a game where each participant must choose between actions which provide private benefits but may also cause losses to others. Participants face varying and randomly occurring circumstances over time. In some circumstances, the net private gain from the harmful action (over that of the harmless action) is greater than the loss imposed on others. The harmful action is then socially efficient in the sense that in the long run everyone's wealth would be greater if everyone chose the harmful action in such circumstances. In other circumstances the situation is reversed. While the net private gain from the harmful action remains positive, it is smaller than the loss imposed on others. From a social point of view, the harmful action should then be avoided in such circumstances because it reduces average per capita wealth.

We compare three legal environments: no legal liability which we refer to as No Law, strict liability and an efficiently designed negligence rule. In the latter, the legal standard of conduct is such that individuals are not held liable if the loss imposed on others is smaller than the net private gain from the harmful action. Liability rules are either perfectly enforced (Strong Law) or only weakly so (Mild Law). Under Strong Law, individuals causing harm are always detected. When the rule is strict liability rule, the individuals detected are forced to fully compensate victims; under the negligence rule, they must do so only if their behavior did not comply with the legal standard of conduct. Under Mild Law, the individuals causing harm are detected only half of the time. The calibration of payoffs is such that Mild Law should be non-deterrent for most individuals; that is, except for those with an exceptionally high degree of risk aversion.

Under either No Law or Mild Law, the prediction under standard preferences is that participants will always choose the harmful action irrespective of circumstances. Under Strong Law, they will always behave efficiently from a social point of view. Our experiment, however, demonstrates a somewhat different pattern of behavior. Under Strong Law, in circumstances where self and group interest conflict, strict liability and the negligence rule efficiently regulate behavior and do much better than No Law, as would be expected if individuals had purely self-regarding preferences. However, under Mild Law, both liability rules still do significantly better than No Law even though the threat of legal sanction is essentially non-deterrent. Moreover, strict liability then does significantly better than the negligence rule even though monetary incentives to avoid causing harm are exactly the same under both rules when private and collective interest conflicts. We investigate how legal sanctions and social preferences might interact to yield this pattern of behavior. We show that behavior is consistent with the assumption that individuals trade-off narrow self-interest and social efficiency, equivalently individuals trade-off private benefits against the losses imposed on others. We show that, when enforcement is non-deterrent, the introduction of legal liability has a norm-activation effect in the sense that individuals put greater weight on social efficiency concerns. This effect varies with the content of liability rules independently of material sanctions. We also show that social disapproval influences behavior and analyze the determinants of disapproval in terms of what it may suggest with respect to the underlying social norms.

The paper proceeds as follows. Section 2 briefly reviews the relevant experimental literature. Section 3 presents the experimental design. Section 4 draws on the theoretical literature on social preferences to develop a simple model from which predictions are derived for our experiment. Section 5 presents the results. Section 6 discusses additional points of interpretation and Section 7 concludes.

2 Literature Review

To be completed...

3 The Experiment

The experiment consists of three phases. The first phase is a series of short games testing for the participants' risk aversion and social preferences. The second phase is the core of the experiment. Subjects play the *Liability Game*, a game in which legal rules and the enforcement policy change across treatments. This game is repeated 10 times with blaming points and is followed by a modified version of the same game in which blaming points are replaced by costly punishment points, also repeated 10 times. Finally, in the third phase, we run a questionnaire for demographics and additional control questions.

3.1 Liability Game

Common Set-up. After the preliminary phase, participants are introduced to the liability game with a new set of instructions. They are told that they are going to play a game that will be repeated 10 times. At each round, they will be randomly and anonymously matched into groups of 4 participants.

Participants start each round with an initial endowment of 20 ECU. At each round they need to choose between two actions, Y and X. Action Y yields an income of 6 ECU and does not affect the other participants' earnings. Action X yields a state-dependent income and reduces by 4 ECU the earnings of each other participant in the group. At the beginning of each round, a random state is drawn for each participant among four possible states, A, B, C and D, with equal probability (i.e., 25%). States are independently drawn and are private information all along the game. The states define a participant's circumstances with respect to the private benefit of action X: action X yields an income of 14 ECU in state A, 16 ECU in state B, 20 ECU in state C, and 22 ECU in state D. **Treatments.** Participants were subjected to five

different treatments defined in terms of the applicable liability rule and the enforcement policy.

- **No Law (NL).** In the first treatment, there is no liability rule. Each participant has to bear the losses caused by the actions X of other participants in the group. A participant's net payoff per period, in addition to the endowment at the start of the round, is therefore (i) the private benefit from one's own action Y or X, (ii) minus the losses caused by the actions X of other participants.
- **Strong Strict Liability (SSL).** In the second treatment, participants are told they will be required to compensate the other group members for the losses caused by their decision to engage in action X. Therefore no one suffers from the other participants' decision to

engage in X. The net period income equals (i) the private benefit from one's own action Y or X, (ii) minus the damages (12 ECU) for compensation if action X was chosen.

- **Strong Negligence Rule (SNR).** In the third treatment, participants are told they will be required to compensate for the losses caused by their action X if they were in state A or B. A participant therefore suffers from the other participants' actions X only if these actions were undertaken in the circumstances C or D. The net period income equals (i) the private benefits from one's own action Y or X, (ii) minus the damages (12 ECU) for compensation if they choose X in state A or B, (iii) minus the losses caused by the other participants' decision to engage in X in the circumstances C or D.
- **Mild Strict Liability (MSL).** The fourth treatment is similar to the second treatment, except that participants who engage in X are made to compensate only with a probability equal to 0.5, henceforth the detection or enforcement probability. The net period income equals (i) the private benefit from one's own action Y or X, (ii) minus the eventual damages (12 ECU) for compensation if they choose X and are detected, (iii) minus the losses caused by the other participants' decisions to engage in X that were not detected.
- **Mild Negligence Rule (MNR).** The fifth treatment is similar to the third treatment, except that participants who engage in X in state A and B are made to compensate only with a probability of one half. The net period income equals (i) the private benefit from one's own action Y or X, (ii) minus the eventual damages (12 ECU) for compensation if the participant chooses X in state A or B and is detected, (iii) minus the losses caused by the other participants' decision to engage in X when either they were in state C or D or they were in the states A or B but were not detected.

Blaming Points and Payoffs. At the end of each period, participants learn the number of other group members who chose action X. In all treatments but *No Law*, participants also know whether each of the other three group members (anonymously identified as player 1, 2 or 3) had to compensate other participants, i.e., was held "legally liable" . Individual actions and states of nature are therefore private information except in so far as actions and states can be inferred from the assignment of liability or from the total number of actions X.

After receiving this information and learning their period payoff, participants have the opportunity to assign blaming points (between 0 and 6) to each other participant in their group. The blaming points are individual. After the assignment of blaming points, a final screen displays to each participant the amount of blaming points the participant received.

In the liability game, the total period payoff of individual i in state $s(i) \in \{A, B, C, D\}$ is given by:

$$\text{payoff}_i = 20 + X_i(R_{s(i)} - 12\ \text{cond}_i) + 6(1 - X_i) - 4\sum_{j \neq i} X_j(1 - \text{cond}_j) \tag{1}$$

where X_i is a dummy variable equal to 1 if the individual i chooses action X and 0 if action Y is chosen; cond_i is a dummy variable if individual i has been required to compensate for the losses due to action X; $R_{s(i)}$ is the income from action X in state $s(i)$.

The variable cond_i is defined as follows:

$$\text{cond}_i = \begin{cases} 0 & \text{in No Law} \\ X_i & \text{in Strong Strict Liability} \\ \mathbb{1}_{A,B} X_i & \text{in Strong Negligence Rule} \\ \mathbb{1}_{\text{det}} X_i & \text{in Mild Strict Liability} \\ \mathbb{1}_{\text{det}} \mathbb{1}_{A,B} X_i & \text{in Mild Negligence Rule} \end{cases}$$

where $\mathbb{1}_{A,B}$ is a dummy variable equal to 1 if individual i is in state A or B and equal to 0 otherwise; $\mathbb{1}_{\text{det}}$ is a dummy variable equal to 1 if individual i is detected after engaging in action X, and is 0 otherwise.

Sanction points and payoffs After the ten rounds of the *Liability Game*, participants are given a new set of instructions for the *Punitive Liability Game* and learn that they will play another 10 rounds of the previous game. The only change compared to the 10 previous rounds is that blaming points are replaced by sanction points. Participants have the opportunity to impose sanction points on other group members. Each point decreases both the participant's and the target's payoff by 0.5 ECU. Each participant can assign up to 6 sanction points to each group member.

In the *punitive* liability game, payoffs are given by:

$$\begin{aligned} \text{payoff}_i \quad = \quad & 20 + X_i(R_{s(i)} - 12 \ \text{cond}_i) + 6(1 - X_i) \\ & -4 \sum_{j \neq i} X_j(1 - \text{cond}_j) - 0.5 \ [\sum_{j \neq i} \text{sanct}_{i,j} + \sum_{j \neq i} \text{sanct}_{j,i}] \end{aligned} \tag{2}$$

where $\text{sanct}_{k,m}$ is the number of sanction points imposed by participant k on participant m.

3.2 Control Questions.

Before playing the game, participants were asked a series of question to insure that the game was well understood. We generated a mock stage game in which we displayed the actions, situations and condemnations of all four participants. Control questions were designed to address all the mechanisms that affect the period payoff. Participants had to fulfill, step by step, a table which required to compute the losses each participant imposed and bore, the compensation each of them gave and received, and their final payoff. In the appendix, we show the control questions for the *Mild Negligence Rule* which was the most demanding regarding the complexity of the mechanisms at play.

3.3 Questionnaire

After the *Punitive Liability Game* participants are asked to fill out a questionnaire on demographics (age, gender) and on preferences and self-perception. These include (i) self-declared political orientation, (ii) attitude with respect to state intervention in the economy, (iii) self-assessed risk aversion, (iv) the extent to which they see themselves as selfish, (v) how much they think others see them as selfish, (vi) the extent to which they feel concerned about the well-being of others, (vii) how much they think others see them as being concerned by the well-being of others. For cross-study comparison purposes, the last four questions were adapted from Angelova et al. (2014).

4 Model and Predictions

Individuals repeatedly interact with strangers under a given liability rule. With purely self-interested agents, the equilibrium prediction is that of a one-shot game at each period of play. When one's circumstances are C or D, the strictly dominant strategy is to choose action X in all legal regimes. When the circumstances are A or B, action X is strictly dominant under No Law while action Y is strictly dominant under Strong Law irrespective of the legal regime.

In the circumstances A and B under Mild Law, action X is chosen by a risk neutral under both legal regimes. By contrast, a sufficiently risk averse might then choose action Y to avoid bearing the liability risk. However, given the initial endowments and the probability of detection, it can be shown that only the extremely risk averse would choose Y. For example, assuming a Constant Relative Risk Aversion utility of wealth function, an individual in the circumstance A who expects all three other individuals in his group to choose X in all circumstances will himself choose Y only if his coefficient of relative risk aversion is above 3.5. This is far outside the normal range for relative risk aversion coefficients, usually estimated to be between 0.5 and 1.5 (see for instance Holt and Laury 2002). We conclude that our Mild Law is non-deterrent except possibly for abnormally risk averse individuals. In what follows, for the case of imperfect enforcement, expected values are taken as an acceptable approximation of certainty equivalents.[1]

Efficiency concerns. We anticipate some departure from pure self-interest. In the present set-up with essentially anonymous interactions, non purely self-interested preferences mean that individuals balance the private benefits of their actions against the losses they impose on strangers. Equivalently, individuals trade-off private gains against total group benefits. As in Charness and Rabin (2002) and others, we postulate that individuals have a utility function of the form:

$$u_i = (1-\lambda)\pi_i + \lambda \sum_j \pi_j$$

where π_j is the total payoff of individual j and λ is the the weight individual i puts on social efficiency as measured by the total group payoff, including that of individual i. If all individuals have λ equal to zero, the model reduces to pure self-interest.

Individual i's payoff can be written as

$$\pi_i = w_i + g_p$$

where w_i is the part of the individual's payoff that does not depend on his action (but may depend on the actions of others) and g_p is the private benefit from this action. Similarly,

$$\sum_j \pi_j = w_i + w_{-i} + g_s$$

where w_{-i} is the part of the other group members' payoff that does not depend on individual i's action and g_s is the social benefit of the action, i.e., the effect on the total group payoff. The individual's utility is therefore

$$u_i = (1-\lambda)(w_i + g_p) + \lambda(w_i + w_{-i} + g_s). \tag{3}$$

The individual chooses the action which maximizes the above expression. Factors affecting w_i and w_{-i} or beliefs about these values are irrelevant in this decision problem.

Let y denote the gain from action Y, x the gain from action X and h the harm caused to others. For action Y, $g_p = y$ and $g_s = 0$. For action X, in expected value, $g_p = x - ph$ and $g_s = x - h$ where p is the probability of having to compensate the harm. Depending on the circumstances, the legal regime and the enforcement policy, p is either zero, one half or unity. Let us denote with capital letters the net consequences of action X compared to action Y, i.e.,

$$G_p = x - y - ph, \quad G_s = x - y - h.$$

Then individual i chooses X if

$$\Delta u_i = (1-\lambda)G_p + \lambda G_s \tag{4}$$

is positive. Note that (4) can be rewritten as

$$\Delta u_i = G_p - \lambda(G_p - G_s)$$

where $G_p - G_s = (1-p)h$ is the expected uncompensated harm imposed on others. Thus, the social efficiency parameter λ can be interpreted as the rate at which the individual trades-off private benefits against the net loss caused to third parties.

[1] In the subsequent analysis, we rejected the hypothesis that the certainty equivalent significantly differs from the expected value.

Obviously, $G_p \geq G_s$. If action X is socially efficient, i.e., $G_s > 0$, then (4) is positive and action X is chosen irrespective of the value of λ, the legal regime and the enforcement policy. If action X is socially inefficient, i.e., $G_s < 0$, and the law perfectly internalizes the harm, i.e., $G_p = G_s$, then (4) is negative and action Y is therefore chosen. It follows that the social concern parameter λ matters only in circumstances where self and group interest conflict (i.e., $G_p > 0$ and $G_s < 0$) and where liability rules are imperfectly enforced or non existent.

For such cases, whether efficiency concerns make a difference will depend on the distribution of the λ's in the population and on the choices facing the individuals. An individual with a sufficiently large λ will choose action Y even under No Law when this is the socially efficient action. For a given circumstance such as A or B, an individual who would not have chosen Y under No Law may well do so when liability rules are introduced, even if enforcement is non-deterrent, because the private gain G_p is decreasing in p while the social gain G_s is unaffected. Also, under No Law or under Mild Law, an individual who would have chosen Y in the circumstance A may well choose X in the circumstance B. This follows from the fact that both G_p and G_s are increasing in the gross benefit x from action X. Our prediction is therefore as follows: (i) in the circumstances A or B, the proportion of agents choosing X will decrease between No Law, Mild Law and Strong Law; (ii) under No Law or Mild Law, the proportion of agents choosing X will increase between the circumstances A and B.

Additional considerations. It may well be that, other things equal, some individuals strictly prefer not to cause harm. Specifically, when both actions are equally efficient, i.e., $G_s = 0$, and the legal system perfectly internalizes the harm caused to others, i.e., $G_p = G_s$, then some individuals strictly prefer action Y and presumably no one strictly prefers action X. This modifies the expression in (4) to

$$\Delta u_i = (1 - \lambda)G_p + \lambda G_s + \delta$$

where $\delta \leq 0$ captures a willingness to pay to avoid causing harm *per se*, i.e., even when the harm is compensated.

So far we have assumed that the λ's are exogenous. A possibility is that the λ's also incorporate reciprocity considerations and ultimately depend on the equilibria. For instance, $\lambda = \lambda_0 + q\theta$ where θ is a reciprocity parameter as in Charness and Rabin (2002) and q is the proportion of agents who "behave well" in the sense that they choose or are expected to choose Y in the circumstances A and B. In the literature on public good games with non-deterrent "legal obligations", it has been emphasized that such obligations may help coordinate conditional cooperators on better equilibria.[2] We will not formally study such equilibria but will allow for the possibility that the phenomenon exists. It may also be that liability rules introduce a sense of obligation or reinforce a "responsibility norm" irrespective of sanctions and expectations.

In any case, we take it that the social concern parameters may well depend on the legal regime. We expect that, at equilibrium, $\lambda_{Law} \geq \lambda_{NoLaw}$, i.e., legal liability does not crowd out social efficiency concerns. In the circumstances A and B, a change from No Law to Mild Law then reduces the proportion of individuals choosing X through two channels: (i) first, the private benefit of action X is reduced as discussed above; (ii) secondly, the law may reinforce social efficiency concerns. The latter effect may itself depend on the format of the liability rule. Compared with the negligence rule, the obligation under strict liability is unequivocal: in principle one is always "responsible" for the harm caused to others. In so far as this matters, we would expect $\lambda_{Strict} \geq \lambda_{Negligence}$, i.e., the impact on social efficiency concerns is at least as large with strict liability as with the negligence rule.

To sum up, the expression in (4) is now modified to

$$\Delta u_i = (1 - \lambda_r)G_p + \lambda_r G_s + \delta \tag{5}$$

where λ_r is the social efficiency parameter (possibly at equilibrium) in the legal regime r where r is either No Law (NL), strict liability (SL) or the negligence rule (NR).

[2] With reference to (3), beliefs about w_i and w_{-i} and therefore about the strategies of others will now be relevant in one's decision problem.

5 Results

Procedures. The experiment was computerized using z-Tree (Fischbacher 2007). We ran 10 sessions (two per treatment) in May, July and September 2016 in Québec (Canada) and Strasbourg (France). Each session included 20 participants, amounting to five groups of four subjects at each round, except for one *No Law* session that included only 16 participants. Overall, 196 participants took part in the experiment. An ECU was convertible to Canadian dollars at 30 ECU = 1 dollar or to Euros at 40 ECU = 1 Euro.

Decriptive Statistics Figure 1 displays the proportion of actions X undertaken in each of the four possible circumstances under every legal system. Clearly the presence of a legal system and the extent to which it is enforced greatly impact decisions when self and group interest conflict, i.e., in the circumstance A and B. In these circumstances, the proportion of actions X is greatest under No Law and reaches the smallest level under Strong Law, both under strict liability and under the negligence rule. Thus, it seems that a perfectly enforced legal system successfully achieves its main objective. By modifying private incentives, liability rules align private benefits with the public interest and therefore reduces the level of socially inefficient negative externalities.

However, the data also partially contradict several of the theoretical predictions based on purely self-interested preferences. First, in the circumstances A and B, there is a significantly smaller proportion of actions X under Mild Law than in No Law. In these cases, it is privately inefficient to undertake Y but about half of the participants choose to do so. Secondly, the proportion of actions X increases with the circumstances: more participants tend to undertake X in situation B compared with A or in situation D compared with C, although standard preferences would predict similar choices. Finally, we also observe that strict liability, both under Mild and Strong Law, yields a smaller proportion of actions X than the negligence rule or than No Law in the circumstances C and D, i.e., in circumstances where it is optimal to choose X. Under standard preferences, all participants would then be expected to always undertake action X.

We proceed as follows. First, we test the explanatory power of the theoretical model proposed in the last section. The goal is to assess how much the individuals' decisions can be explained in a framework combining self-interest and social efficiency concerns together with some additional considerations as discussed in Section 4. Next we use non-structural econometrics to corroborate this analysis while also taking into account additional factors that may impact the decision to undertake X. Finally, we explore the decisions to blame and sanction other participants.

5.1 Structural analysis

We first turn to a simple approach for understanding the data on the basis of the model developed in Section 4. For simplicity, we assume that all participants share the same fixed set of preferences except for an additive noise term in maximizing utility which crudely captures preference heterogeneity among participants. We fit the logit regression

$$P = \frac{e^{\gamma \Delta u}}{1 + e^{\gamma \Delta u}} \tag{6}$$

where P is the probability of undertaking action X, Δu is the difference in utility between the actions X and Y as defined in (5), and γ is the parameter capturing the sensitivity of behavior to differences in utility. The size of γ reflects the explanatory power of the model. Thus, we estimate a binary-response logit with the propensity score

$$\gamma \left[(1 - \lambda_r) G_p + \lambda_r G_s + \delta \right], \, r \in \{NL, SL, NR\}$$

Table 3 shows the regression results for a variety of restrictions on the parameter values. In Model 1, all parameters are constrained to equal zero except γ. This model therefore corresponds to purely self-interested preferences with respect to monetary payoffs. In Model 2, we allow for social concerns with a parameter that does not depend on the legal regime. The parameter is highly significant and the log-likelihood improves markedly. Individuals put non-negligible weight on the group payoff in addition to their own private benefit. In Model 3, we allow the constant term to differ from zero. The term has the expected sign indicating a reluctance to cause harm everything else equal, but the gain in explanatory power is slight and the willingness to pay to avoid causing harm *per se* is small (0.27 ECU).

In the last three models, we allow the social concern parameter to differ between legal regimes. The major gain here is with respect to the value of the parameter under strict liability. In Model 4a, λ_{SL} may differ from $\lambda_{NL} = \lambda_{NR}$. The difference is highly significant. Individuals put greater weight on others' payoff under strict liability compared with No Law and the negligence rule. Model 4b does the same with respect to λ_{NR} but this is much less successful. Finally, in Model 5, all restrictions are removed. The social concern parameters are quantitatively important in all legal regimes and they differ significantly between regimes. In this more flexible model, the reluctance parameter δ essentially vanishes. Figure 2 illustrates the predictions of the model.[3] To summarize, we have the following results.

Result 1 Individuals care about others' payoff. In the absence of legal obligations, there is a trade-off rate of 28% between private benefits and the losses imposed on others.

Result 2 Legal rules crowd in social concerns. The trade-off rate between private benefits and the losses imposed on others increases to 37% under the negligence rule and to 50% under the strict liability rule.

5.2 Determinants of action X

The predictive model developed in Section 4 explains a great deal of the data. About 43% of the individual decisions[4] are explained by expected private benefits, the effect on group payoff, and the impact of legal rules on social concerns. We now turn to non-structural econometrics to consider further refinements. We exploit the repeated choices of each individual by adding individual effects. In addition, we include the control variables of the questionnaire and also consider the impact of social blames and informal sanctions on the decision to undertake X.

The Tables 4 to 7 show estimates of models that seek to explain individual decisions to choose X over Y. The binary dependent variable is analyzed either by logistic regression with individual random effects, by linear probability regression with cluster at the session level, and by multilevel (alternatively called *hierarchical* or *mixed effects*) models with individual, session and treatment effects.[5] The Tables 4 and 5 consider a linear impact of expected private gain on the latent utility to undertake X, while the Tables 6 and 7 introduce a quadratic term.

In all models and all specifications, higher expected private gains are associated with a significantly higher probability to undertake action X rather than action Y. The quadratic estimations suggests that the impact of the marginal (expected) private gain is decreasing. As in the previous section, individuals are concerned with efficiency consequences when deciding to undertake X: the estimated coefficient associated of the social gain variable is significant in all specifications estimated by logistic regression and by the multilevel models, and in the Linear

[3] Allowing δ to differ between legal rules was not significant. We also tested for the role of risk aversion under imperfect enforcement with dummy variables subtracted from the expected private gain. This was also not significant.

[4] On the basis of the Pseudo R^2

[5] The statistical significance of the logistic regressions is the least conservative given that no satisfactory correction for within-session interdependence is available. The Linear Probability Model is more suited to control for within-session dependence but assumes normally distributed errors, which is obviously violated with a binary dependent variable. Multilevel models are the most suitable estimates since they take into account both the binary dimension of the dependent variable and considers random errors at the session (and even treatment) level.

Probability Model when the private gains include a decreasing return element.[6]

Our experiment allows participants to blame or assign small sanction points to their fellow participants at the end of each round. This aims at capturing the effect of social disapproval on behavior. Blaming or informal sanctions have been showed to substitute for or complement formal sanctions. Model 2 suggests that individuals are indeed less likely to undertake X when they have been blamed at the previous round. This effect also holds when we consider informal sanction points rather than blaming points (Model 3). The behavioral impact of social disapproval occurs mainly in situations where individuals were blamed after undertaking X (Models 4 and 5), while disapproval following action Y does not affect subsequent decisions. Last, the proportion of X actions by other participants in the past history of the game does not seem to affect one's decision to undertake action X.

Result 3 Individuals are less likely to reiterate a harmful action when they have been blamed or informally sanctioned at the previous period in situations where they have engaged in the harmful action.

5.3 Blames and informal sanctions

A second objective for introducing blaming and sanction points, which we now deal with, was to identify the determinants of disapproval as a reflection of the underlying social norms. Disapproval should reflect the participants' beliefs as to how one *should* behave. Because our groups are rematched after each round, note that disapproval can only be expressed about essentially anonymous persons, e.g., those who have been found liable.

Participants largely took advantage of the possibility to express disapproval. Figure 3 shows that participants received on average 4 to 8 blaming points per round depending on the treatment considered. Figure 4 suggests, nevertheless, that informal sanctions are very sensitive to the cost of punishment. When a legal regime is in force, social disapproval appears to be mainly concentrated on participants who were held liable, hence who are known to have undertaken X. This holds both when social disapproval is free (Figure 5) and when it is costly (Figure 6). However, there is also much disapproval of individuals who have not been held liable. This is particularly obvious under No Law where by definition no one is ever held liable. Hence "convictions" are not the only determinant of disapproval.

We ran a series of regressions on the decision of individual i to blame or sanction participant j of his group at round t. Recall that, at the end of each round, participant i observes whether participant j was condemned or not and he also observes the total number of actions X for which individuals were not held liable. From the latter, participant i infers a probability that a non-condemned participant j undertook the harmful action. This inference does not depend on the legal regime and in some cases the up-dated probability may very well equal unity. This allows us to disentangle two effects: (i) how one's belief that j engaged in X affects one's disapproval of j; (ii) the effect of a conviction *per se*, in which case of course the belief equals unity. The regression results are displayed in Table 8 for non-costly social disapproval and in Table 10 for costly sanctions. The main driver of disapproval appears to be the belief that other participants engaged in the harmful action. Convictions *per se* constitute an additional determinant but the quantitative effect is much smaller.

Engaging in action X does not necessarily imply socially inefficient behavior. The same observation holds with respect to convictions, except under the negligence rule. This leads us to inquire whether it is the belief that one has *inefficiently* engaged in X that elicits disapproval. For this purpose, we construct a variable [A DEVELOPPER] capturing participant i's belief that participant j engaged in X in the circumstances A or B. The aims is now to disentangle between disapproval of individuals who are believed to have caused harm and individuals who are believed to have caused harm inefficiently. Our results show that efficiency concerns do

[6]We suspect that the lack of significance in the Linear Probability Model without this feature results from the strong correlation between the expected private benefit and the social gain *within* each treatment.

constitute a determinant of disapproval, but again the main driver seems to be the mere fact of causing harm.

Finally, we also observe some tendency for "blind revenge " with respect to blaming. Participants are more likely to blame other participants the more they have been blamed themselves at the previous round. The effect is significant but quantitatively small and is non existent when social sanctions are costly. Although "blind revenge " is more likely to occur when an individual was blamed after undertaking Y as opposed to X, the difference is not statistically significant.

Result 4 The main driver of social disapproval is the belief that a participant has engaged in the harmful action. To a smaller extent, legal liability and the belief that a participant behaved inefficiently also elicit disapproval.

6 Discussion

The above results show great deviations between the predictions under standard preferences and the actual behaviors of participants. Our theoretical model developed above partially explains these deviations, by including efficiency concerns and considering potential *activation effects* of legal rules.

[A COMPLETER]

7 Conclusion

Liability rules mainly aim at deterring behavior that generate negative externalities. An extensive theoretical literature has discussed the merits of strict liability and fault-based legal systems in this respect. This body of research concludes that both types of rules, when perfectly enforced, achieve efficiency by aligning self-interest on collective interest.

We designed an experiment to investigate how agents behave with or without liability rules and when rules are weakly enforced. Our set-up is related to public good games in which participants are randomly matched with strangers and where private and group interests potentially conflict. In our setting, participants must decide between two actions, one of which generate negative externalities. In principle the legal rules considered completely align private and group interests but they are insufficiently deterrent when poorly enforced.

The participants' behavior contradicts the predictions under standard self-interested preferences. First, self-interest is far from explaining the range of observed decisions. The evidence suggests that individuals also care about group payoff. They are willing to partially trade-off private gains against smaller losses imposed on others. Second, our experiment reveals that the weight given to social concerns relative to private benefits is affected by the legal regime. Social concerns are stronger under the strict liability rule and weaker in the absence of any legal rule. The negligence rule stands in-between. The difference may well result from a *norm-activation* effect, i.e. social concerns are reinforced by the normative message conveyed by the legal rule, but the precise channel through which the effect operates warrants further research.

A further contribution of our paper consists in analyzing the role and determinants of social disapproval in this setting. Formal legal liability does not rule out social or normative pressure. As expected, participants are indeed less likely to engage in harmful conduct when they have been blamed or informally sanctioned by other group members. Secondly, individuals tend to disapprove of other group members who have or may be believed to have engaged in actions generating negative externalities. Causing harm rather than harm combined with socially inefficient behavior is the main driver of social disapproval.

References

Ambrus A., Greiner B. (2012). Imperfect Public Monitoring with Costly Punishment: An Experimental Study. American Economic Review 102, 3317-3332.

Angelova V., Armentier O., Attanasi G., Hiriart Y. (2013). Relative Performance of Liability Rules: Experimental Evidence. Theory and Decision, DOI 10.1007/s11238-013-9405-0.

Andreoni (2002)

Becker G. (1968. Crime and Pumishment: An Economic Approach. Journal of Political Economy 76, 169-217.

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

Bénabou R., Tirole J. (2011). Laws and Norms. NBER wp 17579.

Brennan G., Gonzales L. G., Güth W., 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.

Carpenter J.P. (2007). The demand for punishment. Journal of Economic Behavior and Organization, 62, 522–542.

Carpenter J.P., Matthews P.H. (2012). Norm enforcement: Anger, indignation or reciprocity. Journal of the European Economic Association, 10, 555–572.

Charness G., Rabin M. (2002) Understanding Social Preferences with Simple Tests. Quarterly Journal of Economics, 117, 817–869.

Charness G., Rigotti L., Rustichini A. (2016). Social Surplus Determines Cooperation Rates in the One-Shot Prisoner's Dilemma. Games and Economic Behavior, 100, 113-124.

Chaudhuri A. (2011). Sustaining Cooperation in Laboratory Public Goods Experiments: A Selective Survey of the Literature. Experimental Economics, 14, 47-83.

Cooter R. (2000a), Do Good Laws Make Good Citizens? An Economic Analysis. Virginia Law Review 86, 1577-1601.

Cooter R. (2000b), "Three Effects of Social Norms on Law: Expression, Deterrence and Internalization." Oregon Law Review, 79, 1-22.

Dal Bó E., Dal Bó P. (2014). Do the Right Thing: The Effect of Moral Suasion on Cooperation. Journal of Public Economics, 117, 28-38.

Deffains B., Fluet C. (2013). Legal Liability when Individuals Have Moral Concerns. Journal of Law, Economics, and Organization, 29, 930-955.

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

Ellickson R. (1998) Law and Economics Discovers Social Norms. Journal of Legal Studies, 27, 537-552.

Fehr E., Schmidt K. (1999). A Theory of Fairness, Competition and Cooperation. Quarterly Journal of Economics, 114, 817-68.

Fehr E., Gächter S. (2000). Cooperation and Punishment in Public Goods Experiments." American Economic Review, 90, 980-994.

Fehr E., Fischbacher U. (2004). Third-Party Punishment and Social Norms. Evolution and Human Behavior, 25, 63-87.

Fischbacher U. (2007).

Fischbacher U., Gächter S., Fehr E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. Economics Letters, 71, 397–404.

Funk P. (2007). Is there an Expressive Function of Law? An Empirical Analysis of Voting Laws with Symbolic Fines. American Law and Economics Review, 9, 135-159.

Funk P. (2010). Social Incentives and Voter Turnout: Theory and Evidence. Journal of the European Economic Association, 8, 1077-1103.

Galbiati R., Vertova P. (2008). Obligations and Cooperative Behavior in Public Good Games. Games and Economics Behavior 64, 146-70.

Galbiati R., Vertova P. (2014). How Laws Affect Behavior: Obligations, Incentives and Cooperative Behavior. International Review of Law and Economics 38, 48-57.

Grechenig K., Nicklish A., Thöni C. (2010). Punishment Despite Reasonable Doubt – A Public Good Experiment with Sanctions and Uncertainty. Journal of Empirical Legal Studies 7, 847-867.

Harrison G.W., Johnson E., McInnes M. M., Rutström E. E. (2005). Risk aversion and incentive effects: Comment. American Economic Review, 95, 897–901.

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

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

Kube S., Traxler C. (2011). The Interaction of Legal and Social Norm Enforcement. Journal of Public Economic Theory, 13, 639-660.

Landes W.M., Posner R.A. (1987). The Economic Structure of Tort Law. Harvard University Press, Cambridge, MA.

McAdams R.H., Rasmusen E.B. (2007). Norms in Law and Economics. In Polinsky, A. M. and S. Shavell (eds.), Handbook of Law and Economics, Vol. 1, North-Holland.

Markussen T., Putterman L., Tyran J.-R. (2014). Self-Organization for Collective Action: An Experimental Study of Voting on Sanction Regimes. Review of Economic Studies, 81, 301-324.

Masclet D., Noussair C., Tucker S., Villeval M.C. (2003). Monetary and Non-Monetary Punishment in the Voluntary Contribution Mechanism. American Economic Review 93, 366-380.

Polinsky M.A., Shavell S. (2007). The theory of public enforcement of law. In Polinsky A. M. and S. Shavell (eds.), Handbook of Law and Economics, Vol. 1, New York: North-Holland.

Posner E. (2000). Law and Social Norms. Harvard University Press.

Posner R.A. (1972). Economic Analysis of Law, 1st edn. Little, Brown and Company, Boston

Rabin M. (1993). Incorporating Fairness into Game Theory. American Economic Review, 83, 1281-1302.

Raz J. (1980). The concept of a legal system. Oxford University Press, Oxford.

Rege M., Telle K. (2004). The Impact of Social Approval and Framing on Cooperation in Public Good Situations. Journal of Public Economics, 88, 1625-1644.

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

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

Sunstein C.R. (1996). On the Expressive Function of Law. University of Pennsylvania Law Review, 144, 2021–53.

Shavell (2007), Liability for Accidents. In Polinsky A. M. and S. Shavell (eds.), Handbook of Law and Economics, Vol. 1, New York: North-Holland.

Tyran J., Feld L. (2006). Achieving Compliance When Legal Sanctions are Non-Deterrent. Scandinavian Journal of Economics 108, 135-156.

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

A Tables

Variable Name	Description
expGains_{it}	Expected private gains of individual i at round t of undertaking X rather than Y given the situation $k(it)$.
socialGains_{it}	Social contribution to the group welfare of individual i at round t of undertaking X rather than Y given the situation $k(it)$.
recBlames_{it}	Number of blames received by individual i at round t.
recSanctions_{it}	Number of sanction points received by individual i at round t.
X_{it}	Dummy variable equal to 1 if the individual i undertakes action X at round t, 0 if she/he chooses Y.
givBlames_{ijt}	Number of blames given by individual i to participant j at round t.
$\text{givSanctions}_{ijt}$	Number of sanction points given by individual i to participant j at round t.
history_X_{it}	Proportion of actions X that other group members of individual i undertook until round t.
condemned_{jt}	Dummy variable equal to 1 if individual j was condemned at round t, 0 otherwise.
uncompens_{it}	Number of uncompensated accidents (other than those of individual i) for group $g(i)$ at round t.

Table 1: Description of the variables.

Table 2: Descriptive Statistics

Variable	Mean	St. Dev.	Min	Max
expGains	6.505	6.759	-4	16
socialGain	.0133	3.151	-4	4
recBlames	6.262	4.615	0	18
recSanctions	.789	1.823	0	12
X	.7059	.4557	0	1
givBlames	2.087	2.441	0	6
givSanctions	.263	1.032	0	6
history_X	.696	.1839	0	1
condemned	.2616	.4396	0	1
uncompens	.2629	1.032	0	6

Table 3: Structural Econometrics. (N=3,920)

Model	Restrictions	γ	λ_{NL}	λ_{SL}	λ_{NR}	δ	LL
(1)	$\lambda_{NL} = \lambda_{SL} = \lambda_{NR} = \delta = 0$	.299*** (.01)					-1439.1
(2)	$\lambda_{NL} = \lambda_{SL} = \lambda_{NR}; \delta = 0$	.471*** (.019)	.391*** (.022)	.391*** (.022)	.391*** (.022)		-1366.9
(3)	$\lambda_{NL} = \lambda_{SL} = \lambda_{NR}$	.46*** (.019)	.349*** (.033)	.349*** (.035)	.349*** (.035)	-.273* (.158)	-1365.4
(4-a)	$\lambda_{NL} = \lambda_{NR}$	.464*** (.019)	.319*** (.035)	.488*** (.045)	.319*** (.035)	-.141 (.045)	-1357
(4-b)	$\lambda_{NL} = \lambda_{SL}$	.46*** (.02)	.351*** (.035)	.351*** (.035)	.344*** (.044)	-.276* (.159)	-1365.4
(5)	none	.463*** (.019)	.278*** (.044)	.498*** (.046)	.374*** (.044)	-.078 (.162)	-1355

Significance level: *** significant at 1% level; ** significant at 5% level; * significant at 10% level.
LL is the log-likelihood.
For Model 4-a: $H_0 : \lambda_{NL} = \lambda_{SL}$; p<1%.
For Model 4-b: $H_0 : \lambda_{NL} = \lambda_{NR}$; p=84.7%.
For Model 5: $H_0 : \lambda_{NL} = \lambda_{SL}$; p<1%.
$H_0 : \lambda_{NL} = \lambda_{NR}$; p<1%.
$H_0 : \lambda_{SL} = \lambda_{NR}$; p<1%.

Table 4: Regression of the decision to undertake X (linear in expected private gains).

Method	Logit			Linear Probability Model			Multilevel Model		
Model	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
expGains	0.351***	0.344***	0.484***	.0322***	.0341***	.0339***	.2430***	.2724***	.2555***
	(0.0261)	(0.0316)	(0.0456)	(0.00877)	(0.00662)	(0.00859)	(0.0351)	(0.0487)	(0.0655)
socialGains	0.237***	0.199***	0.285***	.0246	.0234	.0206	.3761***	.2868***	.5644***
	(0.0370)	(0.0447)	(0.0607)	(0.0230)	(0.0201)	(0.0241)	(0.0526)	(0.0719)	(0.1026)
recBlames_{t-1}		-0.0499***			-0.00295			-0.0546***	
		(0.0191)			(0.00203)			(0.0193)	
$\text{recSanctions}_{t-1}$			-0.139**			-0.0110***			-0.1416***
			(0.0548)			(0.00379)			(0.0558)
Period	0.0386***	0.0817***	0.143***	0.00254*	0.00521	0.00558	0.0389***	0.0852***	0.1466***
	(0.00977)	(0.0317)	(0.0387)	(0.00153)	(0.00335)	(0.00415)	(0.0098)	(0.0318)	(0.0392)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual	RE	RE	RE	RE	RE	RE	RE	RE	RE
Session				Cluster	Cluster	Cluster	RE	RE	RE
Treatment							RE	RE	RE
N	3920	1764	1764	3920	1764	1764	3920	1764	1764

Significance level: *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

Table 5: Regression of the decision to undertake X (linear in expected private gains). (Cont'd)

Method	Logit			Linear Probability Model			Hierarchical Model		
Model	(4)	(5)	(6)	(4)	(5)	(6)	(4)	(5)	(6)
expGains	0.355***	0.469***	0.325***	0.0334***	0.0334***	0.0312***	0.268***	0.256***	0.244***
	(0.0328)	(0.0458)	(0.0288)	(0.00566)	(0.00623)	(0.00950)	(0.048)	(0.066)	(0.0360)
socialGains	0.208***	0.295***	0.274***	0.0240	0.0234	0.0262	0.319***	0.557***	0.379***
	(0.0454)	(0.0607)	(0.0414)	(0.0186)	(0.0213)	(0.0242)	(0.071)	(0.103)	(0.0541)
recBlames$_{t-1}$ X$_{t-1}$	-0.0770***			-0.00243			-0.086***		
	(0.0225)			(0.00297)			(0.023)		
recBlames$_{t-1}$ Y$_{t-1}$	0.0429			0.00261			0.0399		
	(0.0382)			(0.00456)			(0.038)		
recSanctions$_{t-1}$ X$_{t-1}$		-0.202***			-0.0154***			-0.202***	
		(0.0658)			(0.00433)			(0.0671)	
recSanctions$_{t-1}$ Y$_{t-1}$		-0.0309			0.00662			-0.053	
		((0.0913)			(0.00708)			(0.0935)	
X$_{t-1}$	0.431	0.441*		0.0919***	0.150***		0.462	0.381	
	(0.280)	(0.257)		(0.0299)	(0.0249)		(0.281)	(0.259)	
history_X			1.557***			0.131*			0.418
			(0.596)			(0.078)			(0.670)
Period	0.0699**	0.137***	0.0455***	0.00347	0.00487	0.003**	0.0729***	0.142***	0.0469***
	(0.0272)	(0.0386)	(0.0108)	(0.00260)	(0.00361)	(0.00152)	(0.0274)	(0.0391)	(0.0108)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual	RE	RE	RE	RE	RE	RE	RE	RE	RE
Session				Cluster	Cluster	Cluster	RE	RE	RE
Treatment							RE	RE	RE
N	1,960	1,764	3,724	1,960	1,764	3,724	1,960	1,764	3,724

Significance level: *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

Table 6: Regression of the decision to undertake X (quadratic in expected private gains).

Method	Logit			Linear Probability Model			Multilevel Model		
Model	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
expGains	0.506***	0.470***	0.690***	0.0784***	0.0732***	0.0855***	0.4694***	0.4596***	0.6923***
	(0.0383)	(0.0485)	(0.0662)	(0.00675)	(0.00567)	(0.00727)	(0.0563)	(0.0622)	(0.0835)
$expGains^2$	-0.0148***	-0.0122***	-0.0240***	-0.00353***	-0.00303***	-0.00402***	-0.0139***	-0.0122***	-0.0251***
	(0.00249)	(0.00320)	(0.00453)	(0.000448)	(0.000441)	(0.000449)	(0.0029)	(0.0035)	(0.0049)
socialGains	0.211***	0.191***	0.273***	0.0259***	0.0253***	0.0236***	0.2457***	0.1999***	0.2875***
	(0.0371)	(0.0459)	(0.0619)	(0.00749)	(0.00829)	(0.00754)	(0.0533)	(0.0587)	(0.0769)
$recBlames_{t-1}$		-0.0576***			-0.00542***			-0.0578***	
		(0.0194)			(0.00200)			(0.0194)	
$recSanctions_{t-1}$			-0.159***			-0.0133***			-0.1563***
			(0.0560)			(0.00456)			(0.0567)
Period	0.0409***	0.0858***	0.156***	0.00313**	0.00651**	0.00682**	0.0407***	0.0864***	0.1543***
	(0.00994)	(0.0322)	(0.0399)	(0.00141)	(0.00319)	(0.00340)	(0.0099)	(0.0322)	(0.0399)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual	RE	RE	RE	RE	RE	RE	RE	RE	RE
Session	Cluster	Cluster	Cluster	Cluster	Cluster	Cluster	RE	RE	RE
Treatment							RE	RE	RE
N	3920	1764	1764	3920	1764	1764	3920	1764	1764

Significance level: *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

Table 7: Regression of the decision to undertake X (quadratic in expected private gains). (Cont'd)

Method	Logit			Linear Probability Model			Hierarchical Model		
Model	(4)	(5)	(6)	(4)	(5)	(6)	(4)	(5)	(6)
expGains	0.482***	0.671***	0.488***	0.0692***	0.0776***	0.0784***	0.4557***	0.6737***	0.4691***
	(0.0491)	(0.0670)	(0.0428)	(0.00619)	(0.00787)	(0.00660)	(0.0686)	(0.0814)	(0.0568)
expGains^2	-0.0123***	-0.0232***	-0.0141***	-0.00280***	-0.00352***	-0.00354***	-0.0117***	-0.0247***	-0.0138***
	(0.00317)	(0.00451)	(0.00262)	(0.000437)	(0.000532)	(0.000461)	(0.0036)	(0.0048)	(0.0030)
socialGains	0.197***	0.283***	0.229***	0.0266***	0.0263***	0.0260***	0.2203***	0.2969***	0.2474***
	(0.0464)	(0.0626)	(0.0417)	(0.00930)	(0.00892)	(0.00835)	(0.0648)	(0.0754)	(0.0540)
$\text{recBlames}_{t-1}\ X_{t-1}$		-0.0825***			-0.00563**			-0.0845***	
		(0.0226)			(0.00256)			(0.0228)	
$\text{recBlames}_{t-1}\ Y_{t-1}$		0.0398			-0.000449			0.0392	
		(0.0392)			(0.00504)			(0.0389)	
$\text{recSanctions}_{t-1}\ X_{t-1}$			-0.215***			-0.0196***			-0.2064***
			(0.0667)			(0.00582)			(0.0676)
$\text{recSanctions}_{t-1}\ Y_{t-1}$			-0.0595			-0.00271			-0.0662
			(0.0958)			(0.00725)			(0.0951)
X_{t-1}	0.413	0.377		0.0868***	0.116***		0.4357	0.3715	
	(0.286)	(0.265)		(0.0307)	(0.0322)		(0.2851)	(0.2640)	
history_X			0.883			0.0424			0.4743
			(0.609)			(0.0688)			(0.6646)
Period	0.0717***	0.150***	0.0477***	0.00441*	0.00603**	0.00357**	0.0722***	0.1466***	0.0479***
	(0.0276)	(0.0398)	(0.0109)	(0.00247)	(0.00291)	(0.00145)	(0.0276)	(0.0397)	(0.0109)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual	RE	RE	RE	RE	RE	RE	RE	RE	RE
Session				Cluster	Cluster	Cluster	RE	RE	RE
Treatment							RE	RE	RE
N	1,960	1,764	3,724	1,960	1,764	3,724	1,960	1,764	3,724

Significance level: *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

Table 8: Decision of individual i to blame individual j at period t.

Method	GLS with RE				Hierarchical Model			
Model	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
condemned$_j$	1.191***	1.729***	1.706***	1.705***	1.130***	1.782***	1.759***	1.758***
	(0.184)	(0.197)	(0.186)	(0.187)	(0.068)	(0.0879)	(0.0930)	(0.0930)
condemned$_j\times$ uncompens$_{t-1}$		-0.0963	-0.116	-0.114		0.0470	0.04919	0.0531
		(0.189)	(0.157)	(0.155)		(0.0879)	(0.0946)	(0.0947)
(1-condemned$_j$)$\times$ uncompens$_{t-1}$		0.535***	0.544***	0.545***		0.678***	0.710***	0.710***
		(0.089)	(0.093)	(0.092)		(0.040)	(0.042)	(0.042)
recBlames$_{t-1}$			0.0502***				0.0203***	
			(0.0121)				(0.0068)	
recBlames$_{t-1}\times$ X$_{t-1}$				0.0498***				0.01646**
				(0.0166)				(0.0083)
recBlames$_{t-1}\times$ Y$_{t-1}$				0.0743***				0.0405***
				(0.0106)				(0.0141)
X$_{t-1}$				-0.0416				0.05113
				(0.120)				(0.1006)
Period	0.0536***	0.0495***	0.0234	0.0224	0.05339***	0.04640***	0.02678**	0.0261**
	(0.0193)	(0.0176)	(0.0168)	(0.0161)	(0.0085)	(0.0093)	(0.0109)	(0.0109)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual	RE	RE	RE	RE	RE	RE	RE	RE
Session	Cluster	Cluster	Cluster	Cluster	RE	RE	RE	RE
Treatment					RE	RE	RE	RE
N	5,880	4,800	4,320	4,320	5,880	4,800	4,320	4,320

Significance level: *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

Note for model 4: $H_0 : \beta_{recBlamest-1\times X_{t-1}} = \beta_{recBlamest-1\times Y_{t-1}}$; p>10% for GSL and Multilevel models.

Table 9: Decision of individual i to blame individual j at period t. (NOUVEAU)

Method	GLS with RE					Hierarchical Model				
Model	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
condemned$_j$	1.191***	0.316*	0.286	0.285*	0.269*	1.130**	-0.0376	-0.0821	-0.151	-0.147
	(0.184)	(0.180)	(0.183)	(0.168)	(0.162)	(0.0681)	(0.0893)	(0.092)	(0.097)	(0.0969)
pX		1.493***	1.325***	1.305***	1.315***		1.916***	1.926***	1.975***	1.972***
		(0.211)	(0.195)	(0.188)	(0.181)		(0.099)	(0.099)	(0.105)	(0.105)
pX_AB			0.677	0.568	0.604			0.473*	0.349	0.340
			(0.560)	(0.470)	(0.468)			(0.255)	(0.273)	(0.273)
recBlames$_{t-1}$				0.0574***					0.029***	
				(0.0123)					(0.0061)	
recBlames$_{t-1}\times$ X$_{t-1}$					0.0615***					0.0311***
					(0.0166)					(0.0073)
recBlames$_{t-1}\times$ Y$_{t-1}$					0.0832***					0.0407***
					(0.0114)					(0.0134)
X$_{t-1}$					-0.152					-0.073
					(0.183)					(0.0939)
Period	0.0536***	0.0432**	0.0431**	0.0129	0.0121	0.0534***	0.0399***	0.0401***	0.0147	0.0145
	(0.0193)	(0.0210)	(0.0211)	(0.0202)	(0.0196)	(0.0085)	(0.0082)	(0.0082)	(0.0096)	(0.0096)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual	RE	RE	RE	RE	RE	RE	RE	RE	RE	RE
Session	Cluster	Cluster	Cluster	Cluster	Cluster	RE	RE	RE	RE	RE
Treatment						RE	RE	RE	RE	RE
N	5,880	5,880	5,880	5,292	5,292	5,880	5,880	5,880	5,292	5,292

Significance level: *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

Note for model 5: $H_0 : \beta_{recBlames_{t-1} \times X_{t-1}} = \beta_{recBlames_{t-1} \times Y_{t-1}}$; p>10% for GSL and Multilevel models.

Table 10: Decision of individual i to sanction individual j at period t.

Method	GLS with RE				Hierarchical Model			
Model	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
condemned$_j$	0.211***	0.274***	0.266***	0.266***	0.185***	0.246***	0.243***	0.242***
	(0.0483)	(0.0554)	(0.0643)	(0.0640)	(0.0285)	(0.0389)	(0.040)	(0.040)
condemned$_j \times$ uncompens$_{t-1}$		-0.0793	-0.0851	-0.0956		0.0355	0.027	0.028
		(0.0638)	(0.0747)	(0.0770)		(0.0387)	(0.040)	(0.040)
(1-condemned$_j$)$\times$ uncompens$_{t-1}$		0.0266	0.0241	0.0150		0.0826***	0.074***	0.075***
		(0.0390)	(0.0349)	(0.0398)		(0.0179)	(0.019)	(0.019)
recSanctions$_{t-1}$			0.0173				0.0065	
			(0.0144)				(0.0075)	
recSanctions$_{t-1}\times$ X$_{t-1}$				0.0178				0.0004
				(0.0128)				(0.0086)
recSanctions$_{t-1}\times$ Y$_{t-1}$				0.00326				0.023*
				(0.0218)				(0.013)
X$_{t-1}$				0.128				0.0190
				(0.111)				(0.033)
Period	-0.00938	-0.0159*	-0.0148**	-0.0151**	-0.009***	-0.0178***	-0.0178***	-0.018***
	(0.00871)	(0.00932)	(0.00744)	(0.00742)	(0.0037)	(0.0042)	(0.0049)	(0.0049)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual	RE	RE	RE	RE	RE	RE	RE	RE
Session	Cluster	Cluster	Cluster	Cluster	RE	RE	RE	RE
Treatment					RE	RE	RE	RE
N	5,880	4,800	4,320	4,320	5,880	4,800	4,320	4,320

Significance level: *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

Note for model 4: $H_0 : \beta_{recSanctionst-1\times X_{t-1}} = \beta_{recSanctionst-1\times Y_{t-1}}$; p>10% for GSL and Multilevel models.

Table 11: Decision of individual i to sanction individual j at period t. (NOUVEAU)

Method	GLS with RE					Hierarchical Model				
Model	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
condemned$_j$	0.211***	0.147	0.138	0.110	0.108	0.1851***	0.0308	-0.0026	0.0037	0.0024
	(0.0483)	(0.106)	(0.108)	(0.108)	(0.112)	(0.0285)	(0.0385)	(0.040)	(0.0415)	(0.0415)
pX		0.115	0.0665	0.0633	0.0654		0.261***	0.2653***	0.2533***	0.2544***
		(0.148)	(0.144)	(0.148)	(0.149)		(0.0445)	(0.0445)	(0.0467)	(0.0467)
pX_AB			0.177	0.262	0.266			0.321***	0.3553***	0.3568***
			(0.331)	(0.351)	(0.364)			(0.1049)	(0.1084)	(0.1084)
recSanctions$_{t-1}$				0.0328***					0.0064	
				(0.00841)					(0.0070)	
recSanctions$_{t-1}\times$ X$_{t-1}$					0.0391***					0.0017
					(0.0137)					(0.0079)
recSanctions$_{t-1}\times$ Y$_{t-1}$					0.0190					0.0217*
					(0.0116)					(0.0130)
X$_{t-1}$					-0.0340					0.0029
					(0.108)					(0.0313)
Period	-0.00938	-0.0102	-0.0103	-0.00756	-0.00731	-0.0093**	-0.0112***	-0.0112***	-0.0095**	-0.0096**
	(0.00871)	(0.00975)	(0.00979)	(0.00979)	(0.00965)	(0.0037)	0.0037	(0.0037)	(0.0042)	(0.0043)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual	RE	RE	RE	RE	RE	RE	RE	RE	RE	RE
Session	Cluster	Cluster	Cluster	Cluster	Cluster	RE	RE	RE	RE	RE
Treatment						RE	RE	RE	RE	RE
N	5,880	5,880	5,880	5,292	5,292	5,880	5,880	5,880	5,292	5,292

Significance level: *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

Note for model 5: $H_0 : \beta_{recBlames_{t-1} \times X_{t-1}} = \beta_{recBlames_{t-1} \times Y_{t-1}}$; p>10% for GSL and Multilevel models.

B Figures

Figure 1: Proportion of X actions across treatments and situations.

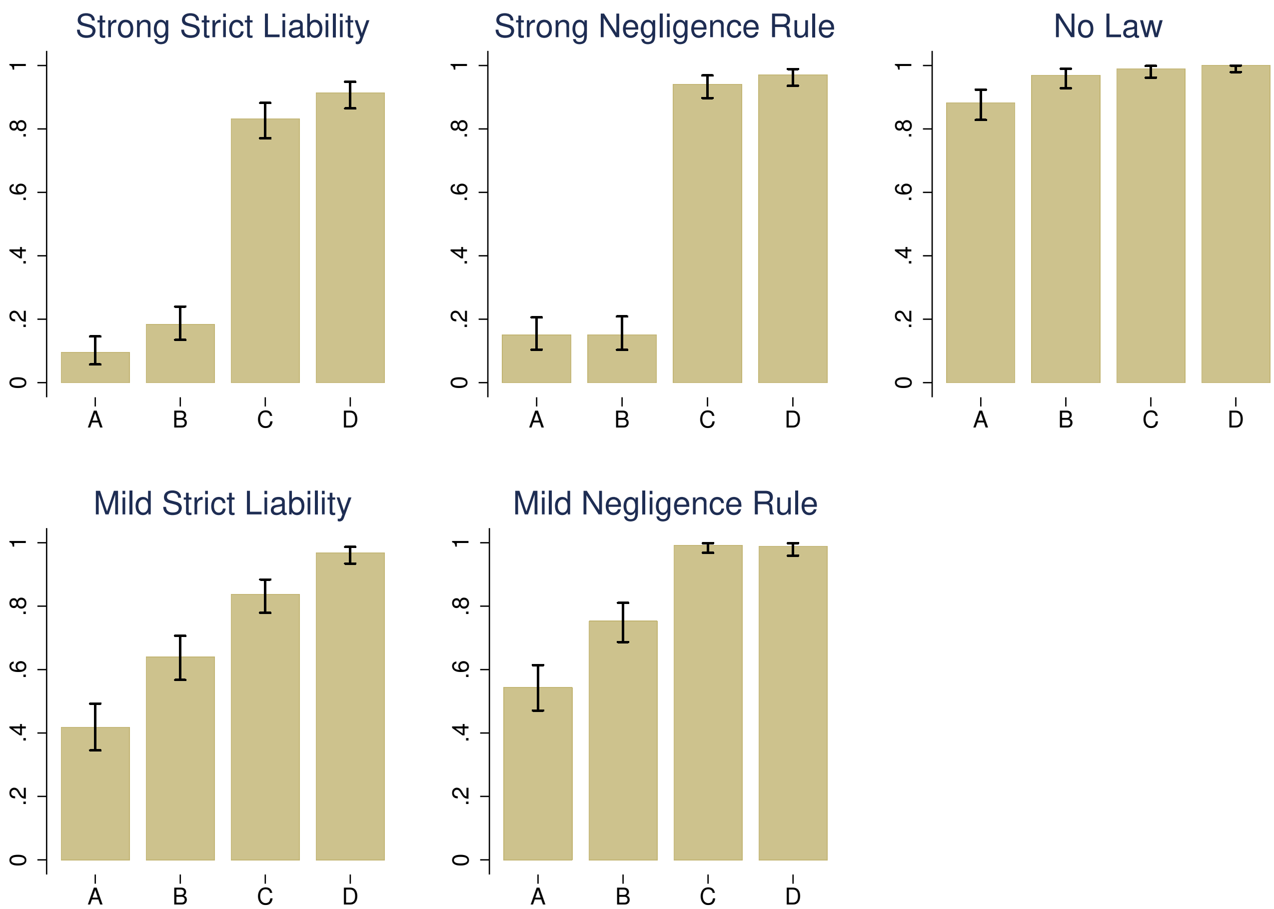


Figure 2: Prediction of behaviors after logit regression (Model 5, table 3).

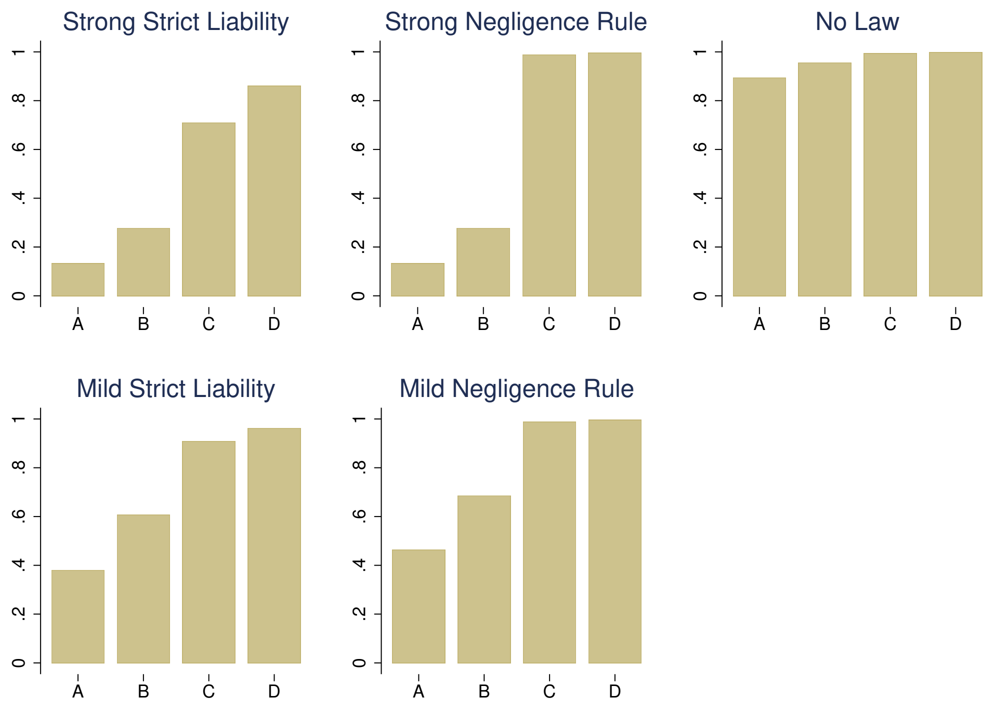


Figure 3:

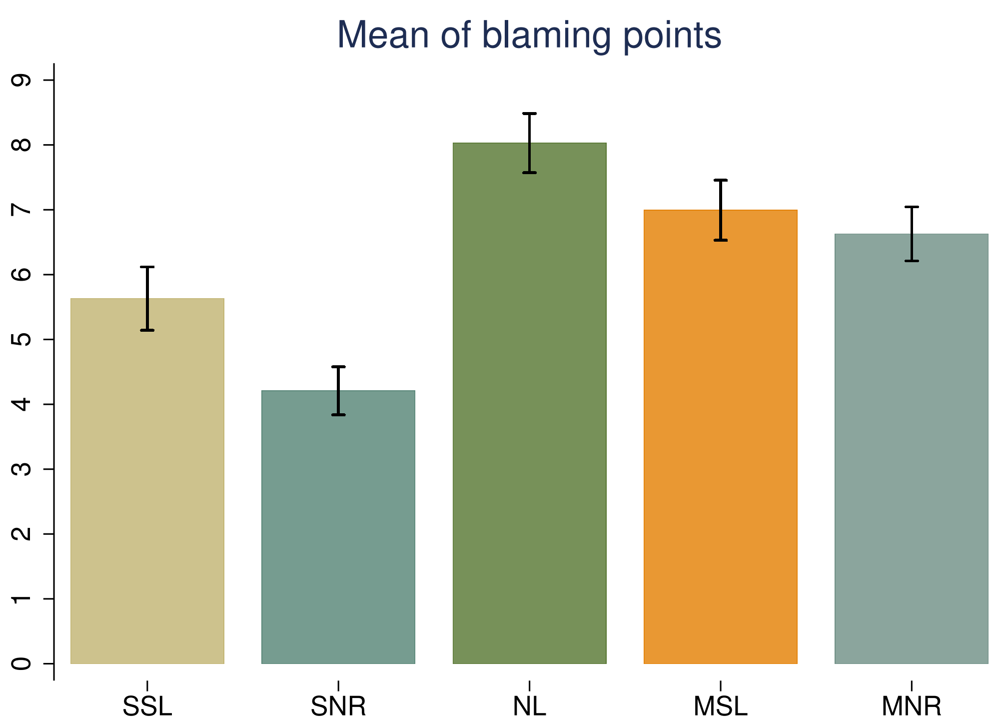


Figure 4:

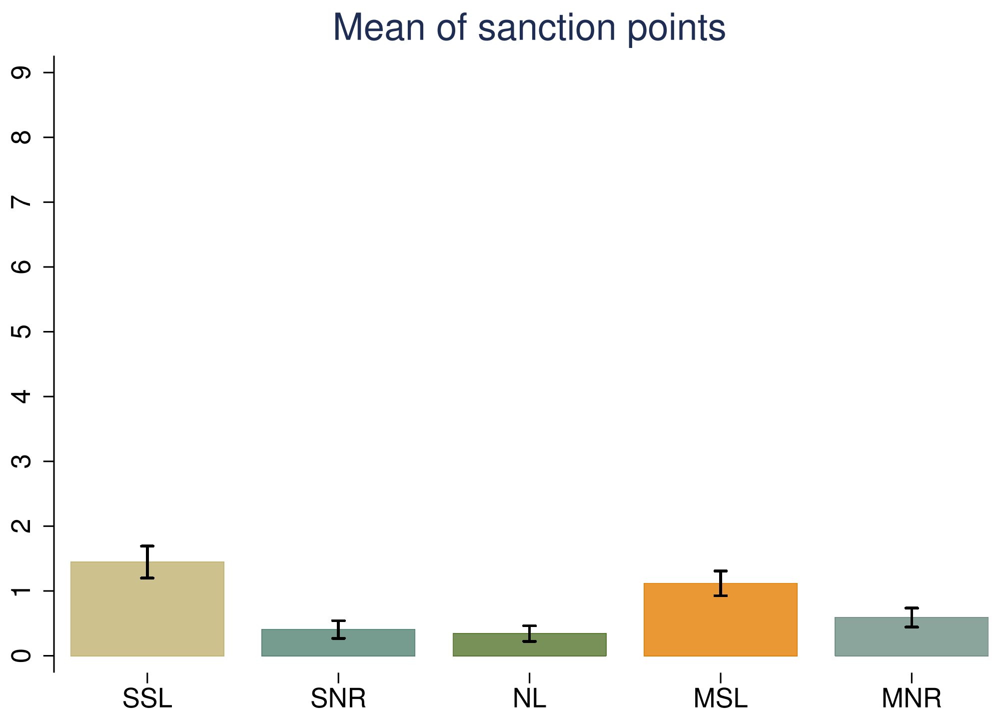


Figure 5:

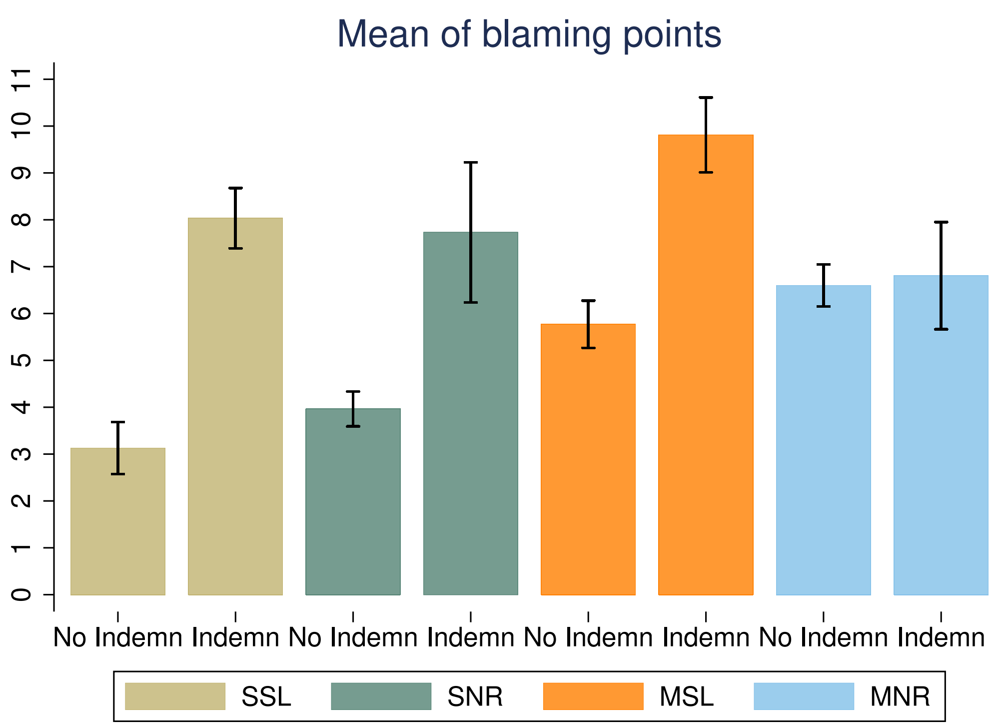
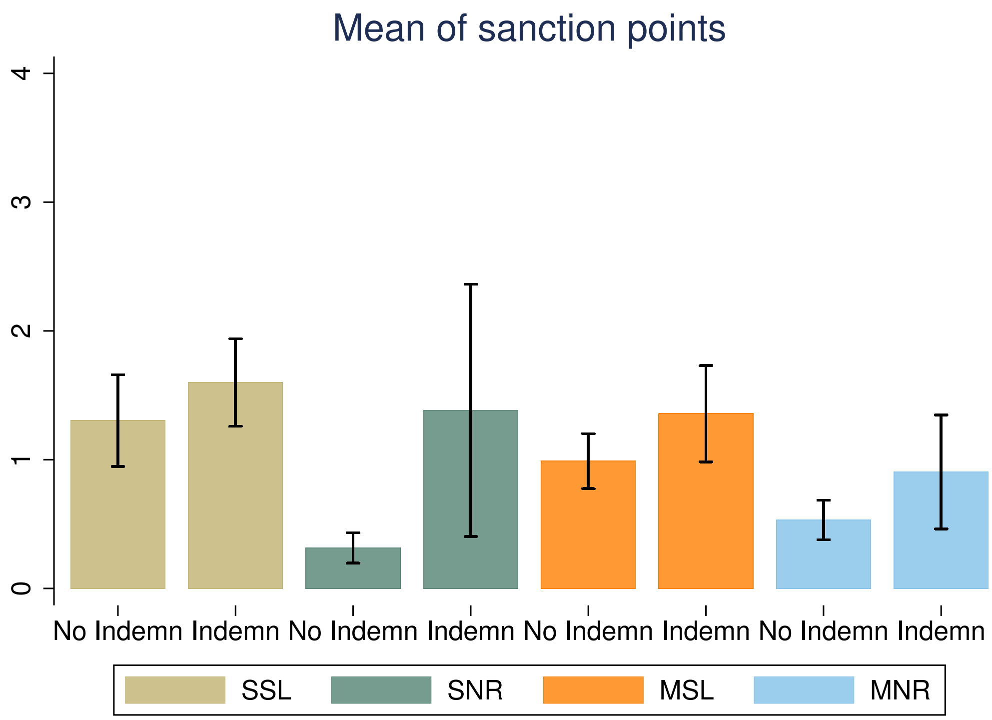


Figure 6:



Mean of sanction points
4
3
2
1
0
No Indemn
Indemn
No Indemn
Indemn
No Indemn
Indemn
No Indemn
Indemn
SSL
SNR
MSL
MNR