

# **On the propensity to settle or litigate in laboratory disputes**

## **Abstract**

This study experimentally investigates the role of the origin of the resources of victims in the propensity to settle or litigate in laboratory disputes. The analysis focuses on victims' decision to use a costly and uncertain procedure, that is, litigation, to recover money from offenders and on offenders' decision to hand over what taken before litigation starts. Results suggest that although the share of offenders who return money increases once they learn that the victims have had to work to gain their resources, the threat of a costly dispute is stronger than any consideration about the origin of victims' resources. In this study, the religiosity and gender of decision makers are included among the control variables because these individual features are traditionally debated as able to affect individuals' attitude in conflicts and their resolutions. In particular, gender seems to play a significant role since main results are driven by female decision makers.

JEL classifications: C91, D63, D74, K40.

Keywords: Litigation; Settlement; Costly second-party punishment; Laboratory experiment.

---

The authors wish to stress that at the time the experiment was carried out, the university in which the lab is located did not require ethical approval for economics experiments; nor did a specific committee exist. Consequently, no ethical approval was required. Questionnaires are anonymous. No confidential information can be related to any specific participant who provided informed consent.

## 1. Introduction

In-court disputes are unpleasant, uncertain, and costly. Meanwhile, out-of-court settlements offer an alternative solution; here, the victim asks the offender for (at least some) compensation without immediately resorting to litigation and does so in the “shadow of the court.” Settlements are considered an efficient solution because they are faster, cheaper, and usually less uncertain than judicial decisions; when a judicial decision is strongly expected to be in favor of or against one of the two litigants, it presents a solid incentive to settle. The possibility of appealing to a court, even in the background and even without actually undertaking it, drives most settlements.

Investigating when and why parties succeed in settlements or fail and go to court, even though litigation is a costly strategy, is interesting from an economic point of view. On the one hand, law and economics scholars typically deal with the problem by assuming rational and self-interested utility-maximizer litigants who are concerned about their own monetary-equivalent expected benefits and costs of actions. In this stream of literature, the litigation-vs.-settlement dynamics is usually modeled as a strategic interaction: the main reason for settlement failure is asymmetric information while high litigation costs favor settlement (Boyd and Hoffman, 2012; Daughety and Reinganum, 2008; Schwartz and Wickelgren, 2009). On the other hand, behavioral economists have discussed the role of cognitive biases and intrinsic motivations in settlement dynamics.<sup>1</sup> Intrinsic and individual motivations might be related to fairness concerns,<sup>2</sup> desire for revenge, taste for punishment, unselfish attitudes, or, conversely, pro-social attitudes, etc.<sup>3</sup> The experiment proposed herein is aimed at understanding whether fairness concerns and litigation costs may have any discouraging effect on the decision to undergo a trial when settlement is feasible.

Remarkably, the literature largely disregards the triggers that can prompt a potential plaintiff to proceed against a defendant even when commencing a lawsuit implies relevant costs and ultimately lead to disadvantageous results. Furthermore, individual motivations affecting the attitude of a potential defendant are neglected in existing studies.<sup>4</sup>

---

<sup>1</sup> Besides confirming most of the theoretical predictions, behavioral law and economics recognizes a role for cognitive biases in the settlement behavior of parties. See Korobkin and Ulen (2000), Robbennolt (2014), and Landeo (2018) who provides a comprehensive survey of the literature. Divergent beliefs able to prevent settlement agreements may originate not only from asymmetric information (see, among others, Sullivan 2016) but also from self-serving bias (Loewenstein et al. 1993; Babcock et al. 1995; Babcock and Loewenstein 1997, Farmer and Pecorino 2002). Similarly, anchoring, and framing effects and reactive devaluation are documented in settlement circumstances by Korobkin and Guthrie (1994), Babcock and Pogarsky (1999) and Pogarsky and Babcock (2001).

<sup>2</sup> Bies (1986) overviews the topic across disciplines. Rawls (1963) provides a seminal contribution on the topic.

<sup>3</sup> Korobkin and Guthrie (1994), Guthrie (1999), Robbennolt (2014) and (2018) overview psychological phenomena affecting litigation and settlement. Landeo (2007) analyses the issue in an experimental setting of strategic litigation. On intrinsic motivations in behavioral law and economics, see Mitchell (2014).

<sup>4</sup> Motivations behind the settlement/litigation dynamics are explored by Hollander-Blumoff (2011, pp. 65-66). On the importance of intrinsic factors including values, principles, and emotions, see Hadfield (2008) and Abrams (2011).

The current study contributes to the literature by presenting an original experiment that seeks to identify the possible reasons why people may prefer in-court litigation or out-of-court settlement. Although many individual factors might be relevant, we opt to refer to the existing behavioral economics literature in selecting a sufficiently consistent set of variables to be studied. This study primarily investigates how the *origin* of victims' resources influences the individual propensity to settle/litigate. When an individual causes damage to another person's wealth, this individual's proneness to restore the victim may credibly depend on the origin of the damaged resources (wages, bequests, received gifts, lottery prizes, etc.).

By following Druckman and Wagner (2017) and Albin and Druckman (2014), we can frame the experimental investigations into resource origin in the debate about the relation between *justice and fairness concerns* and success of settlement agreements. Cappelen et al. (2007) showed that in their ideals of fairness, people care about the *investment* made by their counterparts in their interaction. By considering two alternative origins of victims' resources, the design of the present experiment allows the elicitation of some specific individual concerns of victims and tortfeasors. In particular, the related results can be cautiously used to infer how *fairness* concerns influence litigants' decision making.

The analysis includes some control variables, including gender and religiosity of decision makers. The extant literature, indeed, shows that women and men may have different preferences in socio-economic contexts (Croson and Gneezy, 2009) and are more averse to unfair behaviors, which they are more prone to punish than men (Eckel and Grossman, 1996). Moreover, gender and religiosity are generally, though controversially, associated with different attitudes in conflicts and dispute resolutions. According to a part of the literature, men and women behave significantly differently in litigation (Brahnam et al., 2005; Menkel-Meadow, 2012; Wofford, 2017). Meanwhile, the influence of religiosity on decisions involving fairness issues remains debatable (Henrich et al., 2010).

This study is a novel attempt to find experimental evidence on the settlement-vs.-litigation dynamics by exploring how victims commence an action against offenders and how the latter reacts to requests for restitution on the basis of the origin of the victims' resources.

To mirror a simple *tort* dispute emerging between two *strangers*, our experiment adapts a costly second-party punishment design (Güth et al., 1982). As usual, if a participant hurts a counterparty, the latter can act against the former by means of a costly action. In standard costly punishment experiments (Carpenter and Matthews, 2012; Fehr and Fischbacher, 2004; Henrich et al., 2006; Leibbrandt and López-Pérez, 2012), the punishment reduces the offender's payoff, and the victim bears the cost while never recovering any part of what was lost. Although standard settings facilitate the investigation into the motivations and concerns underlying punishment decisions (Falk and

Fischbacher, 2002; Feess et al., 2015; Sonnemans and van Dijk, 2012), they fail to grasp many of the dimensions of litigation. For this reason, departing from the standard designs, in the present setting the victim might obtain full compensation.

Although limited to a laboratory dispute that mirrors a simple tort case arising between two strangers, this study reveals that offenders are more prone to settle when their victims had to work to gain their money. However, the threat of a costly dispute seems more effective in inducing restitution than the consideration of resources' origin. The power analysis and especially the econometric analyses provided in this study reveal that the conditional treatment effects identified herein are reasonably robust.

## 2. Background and experimental hypotheses

To properly capture the uncertainty of trial and the fact that victims can obtain a restoration of the damage they have suffered by proceeding against the offenders, we adapt the costly probabilistic punishment (see, for example, Garoupa, 2003; Perry et al., 2002; Qin and Wang, 2013) that can be undertaken by the victim.<sup>5</sup> In particular, we define costly punishment as a computerized litigation (Sullivan, 2016) that can lead to a full damage restoration with some probability. This allows us to adapt *uncertainty* and *redress* that characterize litigation to a second-party punishment design. Given the possibility to proceed by means of computerized litigation, parties are nonetheless given the chance to settle the case simply by agreeing on the restoration.<sup>6</sup>

Then, we focus on how the origin of a victim's resources (later interpreted as a way to elicit ideals of fairness) influences the individual propensity to settle or litigate. Given that the extant literature suggests the relevance of further individual characteristics, we include the gender and religiosity of the offender and victim among the list of controls deserving attention. As discussed in Section 4, these two variables are relevant to clearly identify the treatment effect.

Finally, the game proposed in this study is similar to the power-to-take setting (Bosman and van Winden, 2002), where two players interact: the first player may take a part or all of the second player's endowment while the second player may partially or totally destroy one's own initial endowment before the first player takes his or her desired share. Bosman et al. (2005) showed that people who

---

<sup>5</sup> Qin and Wang (2013) show that high probabilities (around 50%) of being punished are more deterrent than high fines. This might be due to the fact that, in general, people tend to subjectively magnify objective probabilities, thus increasing the deterrence effect (Perry et al., 2002 and Garoupa, 2003).

<sup>6</sup> Costly-punishment design has been already used in the experimental literature to simulate legal disputes in the laboratory and investigate behavioral issues in the judgment process (Falk and Fischbacher 2002, Feess et al. 2015, Lewis et al. 2015, Sonnemans and van Dijk 2012). However, these contributions are mainly based on a third-party punishment design that does not allow investigating the attitude of parties towards settlement and litigation.

earn their endowment destroy less often but face larger stakes than those who receive their endowment for free; indeed, the former chooses to destroy the entire endowment more often than the latter, who generally chooses an intermediate share. The experiment presented herein differs from the power-to-take game, that is, in the design used in this study, the victim may proceed to a trial while the offender is allowed to return what was taken.

### *2.1 Origin of victim's resources*

Previous studies have suggested that the origin of the property that is damaged or taken from a counterparty is a key factor that influences the choices of the decision makers involved in experiments. Even if people generally prefer punishment and even when punishment is costly for the punisher (Fehr and Gächter, 2000; Henrich et al., 2006), victims are less punitive when they had to work diligently to earn their initial endowment than when they received their endowment exogenously (Bosman et al., 2005). Results with non-monetary payoffs are qualitatively the same, although in this case, the reaction against the transgressor is larger (Masclet et al., 2003). However, on the basis of frameworks for the provision of public goods, Antinyan et al. (2015) and Muehlbacher and Kirchler (2009) observed that people who have had to earn their resources tend to punish defectors who do not contribute to public good less than those who have received their endowment effortlessly.

The origin of endowment can also affect the decision to misbehave. According to Gravert (2013), participants who earn their resources are more likely to take undeserved payoff than participants with randomly allocated endowment. According to the author, one explanation for this outcome is that individuals who worked to earn their resources feel that they deserve an extra payoff. This sense of deserving can come from the effort exerted in the task, the time spent on the task, or a difference between an individual reference point and the actual outcome.

Cappelen et al. (2007) investigated the role of the initial investment in a dictator game experiment where the contribution phase is preceded by a production phase. Their results showed that in the contribution phase, decision makers tend to care about the *investment* made by their opponents. The authors explained this evidence in terms of the fairness concerns of decision makers, although people have different conceptions of what fairness is.<sup>7</sup>

---

<sup>7</sup> Cappelen et al. (2007) distinguish between *strict egalitarianism* (all inequalities should be equalized), *libertarianism* (each person must receive what he/she produces), and an intermediate ideal of fairness, *liberal egalitarianism*.

Although the extant literature provides ambiguous predictions about how the effort made to obtain some resources may affect the subsequent behavior of people faced with an offense, it mainly points in the same direction: people who exerted effort to gain a resource defend it more strenuously than those who received such resource effortlessly. Thus, on the basis of the findings of Cappelen et al. (2007), Fehr and Gächter (2000), Henrich et al. (2006), and (Masclet et al. 2003), we formulate the following empirical hypothesis:

*H1 Origin of the victim's resources.* If we expect any role of the origin of the victim's resources in litigation, we aim to verify the following:

*H1.1 Offenders who learn that their victims have had to work for their money restore more often than those whose victims have received their endowment from the experimenter.*

*H1.2 In the case of no restitution, victims who have had to work for their money commence litigation more often than those who have received their endowment from the experimenter.*

## *2.2 Control variables*

As mentioned previously, two relevant control variables that must be considered are decision makers' gender and religiosity. These variables are not the focus of the analysis, but in light of their importance in experimental economics and dispute resolution, their role calls for a comprehensive discussion.

Although behavioral differences between women and men are often explained by risk aversion in experimental economics, Filippin and Crosetto (2016) concluded that gender differences systematically correlate with the features of the method used to elicit risk attitude. Other authors have suggested that behavioral differences between women and men may be explained in terms of procedural justice (Kitzmann and Emery, 1993; Lind et al., 1994).

Focusing on specific areas of law, most studies have shown that women and men tend to behave differently. Although gender cannot be considered as a direct indicator of intrinsic motivations and that differences between males and females may arise from the fact that in reality, women may experience the legal system differently from men (i.e., socialization, social stereotypes, and discrimination may be important), the laboratory setting helps to attenuate such concerns. Experiments show that women and men display considerable heterogeneity in their preferences and actions. On the one hand, women seem to be systematically more collaborative in conflicts than men, preferring cooperation over coercion and control (Burke and Collins, 2001; Thomas, 1994). Menkel-Meadow (2012) highlighted that women are less aggressive than men and that they generally look for solutions, which do not imply lawsuits. Wofford (2017) showed that women approach litigation

differently from men: although males and females are generally reluctant to file suits and prefer either mediation or resolution of the dispute outside the court, women usually favor the more cooperative and collaborative methods of dispute resolution over proceeding to trial. However, this gender effect emerges only in certain types of cases and at certain stages of the legal process (Frazier and Hunt, 1998). Then again, experimental evidence has shown that women are more sensitive to misbehavior than men; therefore, when given the opportunity to punish offenders, they take it more often than men (Croson and Buchan, 1999; Croson and Gneezy, 2009; Eckel and Grossmann, 1996; Geniole et al., 2015).

Concerning religiousness, it may affect an individual's sense of fairness and pro-social behavior (Brañas-Garza et al., 2009, 2014). However, the extensive survey by Hoffmann (2013) provides very mixed and often weak evidence about the relation between religion, fairness concerns, and pro-social behavior, such as cooperation, trust, and altruism. Nevertheless<sup>8</sup>, all these dimensions are possibly relevant to the propensity to settle or litigate. As explained by Henrich et al. (2010), participation in a world religion is associated with some measure of fairness and pro-sociality. Among the existing studies on the significant relations between religiousness and pro-social behavior, we consider the works of Ahmed (2009), who found higher levels of altruism in subjects classified as religious; and Benjamin et al. (2016), who clustered subjects into those with religious affiliations (Protestant, Catholic, and Jewish) and those without. The authors examined whether religious priming influences cooperation and found that Catholicism has a negative effect on cooperation. Furthermore, Migheli (2017) showed that when sharing their endowments with unknown people, atheists propose more generous shares than Catholics.

Following these previous behavioral contributions, we expect gender and religiosity variables to be statistically significant, thereby providing interesting evidence of their role in the settlement-vs.-litigation dynamics.

As the standard theoretical literature on settlement (Daughety and Reinganum, 2008; Schwartz and Wickelgren, 2009) shows that higher litigation costs favor settlement, we control for the cost of litigation. Herein, the experimental design, as presented in the next section, allows variations in litigation fees.

### **3. Experiment**

#### *3.1 Design and procedure*

---

<sup>8</sup> About the difference between morality and religiosity, see also McKay and Whitehouse (2015).

Although with caution in terms of external validity, the present experimental design is aimed at mirroring real situations of tort litigation. A tort is a violation where one person causes damage to another person. Violations may result from intentional actions, a breach of duty, or a violation of statutes. Tort lawsuits are the largest category of civil litigation and encompass a wide range of cases; many of them emerge between strangers who do not have any previous relationship. Examples include nuisance and trespass, car accidents (according to our setup, limited to the case wherein the offender intentionally misbehaves), damages and injuries determined by children/animals/goods intentionally and not properly controlled, product liability, and damages caused by people who intentionally decide to misbehave for their advantage while harming another person.

The specific circumstance that we simulate in the laboratory involves an individual who can increase her initial wealth by damaging (the wealth of) another person. If this happens, the victim can ask the offender for full reparation (i.e., restitution) in the shadow of the court. Hence, if the request is denied, a simplified in-court litigation can be commenced by the victim. Litigation is uncertain and costly for both parties, and the parties do not know each other.

The experiment consists of two phases and involves two types of players randomly assigned to their role. We denote the two types of players as subjects *A* and subjects *B*. Phase 1, which involves only subjects *A*, involves the implementation of the treatment related to the origin of the resources of subjects *A* (*effort treatment* vs. *no effort treatment*). Phase 2 involves both types of subjects and allows the observation of their behavior. The translation of the instructions for the participants is provided in the Supplementary Appendix available upon request (SA, hereafter). In what follows, we summarize the experimental procedure.

*Phase 1* Subjects *A* are randomly assigned to one of the two treatments related to the origin of endowment. During this phase, subjects *B* are not present. They are called for a later appointment in the laboratory. Therefore, subjects *B* have no idea about the treatment of phase 1 involving subjects *A*. Appendix 1 provides a summary of the experimental treatments.

- Under the *effort treatment*, subjects *A* work for 30 minutes. At the end of the required task, they earn a sum corresponding to their resources in the second phase of the experiment. The data entry task entails recopying quadruplets of fictitious badge numbers, names, and marks in a form displayed on a computer screen. In the program, the subjects need to recopy one quadruplet at a time and cannot proceed further until they have entered all the information correctly. The performance of each subject *A* is measured as the number of quadruplets correctly recopied in 30 minutes. At the end of the task, the subjects are ranked in tertiles according to their performance.

The subjects ranked in the best tertile receive €15, those in the second tertile receive €10, and those who perform least receive only €5 (screens 1.1–1.5 in the SA).

- Under the *no effort treatment*, subjects *A* do not work but receive their initial endowment directly from the experimenter. Randomly, one-third of the participants receive €15, one-third receive €10, and one-third receive €5 to replicate the distribution of resources of subjects *A* under the effort treatment (screen 3 in the SA).

In sum, the experiment is based on a 2 x 2 between subjects design, with two treatments for the origin of the endowment and other two involving the cost of the litigation fee. The endowment in tertiles does not represent an additional treatment, as it is common to all the potential victims, no matter whether they worked to earn it. Therefore, any effect of the different endowments, which may be present, cancels out when the two groups of subjects of interest (those who worked and those who did not) are compared.

Phase 2 involves both subjects *A* and *B*. *Bs* enter the lab at this stage and receive €8 as an initial endowment. Although subject *Bs* in the *effort treatment* do not see subjects *A* working, they observe that as they enter the lab, *As* are already sitting inside; such a procedure should reinforce their perception that subjects *A* have already performed some tasks. In the *no effort treatment*, subjects *A* and *B* enter the lab together; such a procedure renders—once revealed—the information that *As* have received their endowments from the experimenters and not from completing a task credible for subjects *B*. At the beginning of phase 2, each subject *A* is anonymously paired with a subject *B*.

- The game starts with subject *B*'s choice to take €2 from resources of subject *A*. At this stage of the game (screen 3 in the SA), *Bs* do not know the origin of subject *A*'s resources, but all the players are acquainted with the following information:
  - If subject *B* takes €2 from subject *A*, the latter can start a costly computerized procedure to recover the sum.
  - Each subject knows one's own cost of the computerized procedure but not that of the counterparty. The litigation fee is set at either €1 or €3 for all the subjects.
  - The probability of winning the computerized procedure is communicated to all subjects at the start of the experiment; this probability is set to be equal to 0.5.<sup>9</sup>

---

<sup>9</sup> The probability that the litigation is won by one of the parties is set at 50% to help participants with a very easy-to-understand setting. This means that the computerized procedure cannot acquire specific proofs in favor of/against the plaintiff/defendant and decides the case by simply “flipping a coin”. In more realistic scenarios, where the judge can observe items of evidence, one could assume different litigation risks.

- The origin of subject *A*'s resources is disclosed to subject *B* after the latter's decision to take €2 but before the possible commencement of litigation. At this stage, *B*s are informed about whether their counterparts had to work to gain their money or received their endowment effortlessly. Subjects *B* now have their first chance to return what has been taken. If subjects *B* decide not to return the money, subjects *A* are asked whether they want to threaten subjects *B* by initiating a litigation if the latter declines restoration. Related decisions are then communicated to subjects *B*, who now have another chance to return the sum.
- If subjects *B* maintain their decision against restoration, computerized litigation starts.

Figure 1 summarizes the different steps of phase 2; further details of the instructions for the participants are shown in the SA (screens 4–8).

Finally, a brief standard questionnaire (Lotito et al., 2013) is administered to the participants at the end of each session to collect sociodemographic information, such as gender, religious orientation, and whether the person habitually volunteers.

Computerized sessions were conducted using the software z-Tree (Fischbacher, 2007). The sessions involved 240 undergraduate students (127 males and 113 females) attending different undergraduate degree programs at the University of Milano-Bicocca and Turin. The participants were recruited using ORSEE software (Greiner, 2015). Each session involved 12 students as potential victims and 12 students as potential offenders.<sup>10</sup>

[INSERT FIG. 1 ABOUT HERE]

**Fig. 1** Phase 2: Tree representation of the game between a pair of subjects *A* and *B*, with payoffs in parentheses and frequencies in brackets

### 3.2 Comments on the design

Concerning the initial decision of the potential offender to take money, two comments are needed. First, the experiment design effectively mirrors intentional torts and torts involving a tortfeasor who deliberately did not take precautions to save money. Obviously, completely unintentional accidents are not represented (and often do not imply liability). Second, when deciding to take money, the decision makers are aware that litigation is a possible consequence of their choice. This allows participants to frame their laboratory experience as a situation mirroring a legal dispute.

---

<sup>10</sup> At the time of the experiment (2017), the universities in which the laboratories are located did not require ethical approval for economics experiments, nor did a specific committee exist. Consequently, no ethical approval was requested or required.

In the design, litigation is extremely simplified and remains in the background; nevertheless, it is effective in allowing participants to frame litigation as an uncertain and costly solution. The threat of a costly trial provides economic incentives to settle. Litigation is costly for both parties, and we consider small changes in this cost across sessions to observe the effect of litigation costs on parties' decisions.

Concerning parties' endowments and costs and the related payoffs, they are set in such a way that an interpretable setup is guaranteed. If subjects are risk neutral or averse, with complete information, the strategy ("taking money," "not restoring," "litigating") is dominated by the alternative ones. We opt for this setup to guarantee a positive number of subjects  $B$  who opt to take their counterpart's money and to insulate as much as possible the effects related to the origin of the resources of subjects  $A$ . As in the case of experiments on the ultimatum game (Cabrales and Ponti, 2015; Camerer, 2003; Hernández and Pavan, 2015), where the results depart from the theoretical Nash equilibrium based on agents' rationality and self-interest because of other-regarding concerns, we expect that some  $B$ s decide to restore, and some  $A$ s who are not restored decide to litigate anyway despite the higher costs implied by such choices; significant differences are also expected depending on the origin of the resources of subjects  $A$ . To have some room for litigation, the tortfeasors do not know the litigation costs of the victims: once threatened, litigation might appear as a credible threat to some extent.

Finally, the design proposed in this experiment guarantees that the parties never know each other on purpose; this helps to rule out motivations such as empathy and friendship, which exist in the real world and are likely to affect people's behavior when facing an offense followed by a legal conflict. Such a choice does not mean that the experiment aims to neglect what happens in the real world; rather, it isolates motivations that do not depend on a specific situation while being common to all.

#### 4. Results and discussion

Table 1 presents the descriptive statistics of the variables used as regressors in the analyses.

[INSERT TABLE 1 ABOUT HERE]

The values in Table 1 show that the gender composition of the subsamples is rather homogeneous between the two treatments (45% of females received the *effort treatment* with work while 49% received the *no effort treatment*). On average, the male participants were faster than female participants in performing the data entry task. Although we do not know whether this difference may affect the results, we control for individual performance in the econometric analysis. To test whether the results presented in the next tables might be driven by the uneven distribution of some

characteristics between the samples subjected to the different treatments, Table A2 in Appendix 1 presents an ancillary regression. The table shows that the characteristics relevant to the analyses do not feature any statistical differences between the two groups. Table 2 reports the percentage of potential offenders who actually committed the offense under the effort treatment and no effort treatment and depending on the litigation fee. The figures presented in the table are the percentages of subjects who were given the opportunity to take money from their counterparty.

[INSERT TABLE 2 ABOUT HERE]

About three-quarters of the players with the role of potential offender decided to take a portion of their opponent's resources. The values show that the two pairs of subsamples do not differ in their probability of offending. In particular, concerning the effort and no effort treatments (with and without work task, respectively), the lack of any statistically significant difference between the two subsamples is consistent with the fact that potential offenders make their decision without any knowledge of the origin of the victims' resources.<sup>11</sup>

According to the experimental design, the subjects who took a portion of their opponents' endowment are twice offered the restoration option. On the first occasion, the offenders are simply informed about the origin of the victims' resources. On the second occasion, the offenders can restore to avoid litigation.

Table 3 presents the frequency of restitutions (probability of settlement) in the first of these two rounds. We observe that 13.64% of the offenders decided to return the money upon learning that their opponents had to work to gain their resources. The share of returners under the alternative treatment is 6.52%, and the difference between the two treatments is (weakly) statistically significant. However, controlling for gender, the values reveal that the result is driven by the female subsample: in the effort treatment, women were five times more likely than men to restore the amount taken; under the two treatments, men were equally likely to return the amount taken. As no gender effect is detected under the no effort treatment, evidence seems to suggest that only women's choices are affected by the origin of the victims' resources.

Given the limited sample size, one may wonder whether it is enough to detect the effect of the treatments, if any. Hence, following Batistatou et al. (2014) and Juul and Frydenberg (2021), we calculate the statistical power for the restitution rate after knowing the origin of the victims' resources and for the decision to start litigation. In both cases, the analysis reveals that the sample size is

---

<sup>11</sup> Through a further ancillary regression, we verified that, for our sample, the endowment of the victim does not significantly affect the tortfeasor's decision to take money.

sufficient to detect the treatment effect at a 95% confidence level.<sup>12</sup> However, Gelman and Carlin (2014) warned about the possibility that even if the statistical power is enough to detect the treatment effect, errors of type S or M may affect the results. The results of the test for the unconditional treatment effect (Table 3) reveal that it may be overestimated by about 85%, thus resulting in a false positive outcome. Nevertheless, when the same test is applied to the effect of the treatment conditioned to other variables (e.g., columns 6 and 7 of Table 4), the value of the exaggeration parameter provided by Gelman and Carlin's test decreases to less than 1.001, suggesting that such conditional effects are not false positives and are correctly identified.

[INSERT TABLE 3 ABOUT HERE]

Table 4 reports the logit estimates, where the dependent variable is a dummy that takes a value of 1 if the offender settles (makes restoration) after becoming aware of the origin of the victim's resources. The values fully support hypothesis H1.1, that is, offenders who learn that their victims have had to work for their money return ill-gotten sums more often than those whose victims have received their endowment from the experimenter. This result confirms the main findings presented in Table 3 and 4 is robust to different specifications.<sup>13</sup>

If we believe in the experiment design and in the appropriateness of the effort and no effort treatments to elicit fairness concerns, as in Cappelen et al. (2007), we can cautiously conclude that at this stage, fairness concerns affect the choice of the offenders to restore. Moreover, by following Druckman and Wagner (2017) and Albin and Druckman (2014), one could warily frame the results in the wider debate about the positive relation between procedural justice and success of settlement agreements. As procedural justice concerns the fairness and transparency of the processes that involve decision makers, fairness concerns, as those elicited in the current study, seem to relate also to procedural justice issues. In fact, the knowledge of the origin of the victim's resources allows the offender to assess how the counterparty obtained such resources; being paid for a task is probably perceived as a more transparent and fair way to obtain money.

Controlling for the gender of the offenders, we find that gender has no statistically significant effect; this suggests that men and women have an equal probability of returning what was taken after knowing the origin of the victims' resources. In particular, atheists are more likely than religious people to return what was taken after knowing the origin of the resources of the other person.

---

<sup>12</sup> The analysis is performed using STATA 15.0, using the t-test and the log-ranking methodology for comparing the means of two groups with different variance.

<sup>13</sup> The inclusion of the interaction between gender and the origin of the endowment inflates the odds ratios, which is likely, as it introduces collinearity in the regression.

[INSERT TABLE 4 ABOUT HERE]

Table 5 reports the share of offenders who did not restore after learning the origin of the victims' resources but later decided to return the money to avoid in-court litigation.

[INSERT TABLE 5 ABOUT HERE]

A large number of subjects chose restoration to avoid litigation, and no differences are detected when comparing the effort vs. no effort treatments. This result suggests that at this stage, the offenders' decision to restore is driven by the desire to avoid possible losses resulting from a trial. We also note that the share of subjects who chose restoration to avoid litigation is larger than that of the subjects who chose restoration after being informed about the origin of the victims' resources. Tests on the differences between the shares at these two stages reveal that they are statistically different at a 99% significance level. Here, the effects of gender and religiosity are seldom statistically significant.

In other words, after eliminating the subjects who returned the money upon learning of its origin (for reasons of fairness according to our interpretation), we are left with those whose decision was driven by economic motivation, that is, to save on litigation costs. Indeed, the higher the fee, the higher the probability of the offender returning the amount taken to avoid a costly trial. Although such a behavior is common to men and women, the effect of the fee on the latter, in absolute terms, is greater than that on the former. This seems to suggest that female offenders are more sensitive to this kind of extrinsic motivation than their male subjects.

[INSERT TABLE 6 ABOUT HERE]

Table 6 reports the logit analysis on the probability of restoring to avoid litigation. The results show that the offenders subjected to the higher fee for the litigation were much more prone to restore the victims than those who were subjected to the lower fee. The origin of the endowment maintained its effect on the probability of restoring at this stage; that is, the offenders were more likely to restore the victims who earned the initial endowment in the effort treatment. The two treatment effects are independent of each other in the sense that their interaction does not result in a statistically significant effect. A possible interpretation of the persistence of the effect of the effort treatment at this stage (i.e., after the previous restoring round when the information was released) is that the information about the origin of their counterparts' endowment engendered a less instinctive reaction in some of the offenders than in the offenders who immediately responded to the same stimulus. The odds ratio relative to the time taken to decide seems to support such an interpretation; indeed, although no time

effect emerged in the first round of restitution, the subjects who finally decided to restore in the second round took a longer time to decide than those who did not restore.<sup>14</sup>

Another interesting result is the effect of the size of the victim's endowment: the odds ratios are all statistically significant and larger than zero, thus suggesting that tortfeasors expect that richer victims will be more likely to start litigation. Therefore, to maximize their payoff, offenders prefer to restore to avoid litigation.

Finally, the questionnaires allowed us to obtain further information, including whether the offender habitually volunteers. Volunteering is associated with a lower probability of restoring, especially if the victim had to work to obtain the initial endowment.<sup>15</sup> This control variable is a proxy for the level of pro-sociality of the experimental subjects. However, an analysis of the data reveals that 60% of subjects engaged in some form of volunteering decided to take part of the potential victim's endowment against 80% of the people not engaged in volunteering activities (p-value 0.022). These figures seem to suggest that pro-sociality (proxied by volunteering) affects only the decision whether to take part of the endowment; once this decision is taken, pro-social subjects are less responsive than the others to both the invitation to return what taken and the origin of the endowment. This outcome may reflect different intrinsic motivations between the group of volunteers and that of the others.

Table 7 reports the statistics concerning the decision of the victims to start litigation or not. Clearly, this option only applies to the subjects whose money was taken by the offenders and who were not restored in any of the two restitution rounds. A total of 49 subjects chose to litigate.

[INSERT TABLE 7 ABOUT HERE]

The values in Table 7 suggest that on average, looking at the full sample, the victims who had to work were less likely than victims who did not work to commence litigation. Although this result is weakly statistically significant, it seems to contradict our experimental hypothesis H1.2 and thus deserves additional analysis. Table 7 shows that women are, on average, more prone to litigate than men. For females, such a propensity decreases with the fee. Hence, focusing on women, the effort treatment has some effect on the decision to litigate. Victims who decide to litigate must pay the fee using (part

---

<sup>14</sup> Although one might argue that in this case the offenders have to process two pieces of information (one about the origin of the resources, the other about the probability of starting litigation, it should be stressed that, at this stage, all the offenders have to think about both, therefore, if those who restore need more time, this means that for them the decisional process is slower than for those who decide to keep the money taken. Nevertheless, it is true that we cannot understand whether this longer time is due to processing the information about resource origin, the probability that the victim starts litigating, or both.

<sup>15</sup> The same evidence emerges also in the first round of restitution, where all the offenders engaged in volunteering did not restore victims. As the estimation procedure is based on logit regression, such a result cannot be displayed in Table 4, as, when a variable perfectly predicts failure or success in a maximum likelihood estimation, it cannot be used as a regressor.

of) the residual endowment, therefore the effort spent in the task may render the payment of the fee psychologically more expensive for the victims who earned their endowment than for those who received it from the experimenters without working for it. Female participants seem to be more sensitive to this aspect than males (see Eckel and Grossman, 1996). In the light of Geniole et al. (2015) –who showed that women are more likely to punish than men when offenders damage their resources– women tend to feel more entitled than men to receive money as a compensation for their effort rather than as a gift.

[INSERT TABLE 8 ABOUT HERE]

Given the results shown in Table 7 and especially given that all the female subjects started litigating under the low-fee treatment, we control for the interaction between the fee level and the subjects' gender in the regressions in Table 8. The odds ratios after the logit estimations of the probability of starting litigation after settlement failure are shown; it is noteworthy that as the totality of women started litigating in the low fee treatment, some of the odds ratios presented in column 6 of Table 8 display very large or small figures. Indeed, when one of the individuated categories is that of women under the low fee treatment, any variation with respect to the complementary, empty category (women in the low-fee treatment who did not start litigating) is of extremely large magnitude. The values in Table 8 show that the effort treatment exerted a hardly robust negative effect on the probability of starting litigation; in addition, the associated odds ratios are statistically significant only in two out of the seven specifications presented in the table. Gender has the role expected from the results displayed in Table 7, with women much more willing to start litigating than men, especially in the low-fee treatment. In addition, the last specification of the regression (column 7) includes a dummy capturing whether the subjects habitually volunteer and its interaction with the effort treatment. The interaction term shows a very small and highly statistically significant odds ratio.

Before concluding this work, we need to comment about risk aversion. The choices made by the subjects may also depend on individual risk aversion, which was not elicited in the experiment. The reason for such a choice is that this study focuses on the elicitation of fairness concerns (effort vs. no effort treatments).<sup>16</sup> Obviously, in the present context, risk aversion was very important to the participants' choices and could even help to partially explain some of the evidence. On the one hand, risk-averse offenders may decide to take less and return more often to avoid litigation while risk-averse victims may rarely commence litigation. On the other hand, less risk-averse and risk-seeking

---

<sup>16</sup> Moreover, risk attitudes are a latent construct that can only be indirectly and imperfectly measured. Risk attitude measurement is, by construction, a combination of the latent preferences and the measurement error induced by their elicitation. The question of if, to what extent, and in which direction the observed results are driven by the adopted elicitation tool is interesting and, as yet, unexplored (Filippin and Crosetto, 2014).

subjects may take more, restore less, and litigate more frequently. Nevertheless, the main results of this study are not jeopardized. In fact, risk aversion is an intrinsic characteristic of people, including the participants of the experiment. As the participants were randomly selected, we may assume that the sample distribution of risk aversion in the present experiment mirrors the risk aversion of the population from which the sample was taken.<sup>17</sup> As the subjects were randomly assigned to the different treatments, we may assume that all the individual traits are identically and independently distributed between the two subsamples. Table A2 shows that this assumption holds for the observable variables. Under this assumption, risk aversion may certainly play a role, but on average, it is equal in the two groups of experimental subjects; therefore, the results based on the differences between treatments are unaffected by risk aversion. In other words, the presence of risk aversion in the experiment might have influenced the choices made but not the differences between them. Finally, we believe that although risk aversion may play a role, the general message of the study is not affected by individual risk aversion.

## 5. Conclusions

The experiment was aimed at understanding whether the origin of an endowment has any effect on 1) the probability that an offender, who previously acted so that to reduce the endowment of another subject, engages in restoration and 2) the probability that the person whose endowment was reduced initiates a litigation against the offender if the latter refuses to restore the initial endowment of the offended. In particular, the potential victims might either earn their endowment by performing a task or receive it effortlessly from the experimenters. A second ancillary treatment is concerned about the fee to be paid in litigation; herein, it was set at a low or high level to study how the costs of litigation may affect parties' decisions.

We observe that offenders tend to restore more frequently when the victims worked for their endowment. If the experiment design is appropriate to elicit fairness concerns, we can cautiously conclude that such concerns influence the probability of settlement. The gender analysis of the offenders' rate of restitution reveals that women restore more frequently to worker victims than to non-worker victims. This result may be attributed to the fact that women are particularly sensitive to

---

<sup>17</sup> Halek and Eisenhauer (2001) show that university graduates are generally as risk averse as the rest of the population, and that the effect of human capital on risk aversion is too small to be negligible. Other studies, which inquire into the link between education and risk aversion (for a comprehensive survey see Outreville, 2015) have always framed it in a context of financial choices and are therefore not comparable to the choices of the present experiment. However, if risk aversion differs in levels, but not in distribution, between university graduates and the rest of the population, experimental results based on a sample of the first are generalizable to the entire population.

(this kind of) fairness concerns. Conversely, men are likely to restore to worker and non-worker victims in equal measures.

In the study, the share of offenders who restored to avoid trial was larger than the those who restored once they were aware of the origin of the victims' endowment. That is, the threat of trial seemed to be stronger than any other considerations of fairness, except in the case of female offenders, who seemed to be equally concerned about fairness and the risk of losing the trial.

However, from the perspective of the victims, the origin of their resources did not significantly affect their decision to litigate, but a clear gender effect emerged: women are more likely to litigate than men perhaps because the offense takes the form of a harm to the initial resources held by the victims. In fact, as mentioned previously, women react harsher than men when their resources are harmed, as in the situation implemented in our experiment.

Concerning the observed effects related to the religiosity of the decision makers, atheism seems to favor restitution (after the revelation of the origin of victims' resources). Recall that our setting did not allow the participants to be clustered according to their religion and credibly, although most of the religious participants (Italian students) were Catholics; the effect seems to be aligned with the evidence provided by Benjamin et al. (2016) and Migheli (2017). These authors showed that the Catholic concept of fairness is strictly related to a peculiar way of interpreting distributive justice. In the present design, we can guess that once a Catholic offender decides to take money from a victim, the offender tends to not return the money for reasons related to distributive justice.

In relation to standard extrinsic motivations, we find some expected results. In particular, higher litigation fees induce offenders to more frequently make restoration to victims to avoid litigation. Concerning the probability of starting litigation, the results show that the higher the fee the lower the probability of litigating, in line with the expectations, although the difference between the treatments is not statistically significant and it is driven by the female subjects. Finally, women are less prone to litigate when they earned their endowment by working. Such a result is consistent with Eckel and Grossman (1996), who show that, as the cost of punishment increases, women are less willing to penalise offenders than men are. In fact, starting to litigate has a direct (the fee) and an indirect (the effort to earn the endowment) cost. Women seem to be more sensitive to the second than men are, so explaining their behaviour in the last part of the experiment.

One may wonder whether the results are affected by risk aversion to some extent. However, when the experiment is based on comparisons between two treatments and participants are randomly selected, risk aversion is not an issue. It is reasonable to also assume that risk aversion is randomly and

identically distributed between the subjects across treatments. Therefore, any effect of risk aversion is cancelled out when different treatments are compared.

Although the present design should be interpreted with caution in terms of its external validity, the results of the study show that although relevant (especially for female offenders), fairness concerns and individual factors only partially influence the probability to settle a dispute because few offenders tend to make restoration to their victims because of these concerns while most offenders act under the threat of litigation.

The power analysis reveals that the sample size is sufficient to detect at least the conditional effect of treatment; nevertheless, as for most experiments in economics, ours involved undergraduate students. As in any experiment involving this kind of subjects, concerns about the external validity of the results may be raised. Although the problem cannot be disregarded and one cannot assert that laboratory experiments do not potentially suffer from such a problem, a growing body of literature in economics shows that generally, the preferences revealed in lab experiments with undergraduate students mirror those of larger and more representative samples (e.g., Alm et al., 2015; Galizzi and Navarro-Martinez, 2018; Quaife et al., 2018). As Kessler and Vesterlund (2015) pointed out, the external validity of qualitative results is generally stronger than that of quantitative results even in empirical studies. In other words, although the effect size detected in a sample may differ from the true effect size in the whole population, the direction of the effect identified in an experiment is generally the same as that in the whole population.

A possible further limitation of the results of this study is the fact that we did not measure the perceptions of fairness in the two groups of subjects (the treated and the control). Although we showed that the compositions of the groups are not different according to the observed characteristics, differences in unobserved characteristics may somehow affect the results, as is usually the case in empirical analyses.

## References

- Abrams, K. (2011). “Emotions in the Mobilization of Rights,” *Harvard Civil Rights/Civil Liberties Law Review*, 46: 551–90.
- Ahmed, A.M. (2009) “Are religious people more prosocial? A quasi-experimental study with Madrasah pupils in a rural community in India” *Journal for the Scientific Study of Religion*, 48(2): 368–374.
- Alm, J., Bloomquist, K.M. and McKee, M. (2015). “On the External Validity of Laboratory Tax Compliance Experiments” *Economic Inquiry*, 53(2): 1170-1186.
- Antinyan, A., Corazzini, L. and Neururer, D. (2015). “Public Good Provision, Punishment and the Endowment Origin: Experimental Evidence” *Journal of Behavioral and Experimental Economics*, 56: 72-77.
- Babcock, L., Farber, H.S., Fobian, C. and Shafir, E. (1995). “Forming Beliefs About Adjudicated Outcomes: Perceptions of Risk and Reservation Values,” *International Review of Law and Economics* 15: 298–303.
- Babcock, L. and Loewenstein, G. (1997). “Explaining bargaining impasse: The role of self-serving biases” *Journal of Economic perspectives*, 11: 109—126.
- Babcock, L. and Pogarsky, G. (1999). “Damage caps and settlement: a behavioral approach” *The Journal of Legal Studies*, 28: 341-370.
- Batistatou, Evridiki, Chris Roberts and Steve Roberts (2014). “Sample Size and Power Calculation for Trial and quasi-Experimental Studies with Clustering” *The STATA Journal*, 14(1): 159-175.
- Benjamin, D.J., Choi, J.J. and Fisher, G. (2016). “Religious identity and economic behaviour” *Review of Economics and Statistics*, 98: 617—637.
- Bies, R.J..(1986). Interactional justice: Looking Backward, Looking Forward in Cropanzano, R.S. Ambrose M.L. (Eds.), 2015, *The Oxford Handbook of Justice in the Workplace*. Oxford Handbooks Online.
- Bosman, R., and van Winden, F. (2002). “Emotional Hazard in a Power-to-Take Experiment” *Economic Journal*, 112: 146-169.
- Bosman, R., Sutter, M., and van Winden, F. (2005). “The Impact of Real Effort and Emotions in the Power-to-Take Game” *Journal of Economic Psychology*, 36(3): 407 – 429.
- Boyd, C. and Hoffman, D. (2012). “Litigating Toward Settlement,” *Journal of Law, Economics, and Organization*, 29(4): 898-929
- Brahnam, S.D., Margavio T.M., Hignite M.A., Barrier T.B. and Chin J.M. (2005). “A Gender-Based Categorization for Conflict Resolution” *Journal of Management Development*, 24(3): 197-208.
- Brañas-Garza, P., Espín, A.M. and Neuman, S. (2014). “Religious pro-Sociality? Experimental Evidence from a Sample of 766 Spaniards” *PloS One*, 9(8): e104685.
- Brañas-Garza, P., Rossi, M. and Zaccicever, D. (2009). “Individual’s Religiosity Enhances Trust: Latin American Evidence for the Puzzle” *Journal of Money, Credit and Banking*, 41: 555-566.
- Burke, S., and Collins, K.M. (2001). “Gender Differences in Leadership Styles and Management Skills” *Women in Management Review*, 16: 244–57.

- Cabrales, A. and Ponti, G. (2015). “Social Preferences” in *Experimental Economics – Economic Decisions* (P. Brañas-Garza and A. Cabrales Eds.), London: Palgrave MacMillan.
- Camerer, C.F. (2003). *Behavioral Game Theory*. Princeton: Princeton University Press.
- Cappelen, A.W., Hole, A.D., Sørensen, E.Ø. and Tungodden B. (2007). The Pluralism of Fairness Ideals: An Experimental Approach. *The American Economic Review*, 97(3): 818-827.
- Carpenter, J., and Matthews, P.H. (2012). Norm Enforcement: Anger, Indignation, or Reciprocity? *Journal of the European Economic Association*, 10(3): 555 – 572.
- Croson, R. and Buchan, N. (1999). “Gender and Culture: International Experimental Evidence from Trust Games” *The American Economic Review: Papers and Proceedings*, 89(2): 386-391.
- Croson, R. and Gneezy, U. (2009). “Gender Differences in Preferences” *Journal of Economic Literature*, 47(2): 448-474.
- Daughety, A.F. and Reinganum, J.F. (2008) “Settlement” in *Encyclopedia of Law and Economics* (2<sup>nd</sup> Ed.) Ed. By C. W. Sanchirico, Edward Elgar.
- Eckel, C. and Grossmann, P. (1996). “The Relative Price of Fairness: Gender Differences in a Punishment Game” *Journal of Economic Behavior and Organization*, 30(2): 143-158.
- Falk, A. and Fischbacher, U. (2002). “Crime” in the lab - Detecting social interaction” *European Economic Review*, 46: 859-869.
- Falk, A., Fehr, E. and Fischbacher, U. (2005). “Driving Forces behind Informal Sanctions” *Econometrica*, 73(6): 2017 – 2030.
- Farmer, A., and Pecorino P. (2002). “Pretrial bargaining with self-serving bias and asymmetric information.” *Journal of Economic Behavior & Organization*, 48(1): 163—176.
- Feess, E., Schildberg-Hörisch, H., Schramm, M., and Wohlschlegel, A. (2015). “The impact of fine size and uncertainty on punishment and deterrence: Theory and evidence from the laboratory.” SFB/TR 15 Discussion Paper, No. 526.
- Fehr, E. and Gächter, S. (2000). “Cooperation and Punishment in Public Goods Experiments” *The American Economic Review*, 90(4): 980 – 994.
- Fischbacher, U. (2007). “z-Tree: Zurich toolbox for ready-made economic experiments.” *Experimental Economics*, 10 (2007): 171–178.
- Frazier, P., and Hunt, A.J.S. (1998). “Research on Gender and the Law: Where Are We Going, Where Have We Been?” *Law and Human Behavior*, 22: 1–16
- Filippin, A. and Crosetto, P. (2016). “A reconsideration of gender differences in risk attitudes.” *Management Science* 62: 3138—3160.
- Galizzi, M.M. and Navarro-Martinez, D. (2018). “On the External Validity of Social Preference Games: a Systematic Lab-Field Study” *Management Science*, 65(3): 976-1002.
- Garoupa, N. (2003). “Behavioral Economic Analysis of Crime: a Critical Review” *European Journal of Law and Economics*, 15(1): 5-15.

- Gelman, A. and Carlin, J. (2014). “Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors” *Perspectives on Psychological Science*, 9(6): 641-651.
- Geniole, S.N., Cunningham, C.E., Keyes, A.E., Busseri, M.A. and McCormick, C.M. (2015). Costly Retaliation is Promoted by Threats to Resources in Women and Threats to Status in Men” *Aggressive Behavior*, 41: 515-525.
- Gravert, C. (2013). “How Luck and Performance Affect Stealing” *Journal of Economic Behavior & Organization*, 93: 301-304.
- Greiner, B. (2015). “Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE” *Journal of the Economic Science Association*, 1(1): 114-125.
- Güth, W., Schmittberger, R. and Schwarze, B. (1982). An Experimental Analysis of Ultimatum Bargaining. *Journal of Economic Behavior & Organization*, 3(4): 367 – 388.
- Guthrie, C. (1999). Better settle than sorry: The regret aversion theory of litigation behaviour. *University of Illinois Law Review*, 1999: 43-90.
- Hadfield, G. K. (2008). “Framing the Choice Between Cash and the Courthouse: Experiences with the 9/11 Victim Compensation Fund,” *Law & Society Review*, 42: 645–82.
- Halek, M. and Eisenhauer, J.G. (2001). “Demography of Risk Aversion” *The Journal of Risk and Insurance*, 68(1): 1-24.
- Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., Ziker, J. (2010). Markets, Religion, Community Size, and the Evolution of Fairness and Punishment. *Science*, 327(5972), 1480-1484. <https://doi.org/10.1126/science.1182238>
- Hernández, P., and Pavan M. (2015). “Game Theory: Basic Concepts” in *Experimental Economics – Economic Decisions* (P. Brañas-Garza and a. Cabrales Eds.), London: Palgrave McMillan.
- Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J.C, Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F. Tracer, D. and Ziker, J. (2006). Costly Punishment across Human Societies. *Science*, 312: 1767-1770.
- Hoffmann, R. (2013). The experimental economics of religion. *Journal of Economic Surveys*, 27(5): 813—845.
- Hollander-Blumoff, R. (2011). “Intrinsic and Extrinsic Compliance Motivations: Comment on Feldman” *Washington University Journal of Law & Policy* 35: 53-67.
- Juul, S. and Frydenberg, M. (2021). *An Introduction to Stata for Health Researchers*. College Station: STATA Press.
- Kessler, J.B. and Vesterlund, L. (2015). “The External Validity of Laboratory Experiments: The Misleading Emphasis on Quantitative Effects” in (Fréchette, G.R. and Schotter A. Eds.) *Handbook of Experimental Economic Methodology*, Oxford: Oxford University Press.
- Kitzmann, K. M., and Emery, R. E. (1993). “Procedural justice and parents' satisfaction in a field study of child custody dispute resolution” *Law and Human Behavior*, 17: 553-567.

- Korobkin, R.B. and Guthrie, C. (1994). “Psychological Barriers to Litigation Settlement: An Experimental Approach” *Michigan Law Review*, 93: 107-192.
- Korobkin, R.B. and Ulen, T.S. (2000). “Law and behavioral science: Removing the rationality assumption from law and economics” *California Law review*, 88: 1051-1144.
- Landeo, C. (2007). “Split-awards and disputes: An experimental study of a strategic model of litigation.” *Journal of Economic Behavior & Organization* 63(2): 553-572
- Landeo, Claudia M. (2018) “Law and economics and tort litigation institutions: theory and experiments”, in Joshua C. Teitelbaum and Kathryn Zeiler (Eds.) *Research Handbook on Behavioral Law and Economics*, Edward Elgar, 247.
- Leibbrandt, A. and López\_Pérez, R. (2012). “An Exploration of Third- and Second-Party Punishment in Ten Simple Games” *Journal of Economics Behavior and Organization*, 84(3): 753 – 766.
- Lewisch, P. Ottone, S. and Ponzano, F. (2015) Third-party punishment under judicial review: An economic experiment on the effects of a two-tier punishment system. 11 *Review of Law & Economics* 209—230.
- Lind, E. A., Huo, Y. J. and Tyler, T R. (1994). “...And justice for all: Ethnicity, gender, and preferences for dispute resolution procedures” *Law and Human Behavior*, 18: 269-290.
- Lotito, G., Migheli, M. and Ortona, G. (2013). “Is Cooperation Instinctive? Evidence from the Response Times in a Public Goods Game” *Journal of Bioeconomics*, 15(2): 123 – 133.
- Loewenstein, G., Issacharoff, S., Camerer, C. and Babcock, L. (1993). “Self-serving assessments of fairness and pretrial bargaining” *The Journal of Legal Studies* 22: 135—159.
- Masclet, D., Noussair, C., Tucker, S. and Villeval, M. (2003). “Monetary and Nonmonetary Punishment in the Voluntary Contribution Mechanism” *The American Economic Review*, 93(1): 366 – 380.
- McKay, R. and Whitehouse, H. (2015) Religion and morality. 141 *Psychological Bulletin*, 447-473.
- Menkel-Meadow, C. (2012). “Women in Dispute Resolution: Parties, Layers and Dispute Resolvers” *Dispute Resolution Magazine*, 18: 4-11.
- Migheli, M. (2017). “The *Gospel* and Economic Behaviour: Experimental Evidence from a Trust Game” *Annals of Public and Cooperative Economics*, 88(1): 33-45.
- Mitchell, G. (2014). Alternative Behavioral Law and Economics. In Zamir, E. and Teichman, D. (Eds.) *The Oxford Handbook of Behavioral Economics and the Law*, Oxford University Press, New York, pp. 167-191.
- Muehlbacher, S. and Kirchler, E. (2009). “Origin of Endowments in Public Good Games: The Impact of Effort on Contributions” *Journal of Neuroscience, Psychology and Economics*, 2(1): 59-67.
- Outreville, J.F. (2015). “The Relationship between Relative Risk Aversion and the Level of Education: A Survey and Implications for the Demand for Life Insurance.” *Journal of Economic Surveys*, 29(1): 97-111.
- Perry, O., Erev, I. and Haruvy, E. (2002). “Frequent Probabilistic Punishment in Law Enforcement” *Economics of Governance*, 3: 71-86.
- Pogarsky, G. and Babcock, L. (2001). “Damage Caps, Motivated Anchoring, and Bargaining Impasse.” *The Journal of Legal Studies* 30: 143-159.

- Qin, X. and Wang, S. (2013). “Using and Exogenous Mechanism to Examine Efficient Probabilistic Punishment” *Journal of Economic Psychology*, 39: 1-10.
- Quaife, M., Terris-Prestholt, F., Di Tanna G.L. and Vickerman, P. (2018). “How well Do Discrete Choice Experiments Predict Health Choices? A Systematic Review and Meta-Analysis of External Validity” *European Journal of Health Economics*, 19: 1053-1066.
- Rawls, J. (1963) “The Sense of Justice” *The Philosophical Review*, 72(3): 281-305.
- Robbennolt, J.K. (2018). “Attorneys, Apologies, and Settlement Negotiation” *Harvard Negotiation Law Review*, 13: 349.
- Robbennolt, J.K. (2014). *Litigation and Settlement*. In Zamir, E. and Teichman, D. (Eds.). *The Oxford Handbook of Behavioral Economics and the Law*, Oxford University Press, New York, pp. 623–642.
- Schwartz, W.F. and Wickelgren, A.L. (2009). Advantage Defendant: Why Sinking Litigation Costs Makes Negative-Expected-Value Defenses but Not Negative-Expected-Value Suits Credible” *Journal of Legal Studies*, 38: 235-253.
- Sonnemans, J. and van Dijk, F. (2012). “Errors in Judicial Decisions: Experimental Results” *Journal of Law, Economics, and Organization*, 28(4): 687-716.
- Sullivan, S.P. (2016). “Why Wait to Settle? An Experimental Test of the Asymmetric-Information Hypothesis” *Journal of Law and Economics*, 59(3): 497-525.
- Thomas, S. (1994). *How Women Legislate*. Oxford University Press, New York.
- Tversky, A. and Kahneman, D. (1991). “Loss Aversion in Riskless Choice: A Reference-Dependent Model” *Quarterly Journal of Economics*, 106(4): 1039-1061.
- Wofford, C. B. (2017). “The Effect of Gender and Relational Distance on Plaintiff Decision Making in the Litigation Process” *Law & Society Review*, 51: 966-1000.

## Tables

**Table 1. Descriptive statistics.**

|         | Potential victims in the treatment with effort | Potential victims in the treatment without effort | Atheists | Work performance <sup>1</sup> | Average time to decide whether restoring to avoid trial <sup>2</sup> |
|---------|------------------------------------------------|---------------------------------------------------|----------|-------------------------------|----------------------------------------------------------------------|
| Males   | 36                                             | 22                                                | 37       | 45.08                         | 49.64                                                                |
| Females | 24                                             | 36                                                | 25       | 34.04                         | 50.82                                                                |
| Total   | 60                                             | 58                                                | 62       | 40.11                         | 50.19                                                                |

|         | Potential tortfeasors in the treatment with effort | Potential tortfeasors in the treatment without effort | Atheists | Average time to decide whether resroting (first time) | Average time to decide whether resroting (second time) |
|---------|----------------------------------------------------|-------------------------------------------------------|----------|-------------------------------------------------------|--------------------------------------------------------|
| Males   | 30                                                 | 37                                                    | 37       | 66.22                                                 | 71.07                                                  |
| Females | 30                                                 | 21                                                    | 26       | 63.29                                                 | 70.41                                                  |
| Total   | 60                                                 | 58                                                    | 63       | 64.96                                                 | 70.79                                                  |

1 Average number of quadruplets (name, surname, id number and mark) correctly copied

2 This is the time elapsed between when the screen proposing this possibility appeared and when the subject confirmed her decision, by pushing on "enter" on her keyboard. The time is measured in seconds

**Table 2. Percentage of tortfeasors**

|                     | Full sample | Significance | Males | Females | Sig. M | Sig. F. |
|---------------------|-------------|--------------|-------|---------|--------|---------|
|                     | 74.57       |              | 71.64 | 78.43   |        |         |
|                     |             |              | -     |         |        |         |
| Effort treatment    | 72.5        | -            | 68.18 | 77.78   |        |         |
|                     |             |              | -     |         | -      | -       |
| No effort treatment | 74.14       |              | 72.88 | 75.44   |        |         |
|                     |             |              | -     |         |        |         |
| High litigation fee | 70.83       | -            | 66.67 | 75.76   |        |         |
|                     |             |              | -     |         | -      | -       |
| Low litigation fee  | 80.43       |              | 78.57 | 83.33   |        |         |
|                     |             |              | -     |         |        |         |

-: non significant at any conventional level

The figures in the table represent percentage points of subjects who were given the possibility to take money from the counterpart.

Mann-Whitney two-tailed tests were used to assess statistical significance.

**Table 3. Percentage of tortfeasors who restore after knowing the origin of the counterpart's endowment.**

|                     | Full sample |   | Males | Females | Sig. M | Sig. F. |
|---------------------|-------------|---|-------|---------|--------|---------|
|                     |             |   | 6.00  | 15.00   |        |         |
|                     |             |   | *     |         |        |         |
| Effort treatment    | 13.64       | * | 4.76  | 21.74   |        |         |
|                     |             |   | **    |         | -      | *       |
| No effort treatment | 6.52        |   | 6.90  | 5.88    |        |         |
|                     |             |   | -     |         |        |         |

Significance levels: \*\* 95%, \* 90%, - non significant

Mann-Whitney two-tailed tests were used to assess statistical significance.

**Table 4. Decision to restore, after knowing the origin of the counterpart's endowment.**

|                                             | (1)                   | (2)                  | (3)                   | (4)                   | (5)                   | (6)                   | (7)                   |
|---------------------------------------------|-----------------------|----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| Effort Treatment                            | 2.263<br>(1.027)*     | 1.992<br>(0.956)*    | 4.444<br>(4.324)      | 2.538<br>(1.094)**    | 2.547<br>(1.140)**    | 10.29<br>(9.483)**    | 8.439<br>(6.998)**    |
| Male                                        |                       | 0.398<br>(0.257)     | 1.185<br>(1.518)      | 0.822<br>(0.532)      | 0.788<br>(0.525)      | 1.157<br>(1.484)      | 1.247<br>(1.590)      |
| Male x Effort Treatment                     |                       |                      | 0.152<br>(0.241)      |                       |                       | 0.128<br>(0.203)      | 0.127<br>(0.218)      |
| Atheist                                     |                       |                      |                       | 5.116<br>(4.172)**    | 4.884<br>(4.092)*     |                       |                       |
| High Fee                                    |                       |                      |                       |                       | 0.720<br>(0.333)      | 1.571<br>(1.450)      | 1.473<br>(1.297)      |
| High fee x Effort Treatment                 |                       |                      |                       |                       |                       | 0.259<br>(0.242)      | 0.305<br>(0.284)      |
| Time to take the decision whether returning |                       |                      |                       |                       |                       |                       | 0.984<br>(0.0159)     |
| Constant                                    | 0.0698<br>(0.0282)*** | 0.114<br>(0.0514)*** | 0.0625<br>(0.0602)*** | 0.0270<br>(0.0206)*** | 0.0331<br>(0.0283)*** | 0.0481<br>(0.0439)*** | 0.0548<br>(0.0451)*** |
| Observations                                | 90                    | 90                   | 90                    | 90                    | 90                    | 90                    | 90                    |
| Log-pseudolikelihood                        | -28.616               | -27.819              | -27.144               | -20.714               | -20.638               | -26.623               | -26.330               |
| Pseudo R-squared                            | 0.022                 | 0.049                | 0.072                 | 0.110                 | 0.113                 | 0.090                 | 0.100                 |

Odds ratios after logit estimation; standard errors, clustered by session, in brackets

Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Table 5. Percentage of tortfeasors who restore to avoid litigation**

|                     | Full sample | Males | Females | Sig. M | Sig. F. |
|---------------------|-------------|-------|---------|--------|---------|
| Effort treatment    | 34.21       | 40.00 | 27.78   |        |         |
| No effort treatment | 31.71       | 28.00 | 37.50   |        |         |
| High litigation fee | 35.96       | 35.56 | 36.36   |        |         |
| Low litigation fee  | 23.07       | 29.72 | 14.29   |        |         |

-: non significant at any conventional level

Mann-Whitney two-tailed tests were used to assess statistical significance.

**Table 6. Decision of restoring to avoid the litigation (and after knowing the origin of the counterpart's endowment)**

| VARIABLES                                   | (1)                  | (2)               | (3)                | (4)                   | (5)                 | (6)                  | (7)                           |
|---------------------------------------------|----------------------|-------------------|--------------------|-----------------------|---------------------|----------------------|-------------------------------|
| Counterpart worked                          | 1.052<br>(0.550)     | 1.003<br>(0.476)  | 0.763<br>(0.279)   | 15.016<br>(20.016)**  | 10.990<br>(14.126)* | 12.420<br>(14.646)** | 34.621<br>(72.629)*           |
| Male                                        |                      | 1.221<br>(0.777)  | 1.243<br>(0.758)   | 1.401<br>(0.809)      | 0.741<br>(0.704)    | 0.747<br>(0.702)     | 0.239<br>(0.321)              |
| Counterpart worked X Male                   |                      |                   |                    | 0.331<br>(0.422)      | 0.439<br>(0.587)    | 0.428<br>(0.578)     | 0.042<br>(0.094)              |
| High fee                                    | 1.999<br>(0.878)     | 2.215<br>(0.959)* | 7.342<br>(5.646)*  | 11.653<br>(10.214)*** | 9.342<br>(8.810)**  | 7.293<br>(7.170)**   | 5.702<br>(5.897)*             |
| High fee x Atheist                          |                      |                   | 0.0418<br>(0.049)  | 0.028<br>(0.040)**    | 0.028<br>(0.046)**  | 0.030<br>(0.049)**   |                               |
| High fee X Counterpart worked               |                      |                   |                    |                       |                     | 0.649<br>(0.555)     | 1.089<br>(2.252)              |
| Atheist                                     |                      | 0.895<br>(0.640)  | 4.605<br>(4.129)*  |                       | 3.689<br>(4.478)    | 3.795<br>(4.622)     |                               |
| Time to take the decision whether returning |                      |                   |                    |                       |                     |                      | 1.496<br>(0.152)***           |
| Counterpart's endowment                     |                      |                   |                    |                       | 1.192<br>(0.097)**  | 1.197<br>(0.106)**   | 1.508<br>(0.331)*             |
| Volunteering                                |                      |                   |                    | 0.156<br>(0.181)      | 0.120<br>(0.166)    | 0.138<br>(0.185)     | 0.057<br>(0.065)**            |
| Volunteering x Counterpart worked           |                      |                   |                    | 0.039<br>(0.054)**    | 0.029<br>(0.044)**  | 0.033<br>(0.049)**   | 0.192<br>(0.187)*             |
| Constant                                    | 0.313<br>(0.0974)*** | 0.294<br>(0.194)* | 0.159<br>(0.122)** | 0.267<br>(0.213)*     | 0.074<br>(0.079)**  | 0.075<br>(0.081)**   | 1.03*10-15<br>(9.02*10-15)*** |
| Observations                                | 75                   | 75                | 75                 | 71                    | 71                  | 71                   | 71                            |
| Log-pseudolikelihood                        | -49.060              | -41.214           | -37.690            | -35.013               | -33.022             | -32.974              | -17.640                       |
| Pseudo R-squared                            | 0.020                | 0.028             | 0.113              | 0.174                 | 0.221               | 0.222                | 0.612                         |

Odds ratios after logit estimation; standard errors, clustered by session, in brackets

Significance levels: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Table 7. Percentage of victims who decided to litigate.**

|                     | Full sample | Males        | Females | Sig. M | Sig. F. |
|---------------------|-------------|--------------|---------|--------|---------|
| High litigation fee | 61.76       | 60.00        | 63.13   | -      | ***     |
| Low litigation fee  | 70.83       | 36.36<br>*** | 100.00  |        |         |
| Effort treatment    | 58.06       | 52.94        | 64.29   | -      | **      |
| No effort treatment | 74.07       | 44.44<br>*** | 88.88   |        |         |

Significance levels: \*\*\* 99% \*\* 95%, \* 90%, - non significant

Mann-Whitney two-tailed tests were used to assess statistical significance.

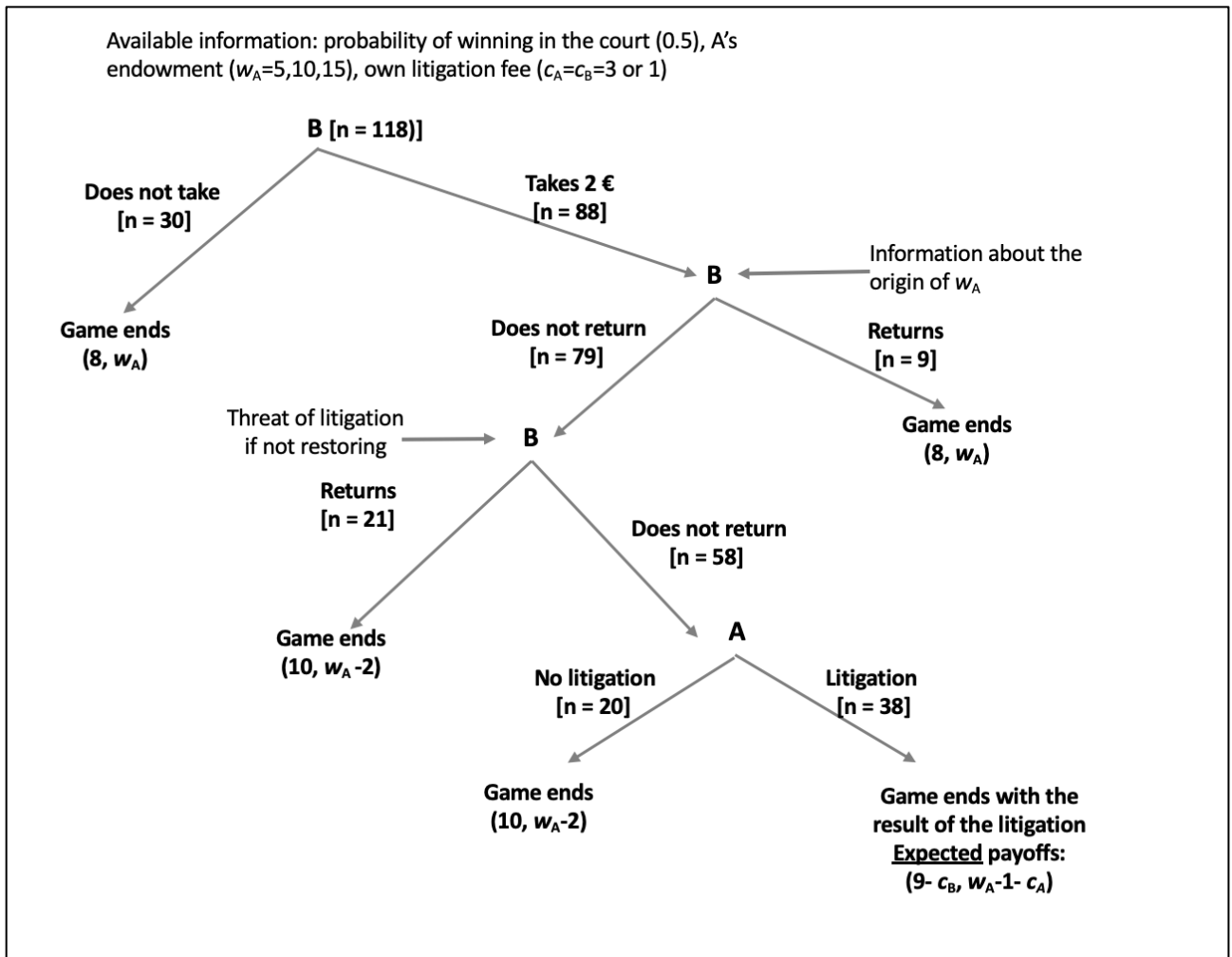
Table 8. Decision of initiating the trial.

|                                 | (1)                 | (2)                 | (3)               | (4)                 | (5)                 | (6)                         | (7)                       |
|---------------------------------|---------------------|---------------------|-------------------|---------------------|---------------------|-----------------------------|---------------------------|
| Effort treatment                | 0.485<br>(0.233)    | 0.600<br>(0.345)    | 0.581<br>(0.344)  | 0.225<br>(0.230)    | 0.241<br>(0.224)*   | 0.263<br>(0.165)**          | 0.486<br>(0.448)          |
| Male                            |                     | 0.307<br>(0.244)    | 0.294<br>(0.221)* | 0.100<br>(0.106)**  | 0.101<br>(0.105)**  | 7.10e-09<br>(7.56e-09)***   | 2.11e-09<br>(3.14e-09)*** |
| Effort treatment x Male         |                     |                     |                   | 6.250<br>(8.278)    | 5.923<br>(7.839)    | 4.641<br>(2.791)**          | 5.269<br>(4.827)*         |
| High fee (dummy)                |                     |                     |                   |                     | 0.700<br>(0.250)    | 5.87e-08<br>(4.87e-08)***   | 4.33e-08<br>(4.44e-08)*** |
| High fee (dummy) x Male         |                     |                     |                   |                     |                     | 4.951e+07<br>(4.600e+07)*** | 4.97e+08<br>(7.42e+08)*** |
| Endowment                       |                     |                     | 1.090<br>(0.061)  |                     |                     | 1.086<br>(0.0664)           | 1.058<br>(0.136)          |
| Volunteering                    |                     |                     |                   |                     |                     |                             | 5.250<br>(7.548)          |
| Volunteering x Effort treatment |                     |                     |                   |                     |                     |                             | 0.005<br>(0.010)***       |
| Constant                        | 2.857<br>(0.984)*** | 4.546<br>(2.089)*** | 1.969<br>(1.398)  | 8.000<br>(5.528)*** | 9.680<br>(5.481)*** | 2.832e+07<br>(2.415e+07)*** | 5.73e+07<br>(9.81e+07)*** |
| Observations                    | 58                  | 58                  | 58                | 58                  | 58                  | 58                          | 51                        |
| Log-pseudolikelihood            | -36.534             | -34.466             | -33.717           | -33.340             | -33.172             | -28.335                     | -18.975                   |
| Pseudo R-squared                | 0.022               | 0.078               | 0.098             | 0.108               | 0.112               | 0.242                       | 0.415                     |

Odds ratios after logit estimation; standard errors, clustered by session, in brackets.

Significance levels: \*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

Figure 1



## Appendix 1: Summary of the treatments and balancing check.

| Table A1 - Treatments      |                                                                                            |                                                                                           |
|----------------------------|--------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------|
|                            | <i>High fee treatment</i>                                                                  | <i>Low fee treatment</i>                                                                  |
| <i>Effort treatment</i>    | Players A work to gain their endowment.<br>Litigation fee for players B is 3 euros.        | Players A work to gain their endowment.<br>Litigation fee for players B is 1 euro.        |
| <i>No effort treatment</i> | Players A do not work to gain their endowment.<br>Litigation fee for players B is 3 euros. | Players A do not work to gain their endowment.<br>Litigation fee for players B is 1 euro. |

**Table A2. Probit estimates for stealing the 2 euro**

| VARIABLES                 | Coefficients      | Marginal effects  |
|---------------------------|-------------------|-------------------|
| Male                      | 0.525<br>(0.955)  | 0.148<br>(0.251)  |
| Work (1 = yes)            | -0.366<br>(0.516) | -0.065<br>(0.120) |
| Male x Work               | 0.154<br>(0.636)  | 0.031<br>(0.136)  |
| Atheist                   | 0.325<br>(0.644)  | 0.059<br>(0.134)  |
| Atheist x Male            | -0.508<br>(0.754) | -0.143<br>(0.294) |
| Victim's endowment        | 0.128<br>(0.093)  | 0.028<br>(0.022)  |
| Male x Victim's endowment | -0.110<br>(0.104) | -0.024<br>(0.020) |
| Constant                  | 0.196<br>(0.758)  |                   |
| Pseudo R-squared          | 0.1045            |                   |
| Observations              | 98                | 98                |

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Standard errors clustered by session

## Supplementary Appendix: Experimental instructions

Available upon request