WORKING PAPERS

Sanctions that Signal: an Experiment.

Roberto Galbiati
Karl Schlag
Joël van der Weele

March 2011

Working Paper No: 1107

DEPARTMENT OF ECONOMICS

UNIVERSITY OF VIENNA

Sanctions that Signal: an Experiment.*

Roberto Galbiati† Karl Schlag‡ Joël van der Weele§

March 19, 2011

Abstract

Sanctions are a means to provide incentives towards more pro-social behavior. Yet their implementation can be a signal that past behavior was undesirable. We investigate experimentally the importance of the informational content of the choice to sanction. We place this in a context of a coordination game to focus attention on beliefs and information and less on intrinsic or pro-social motivations. We compare the effect of sanctions that are introduced exogenously by the experimenter to that of sanctions which have been actively chosen by a subject who takes the role of a fictitious policy maker with superior information about the previous effort of the other players. We find that cooperative subjects perceive actively chosen sanctions as a negative signal which eliminates for them the incentive effect of sanctions.

Keywords: Sanctions, beliefs, expressive law, deterrence, coordination, minimum effort game.

JEL-codes: C92, D83, K42

*For valuable comments and discussions we wish to thank numerous seminar and conference participants at Granada (IMEBE), Erasmus University Rotterdam, Heidelberg University, Sciences Po, University of Bologna, Paris School of Economics, Max Planck Institute in Bonn, Lyon (ESA), the European University Institute, Siena (Labsi), as well as James Andreoni, Uri Gneezy, Mark Le Quement, Rosemarie Nagel, Emeric Henry, Steffen Huck and Rick van der Ploeg. Karl Schlag greatfully acknowledges financial support from the Department of Economics and Business of the Universitat Pompeu Fabra, from the Universitat Pompeu Fabra, Grant AL 12207, and from the Spanish Ministerio de Educacion y Ciencia, Grant MEC-SEJ2006-09993. Joël van der Weele thanks the Nuffic grant authorities.

†CNRS-EconomiX, Paris and Department of Economics Sciences Po Paris. E-mail: galbiatir@gmail.com

‡University of Vienna. E-mail: karl.schlag@univie.ac.at

§Goethe University, Frankfurt. E-mail: vanderweele@econ.uni-frankfurt.de

1 Introduction

Many economic and social interactions involve situations with multiple equilibria, some more efficient than others. An important task of policy makers is to induce cooperative behavior associated with more efficient equilibria. Typically, economists focus their policy analysis on the use of penalties or rewards that alter agents' behavior by changing the material trade-offs that agents are facing - an 'incentive effect'. However, material trade-offs are not the only determinant of behavior. Equilibrium play also critically depends on the beliefs and expectations that agents in a society have about the behavior of others. Policies that change these beliefs may induce shifts in behavior - a 'belief effect'.

Such belief effects may counteract the incentive effect. In particular, the discretionary introduction of a policy, rule or system of incentives may lead people to infer something about the reasons for the introduction, and may be a signal that other people are not behaving well. For example, to increase punishment for some crime can inform the public that this crime is prevalent. To financially intervene in a country can inform the rest of the world that their distress is worse than believed. To increase monitoring on immigrant groups may have a stigmatizing effect and lead people to believe that these groups have bad intentions. Such belief effects may reduce the motivations of people to behave cooperatively, and counteract the incentive effect of the policy.

It is this tension between incentive and belief effects that we study in this paper. Our experimental setup is a two person minimum effort game: a coordination game with many Pareto ranked equilibria, based on Goeree and Holt (2001, 2005). Each player chooses a level of costly effort, and is rewarded according to the minimum of the efforts of all players in the group. The more efficient equilibria result only if all players play individually risky strategies. Doubt about the other player's willingness to play such a strategy may result in inefficient outcomes. Thus, the game is particularly suitable as a workhorse, because there are multiple equilibria and players' efforts are strategic complements.

In all experimental treatments agents were matched in groups of three, where the third player was a 'principal' who benefitted proportionally to the minimum effort chosen by the other two in the group. The subjects played the minimum effort game twice, but the principal was the only one to be informed of the outcome of the first round before the second round was played. This information structure was common knowledge. Apart from effort choices, we also elicited the subjects' beliefs about the effort of the other player.

In the context of this game, we consider the following research questions about the effects of sanctions:

1. Can the introduction of incentives associated with small, non-deterrent[1] sanctions induce desired behavior and make agents more *optimistic* about other players' actions?

2. In situations of imperfect information about the past behavior of other group members, can the introduction of sanctions make agents more *pessimistic* about the actions of others by implicitly signaling that other players have not been cooperating? If so, does this reduce the effectiveness of sanctions?

Question 1 addresses the direct incentive effect of changing material trade-offs, but also a *forward looking* belief effect. Since efforts are compliments, anticipation of the other agent's reaction to the incentive effect will increase the own motivation to exert effort.

To answer Question 1, we compare a control treatment without sanctions with a sanction treatment. In a control treatment no sanctions were introduced between rounds, and consequently the second round was the same as the first. In the treatment, a mild sanction F was introduced for both players in the group, that lowered the earnings of a subject if she selected low effort (and also carried a small fixed cost for the principal). Because the sanction was introduced by the experimenter unconditional on past effort choices by the subjects, we call this the 'exogenous sanction' treatment.

Our hypotheses for the effect of such sanctions are based on a simple formal model, explained in Section 4. Because there are many Nash equilibria in this game, standard analysis does not help us very much. Instead, we use a model similar to level-k reasoning (Nagel 1995, Costa-Gomez and Crawford 2006), where we assume that people think of their partner as either a pessimistic type, who believes the partner will play low effort, or an optimistic type, who believes their partner will play high effort. This model predicts that sanctions should increase effort both through an incentive and through a belief effect. The first result of the paper is that we do indeed find both effects in the data, and so our answer to Question 1 is affirmative.

[1]By 'small', 'non-deterrent' or 'mild', we mean that sanctions do not make playing the highest possible effort a dominant strategy.

Question 2 addresses *backward looking* belief effect, or signaling effect, of sanctions. To answer it, we introduced an additional treatment. Before the second round of the minimum effort game was played, the principal could decide whether or not to introduce the same sanction F as above, at a small cost to his own earnings. Because the principal had observed first round behavior and could condition the sanction on this behavior, we call this the 'endogenous sanction' treatment.

Our model predicts that in this treatment there exists an equilibrium in which the principal will sanction if and only if there is at least one player who plays relatively low effort. Therefore, a player who played relatively high effort in the first round, but nevertheless observes a sanction, will learn that his partner is a pessimistic type who is likely to continue to play low effort: sanctions are 'bad news'. This means that for these players the incentive effect of the sanction will be counteracted by the backwards looking belief effect. By contrast, people who initially played low effort will not get any information from a sanction, because in equilibrium it would have been introduced independently of their partner's behavior. In this way, our design is able to identify a signaling effect of sanctions.

Our second result is that the data are consistent with a backward looking belief effect for those who played relatively high effort in the first round. Moreover, this effect is strong enough that for these players it eliminates the incentive effect of the endogenous sanction. In accordance with the theory, for those who played low first round effort the effect of an endogenous sanction is the same as that of the exogenous sanction.

To our knowledge this is the first paper that looks empirically at the tension between the signaling effect and incentive effect of sanctions. The main message is that the effectiveness of sanctions depends on the context in which they are introduced. On the one hand, people recognize the incentive effects that sanctions will have on others, which increases their effectiveness. On the other hand, when information about the behavior of others is limited, as is the case in modern large-scale societies or firms, the introduction of sanctions may cause pessimism by drawing attention to past misbehaviors. This is especially true for those that are optimistic and behave cooperatively. This finding implies a difficult balancing act that a government or principal must perform: It must try too keep the optimists optimistic, while at the same time encouraging the pessimists to change their behavior.

A final contribution of this paper is the use of nonparametric tests developed by Schlag

(2008) that are able, unlike the Wilcoxon-Mann-Whitney test (WMW test, Wilcoxon 1945, Mann and Whitney 1947), to identify how distributions differ. No distributional assumptions are made and levels of significance are mathematically correct for the given sample sizes. These tests allow to identify statistically significant evidence that outcomes in one sample tend to be higher than they are in the other.

The rest of the paper is organized as follows. In the next section we relate our paper to an existing theoretical literature on the signaling effect of sanctions in economics, and the literature on 'expressive law'. In Section 3 we outline the details of the experiment. In Section 4 we formulate explicit testable hypotheses, based on a formal model explained in more detail in Appendix A. Section 5 presents the results, followed by the conclusion in Section 6.

2 Literature

Our experimental study relates to several strands of the literature. First, it is closely related to a theoretical literature in economics on the potential signaling that occurs when imposing a sanction, or more generally when introducing some policy. Bénabou and Tirole (2003) show how the choice of incentives can provide information to an agent about his or her type. Friebel and Schnedler (2007), Sliwka (2007) and Van der Weele (2011) investigate how the choice of incentive policies can signal information of the policy maker about the relative prevalence of different types in the population to the imperfectly informed agents. In each of these papers, the signaling effect of sanctions depends on the existence of agents with different preferences. In equilibrium, sanctions are a signal that there are many selfish types around, which reduces the motivation of the agents to exert effort, either because of conformistic preferences or because of complementaries in technology. The common finding in this literature is that the signaling effect of sanction leads the principal to use sanctions less often relative to situation where the agents are perfectly informed. We complement this theoretical literature by showing that the signaling effect can also obtain in a setting in which all agents have identical preferences, and only differ according to their beliefs.

Our paper is also related to field experiments on crowding out effect of sanctions (see Frey and Jegen (2001) and Bowles (2008) for a survey). A well-known instance of this literature is Frey and Oberholzer-Gee (1997), who empirically show that a monetary incentive lowers acceptance rates of nuclear repository wastes. Another much cited paper is Gneezy and Rustichini (2000), who showed that a fine for picking up children late from a day-care center actually increased late-coming. Although potentially the mechanism we are looking for may be at work, these papers cannot differentiate it from potential alternative explanations, such as a direct impact of incentives on preferences, or the idea that a fine is a signal about some relevant characteristic of the principal. By contrast, our explicit distinction between exogenous and endogenous sanctions and the use of a between subject design allows us to identify that sanctions carry signals about the past behavior of other agents.

Thirdly, we contribute to the experimental literature on the effect of incentives in coordination games (see Devetag and Ortman (2007) for a survey of experimental results in coordination games). Goeree and Holt (2001, 2005) find in a between subject design that effort levels in a minimum effort game are higher when effort costs are lower. By contrast, our Question 1 refers to the effect of the introduction of sanctions, hence refers to a within subject design. Given our findings, one may conjecture that an exogenous lowering of effort costs will increase effort.

Fourth, our work relates to the literature in legal scholarship on the *focal point theory of law* (McAdams 2000 and McAdams and Nadler, 2003). According to this theory, laws can be used to coordinate expectations on a beneficial equilibrium, a phenomenon we referred to above as the forward looking belief effect of sanctions. In an experiment by Bohnet and Cooter (2001), penalties for choosing the inefficient strategy in a coordination game induce more people to choose the efficient strategy. However, to isolate the forward looking belief effect the payoffs were changed and a lump sum transfer was made to compensate the expected cost of the penalty. In this paper we look at a more conventional penalty that is not offset by a lump-sum transfer. We confirm the existence of a forward looking belief effect, and establish the existence of an incentive effect.

A final strand of literature in law that is of interest to our paper is that on 'expressive law' (e.g. Sunstein 1996, Cooter 1998). The idea here is that laws express the reigning norms in a society, and can discipline people by showing them what the majority of people deem to be 'appropriate'. Tyran and Feld (2006) test this reasoning in a public good game experiment that compares the effects of exogenously introduced 'mild' sanctions with the same sanctions introduced endogenously through group voting. The authors show that endogenous sanctions are more effective in raising contributions, because the preceding vote signals that there are many people who want to cooperate. This setup is quite different from ours where sanctions are implemented by an outsider who is more informed about previous behavior. In this context the introduction of sanctions sends a *negative* rather than a positive signal about the intentions of others.

To our knowledge, Xiao (2010) is the only other paper that studies signaling effect of sanctions by superiorly informed third parties. In a sender-receiver game, an outside 'enforcer' can punish a sender who sends false messages to the receiver. The payoffs of the enforcer does not depend on whether the sender deceives, and so it is not so surprising that one of the results of the experiment is that sanctions become a signal of the sender's deception. By contrast, in our study, the interests of the players and the enforcer are aligned and thus sanctions have a dual role of both signaling information and enhancing coordination. It is the tension between these two roles that is the focus of this paper.

At this point we would like to mention that all experimental papers above build their claims by verifying significant differences in population means using the WMW test, which cannot uncover such evidence without making additional assumptions.

3 Experimental Setup

The study of sanctions comes up in settings that can often be described as either a coordination game or a Prisoners' dilemma. We chose a coordination game as an object of study, because in such games the rational choice depends only on the beliefs about the actions of the other player(s) in the game. This allows us to isolate the sanctions' effects on behavior that derive from the change in a subject's belief, and we can disregard issues to do with social preferences and/or dominant strategies that usually play a role in Prisoners' dilemmas.

3.1 The Experimental Game

As the coordination game underlying our experiment we choose the minimum effort game of Goeree and Holt (2001, 2005). Large action spaces help capture variation in players' beliefs. In this game, two players simultaneously choose an action, to be interpreted as an effort level, between 110 and 170 (the bounds are chosen such that there are no clear focal points). Subjects' payoffs are equal the minimum of these two efforts, minus the amount of their own effort times a cost parameter $k \in [0, 1]$, which is the same for both players.

While in the original setting by Goeree and Holt (2001) the game is played only once, in our experiment the game is played twice where treatments differ according to what happens in the second round. In some treatments a value F is subtracted from the payoffs in the second round, where $F = 0.5 \cdot (170 - e_i)$. The subtraction of F can be interpreted as a sanction, deviations from the maximal effort (170) are punished proportionally. In fact, the sanction is mild as the game remains a coordination game with the same set of pure strategy Nash equilibria.

Another difference to Goeree and Holt (2001) is that we include a third player so that the game becomes a three player game. Depending on the treatment, the third player is either active or inactive. When active, the third player can choose before the start of the second round whether or not to introduce a sanction for both players in the group. When inactive, the third player does not make any choice, instead the choice of whether to introduce a sanction is made by the experimenter. Regardless of her activity status, player 3 receives a payoff proportional to the minimum effort chosen by the other two players. Note that player 3 is present in each treatment to maintain the same context.

In sum, payoffs in round 1 are determined as follows:

$$\pi_i (e_i, e_{-i}) = \min \{e_i, e_{-i}\} - 0.85 \cdot e_i \quad \text{for } i = 1, 2,$$
$$\pi_3 (e_1, e_2) = 0.25 \cdot \min \{e_1, e_2\},$$

where $\pi_i (e_1, e_2)$ is the payoff of player i, $e_i \in [110, 170]$ is the effort level chosen by player i, $i = 1, 2$, and $k = 0.85$ is the cost of effort. Payoffs in round 2 are given by the following equations:

$$\pi_i (e_i, e_{-i}, s) = \min \{e_i, e_{-i}\} - 0.85 \cdot e_i - s \cdot 0.5 \cdot (170 - e_i) \quad \text{for } i = 1, 2,$$
$$\pi_3 (e_1, e_2, s) = 0.25 \cdot \min \{e_1, e_2\} - 4s,$$

where 4 is the cost of introducing a sanction for the third player and $s \in \{0, 1\}$ reflects whether a sanction was introduced ($s = 1$) or not ($s = 0$).

An important element of the experimental design is the information structure. Players do not know before the first round that there will be a second round. They are informed of this only after the first round has concluded. Furthermore, players 1 and 2 do not observe each other's effort levels, nor do they learn their own payoffs before both rounds are over. When active, player 3 is informed about the effort levels of players 1 and 2 in round 1 while these two players only observe before round 2 starts whether or not player 3 has chosen to introduce a sanction.

3.1.1 Parameters, Treatments, and Procedures

We wanted effort choices to be low in round one in order to give player 3 an incentive to introduce a sanction when she is active. Following Goeree and Holt (2001), effort levels tend to be low when the cost of effort is high, hence we decided to set the cost of effort k equal to 0.85. Sanctioning should be moderately costly for player 3 so that there is sufficient diversity in the choices of player 3. Accordingly we chose to set the cost for the third player of introducing a sanction equal to 4, which is comparable to reduction of $4/0.25 = 16$ in the minimal effort of players 1 and 2.

We now describe the treatments. Treatments only differ in the second round, in the first round they are all the same: players 1 and 2 play the minimum effort game and player 3 is inactive. In the control treatment there is no sanction in the second round, and player 3 is inactive. That is, the second round is conducted exactly as the first. In particular subjects are not aware of the fact that sanctions are introduced in other treatments. We refer to this treatment as the exogenous no-sanction (ExNS) treatment. In the treatment we refer to as the exogenous sanction (ExS) treatment the sanction is introduced in the second round. Players one and two are told that the term F is additionally subtracted from their payoffs, player three remains inactive as in ExNS. We use the term *exogenous* to indicate that the choice to (not) introduce a sanction is not conditional in any way on previous decisions by the subjects. This was clear to the subjects because all subjects that belonged to the same session received the same instructions and this was common knowledge as instructions were read publicly. In the third treatment, player 3 who is present but inactive in the other two treatments, receives an active role. After round 1, player 3 observes the effort levels chosen by players 1 and 2 in the first round, whereupon she is asked to decide whether to a) introduce the sanction F to the payoffs of players 1 and 2 at a cost 4 to her own payoffs, or b) to leave the payoff structure unaltered. After player 3 has taken her decision, players 1 and 2 are informed of it and choose their effort levels. We refer to this treatment as the endogenous treatment. We slightly abuse common terminology and refer by the endogenous sanction treatment (EnS) to the case where player 3 introduced the sanction while we speak of the endogenous no-sanction treatment (EnNS) if no sanction was introduced.

Because the experiment features just two rounds of play, it was very important that people understood the game correctly from the start. For this purpose we ran a tutorial before the first round. In the tutorial, participants had 5 minutes to choose hypothetical effort choices of players 1 and 2 and to calculate their payoffs resulting from these choices. The tutorial took place before assigning subjects to a role, so that also players 3 would have practice in the calculation of payoffs of players 1 and 2. In addition to this tutorial, the input screens in the actual experiment provided subjects with the possibility to calculate their payoffs from a given choice. That is, after entering and before confirming their choices, subjects could enter a hypothetical choice of the other player and let the computer calculate their payoffs resulting from these choices.

The experiment was conducted in several sessions at the economics lab of the university of Siena, Italy between May and November 2007. Within each session subjects were matched into groups of 3, faced the same treatment and played the two round game described above a single time. Before playing the game the instructions were read out loud and the tutorial was conducted. The instructions and the input screen are provided in appendix C. Subjects entered their effort and belief choices on a computer that was running on the software z-Tree (Fischbacher 2007). The subjects received a show up fee of 1 euro, their earnings were in tokens as specified above, which were converted into Euro's at the end of the experiment at an exchange rate of 10 tokens = 0.75 Euro.

3.1.2 Elicitation of a Belief Interval

Apart from the effort choices, we are interested in the effect of sanctions on players' beliefs. Therefore, in each round we ask each player who chooses an effort about her beliefs about the effort of the other player in that round. Rather than elicit a point belief, we decided to elicit an *interval*. More precisely, players 1 and 2 have to specify a range (i.e. a lower bound L and an upper bound U) in which the effort chosen by the other player is believed to fall. Elicitation is remunerated as follows:

$$\pi_i(L,U,e_{-i}) = \begin{cases} 0 & \text{if } e_{-i} \notin [L,U] \\ 0.15 \cdot (60 - (U - L)) & \text{if } e_{-i} \in [L,U] \end{cases}$$

where $\pi_i(L,U,e_{-i})$ is the payoff of player i who specifies a range $[L,U]$ when e_{-i} is the effort chosen by the player matched with player i. Note that the payoff is zero when the effort of the other player falls outside the specified range. Accuracy is rewarded as the payoff of a correct guess, the payoff received when the effort of the other lies within the range, is decreasing in the width of the interval. Thus, a smaller range increases the payoff of a correct guess but also increases the risk of not being correct and hence of not earning additional tokens. As such, this elicitation method is both simple and can capture dispersion in beliefs as measured by $U - L$.

We opted against eliciting point beliefs as concern in the minimum effort game is less for mean or median effort than for the lower tail of the belief distribution. In fact, Schlag and van der Weele (2009) have investigated the predictive power of this rule. They find that any risk neutral or risk averse rational decision maker who is incentivized by the above rule and who has single peaked beliefs about the effort of the other player will make a correct guess at least 50% of the time. In other words, $[L, U]$ will contain at least 50% of the mass of her belief distribution.

In our analysis we focus on the lower bound L of the belief interval as the lower tail of the belief distribution is what matters in a minimum effort game. In the following, the term belief refers to the value of L.

4 Hypotheses

The hypotheses we present in the this section are based on a formal model. We believe the intuitions from this model are relatively straightforward, and do not want to clutter the analysis at this point, so we relegate a formal treatment to Appendix A. Note that the model makes some simplifying assumptions regarding risk neutrality, and the distribution of levels of rationality to keep the complexity of analysis within the scope of this paper.

Because there are many Nash equilibria in the experimental game, standard analysis does not bring us very far. Instead, we rely on a model that resembles k-level thinking models (Nagel 1995, Costa Gomes and Crawford 2006). Specifically, we assume that each player assumes that the other player (also referred to as partner) is best responding to one of two distinct belief distributions that are fixed over time. Accordingly, players believe that their partner chooses one of two effort levels which we refer to as 'high' effort and 'low' effort.

Consider first choices in round one. Following the fact that subjects did not know that there would be a second round, we ignore strategic considerations with respect to round two. For simplicity, players are assumed to be risk neutral. Therefore, in round one, a player will choose between the same two effort levels that the partner is believed to be choosing. In particular, the high effort is chosen if and only if the probability that the partner is choosing this effort is sufficiently high.

Consider now round two. As choices will depend on the treatment, we first look at exogenous sanctions. The sanction reduces the marginal cost of effort which results in both an incentive and a belief effect. The incentive effect describes the change in behavior that results if beliefs would remain unchanged. Lower marginal effort costs means that some players have an incentive to switch from low to high effort while none will switch from high to low effort. The belief effect refers to change in behavior driven by players changing their beliefs about their partner's behavior. As we explained above, players believe that their partner maintains the same beliefs as he or she had in round one. Given the reduction in marginal cost of effort, players anticipate that their potential partner will exert higher effort (through the incentive effect) which additionally increases their own incentive to increase their effort. Because each player assumes that the partner best-responds to fixed belief distributions, no further iterations in strategic reasoning are necessary.

It follows that an exogenous sanction will increase both own effort and the beliefs about the effort of the other player. Under a mild assumption on the consistency of beliefs in round one (as specified in the appendix), the average effort of the players increases more than their average beliefs do. The reason for this is that effort is increased through both the incentive and the belief effect, whereas beliefs are only affected by the latter. This leads to our first hypothesis.

Hypothesis 1 *Exogenous sanctions increase beliefs and efforts. This effect is more pronounced for effort levels than it is for belief levels.*

Consider now the treatment with endogenous sanctions. We wish to determine when the third player, or principal, will choose to sanction and how players one and two will react to the non-introduction of sanctions. We focus on Perfect Bayesian Equilibria that do not depend on the prior beliefs of the principal about the beliefs of the agents. We find three candidate equilibrium behaviors for the third player: "always sanction", "never sanction" and "sanction when at least one player exerts low effort". For instance, unconditional sanctioning is best if lower effort costs raise partner's effort sufficiently to offset the cost of sanctioning for the third player. For this to happen, the third player has to anticipate an increase in effort by at least $4/0.25 = 16$ points as a result of the sanction. The equilibrium involving "sanction when at least one player exerts low effort" exists when (i) a player with low effort is sufficiently responsive to a sanction and (ii) the difference in effort between low and high effort players is sufficiently large. The intuition behind these equilibrium conditions will become clearer below.

Common to these three equilibria is that low efforts are always sanctioned if sanctions are chosen. This leads to our next hypothesis.

Hypothesis 2 *In the endogenous treatment, the likelihood of sanctions being imposed by the principal is decreasing in the minimal effort chosen in the first round.*

We now turn to the reaction of players one and two to the sanctioning choice, which naturally depends on the principal's equilibrium policy. In the "always sanction" or "never sanction" equilibrium, no information about partner behavior is transmitted to the players. In this case, endogenous sanctions have the same effect as exogenous sanctions. Information about partner behavior is transmitted only under "sanction when at least one player exerts low effort", and it is only transmitted to a player who exerted high effort. In this equilibrium, absence of a sanction reveals to a player with high effort that the matched partner exerted high effort while a sanction indirectly informs the player that his or her partner chose low effort. On the other hand, players that exerted low effort are sanctioned regardless of the behavior of their partner. Thus, sanctions are 'bad news' for someone who chose high effort in the first round and carry no news for those that chose low effort.

In summary, we can say that a sanction is never good news about the effort of the partner, and is bad news when own effort is high and the third player is believed to "sanction if at least one player exerted low effort". Careful inspection of the equilibrium behavior of agents leads to the following hypotheses.

Hypothesis 3 *a) For those that chose a low effort in the first round, the change in efforts and beliefs under endogenous sanctions will be similar to the change under exogenous sanctions.*

b) For those that chose a high effort in the first round, the change in efforts and beliefs will be larger under exogenous sanctions than under endogenous sanctions.

Note that the model also predicts that effort among those not sanctioned will not change. However, we do not formally test this hypothesis as our primary interest is the effect of sanctions.

5 Results

In this section we present the results of our experiment, both descriptively and statistically, guided by the general research question outlined in the introduction and more formally identified by the hypotheses formulated in Section 4. Formal statistical analysis plays an important role in uncovering the findings. We choose not to add any distributional or parametric assumptions and to build on methodology that enables correct statistical inference for the given sample sizes. In particular, we do not assume that errors are normally distributed, neither do we rely on asymptotic theory that assumes that samples are sufficiently large. Using methods that are correct is also referred to as 'exact statistics'. The term 'correct' refers here to the use of statistical tests that can formally be proven to have the level of significance that they are claimed to have.

Samples gathered in our experiment are small. Instead of using mean tests we consider new nonparametric tests that have been specifically designed for small samples (Schlag 2008). The power of these tests stems from the fact that they are based on ordinal comparisons and hence are less sensitive to unobserved values. Instead of testing whether there is a difference in means we investigate which of two random observations is likely to be larger. Formally, given two random variables Y_1 and Y_2 we test the null hypothesis that $\Pr(Y_2 > Y_1) \leq \Pr(Y_2 < Y_1)$. A rejection presents significant evidence that data drawn from Y_2 tends to be larger than data drawn from Y_1. This so-called test of stochastic inequality is explained in more detail in Appendix B. All our results are consistent with those that can be obtained using the WMW test. Note however, that contrary to conventional wisdom, the WMW test is not an exact test for comparing means unless one is willing to add distributional assumptions.[2]

[2]Counter examples are easily presented, for a simulation study showing this we refer to Forsythe et al. (1994). The WMW test is exact for testing identity of two distributions. A rejection provides significant evidence that the two distributions are not identical without providing results on how they differ.

5.1 Initial Observations

The number of participants in the experiment was 243: 45 in treatment 1 (ExNS), 51 in treatment 2 (ExS), and 147 in treatment 3 where the principal decided to introduce a sanction in 29 out of 49 groups. Hence, we had 29 observations of EnS and 20 of ExNS. Each experimental session lasted roughly 35 minutes and the subjects earned 7.5 euros on average[3]. As we already mentioned, in order to maximize the understanding of the game participants played a 5 minutes tutorial before starting the experiment and being assigned to a role. As an indication of whether people understood the game, we checked whether there were 'anomalous observations': people who specified an effort choice above the upper bound of their belief interval. We found just 6 such observations. Average effort in round one across all treatments was 145 with a large clustering of observations around 170 and a smaller cluster around 110[4]. Figure 1 shows a histogram of first round effort choices.

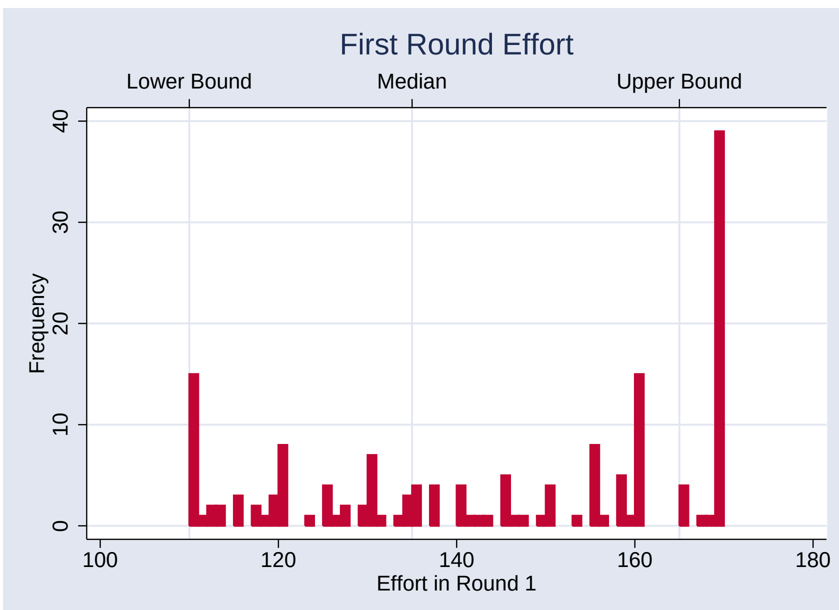

Figure 1: Histogram of first round effort choices of all subjects.

[3]If this seems little, remember that the incentives were concentrated on only two (effort) choices. At each of these choices there was thus relatively a lot at stake.

[4]The average effort levels are higher than those in Goeree and Holt (2001) who implemented a cost of effort of 0.9. Reasons may be that the cost of effort is slightly lower in our setup and that in the instructions we did not use the word "cost" when referring to k.

There is a high correlation between beliefs (as identified by the lower bound L of the elicited belief interval) and efforts in the first round of each treatment, as one would expect in a minimum effort game. The correlation coefficient is 0.85, which is significant at the 1% level.[5]

Treatment 1, in which there is no active principal and where there is no sanction, constitutes the benchmark for comparison with the other treatments. Play within this benchmark is of interest in its own right and relevant when comparing to these other treatments. Table 1 contains the average effort and beliefs in each of the two rounds where mean ExNS1 (ExNS2) denotes the mean of the first (second) round variables in the exogenous no-sanction treatment. The last column contains the estimated stochastic

	n	Mean ExNS1	Mean ExNS2	Stochastic Difference ExNS1 vs ExNS2
Effort	23	133	137	0.22
Belief	29	134	138	0.34^{*}

Table 1: Mean efforts, mean beliefs, and stochastic difference between round 1 (ExNS1) and 2 (ExNS2) in the exogenous no-sanction treatment. The results refer to two-sided test against the null-hypothesis that changes between rounds are 0. * Denotes significance at 10%, ** denotes significance at 5%, *** denotes significance at 1%.

difference between rounds one and two, formally this is the estimate of $\delta = Pr(Y_2 > Y_1) - Pr(Y_2 < Y_1))$ where Y_i is the effort level of a random subject in round i, $i = 1, 2$. Testing the null hypothesis that $\delta = 0$ is here equivalent to performing a sign test as the data is given in matched pairs. We find no marginally significant difference. Similarly, the WMW test does not detect any statistically significant difference at 10%. On the other hand we do find some marginally significant evidence that the beliefs tend to be higher in the second round. However, as we can see from Table 1, the changes in beliefs are so small that people are not sufficiently optimistic to change their effort levels by much.

Before moving to the analysis of our hypotheses we mention some issues that arise when analyzing changes between rounds. Recall that our hypotheses do not refer to

[5]The significance is based on an exact test of Schlag (2008) which has as null hypothesis that the covariance is less than 0. Note that this is not the null hypothesis underlying the Spearman rank correlation test (Spearman, 1904).

behavior of players 1 and 2 in round two but instead to change in their behavior between rounds one and two.

First of all, the observations for the group members in the third treatment are not necessarily independent. The effort decision of a subject in the first round can influence the sanctioning decision of the third player. This in turn can influence the effort and beliefs of the other subject in the second round. In our statistical testing we correct for this dependence by considering the average effort level within each group, thus analyzing the data at group level.

Second, interpreting changes in efforts and belief intervals as reaction to the experimental setting is not straightforward. We predict that people will move up their effort in a reaction to the introduction of exogenous sanctions. However, people who chose effort close to 170 in the first round will not be able to move up their effort any further, which means that these subjects do not have the ability to respond to incentives. We would like to restrict attention to those subjects who actually have room to respond. In practice this means that we restrict attention to those with first round effort below an upper bound of 165 (indicated in Figure 1) in our analysis of efforts and first round beliefs below 165 in our analysis of beliefs (thereby excluding 39 and 11 observations respectively).[6] Note that those with first round effort near 110 also face a constraint, but this is less problematic since we hypothesize that people move up in reaction to incentives. In fact, subjects do not seem to be constrained. We find that no subjects with low effort (see below) in the first round decreased their effort and only 3 subjects who had low beliefs in the first round decreased their beliefs in round two.

Third of all, our hypotheses instruct us to differentiate between high and low effort. As cutoff between the two regions we consider the sample median effort of the remaining subjects in round one which was 135. Thus, in the remainder we define high effort players as those who chose effort in the first round in $\{135, ..., 165\}$ (i.e. above the median), and low effort players as those chose first round effort in $\{110, ..., 134\}$ (i.e. below the median).

[6]The results do not depend on the choice of 165 as a threshold, but hold for any upper bound between 165 and 168, when we take the median of the associated sample as a threshold between high and low effort players.

5.2 The Incentive Effect of Sanctions

In order to separate the incentive effect of non-deterrent sanctions (Question 1) from their signaling effect we designed treatment 2 (ExS) in which sanctions are imposed unconditionally by the experimenters. To identify such an incentive effect we compare behavior in this treatment to that in treatment 1 (ExNS) so as to control for changes in behavior between rounds that occur when subjects face the same environment a second time.

Figure 2 shows the change in the mean belief and mean effort between round 1 and round 2 for the two treatments ExNS and ExS. The number on top of the bar indicates the number of independent observations. This number differs within the same treatment between beliefs and efforts due to the different number of subjects excluded who had first round choices (beliefs or efforts) too close to 170.

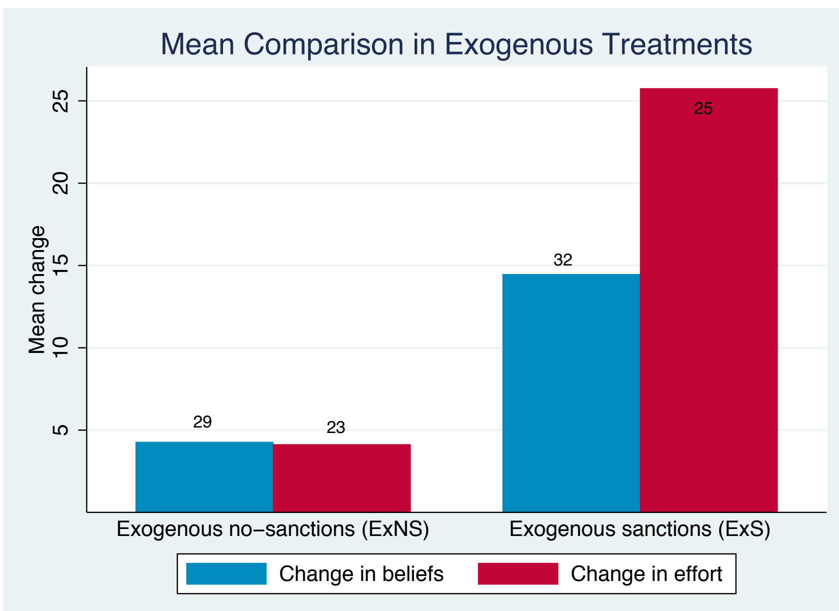

Figure 2: The change in beliefs and sanctions for the whole sample, except those who chose first round efforts $\in \{166, 167, ..., 170\}$ or first round beliefs $\in \{166, 167, ..., 170\}$. (Number of independent observations for each sample at the top of the bar).

In Table 2 we report the estimated stochastic differences between ExNS and ExS with their respective significant levels. To indicate changes between the two rounds of a treatment X we use the notation dX. For instance, the estimated stochastic difference between the change in effort under exogenous no-sanctions (ExNS) and the change in effort

	Stochastic Difference dExNS vs. dExS
Effort	0.64^{***}
Belief	0.31^{*}

Table 2: Values of stochastic difference between *changes* in the exogenous no-sanction (ExNS) treatment and *changes* in the exogenous sanction (ExS) treatment for a one-sided test against the null-hypothesis that the changes in ExS are smaller or equal than in EnNS. * Denotes significance at 10%, ** denotes significance at 5%, *** denotes significance at 1%.

under exogenous sanction (ExS) equals 0.64. We are able to reject the null hypothesis that the stochastic differences is nonpositive at level 1%. Thus, we have strongly significant evidence that exogenous sanctions make efforts change more than they would without sanctions. The impact of sanctions on change in beliefs is only marginally significant. Looking at the estimated change in means (Figure 2) and stochastic differences (Table 2) we observe that sanctions have a stronger impact on efforts than they do on beliefs. Testing this formally, we find that the null hypothesis that the change is equal or higher for beliefs is rejected at the 10% level, with the stochastic difference being 0.35.

Summary 1 *Regarding Hypothesis 1 we find a strongly significant incentive effect of sanctions on effort and a marginally significant forward looking belief effect. There is marginally significant evidence that the effect is stronger for effort than it is for beliefs.*

5.3 The Signaling Effect of Sanctions

We now come to the second main objective of this paper, to investigate the signaling effect of sanctions. For this we investigate behavior in treatment 3, using treatments 1 and 2 as controls. Signaling occurs when the principal conditions the choice of whether to sanction on the effort levels of the two players chosen in round one. Signaling has an effect when subjects make inferences about the effort of the other player when observing the choice of the principal whether or not to sanction. Note that the information contained in the choice of the principal need not be consistent with how subjects interpret why the principal chose to or not to sanction. In particular, this means that we have to separately analyze (i) the sanctioning choice of the principal and (ii) how subjects react to the choice of the principal. To separate the informational content of the sanction in (ii) from its incentive effect we will compare behavior to that in treatments 1 and 2 where the principal is inactive.

5.3.1 The Choice of When to Sanction

We wish to uncover regularities in the sanctioning choice of the principal. In particular we are interested in whether, as predicted by Hypothesis 2, that sanctions are more likely when the minimal effort in round one is low. In order to test Hypothesis 2, we compare the minimum first round effort in the sanctioned groups to the minimum first round effort of non-sanctioned groups.

We use the WMW test because we are interested in uncovering that the choice to sanction is not independent of the minimal effort choice, hence that the two distributions of minimal effort are not identical. However, we cannot find (marginally) significant evidence that the distributions of minimal effort are different in the groups where sanctions are imposed as compared to the group without sanctions imposed (the p-value is 0.63). Of course the samples are small, so the test is not very powerful. However, the descriptive data in Table 3 do not point at large differences either.

	Mean of Min. Group Effort	# Below 165	# Above 166
No Sanction	138	17	3
Sanction	135	28	1

Table 3: Descriptive data on first round minimum effort of sanctioned and non-sanctioned groups. The columns show the mean, and the number of groups with minimum effort below 165 and above 166.

Summary 2 *We have no statistically significant evidence in favor of Hypothesis 2, that sanctioned groups had lower minimum effort. The descriptive statistics similarly indicate a lack of a clear pattern. It is as if sanctions have been imposed independently of the minimum effort.*

Hence we cannot confirm Hypothesis 2. However, note that this result does not contradict our theoretical framework outlined in Appendix A, since this admits equilibria in which sanctions are chosen independent of effort levels. An explanation for the lack of pattern may be the established tendency of people to sanction "too often" in economic experiments (Fehr and Rockenbach 2003).

5.3.2 Reacting to the Information Perceived in the Choice of the Principal

We now turn to investigate whether subjects perceive an informational content in the choice of the principal. Although our analysis above indicates that there is no clear informational content, subjects may still believe that sanctions were imposed as a reaction to low minimum effort levels. Specifically, subjects may follow the same reasoning that led us to formulate Hypothesis 2. If this is the case, sanctions may still influence beliefs about the other group member. Following Hypothesis 3 the predicted effect depends on whether a subject chose low or high effort in round one.

Consider the behavior of low effort players. Figure 3 presents the mean changes in beliefs and effort for people who played low effort in the first round.

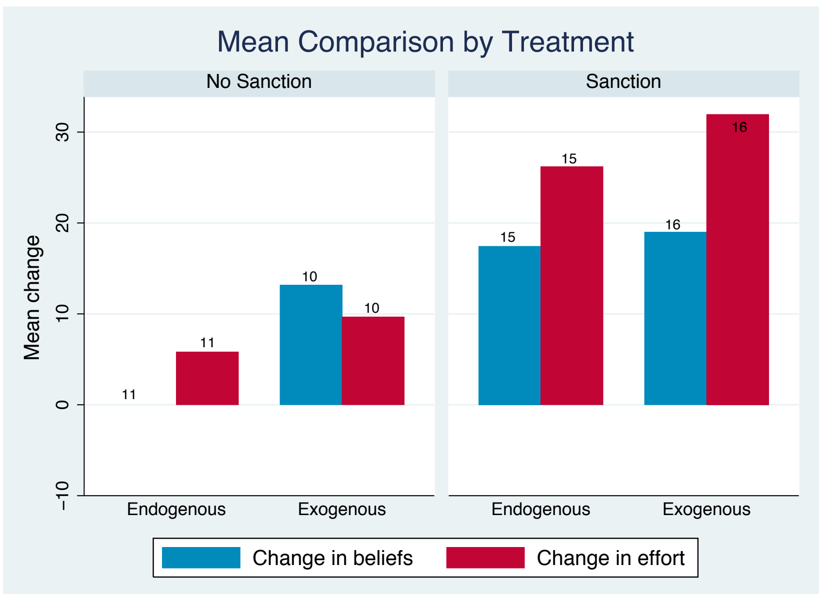

Figure 3: Means of changes in beliefs and effort across treatments, for those who played low effort ($\in \{110, 111, ..., 134\}$) in the first round. (Number of independent observations for each sample at the top of the bar).

Figure 3 reveals no large differences between the exogenous and endogenous sanction treatments. The results in Table 4 show that we cannot reject the null hypothesis of identical distributions in the exogenous and endogenous treatments, using the WMW test, both for effort and beliefs. Hence we cannot reject Hypothesis 3*a*).

	WMW p-values dEnS vs. dExS
Effort	0.29
Belief	0.97

Table 4: p-values of the two-sided Wilcoxon Mann-Whitney test for the null hypothesis that the effects of the exogenous and endogenous sanction treatment are identical among those that chose low effort ($\in \{110, 111, ..., 134\}$) in the first round. * Denotes significance at 10%, ** denotes significance at 5%, *** denotes significance at 1%.

Note that the above does not confirm Hypothesis 3*a*), that there is no effect for low effort players. Confidence intervals could be useful if samples were substantially larger. To obtain some indication in line with Hypothesis 3*a*) we contrast the stochastic difference of the change in behavior of these two treatments and report results in Table 5. Estimates

	Stochastic Difference	
	ExS1 vs. ExS2	EnS1 vs. EnS2
Effort	1^{***}	1^{***}
Belief	0.5^{**}	0.8^{***}

Table 5: Estimates of stochastic difference between round 1 and round 2 of treatments ExS and EnS, for those who played low effort ($\in \{110, 111, ..., 134\}$). Results are for a one-sided test against the null-hypothesis that choices in the second round are lower or equal to that in the first round. * Denotes significance at 10%, ** denotes significance at 5%, *** Denotes significance at 1%..

and levels of significance are very similar. Effectivity of sanctions on low effort players does not seem to depend on whether the sanctions were endogenous or exogenous.

Summary 3 *Regarding Hypothesis 3a, for subjects who made low efforts in the first round we find no statistically significant evidence that endogenous and exogenous sanctions have different effects on either efforts or beliefs. There is some indication that lack of significant* *difference is not due to small sample sizes but that in fact behavior is the same under endogenous and exogenous sanctions.*

Consider now the behavior of high effort players. In Figure 4 we report average changes in efforts and beliefs across treatments for subjects who played high efforts in the first round. Looking at Figure 4 it seems like the exogenous sanctions are more effective than

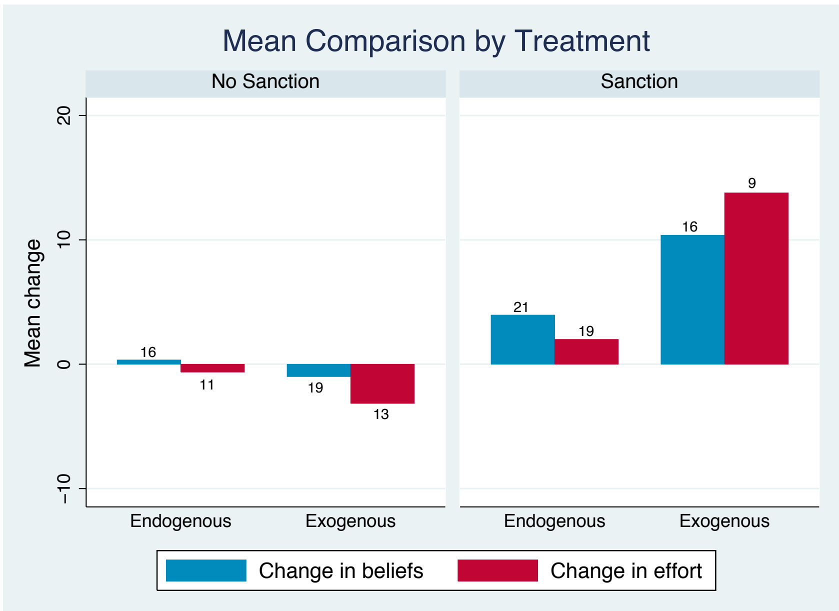

Figure 4: Means of changes in beliefs and effort across treatments, for those who played high effort ($\in \{135, 136, ..., 165\}$) in the first round. (Number of independent observations for each sample at the top of the bar).

the endogenous ones for those who played high effort. Our statistical analysis based on stochastic differences, reported in Table 6, confirms this. We observe significant evidence that exogenous sanctions are more effective in raising effort than endogenous sanctions. There is marginal significant evidence that beliefs tend to change more under exogenous than under endogenous sanctions. One wonders whether endogenous sanctions have any effect at all. To find out we test if there is a difference between the endogenous sanction treatment and the baseline treatment (ExNS). In the first column of Table 7 we report the p-values of the WMW test for this comparison (stochastic differences are similarly insignificant). There is no statistically significant evidence that endogenous sanctions are effective on high effort players. However, the sample sizes are small, so it is possible that

	Stochastic Difference dEnS vs. dExS
Effort	0.66^{**}
Beliefs	0.39^{*}

Table 6: Estimates of stochastic difference between the exogenous and endogenous sanction treatments for those who played high effort ($\in \{135, 136, ..., 165\}$) in the first round. Results are for a one-sided test against the null-hypothesis that dExS is lower than or equal to dEnS. * Denotes significance at 10%, ** denotes significance at 5%, *** denotes significance at 1%.

	WMW p-value dExNS vs. dEnS	Stochastic Difference dExNS vs. dExS
Effort	0.35	0.78^{***}
Belief	0.49	0.48^{**}

Table 7: Comparison of the baseline (ExNS) treatment and the sanction treatments for those who played high effort ($\in \{135, 136, ..., 165\}$) in the first round. Results in the first column are for a two-sided WMW test that EnS do not differ from ExNS, those in the second column are for a one-sided test against the null-hypothesis that dExS is lower than or equal to dExNS. * Denotes significance at 10%, ** denotes significance at 5%, *** denotes significance at 1%.

we would not be able to reject the null hypothesis of equal distributions, even if the actual difference is quite large. To counter this criticism, the last column of Table 7 shows the same stochastic difference when the sanction is imposed exogenously. In this case, despite the small sample sizes, we have statistically significant evidence that exogenous sanctions are effective among the high effort players.

Summary 4 *Regarding Hypothesis 3b, for subjects who chose high effort in the first round, endogenous sanctions are significantly less effective in raising efforts and beliefs than exogenous sanctions. In fact, the effect of endogenous sanctions cannot be distinguished from the effect of not introducing a sanction at all.*

5.4 Sanctions and Uncertainty

Naturally, uncertainty about the action of the other player plays a role in the choice of effort of each player. Although our model is not rich enough to yield hypotheses in this regard, we now look empirically at the effect of the introduction of an exogenous or endogenous sanction on this uncertainty.

One of the reasons we asked the participants to specify an interval rather than a point belief was that the elicited interval provides some indication of the uncertainty about the behavior of the other player. Schlag and van der Weele (2009) show that for given risk preferences, changes in the width of the interval correspond to changes in the dispersion of the distribution. This makes $U - L$ a proxy for the changes in uncertainty of a given subject. Figure 5 shows the changes in the width of the belief interval for those who chose L in $\{110, 111, ..., 165\}$ in the first round. As Figure 5 shows, uncertainty did not

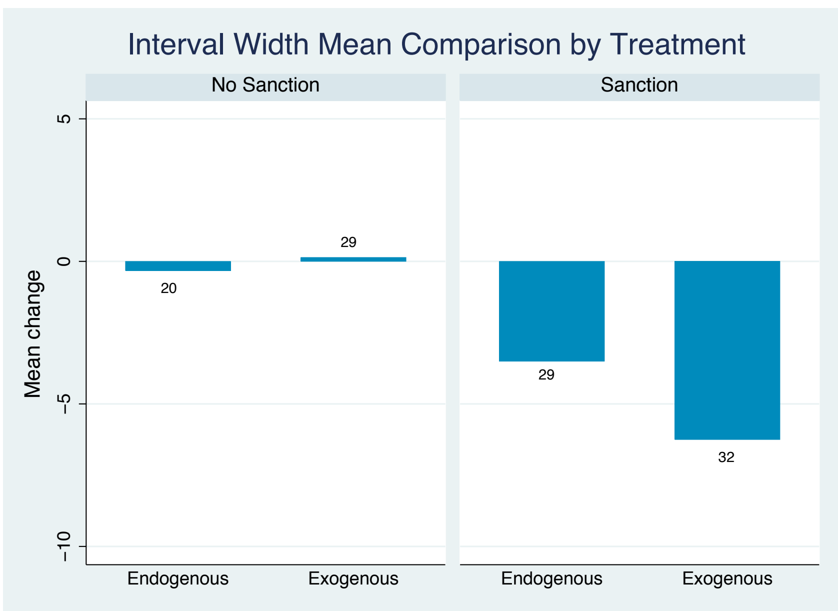

Figure 5: Means of change in the width of the interval across treatments, for those who chose the lower belief interval in the first round in ($\in \{110, 111, ..., 165\}$) in the first round (number of independent observations for each sample at the top of the bar).

change between rounds in both no-sanction treatments, while uncertainty went down in both sanction treatments. We can confirm this result with statistical analysis. Table 8 presents the estimates of stochastic difference between the first and the second round interval width in all treatments. In both cases with no sanctions a test of stochastic

	Stochastic Difference			
	EnNS1 vs. EnNS2	ExNS1 vs. ExNS2	EnS1 vs. EnS2	ExS1 vs. ExS2
Interval Width	0.050	0.0	-0.28^{*}	-0.38^{**}

Table 8: Estimates of stochastic difference between the round 1 and round 2, for those who chose the lower belief interval in the first round in ($\in \{110, 111, ..., 165\}$) in the first round. Results in column 1,2 are for a two-sided test against the null hypothesis that the change is 0, results in column 3,4 for a one-sided test against the null hypothesis that the change in the interval width is either positive or 0. * Denotes significance at 10%, ** denotes significance at 5%, *** denotes significance at 1%..

inequality cannot reject the null hypothesis that the distributions in the two rounds are equal at the 10% level. By contrast, we find that there is significant evidence that the interval decreases under exogenous sanctions and marginally significant evidence that the interval decreases under endogenous sanctions. This reinforces our conclusion that sanctions facilitate coordination partly by reducing uncertainty about the behavior of others.

If sanctions were to have a signaling effect, we would expect for those subjects who chose high effort in the first round, that the reduction in uncertainty is smaller under endogenous sanctions than under exogenous sanctions. Testing the direction of the effect with stochastic inequality, we find a strongly significant decrease in uncertainty at 1% in the exogenous sanction treatment, while under endogenous sanctions it is no longer significant.

Summary 5 *There is statistically significant evidence that uncertainty about the choice of the other player is reduced when sanctions are imposed. This is also true within the subset of high effort players when the sanction is exogenous. No statistical evidence of change in uncertainty is found in absence of sanctions or among high effort players when the sanction is introduced endogenously.*

6 Discussion and Conclusion

The results of our experiment allow us to conclusively answer our two questions. Sanctions have a positive effect on effort levels and beliefs about others' effort level for those that chose low effort in the first round. For those that chose high effort their effect depends on whether sanctions were imposed exogenously or endogenously. In particular, if sanctions are endogenous then high effort players are more pessimistic, as compared to exogenous sanctions, which practically eliminates the effectiveness of the sanction.

We find that the theory that led to our hypotheses is the most plausible explanation for this result. The endogenous introduction of sanctions signals to subjects with high effort in round one that their partner did not 'cooperate', in the sense that she selected low effort. This make them more pessimistic about the effort of their partner in the next round which overrides the incentive effect. The signaling effect also explains why the effect of the sanction does not depend for those with low effort in the first round on the endogeneity. For them there is no signal, as the sanction may be aimed at them rather than at the other player in the group.

Note that there is another possible explanation for the difference between endogenous and exogenous sanctions. Even though the sanction does not operate retroactively, Players could interpret the introduction of an endogenous sanction as an unkind act by the principal. If agents have reciprocal preferences, they may retaliate by reducing their effort in order to lower the payoffs of the principal. We find this explanation less convincing for several reasons. First, the interests of the principal and the agents are completely aligned, and sanctions are costly. It would therefore be strange to interpret the introduction of a sanction as anything else than an attempt to raise the minimum effort, which is in the interest of all the players. Second, the nature of sanctions is such that it hurts those who play low effort the most. Therefore, one would expect low effort players rather than high effort players to retaliate against the principal, which is inconsistent with the data. Third, we consider the following indication of retaliatory punishment: subjects who choose an effort level below the lower bound of their belief level after being sanctioned by the principal (but not in the first round). We find this behavior in only 2 out of 58 cases, one of which represents a change of only 2 points. Thus, it is hard to reconcile this evidence with the hypothesis of retaliatory punishment against the principal.

Turning to the real world, the results of the paper have implications for both public policies and manager-employee relationships in firms. As pointed out by Brandts and Cooper (2006), coordination failure can cause corporations and other organizations to become trapped in unsatisfactory situations both for managers and employee. 'Mild law' has been suggested as a way to improve efficiency in coordination environments by influencing expectations (McAdams 2000). However, our experiment shows that mild law has drawbacks that are associated with the signaling effect. Moreover, the experiment shows that these effects can be quite substantial.

Our experiment leads to policy conclusions that complement those emerging from the study of Tyran and Feld (2006). In their study, a voting procedure for the introduction of a mild sanction gives people the opportunity to send a public signal that they are willing to cooperate. This in turn leads to increased cooperation. Their sanctions are most effective when chosen endogenously. In our experiment, the introduction is under the discretion of a third player who has observed past play of the game. This setup reflects more closely the arrangements of a society where people make the laws through representatives, rather than directly. Our study suggests that 'mild law' may not be the best instrument in this case, because it does not compensate for this signaling effect by providing adequate incentives for efficient behavior. In our study sanctions are more effective when introduced exogenously. In fact, our results suggest that policies introduced by voting will also have a signaling effect, as in the real work the decision to put the vote on the agenda in the first place is endogenous.

There some reason to think that the signaling effect may be greater in the real world than in our experiment. In the experiment, the signaling effect was not present for low effort players, because the groups were so small that the sanction was likely to reflect their own behavior. However, in real life, relevant communities consist of more than two people, and sanctions are likely to be implemented only after misbehavior of a substantial fraction of the group. This means that even people who play low effort may interpret the sanction as a signal of misconduct of others. Assuming some external validity of the experiment, one can conclude that a sanctioning authority needs to attain a careful balance between correcting the behavior of deviants or pessimists and maintaining the optimistic beliefs of cooperators.

How to attain such a balance is an interesting further research question that goes beyond the aim of this paper. One possibility is to implementing deterrent laws that would presumably override the signaling effect. However, such the enforcement may be costly to implement and the execution of harsh penalties may not be in line with prevailing norms of proportionality. Another possibility to investigate is whether appropriate framing of the introduction of a law can mitigate the signaling effect. In the tradition of experimental economics, this paper has tried to use neutral framing, replacing "effort" with "a number", and "sanction" with "subtraction". In real life however, a policy maker could attempt to surround the introduction of sanctions by soothing or stimulating messages. For example, one may say the actual number of people who deviate from the efficient strategy is small, or express the expectation that they will conform to the sanction. However, it is theoretically unclear why such cheap talk would be effective. The experiments by Brandts and Cooper (2008) and Van Huyck et al. (1992) incorporate the possibility of a principal to send written messages and suggestions to the agents. These studies could be combined with the asymmetric information structure in this paper in order to study this issue.

Last but not least, we hope to promote use of new tests that are exact but do not, like the WMW test, impose additional distributional assumptions. We think these tests are an important addition to the toolbox of economists working with small data sets.

7 References

Benabou, Roland and Tirole, Jean (2003). "Intrinsic and Extrinsic Motivation", *Review of Economic Studies*, 70, 489-520.

Bohnet, Iris, and Robert Cooter (2001). "Expressive Law: Framing of Equilibrium Selection?" Mimeo, Harvard University and UC Berkley.

Bowles, Sam (2008). "Policies Designed for Self-Interested Citizens May Undermine the Moral Sentiments: Evidence from Economic Experiments", *Science*, 320.

Brandts, Jordi and Cooper, David (2008). "It's What You Say Not What You Pay", forthcoming in *Journal of the European Economic Association*.

Brandts, Jordi and Cooper, David (2006). "A Change Would Do You Good: An Experimental Study on How to Overcome Coordination Failure in Organizations", *American Economic Review*, 96(3), 669-693.

Brunner, Edgar and Munzel, Ullrich (2000). "The Nonparametric Behrens-Fisher Problem: Asymptotic Theory and a Small-Sample Approximation", *Biometrical Journal*, 42, 17-25.

Cliff, Normann (1993). "Dominance Statistics: Ordinal Analyses to Answer Ordinal Questions", *Psychological Bulletin* 114, 494-509.

Cooter, Robert (1998). "Expressive Law and Economics", *Journal of Legal Studies*, 27, 585-607.

Costa-Gomes Miguel and Crawford, Vince (2006). "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study", *American Economic Review*, 96, 1737-1768.

Devetag, Giovanna and Ortmann, Andreas (2007), When and Why? A Critical Survey on Coordination Failure in the Laboratory, Experimental Economics, 10, 331-344.

Fehr, Ernst and Rockenbach, Bettina (2003), "Detrimental effects of sanctions on human altruism", *Nature*, 422, 137-140.

Fischbacher, Urs (2007). "z-Tree: Zurich Toolbox for Ready-made Economic experiments", *Experimental Economics* 10(2), 171-178.

Forsythe, Robert, Horowitz, Joel L., Savin, N. E., and Sefton, Martin (1994). "Fairness in Simple Bargaining Experiments", Games Economic Behavior, 6, 347-369.

Frey, Bruno S. and Jegen, Reto (2001). "Motivation Crowding Theory", *Journal of Economic Surveys*, 15(5), 589-611.

Frey, Bruno S., & Oberholzer-Gee, Felix (1997). "The Cost of Price Incentives: an Empirical Analysis of Motivation Crowding-Out" *American Economic Review*, 87(3), 746-755.

Friebel, Guido and Wendelin, Schnedler, 2007. "Team Governance: Empowerment or Hierarchical Control," IZA Discussion Papers 3143.

Gneezy, Uri and Rustichini, Aldo (2000). "A Fine is a Price", *Journal of Legal Studies*, 29, 1-17.

Goeree, Jacob and Holt, Charles A. (2005). "An Experimental Study of Costly Coordination", *Games and Economic Behavior*, 51, 349-364

Goeree, Jacob and Holt, Charles A. (2001). "Then Little Treasures of Game Theory and Ten Intuitie Contradictions", *American Economic Review*, 91(5), 1402-22.

Mann, H. B., and Whitney, D. R. (1947), "On a Test whether one of two random variables is stochastically larger than the other," *Annals of Mathematical Statistics*, 18, 50–60.

McAdams, Richard H. (2000). "A Focal Point Theory of Expressive Law", *Virginia Law Review*, 86, 1849-1731.

McAdams, Richard H. and Nadler, Janice (2003). "Testing the Focal Point Theory of Legal Compliance: The Effect of Third-Party Expression in an Experimental Hawk/Dove Game", *Journal of Empirical Legal Studies*, 2(1), 87-123.

Nagel, Rosemarie, (1995). "Unraveling in Guessing Games: An Experimental Study", *American Economic Review*, 85(5), 1313-1326.

Schlag, Karl H. (2008). "A New Method for Constructing Exact Tests without Making any Assumptions", Working Paper 1109, Department of Economics and Business, Universitat Pompeu Fabra.

Schlag, Karl H. and Joël J. van der Weele, (2009) "Efficient Interval Scoring Rules", Working Paper 1176, Department of Economics and Business, Universitat Pompeu Fabra.

Sliwka, Dirk (2007). Trust as a Signal of a Social Norm and the Hidden Costs of Incentive Schemes " *American Economic Review*, 97 (3), 999-1012

Spearman, C. (1904), "The Proof and Measurement of Association Between Two Things," *American J. Psychol.*, 15, 72-101.

Sunstein, Cass R. (1996). "On the Expressive Function of Law", *University of Pennsylvania Law Review* 144(5), 2021-2053.

Tyran, Jean-Robert and Feld, Lars P. (2006). "Achieving Compliance when Legal Sanctions are Non-deterrent," *Scandinavian Journal of Economics*, 108(1), 135-156.

Van Huyck, John, Gillette R. and Battalio, Raymond (1992) "Credible Assignments in Coordination Games", *Games and Economic Behavior*, 4, 606-626.

Van der Weele, Joel (2011). "The Signaling Power of Sanctions in Social Dilemmas", *Journal of Law, Economics and Organization.*

Vargha, András and Harold D. Delaney (1998). "The Kruskal-Wallis Test and Stochastic Homogeneity", *Journal of Educational and Behavioral Statistics*, 23(2), 170-192.

Wilcoxon, Frank (1945), "Individual Comparisons by Ranking Methods," *Biometrics Bulletin* 1, 80–83.

Xiao, Erte (2010), "Profit seeking punishment corrupts norm obedience," Mimeo Carnegie Mellon University.

Appendix A: A Simple Model and Hypotheses

In this section we present a simple model of behavior for the game specified in Section 3.1. A summary is provided in Section 4.

We first predict behavior for the case where the third player does not have an active role. To generate predictions we make the following assumptions. Players believe that they are more sophisticated than their opponents (who we also call partners) and best respond to anticipated effort of their partner. Partners are believed to best respond to a given belief distribution of effort levels where these beliefs do not change over time. Thus, the sophistication of players is as in the models of level k thinking or cognitive hierarchy (Nagel (1995) and Costa Gomes and Crawford (2006)). In the terminology of these models, all players in our paper belong to level 2.[7]

Partners best respond to one of two different belief distributions G_h and G_l, accordingly some choose high while others choose low effort, respectively they are referred to as high types and low types. Each player assesses a probability or belief p that her partner is the high type. Players and partners are risk-neutral. Finally, players choose their effort levels in the first round as if there was no second round, so completely myopically. This is in accordance with the experimental setup, where people did not know in the first round that there would be a second round.

7.1 First Round Effort

To determine her effort in the first round, a player will first calculate the optimal effort of the high type and of the low type partner and then choose a best response on the basis of the probability p of meeting the high type. Denote the optimal effort levels of the high and low type when there is no sanction by $e_h(0)$ and $e_l(0)$ respectively, where 0 indicates that there is no sanction (in later sections a 1 will indicate that a sanction has been imposed). So $e_h(0) \in \arg\max_e \left(\int \min\{e, e'\} \, dG_h(e') - ke\right)$ where the cost of effort k was equal to 0.85 in the experiment. We assume that G_h and G_l are such that $e_h(0) > e_l(0)$. According to our assumptions, each player believes with probability p that she faces a partner who chooses $e_h(0)$ and with probability $1-p$ a partner who chooses $e_l(0)$. Let

[7] At the cost of substantial additional complexity, one could assume more sophisticated distribution of rationality levels. Specifically, specifying higher levels of rationality would lead to more complex belief effects. We believe that the data do not justify the cost of such an analysis.

e_p^r denote the optimal effort level of a player with belief p in round r, $r = 1, 2$. Note that $e_p^1 = e_l(0)$ if $p = 0$ and $e_p^1 = e_h(0)$ if $p = 1$. Taking into account that $e_p^1 \in [e_l(0), e_h(0)]$ holds for all $p \in [0, 1]$ we can write the expected utility of a player with belief p who exerts effort $e \in [e_l(0), e_h(0)]$ as $Eu = p(e - ke)) + (1 - p)(e_l(0) - ke)$ and obtain

$$\frac{d}{de}Eu = p - k.$$

So if $p > k$ then $e_p^1 = e_h(0)$, if $p < k$ then $e_p^1 = e_l(0)$.

7.2 The Effect of an Exogenous Sanction

We now consider choice of effort in round 2 when an exogenous sanction has been imposed. Imposing a sanction means to subtract $k_1(170 - e)$ from the payoff for some given $k_1 > 0$. In the experiment we set $k_1 = 0.5$. This change in payoffs influences effort choices of the level 1 player. Let $e_v(1)$ be the optimal effort of type $v \in \{h, l\}$ when there is a sanction. So $e_h(1) \in \arg\max_e \left(\int \min\{e, e'\} dG_h(e') - ke - k_1(170 - e)\right)$. Note that $e_v(1) \geq e_v(0)$ for $v \in \{h, l\}$, i.e. partners (are believed to) exert more effort after a sanction has been imposed.

Expected utility of a player who exerts effort $e \geq e_l(1)$ is now

$$Eu = p(e - ke - k_1(170 - e)) + (1 - p)(e_l(1) - ke - k_1(170 - e)).$$

Hence, $\frac{d}{de}Eu = p - (k - k_1)$. If $p > k - k_1$ then $e_p^2 = e_h(1)$, if $p < k - k_1$ then $e_p^2 = e_l(1)$. Thus, all players exert weakly more effort after an exogenous sanction has been introduced. We can decompose this change in effort into two effects. First, there is an *incentive effect*, because a sanction effectively reduces the cost of effort k and thus gives incentives for higher effort. Specifically, any player with $p \in (k - k_1, k)$ chooses the effort of the high type in round 2 while they choose the effort of the low type in round 1. Players with $p < k - k_1$ and $p > k$ choose the same effort in round 2 as they do in round 1. Second, there is a *forward looking belief effect* because introducing a sanction leads to a belief that partners will choose a higher effort as they too face lower effort costs. The forward looking belief effect additionally raises the effort levels of the players.

We now compare the effect of a sanction on beliefs to their effect on effort levels. For this we assume that players are drawn from some distribution such that G_p describes the distribution of p. To simplify exposition assume that G_p has no point masses and full support on $[0,1]$. Then the expected beliefs (in terms of the expected effort of a partner) in round one equals

$$\int \left(pe_h(0) + (1-p)\, e_l(0)\right) dG_p(p),$$

while the expected effort (of a player) in round one is equal to

$$\int e_p^1 dG_p(p) = G_p(k)\, e_l(0) + (1 - G_p(k))\, e_h(0).$$

In order to make efforts and beliefs comparable in round two we impose a mild consistency requirement, namely that expected beliefs equal expected effort in the first round. Following the above this means that $\int p dG_p(p) = 1 - G_p(k)$.

In round 2, invoking consistency, we find that expected beliefs equal

$$\int \left(pe_h(1) + (1-p)\, e_l(1)\right) dG_p(p) = G_p(k)\, e_l(1) + (1 - G_p(k))\, e_h(1),$$

and that expected effort equals

$$G_p(k - k_1)\, e_l(1) + (1 - G_p(k - k_1))\, e_h(1).$$

Comparing these two terms we conclude for round 2 that expected effort is higher than expected belief. Given that these two expressions are equal by assumption equal in round 1 we obtain the following result.[8]

Result 1 *Exogenous sanctions increase both beliefs and effort where effort increases more than beliefs.*

Note that in the treatment where no sanction is introduced in round 2, payoffs and beliefs remain unchanged and hence $e_p^2 = e_p^1$, efforts remain unchanged as well.

[8] Without the consistency requirement, this is is not necessarily true. As a counter-example, consider the case where high type partners have point beliefs and hence do not respond to lower effort costs. Assume furthermore that beliefs are such that both players choose high effort in the first round. As the effort of the high type partner remains unchanged in round two, players' effort remains unchanged too. Yet if some probability is put on the low type partners and if these respond to changes in effort cost, we find that beliefs move more than effort. However, this scenario occurs only if beliefs are inconsistent in the sense that first round efforts are higher than first round beliefs.

7.3 The Effect of an Endogenous Sanction

Next we investigate behavior when it is the third player, who we refer to as principal, who chooses whether or not to sanction. The principal's payoffs are given by $0.25 \min\{e_1, e_2\} - cs$ where in our experiment we set $c = 4$. Note that $c/0.25 = 4c$ is the cost of sanctioning in units of efforts. The principal is risk neutral and has a prior G_p over the possible values of belief p held by the players.

We develop some notation. Let $e_p(s)$ be the optimal effort given belief p where $s = 1$ ($s = 0$) indicates that a sanction has been imposed (has not been imposed). Let p_i be the belief of player i, $i = 1, 2$. Let $p_m = \min\{p_1, p_2\}$ and $p_x = \max\{p_1, p_2\}$. Let $s^* : [110, 170]^2 \to \{0, 1\}$ be such that $s^*\left(e^1_{p_1}, e^1_{p_2}\right)$ is the choice of the principal of whether or not to sanction conditional on observed effort level $e^1_{p_i}$ of player i in round 1, $i = 1, 2$.

Choices in the first round are myopic as players do not anticipate that there will be a second round. We will not consider deviations from such play. Thus, the principal will observe only effort choices belonging to $\{e_l(0), e_h(0)\}^2$ and only needs to condition on these. We call $e_l(0)$ and $e_h(0)$ a low and a high first round effort respectively. We will consider only sanctioning strategies where sanctioning choices do not depend on player indices but only on effort levels. Thus we can identify $s^* : [110, 170]^2 \to \{0, 1\}$ with $s^* \in \{0, 1\}^3$ where s^*_1, s^*_2 and s^*_3 are the sanctioning choices conditional on the first round events $\{(e_l(0), e_l(0))\}$, $\{(e_l(0), e_h(0)), (e_h(0), e_l(0))\}$ and $\{(e_h(0), e_h(0))\}$ respectively.

We will make predictions that satisfy the following requirements.

1. The strategies of the principal and the two players can be supported as a Perfect Bayesian Equilibrium (PBE). Out of equilibrium actions of the principal do not change the belief of a player about her partner's effort.

2. The PBE does not depend on the specific form of the prior of the principal.

3. The PBE can be sustained for a non-degenerate interval of values of c.

We make some comments before we turn to the analysis. Given the assumptions above, a PBE is uniquely characterized by the sanctioning function $s^*(,)$ of the principal. Following (2) the equilibrium candidate must be optimal, regardless of the beliefs over p_1 and p_2. This implies that it will be sufficient to evaluate deviations from an equilibrium candidate using a degenerate prior of the principal, i.e. when the principal is (almost) sure about p_1 and p_2. If the principal does not want to deviate under any degenerate prior, she will also not want to do so under more general priors. To see this, it suffices to note that expected payoffs of a deviation under a general prior are just a convex combination of payoffs under some degenerate priors, and therefore cannot be strictly higher.

There are $2^3 = 8$ candidates for a PBE. In two of these the principal's choices are unconditional: $s^* = (1, 1, 1)$ and $s^* = (0, 0, 0)$. To "always sanction", i.e. $s^* = (1, 1, 1)$, can be supported if and only if $e_{p_m}(1) - 4c \geq e_{p_m}(0)$ holds for all p_m. Here we use our requirement that beliefs p_i do not change when the principal chooses the out of equilibrium action to not sanction. Necessary and sufficient conditions are given by $e_h(1) - 4c \geq e_h(0)$ and $e_l(1) - 4c \geq e_l(0)$. When investigating "never sanction", i.e. $s^* = (0, 0, 0)$, special attention must be given to a player with $p \in (k - k_1, k)$. A sanction would induce this player to switch from low to high effort, this is not in the interest of the principal if $e_h(1) - 4c \leq e_l(0)$. In fact, this is a necessary and sufficient condition for supporting "never sanction".

An intuitive conditional strategy is given by $s^* = (1, 1, 0)$ where the principal sanctions if and only if at least one of the two players chose a low effort in the first round. The conditions supporting this as a PBE emerge when considering three subcases. When both players chose low first round effort then s^* prescribes to sanction is best when $e_{p_m}(1) - 4c \geq e_l(0)$ holds for all p_m, hence when $e_l(1) - 4c \geq e_l(0)$. When both exerted high effort in the first round, then $s^* = 0$ which is best if $e_h(0) \geq e_l(1) - 4c$. Finally, consider the case where one player had a low and the other a high first round effort. Then $s^* = 1$ which yields outcome $e_l(1) - 4c$ as the player with high first round effort now believes that her opponent is of low type. Not sanctioning causes the player with low first round effort to choose $e_{p_i}(0) = e_l(0)$ and the one with high first round effort to choose $e_h(0)$, which is worse if $e_l(1) - 4c \geq e_l(0)$. Together this means that $s^* = (1, 1, 0)$ can be supported if and only if $e_h(0) \geq e_l(1) - 4c \geq e_l(0)$. Note that in this equilibrium, sanctions are "bad news" in the sense that playing $s^* = 1$ will alert a high effort player to the fact that her opponent played low effort.

The five remaining strategies can all be ruled out by our requirements $1 - 3$. It is easy to show that $s^* = (0, 1, 0)$, $s^* = (0, 1, 1)$, $s^* = (0, 0, 1)$ cannot be supported at PBE. Moreover, one can rule out $s^* = (1, 0, 0)$ and $s^* = (1, 0, 1)$ using requirement 3. We summarize:

Proposition 1 *The following values of s^* are the only ones that can be supported as a PBE for all G_p for a nondegenerate set of c: (i)* $(1, 1, 1)$ *if* $e_v(1) - 4c > e_v(0)$ *for*

$v \in \{l, h\}$, *(ii)* $(1, 1, 0)$ *if* $e_h(0) > e_l(1) - 4c > e_l(0)$, *(iii)* $(0, 0, 0)$ *if* $e_h(1) - 4c < e_l(0)$.

Proposition 1 implies that there is no unique prediction for whether or not the principal will sanction low types or whether or not she will sanction high types. With respect to players 1 and 2, their efforts remain unchanged if there is no sanction. Players with low first round effort who are sanctioned increase effort in the same way as under an exogenous sanction, because the sanctions do not change the belief about the type of player she is facing. However, the predicted change in effort of a player with high first round effort is ambiguous. She will increase effort in case (i), but when sanctions are "bad news" as in case (ii), she may reduce effort.

Appendix B: Stochastic Difference and Inequality

In the following we present a new exact nonparametric test for comparing outcomes based on two independent samples due to Schlag (2008). It is specifically designed to uncover with small samples how two distributions differ without adding distributional assumptions. Previous such tests can only identify that the two distributions differ but not how they differ, respective the Wilcoxon-Mann-Whitney test see Forsythe *et al.* (1994).

Given two random variables Y_1 and Y_2, $\delta(Y_1, Y_2) = \Pr(Y_2 > Y_1) - \Pr(Y_2 < Y_1)$ is called the *stochastic difference* of Y_1 verses Y_2 (Cliff, 1993). The stochastic difference can be estimated by computing the sample analogues. Consider first the case of matched pairs where data is given by joint observations of Y_1 and Y_2. The estimate is calculated by ignoring all pairs in which $Y_1 = Y_2$ and then taking the difference between the empirical frequency of pairs with $Y_2 > Y_1$ and of pairs in which $Y_2 < Y_1$. Now consider the case in which there are two independent samples, one associated to each variable. Here one can estimate δ by considering the frequency of $Y_2 > Y_1$ among all possible pairs and subtracting from this the frequency in which $Y_2 < Y_1$ among all these pairs. The resulting estimates are unbiased.

If $\delta(Y_1, Y_2) > 0$ then one says that Y_2 tends to yield larger outcomes than Y_1. We wish to identify significant evidence that Y_2 tends to yield larger outcomes than Y_1. So we wish to test the null hypothesis $H_0 : \delta(Y_1, Y_2) \leq 0$ against the alternative hypothesis $H_1 : \delta(Y_1, Y_2) > 0$ for a given specified level α. Following Vargha and Delaney (1998) we call this a test of *stochastic inequality* (see also Brunner and Munzel, 2000).

Assume that data has the form of *matched pairs* as given by n independent observations of (Y_1, Y_2). Then this test reduces to a sign test. One uses a binomial test to test whether the probability that $Y_2 > Y_1$ conditional on $Y_2 \neq Y_1$ is $\leq 1/2$.

Assume instead that data is given by two independent samples of Y_1 and of Y_2. Let n_i be the number of observations of Y_i, $i = 1, 2$. The new test proceeds as follows. Randomly match one observation from each sample to generate $\min\{n_1, n_2\}$ matched pairs. Act as if these matched pairs are the original sample and determine the rejection probability of the randomized version of the sign test that has size $0.2 \cdot \alpha$. Repeat the last two steps, matching and evaluating the rejection probability, infinitely often and record the average rejection probability. Finally, reject the null hypothesis if the average rejection probability is above 0.2. Note that the factor 0.2 used to reduce the size of the randomized test is equal to the threshold used to translate the randomized recommendation into a deterministic recommendation.

Appendix C: Instructions [NOT FOR PUBLICATION]

We report instructions for the endogenous sanction treatment, which are the most comprehensive. Original instructions were in Italian.

Instructions for the first round

Introduction

Welcome! You are going to take part in an experimental study of decision making. Please follow these instructions carefully. You will be paid according to your performance. At the end of the experiment we will tell you how much you earned.

Once everyone is seated we will formally start the experiment by reading the instructions. After this reading you will have the opportunity to ask us questions about the procedure. However at no time may you communicate with any of the other participants of your session. Please also refrain from talking to others about your experience until tomorrow in order not to influence others taking part in our experiment. Please turn off your mobiles in case they are still switched on.

Matching and assignment to a role

The computer will assign you by chance (i.e. at random) to a group consisting of three participants. You will not know the identity of the other two in your group and they will not know your identity. The computer will also assign a role to each in this group. Two of this group (from now on: player 1 and player 2) will have to take a decision as described below, the third (from now on: player 3) will be inactive but still will earn some money.

Decisions and Earnings

During the experiment any choice will lead to some earnings expressed in tokens. Total earnings at the end of the experiment are determined by the sum of all earnings and will then be converted into money at the exchange rate of

1 token $= 7,5$ Eurocents (or equivalently: 100 tokens$= 7,5$ Euro)

It will not be possible to have negative earnings at the end.

Player 1 and Player 2

Players 1 and 2 will simultaneously each be asked to make two decisions: to choose a number and to make a guess about which number the other player chooses. Both decisions have to be entered into a decision screen that is described in more detail below. Neither player will observe the decisions of the other player.

Choosing a Number

Both player 1 and player 2 have to choose a number. This number can be any number between and including 110 and 170 (fractions or decimals not allowed).

The earnings in tokens of either player 1 or player 2 from choosing a number are determined as follows. A player receives the lower of the two numbers chosen by player 1 and player 2 minus 85% of their own number.

This has the following implications:

- Assume players 1 and 2 chose the same number. Then a player will receive his/her own number (since both numbers are equal, this is also the lowest number) minus 85% of his/her own number.

- Assume that players 1 and 2 chose different numbers. Then, the player who chose the lower number could have increased his earnings by choosing a slightly higher number. However, the player who chose the higher number could have increased his earnings by choosing a slightly lower number.

The following mathematical representation will not be read out loud.

Suppose (among players 1 and 2) that one of them chooses the number Y and the other chooses the number Z.

If $Y = Z$ then the player who chose Y receives $Y - 0.85 \times Y$.

If $Y < Z$ then the player who chose Y receives $Y - 0.85 \times Y$.

If $Y > Z$ then player who chose Y receives $Z - 0.85 \times Y$.

In addition, players 1 and 2 first receive a fixed amount of 35 tokens.

Guessing the other's choice

In addition to specifying a number, both player 1 and player 2 are asked to make a guess about the number chosen by the other player. The guess is made by specifying a range (given by its lower bound L and its upper bound U) in which the other player's choice is believed to belong.

The earnings in tokens of either player 1 or player 2 from making this guess are determined as follows. A wrong guess (the actual number chosen by the other player falls outside the specified range) yields nothing. A correct guess (the actual number chosen by the other player lies within the specified range) yields 15% of the difference between 60 and the width of the range $U - L$. Therefore the smaller the specified range, the higher the earnings if the guess is correct. However, a smaller range also increases the risk that the guess is not correct, in which case no tokens are earned.

(The following mathematical representation will not be read out loud:

If the number Z chosen by the other player lies in the range (it is greater than or equal to L and less than or equal to U) then the player who has chosen L and U gets $0.15 \times (60 - (U - L))$ tokens if this number Z does not lie within the range then the player who has chosen L and U gets nothing.)

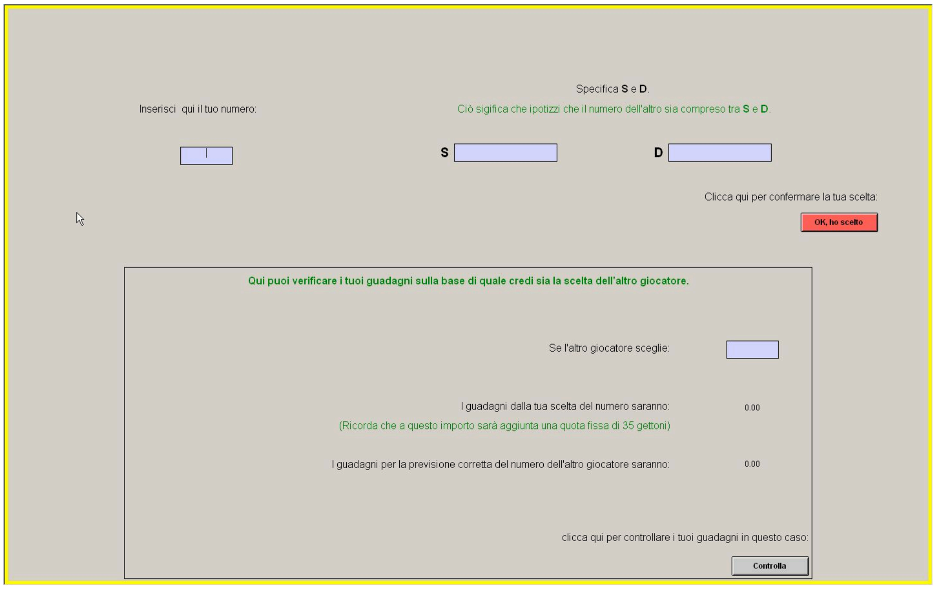

Figure 1: Input screen in the first round.

Player 3

Player 3 does not make any decision during the experiment and earns an amount of tokens equal to 25% of the smaller of the two numbers chosen by players 1 and 2.

A more mathematical representation of this statement will not be read out loud:

Tokens earned by player three = $0.25\times$ (smaller of the two numbers chosen by player 1 and player 2)

Tutorial

Before the experiment starts, so before roles are assigned, all participants have the possibility to practice and to get used to the structure of the game. To this end, you will participate in a tutorial round, where you will see the decision screen as described above. You will have 5 minutes to enter as many different values as you like for both your own number and your guess, and the other player's hypothetical number. You can then use the check button to see what your earnings from these numbers and your guess would be. You are encouraged to verify the calculation behind the earnings of both the number choice and the guess. The values entered in this tutorial have no influence on your earnings and will not be recorded. After 5 minutes the tutorial will stop and the experiment will start.

Final Remarks

During the experiment, you are not permitted to speak or communicate with the other participants. If you have a question while the experiment is going on, please raise your hand and one of the experimenters will come and answer it.

At this time, do you have any questions about the instructions or procedures? If you have a question, please raise your hands and one of the experimenters will come to your seat to answer it.

Instructions for the second round

Introduction

Now we run a second and final experiment. Earnings will be added to your previous earnings. After this new experiment everything is over and your total payment will be calculated.

This new experiment is very similar to the previous one up to some changes we highlight.

Matching and roles

All participants are matched with the same people as before and keep the roles they had before.

Decisions and Earnings

IN CONTRAST to the previous experiment, player 3 now also makes a decision.

Player 3

At the start of the experiment, before player 1 and 2 make any decisions, player 3 observes the numbers chosen by players 1 and 2 in the previous experiment. After having observed these numbers, player 3 makes a decision that determines how earnings of players 1 and 2 are calculated in this new experiment. The outcome of this decision is observed by players 1 and 2 before they make their choices. Player 3 has the following two choices:

a) NOT CHANGE: To choose "not change" means that the earnings of all players are as in the previous experiment. In particular, player 3 earns 25% of the smaller of the two numbers chosen by players 1 and 2.

b) CHANGE: To choose "change" means that earnings in tokens of all players are changed as follows. Players 1 and 2 receive the lower of the two numbers chosen minus 85% of their own number minus 50% of the difference between 170 and the player's own chosen number. That is, relative to the previous experiment, there is an extra amount subtracted to your earnings that is larger the smaller your number is. Player 3 earns 25% of the smaller of the two numbers chosen by players 1 and 2 minus 4. The terms that are new as compared to the previous experiment have been underlined.

This mathematical illustration will not be read out loud:

Suppose player 3 chooses "change" and (among players 1 and 2) that one of them chooses the number Y and the other chooses the number Z.

If $Y = Z$ then the player who chose Y receives $Y - 0.85 \times Y - 0.5 \times (170 - Y)$.

and player 3 receives $0.25 \times Y - 4$.

If $Y < Z$ then the player who chose Y receives $Y - 0.85 \times Y - 0.5 \times (170 - Y)$.

and player 3 receives $0.25 \times Y - 4$.

If $Y > Z$ then player who chose Y gets $Z - 0.85 \times Y - 0.5 \times (170 - Y)$

and player 3 receives $0.25 \times Z - 4$.

Regardless of the choice of player 3, player 1 and 2 also receive a fixed amount of 35 tokens.

Player 1 and Player 2

As in the previous experiment, players 1 and 2 make two decisions: choose a number and make a guess by specifying a range. Earnings from making the guess are as in the previous experiment, earnings from choosing a number are specified above.

Input Screen here

Final Remarks

If you have any questions then please ask them now. Please do not log off the computer when the experiment is over.