Corruption and externalities:
Assessing the role of intentions

Carlos Maximiliano Senci

Instituto de Investigaciones Económicas y Sociales del Sur (IIESS). Universidad Nacional del Sur. CONICET
maximiliano.senci@uns.edu.ar

Abstract

Experimental studies model corruption as reciprocal behavior that generates negative externalities for third parties. Results on how negative externalities factor in individuals' corruption-related decisions have been mixed. While (imputed) intentions have been proven to have an impact on participants' social preferences and kindness evaluations in a wide array of situations, little is known about whether and how the (imputed) intentions of third parties may enter participants' deliberation in corruption contexts. The assumed differential evaluation of externalities according to whether they are borne by active players or dummy players has important implications for real-life corruption. In many instances of corruption negative externalities are indeed incurred by (sets of) 'dummy players', such as 'society at large'. In other cases, however, more active players (such as the political head of a governmental department) are the target of negative externalities. Reciprocity models would predict that kindness evaluations concerning third parties would enter the considerations of partners to the corrupt transaction in the latter but not the former case. This important prediction, however, has not been empirically tested. We find that intentions of third parties do not affect behavior, but affect how kind individuals are perceived by others.

Keywords: corruption; externalities; reciprocity; intentions; experiment

Resumen. *Corrupción y externalidades: evaluar el papel de las intenciones*

Los estudios experimentales modelan la corrupción como una forma de comportamiento establecido sobre la base de relaciones de reciprocidad que genera externalidades negativas para terceros. Los resultados acerca de cómo las externalidades negativas influyen en las decisiones relacionadas con la corrupción de los individuos no son concluyentes. Si bien se ha demostrado que las intenciones (imputadas) tienen un impacto en las preferencias sociales y las evaluaciones de amabilidad de los participantes en una amplia gama de situaciones, poco se sabe si los participantes toman en consideración y de qué manera las intenciones (imputadas) de terceros en contextos de corrupción. La supuesta diferencia al evaluar las externalidades en función de si afectan a jugadores activos o a jugadores pasivos tiene implicaciones importantes para la corrupción en la vida real. En muchos casos de corrupción, las externalidades negativas afectan a (conjuntos de) «jugadores pasivos», como «la sociedad en general». En otros casos, sin embargo, jugadores más activos (como el jefe político de un departamento gubernamental) son objeto de externalidades negativas. Los modelos de reciprocidad predicen que las evaluaciones de amabilidad respecto a los terceros que son objeto de externalidades negativas entrarían en las consideraciones de las partes en una transacción corrupta en el último caso, pero no cuando se trata de jugadores pasivos. Esta importante predicción, sin embargo, no ha sido empíricamente probada. Encontramos que las intenciones de terceros no afectan al comportamiento directamente sino que afectan a la forma en que se percibe la amabilidad de los demás.

Palabras clave: corrupción; externalidades; reciprocidad; intenciones; experimento

Summary

1. Introduction

There is a vast literature in sociology and economics that addresses socially beneficial aspects of cooperation (Ostrom et al., 1994; Kollock, 1998; Fehr and Gächter, 2000). However, much less has been explored about how cooperation unfolds when it entails negative externalities to third parties. In a broad array of situations, people benefit at the expense of others (market transactions, sports, access to grants, etc.). In those contexts, however, harm to others may be more or less tolerable, depending on the prevailing norms.

Corruption affords an example in which cooperation between corrupt parties generally harms unidentified and passive others (see Abbink et al., 2002; also Weisel and Shalvi, 2015). In many situations those harmed by corrupt people have previously deposited their trust in them, that is, they have placed themselves, or their time, resources and confidence at the expense of others.

These kinds of situations are recurrent in the political arena. Union workers may vote for a delegate who instead of supporting a strike to protect the interests of fellow worker members may unilaterally collude with the employer and back up his/her interests, thus undermining support for the strike. Similar examples may involve governments who decide to spend money on procurement with officials and businesspeople having the opportunity to collude. Or it might be voters in a referendum who grant more power to the executive at the risk of subsequent collusion with lobby groups.

These types of situations exemplify the well-known definition of corruption as the abuse of entrusted power for private gain (Transparency International, 2012). In order to capture the features of these situations, we propose a three-player bribery game in which two players can act corruptly to increase their payoffs to the detriment of a third player. The game is sequential and has the following features: the first player may choose to trust, thereby increasing the payoffs of two other players at the risk of collusion between them. These two players enter in a coordination game after the first player's trusting move, and may decide individually either to reciprocate the first player's trusting move or collude to share the surplus while imposing a cost on the first mover. We compare this situation with a similar one in which the first player does not initiate the game, and therefore cannot generate any surplus, while she can still incur in the costs associated to coordination by the other two players.

A well-established result in the literature on corruption (see Abbink et al., 2002; Abbink, 2004; Barr and Serra 2009, Lambsdorff and Frank, 2010; Jacquemet, 2012), though not uncontested (see Gneezy et al., 2018), holds that reciprocity helps sustain collaboration in such collusive bribery settings. Consistent with intentions-based models, in corruption contexts reciprocity is found to be a function of the bribe amount passed by one party: higher bribes offered imply higher reciprocity levels. The supposed mechanism behind this is that people judge higher transfers relative to lower ones as kinder (Falk and Fischbacher, 2006). Therefore, higher transfers spur higher levels of reciprocal behavior.

In line with this mechanism, most models of reciprocity predict that people would treat those who were kinder to them in the first place more kindly (Rabin, 1993; Levine, 1998; Dufwenberg and Kirchsteiger, 2000; Falk and Fischbacher, 2006). Moreover, this implies that people are in general expected to be kinder to 'active players' (i.e., individuals whose actions directly affect one's well-being) than to dummy players (i.e., individuals whose actions are immaterial for one's well-being and who therefore cannot behave 'kindly'; for a discussion of this topic see López-Perez, 2008). In experimental corruption games, negative externalities are generally imposed upon dummy players. For example, costs associated with corrupt collaboration are sometimes implemented as a drastic decrease in the experimenters' donation to an ONG (see Lambsdorff and Frank, 2010).

Depending on whether they are borne by active players or dummy players, the assumed differential evaluation of externalities has important implications

for real-life corruption. In many instances of corruption negative externalities are indeed incurred by (sets of) 'dummy players', such as 'society at large.' In other cases, however, more active players (such as union workers voting for a delegate or citizens granting more power to the executive branch of government in a referendum) are the target of negative externalities. Reciprocity models would predict that kindness evaluations concerning third parties would enter the considerations of the partners to the corrupt transaction in the latter but not the former case. This important prediction, however, has not been empirically tested. This paper tests the prediction that corrupt collaboration may be thwarted by reciprocity concerns towards the third party bearing the externalities, and that these reciprocity concerns are affected by whether the third party is an active player or a dummy. We contend that individuals' sensitivity to externalities may be mediated by the third party's perceived intention. If participants have social preferences and give due weight to externalities, they will refrain from collaboration with the corrupt partner. The weight they assign to these externalities is hypothesized to depend on whether or not the third party is an active player, and if so, on her perceived kindness.

To empirically study whether intentions toward active and dummy third parties could hinder corruption, we exposed participants to a three player game in which the third party potentially suffering from the externalities caused by corruption is either a dummy player (and therefore can signal no intentions) or an active player (capable of signaling intentions). This allows us to investigate participants' choices in our "bribery game", and to assess whether the third party's perceived intentions have an impact on reciprocal behavior between the briber and the player being bribed.

The remainder of the paper is organized as follows. In the next section we discuss the related experimental literature touching upon corruption in particular. Our experimental design is presented in section 2, followed by the behavioral predictions in section 3. In section 4 we present our data and results and discuss them. Section 5 concludes.

2. The study

The objective of the study was to assess whether the effect of externalities on a third party could prevent the establishment of corrupt relationships. The third party could have either a passive or an active role. As a consequence, to evaluate the differential impact of imputed intentions, beliefs about other players' kindness were measured.

2.1. Experimental design

2.1.1. No_trust treatment

Our experiment consists of two treatments: a No_trust treatment (NT henceforth) and a Trust treatment (TT henceforth). In the NT, Role A player is a passive player devoid of any strategic role, and therefore cannot actively

influence other players' decisions. However, A is an actually participant present in the room. Role B player decides between IN or OUT. If she chooses OUT, the game ends and each player payoffs' are (20, 20, 20). If she chooses IN, she passes the play to C, who in turn has to choose between IN or OUT. If C selects OUT, then each player get 20 points, while if she selects IN, then player A gets 10, and the other players get 25 each (see Figure 1).

2.1.2. Trust treatment

Our TT differs from the NT in that Player A is an active player moving first in the sequence (see Figure 2). Player A has the option to opt out of the game (OUT), assuring for himself a payoff of 15 EMUs,[1] or passing-the-play (IN) to player B, who could either a) share the surplus generated by A's action choosing OUT, or b) pass the play to role C player. In the latter case, payoffs depend ultimately on C's action. If C chooses OUT, she thereby rejects the corrupt transaction proposed by B, and this action generates the most equal outcome across all three players (20, 20, 20). Note that besides being the most equal, this outcome is a Pareto improvement compared to the sub-game perfect equilibrium. If B and C 'cooperate in corruption' (i.e., both choose IN), A is always worse-off, getting her minimum payoff in the game. This represents a negative externality for player A.

  Note that in our set-up each individual has a 'fair option' in her strategy space. Therefore, there is a straightforward way to evaluate the kindness of any action by comparison with the available fair option, i.e., B and C players always have an available fair action (which is to choose OUT). For player A, option OUT is always better than IN, and therefore participants in role A would presumably judge the OUT choices of B and C as kinder. B and C, unlike A, would probably judge the IN choices rather than the OUT choices made by participants in other roles as kinder, since IN choices give them a higher payoff.

2.1.3. Procedure

We conducted the experiment at the Sociological Laboratory of the Department of Sociology at the University of Groningen in the Netherlands during December 2012 and early 2013. Subjects were invited from a database of approximately 1200 registered volunteer students from different fields using the ORSEE recruitment software (Greiner, 2015). The recruitment and the experiment complied with the ethical guidelines set out by the Sociological Laboratory (http://www.gmw.rug.nl/~orsee/public/privacy.php). In total, 10 groups of 3 participated in the NT and 9 groups in the TT. We ran three sessions with 15 participants and one session with 12. Subjects participated in only one of the treatments. The experiment was programmed and conducted in z-Tree (Fischbacher, 2007). The average earnings were €10.66 per person

---

1.   Payoffs throughout the paper are expressed in EMUs (experimental monetary units) which were
     exchangeable for euros at the end of the session at the exchange rate of 1 euro for 50 EMUs.

in the TT (€6.66 plus a €4 show up fee) and €12 in the NT (€8 plus a €4 show up fee). Each session lasted about 45 minutes.

Upon arrival at the lab, participants were seated at different computer terminals. They were told that they would participate in a decision-making experiment and that communication would not be allowed during the study. Participants were further told that their earnings in the experiment would depend on their own choices and the choices of others. It was made clear that their choices would remain anonymous and that final cash payout at the end of the study would depend on how many points they had earned during the experiment. Before starting the experiment, the subjects were made aware that the experiment did not involve deception of any form.

Following an introduction by an assistant, participants were provided the instructions, which were completely programmed in z-Tree (see Appendix II for the TT and NT instructions). At the beginning of each round, the participants were randomly divided into groups of three (3) people. In each round, after the groups were formed, the computer randomly assigned an ID to each participant in a group. There were 3 IDs: Role A, Role B and Role C. These IDs were also reassigned in each round, so the same participant interacted in different roles (A, B or C) for 20 periods. Groups were randomly re-matched after every round, meaning that the participants were re-allocated to a different group after each round (random rematching). The participants also knew (i.e., it was common knowledge) that each participant could be assigned to a different (or the same) role in each round and that she would be interacting within a newly formed triplet. This feature of the design was chosen in order to prevent participants' decisions from being affected by the anticipation of possible future monetary gains.[2]For our specific research purposes, however, not allowing for reputation formation to influence trusting behavior permits us to avoid possible confounding explanations. Participants were fully informed about all details of the procedure. In both treatments, participants were informed of all the decisions made in their group and the points earned by each member at the end of each period. We deliberately avoided using in-context terms in the instructions. Following standard practice in experimental economics, we used neutral language that avoided any reference to trust, cooperation and/or externalities.

Attributions of kindness were measured in each round by asking participants to rate the kindness of the other participants in her group (How unkind/kind was Participant X to you?) on a seven-point Likert scale ranging from

---

2. For the sake of clarity, the design does not preclude the possibility that the same participants interact repeatedly over the rounds. The current design also allows for a less direct form of reputation building, or perhaps norm initiation. For instance, an A participant may choose IN in round 1 in the hopes that their B and C participants will be more likely to follow suit when they themselves play role A later on, indirectly benefitting the current A player in later rounds when she plays role B or C if she sees more play of IN by the new A player. However, this is an empirical matter and we found no evidence for this hypothesis. Quite on the contrary, players A became more distrustful over the rounds (see results section).

1 = *very unkind* to 7 = *very kind*. This method is similar to the method used by Falk and Fischbacher (2006). Earnings were computed by summing up the points participants earned in each round at the end of the session, which were converted into cash at an exchange rate of €1 for 50 points. After completing the task, the participants were debriefed and their earnings as well as their show-up fees were placed in envelopes, each of which was picked up privately to ensure they did not know what others had earned.

3. Behavioral predictions

The main purpose of the manipulation was to test whether differences between IN choices are the result of reciprocity concerns. The experimental set-up allows us to investigate A's status (as an active or passive player) by comparing the frequency of IN actions undertaken by role B and C players in the NT *vis-a-vis* the TT. In addition, by comparing kindness perceptions we can assess whether differences in choices might be the result of reciprocity-driven preferences. For instance, a higher level of OUT choices of participants in role B might be the result of positive reciprocity toward A.

In each round, player A finds it in his/her interest to go OUT rather than passing the play to the other players, precisely because he/she knows that it will not be in the interest of the other players to reciprocate him/her by choosing OUT. In a model of selfish utility maximization, both B and C will choose to play IN, and thus corrupt transactions will always be struck between B and C to the detriment of the third party A. Both B and C are then better off if they cooperate with each other. But cooperation is clearly not a Pareto improvement compared to when A chooses OUT, as A is worse off, even though B and C are better off. This feature of our set-up provides us with a backward induction argument to derive the unique sub-game perfect equilibrium for the TT. Paradoxical as it may seem, assuming that individuals are guided by their self-interest implies a prediction of zero-level corruption: the unique sub-game perfect equilibrium involves C and B playing IN and A playing OUT. The conclusion of this backward induction argument is that A should go OUT, which immediately brings the game to an end.

This game-theoretical solution rests upon the assumption that individuals are solely motivated by their own payoffs and expect others to be so motivated as well. For this reason, it does not matter whether player A is an active or a passive player, or whether she can signal fair intentions. A money-maximizing B will always try to strike a corrupt deal by passing the play to C, while C will always choose the option that most favors both B and C. From the perspective of player A, the TT resembles a trust game. By trusting player B (and perhaps player C in the event B proves untrustworthy) through choosing IN, player A can potentially earn a higher payoff than she earns when choosing OUT. From experimental studies employing the trust game we know that despite the sub-game perfect equilibrium prediction of no trust placed, many first movers (players A in our game), choose to trust (IN in our game) (Berg et al., 1995;

McCabe et al., 2003).Let us call this phenomenon the *Trust argument*. Therefore, we expect a non-negligible level of INchoices by role A players in the TT of our experiment. Hence, we expect a sufficient number of moves by player B in the TT to be able to compare the TT to the NT in terms of B's behavior.

The *Trust argument* provides us with a different argument, since we expect individuals to consider those who pass the play to be *kinder* than those who do not. More precisely, player A's trusting behavior should be compensated since it enhances the surplus at a potential own cost. The presence of a Role A player as an active player should therefore impair, to some degree, the striking of corrupt deals between B and C. Therefore, conditional on the Role A player having chosen IN (following kind behavior by A), we expect a lessening of corruption (of IN choices) in the TT, where the presence of a benevolent player A who can signal fair intentions should enhance the awareness of externalities, compared to the NT. We expect this effect to apply to both role B and role C players, and so we derive the following hypothesis.

Hypothesis 1: Role B and Role C players are less likely to choose IN in the TT than in the NT.

However, since (when given the move) final payoffs hinge ultimately on role C's action, we hypothesize that this feature of the decision-making process could induce an enhanced feeling of responsibility on the part of C (reversely, a diminished feeling of responsibility on the part of B). Recent experimental literature shows that morals are malleable, and specific features of the environment, such as being pivotal for the outcomes, have a bearing on individuals' choices (Falk and Szech, 2016). Pivotality also relates to the effect of delegation. In a series of papers, for instance, it has been found that in Dictator Games, the delegation of an unfair decision helps to reduce the delegator's responsibility (see Coffman, 2011; Bartling and Fischbacher, 2012). Bartling and Fischbacher (2012) found that dictators were less likely to be punished when they decided an unfair allocation through delegation rather than directly. According to these authors, delegation allows blame to be shifted, inducing a shifting of the *locus* of responsibility. This insight was also applied to the study of intermediaries in a bribery experiment (Drugov et al., 2014). Allegedly, by shifting the blame, delegation to an intermediary could reduce moral costs associated to bribery. Although in our game we do not include a measure of responsibility attribution, following the idea of this study, we can expect that players in role B will be more likely to choose IN than role C players since responsibility for final payoffs rests ultimately on C's shoulders and B can interpret her own decision as a delegation to C. We thus derive the following hypothesis:

Hypothesis 2: Role C players are less likely to choose IN than role B players.

Based on the above arguments regarding C's responsibility for final payoffs, and considering the fact that in the TT, C's pivotality might be more salient than it is in the NT, we also expect that:

Hypothesis 3: Role C players are more likely to choose IN in the NT than in the TT.

Considering kindness attributions as stated by participants, we propose the following general hypothesis. Following the trust argument, the effect of A's behavior on B's behavior in the TT is mediated by B's kindness evaluation of A's action. Therefore, we pose the next hypothesis.

Hypothesis 4: In general we expect players to be more likely to choose OUT the kinder they find others.

4. Data and variables

The experiment comprised 4 sessions, with 5 triplets in three sessions, and 4 triplets in the remaining one. Therefore, we obtained observations on 57 subjects: 30 participants in the NT and 27 participants in the TT. Since each participant played 20 rounds alternating randomly between different roles, we had a total of 57 x 20 = 1140 observations. Note that in the TT, B could only make a decision if A chose IN. Similarly, C could only make a decision if both A and B chose IN. Thus, not every participant got the turn in each round, so the number of actual decisions amounts to 760.

In order to test the hypotheses, we estimate B's and C's probabilities of choosing IN depending on the experimental treatment and the attributions of kindness. The dichotomous variable Action takes the value of 1 if the actor chooses IN and 0 otherwise. The variable Round was defined by centering the original rank numbers of the 20 rounds as Round = Round rank number – 10.5. So, the variable round ranges from -9.5 to 9.5. Intercepts can then be interpreted as the average number of units of IN decisions in the 'average' round.

4.1. Results

Since we have a repeated measures design with a dichotomous dependent variable (i.e., the decisions to play IN or OUT), we used a multilevel logistic regression model to analyze the data. Multilevel modeling takes into account the hierarchical data structure (Snijders and Bosker, 2012) and allows the handling of data when observations are not stochastically independent. The experiment included 57 participants who each played 20 one-shot interactions, so that a total of 1140 observations (Level 1) were nested within 57 participants (Level 2). The experimental conditions are thus level-2 variables.

We use the Markov Chain Monte Carlo (MCMC) estimation, since this procedure yields less biased estimates than standard approximate estimation techniques. In addition, MCMC estimation yields the deviance statistic, which allows us to use approximate likelihood ratio tests to compare nested statistical models. The difference in the deviance statistics of two nested models approximately follows a chi-square distribution with degrees of freedom equal to

**Figure 1.** Illustration of the game in extensive form of the TT
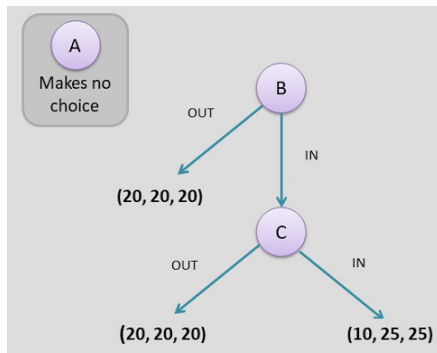
Source: Own elaboration.

the difference in the number of parameters of the models compared (see Hox, 1995). Significant values indicate that the model with the lower deviance is an improvement over the model with the higher deviance.
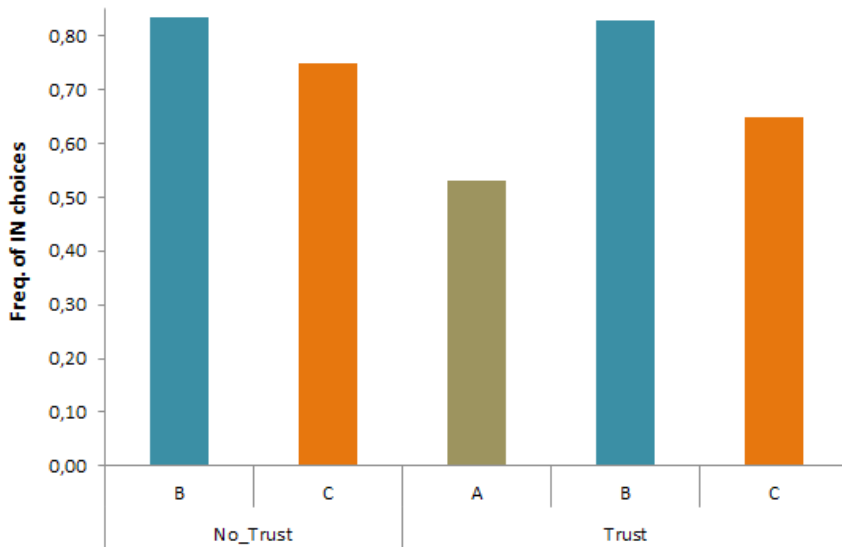
4.2. Descriptive statistics

In the TT, role A players chose IN 53.3% of the time (see Figure 3). When given the turn, Role B players chose IN 82.8% of the time, while role C players made the INchoice 65.4 % of the time.

Figure 2. Illustration of the game in extensive form of the NT
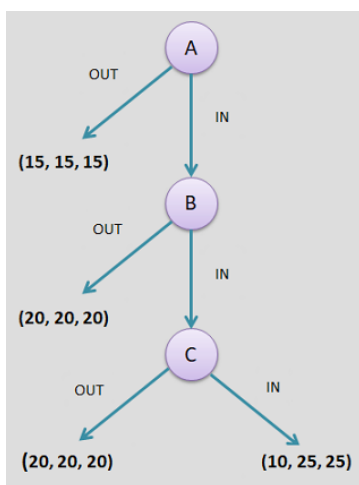
Source: Own elaboration.

Figure 3. Frequency of IN choices plotted against Role and Treatment

In the NT, the percentage of role B players choosing IN was 83.5%, while role C players chose IN 75.3% of the time. Note that the difference in behavior between B and C is larger in the TT than in the NT.

As expected, the number of IN choices in the TT was far from negligible. Indeed, corrupt deals are struck at high levels, regardless of treatments. Moreover, the descriptive statistics go in the direction predicted by hypotheses1 and 4, since role B and C players chose IN fewer times in the TT condition.

Consistent with the descriptive statistics from graphical inspection of Figure 3 we can infer that the frequency of C players choosing IN is lower than that of role B players. Although the figure might suggest otherwise, differences between treatments turned out to be non-significant for role C players.

4.3. Statistical analysis

Multilevel models were fitted using MLwiN software (Rasbash et al., 2009). Following standard practice empty models were fitted first, yielding measures of clustering in the data. Then we fitted random intercept models with fixed effects for the predictor variables. A fixed effect for Round-center was included in all models. As outcomes were binary, we used the logit function.

We use the parameter estimates to evaluate the significance of individual coefficients. In this way, we can test the null hypothesis for a particular estimated parameter using the Wald criterion, which states that an estimate is sig-

nificantly different from zero at α = 0.05 if it exceeds 1.96 times its associated standard error (e.g., Hox, 1995). In other words, if the ratio of an estimated parameter over its standard error exceeds 1.96, then the H0 regarding that parameter is rejected. The parameter should then be included in an adequate model of the data.

4.4. Regression results

The level of corruption was measured taking as our dependent variable IN choices made by B and C. We have allowed the probability of choosing IN to vary across participants (random intercept), but we have assumed that the effects of the explanatory variables are the same for each participant.

The degree of clustering in the data can be summarized by the intra-class correlation coefficient (ICC) and the variance partition coefficient (VPC). The intra-class correlation is a measure of the dependency in the data at the same individual and can have a value in the interval [0, 1]. Normally the formula to calculate the ICC is ICC = Between Group Variance/(Between Group Variance + Within Group Variance) ().In this case the decision-level variance is standardized to the logistic variance of $2/3 \approx 3.29$. In our study our between group variance is 4.2 (See Table 1). Thus ICC = 4.2/4.2 + 3.29 ≈ 0.56. In our model, since we have a binomial distribution, the ICC and VPC are identical and measure the within-subject expected correlation between observations. To put it another way, the ICC measures the proportion of total variance in the outcome variable that is explained by differences between individuals. Interpreted as VPC it represents the proportion of the variation in the propensity to choose IN that lies between subjects. The ICC is .56, which means that the expected correlation between two decisions by the same subject is .56. The value .56 signifies that propensity to choose IN varied greatly between individuals. Interpreted as VPC we would say that 56% of the variation in the probability to choose IN lies between subjects. This indicates that 56% of chances of choosing IN is explained by between subjects differences, which validates our approach using multilevel modeling (for more details on how to calculate the ICC see Leckie and Charlton, 2012; Sommet and Morselli, 2017).

4.4.1. Treatment effects

The core results of the experiment are summarized in Figure 3. To test hypotheses 1 and 3, we inspect the coefficients for variable treatment in Model 3 (see Table 1), in which we added a level-2 variable for treatment (NT as reference category). According to hypotheses 1 and 3, we would have expected fewer IN choices by role B and C players in the TT than in the NT. In line with our core hypotheses, the descriptive statistics suggested that IN choices were slightly higher in the NT. However, adding the treatment variable did not improve the model fit (^{2} = 0.29, (SE) = -0.373 (0.678), p > 0.59). Note that this result contradicts hypotheses 1 and 3. For role B and role C players there is no significant statistical treatment difference in the propensity to choose

Table 1. MCMC estimates of random intercept multilevel logistic regression coefficients for the interactions of treatment with a dummy variable for Role C player that takes the value of 1 when the participant is assigned role C and 0 otherwise. We model the probability for the players to choose IN (IN = 1; OUT = 0). Reference category is role B player.

	Empty model Coeff. (SE)	Model 1 Coeff. (SE)	Model 2 Coeff. (SE)	Model 3 Coeff. (SE)	Final model Coeff. (SE)
Intercept	2.00 (0.337)*	1.694 (0.401)*	2.149 (0.467)*	2.331 (0.597)*	2.107 (0.531)*
Round		0.037 (0.022)	0.044 (0.023)	0.043 (0.024)	0.046 (0.024)
Role C			-0.918(0.28)*	-0.920 (0.288)*	-0.709 (0.339)*
Treatment				-0.373 (0.678)	0.035 (0.747)
Treat * Role C					-0.625 (0.619)
Random Part					
Level-2 Variance	4.200 (1.443)				4.783 (1.749)
ICC	0.56				0.60
Deviance	381.47	379.13	**368.51**	368.22	368.16
N. of obs.	545	545	545	545	545

*Absolute value of (Coefficient/S.E.) larger than 1.96.

Source: Own elaboration.

IN. This means that we must reject a main hypothesized treatment difference, since players' behavior was not affected by the presence of an active third party compared to a passive one.

Hypothesis 2 states that Role C players are less likely to choose IN (than role B players) due to their role pivotality. It turned out that, in line with hypothesis 2, Role C players are significantly less likely than role B players to choose IN, *regardless of treatment*. Model 2 supports this result (see Table 1). To test whether C players were more likely to choose IN in the NT, we compared the empty model with no predictors to the model with a dummy variable for role C (taking B as a reference category). The model with Role C significantly improved the fit ($\chi^2 = 10.62$, $\beta(SE) = 0.918(0.28)$, $p < 0.01$). It seems that being the last player to get the turn had an effect on choices. Presumably, as we hypothesized, the fact that the determination of final payoffs ultimately rests on C's shoulders enhances C's sense of responsibility and heightens the awareness of the externalities imposed upon A. Note also that the coefficient for the interaction term *Treat x Role C* is -0.625 (SE = 0.619), which is not significant. This means that role C players' behavior, contrary to hypothesis 4, did not differ between treatments.

In a second analysis, we extended the analyses to incorporate the kindness attributions measured by the responses to the kindness scale.

4.4.2. Analysis of kindness attributions

In previous analyses we investigated role B and C behavior as a function of their role and treatment variations. In this section we report similar models incorporating the effect of kindness attributions, that is, how role B and C players perceived A's kindness as stated in their responses to the kindness scale.

Table 2. MCMC estimates of random intercept multilevel logistic regression coefficients of roles * kindness for A. The variable Kind A represents role C players' evaluation of the kindness of role A subjects on a seven-point Likert-type scale. The action IN is our dependent variable.

	Model 0 Coefficient (SE)	Model 1 Coefficient (SE)	Model 2 Coefficient (SE)	Model 3 Coefficient (SE)	Model 4 Coefficient (SE)	Model 5 Coefficient (SE)
Intercept	1.930 (0.584)	1.914 (0.546)	2.724 (0.735)	-2.703 (2.28)	-2.653 (2.518)	-3.515 (3.574)
Round Center		0.011 (0.040)	0.009 (0.043)	-0.023 (0.047)	-0.016 (0.046)	-0.021 (0.048)
Role C			-1.375 (0.537)*	-1.735 (0.581)*	0.110 (3.448)	
Kind A				0.898 (0.364)*	0.901 (0.413)*	0.761 (0.547)
Role C * Kind A					-0.294 (0.556)	
Role B						0.413 (3.747)
Role B * Kind A						0.217 (0.598)
Deviance	131.117	131.998	**124.545**	**120.165**	120.470	120.783
Number of obs.	178	178	178	178	178	178

* Absolute value of (Coefficient/SE) larger than 2.

Source: Own elaboration.

As a manipulation check, we first investigated whether kindness attributions were actually explained by other players' actions.

Indeed, participants' actions largely explain kindness attributions. Tables A.1, A.2 and A.3 in Appendix A show the participants' kindness attributions. In the case of B and C, Model 2 in Table A.1 and model 2 in Table A.2 show that the action of A largely improved the model fit ($p < 0.001$ in both cases). The same holds true for the actions of B. A likelihood ratio test comparing model 1 and 2 showed that including the variable for the kindness of B (as perceived by C) significantly increased the model fit ($\chi^2 = 82.93$, $p < 0.001$) (see Table A.3 in Appendix A). Thus, behavior affected kindness attributions, and IN choices were evaluated as "kinder" than OUT choices. The lesson we draw from these results is that the "kindness attribution" component of the intention-based argument of reciprocity seems to hold water.

Naturally, we were mainly interested in the other component of the intention-based argument, that is, how perceptions of kindness influenced behavior in our experiment. We found that participants' decisions were correlated with how they perceived others' actions in terms of their kindness. For the TT, we can analyze the extent to which A's perceived kindness affected the behavior of role B and C players. Table 2 shows the degree to which B's IN choices were influenced by A's perceived kindness. We report random intercept models to analyze the effect of the perceived kindness of A on B's action. In contrast to hypothesis 4, we found no evidence that *the kinder A is perceived to be, the less likely B and C are to choose IN*. Note that the coefficient for the interaction of kindness of A and role B is 0.217 (SE = 0.59), which turned out to be non-significant (in Table 2). The coefficient for the interaction of kindness of A and role C is -0.29 (SE = 0.55) and also turned out to be non-significant.

We showed that C's actions were not influenced by how they perceived others. Based on the intention-based argument of reciprocity theories, what we expected (hypothesis 4) is that the kindness of A, by choosing IN, would lower the likelihood of C choosing IN. Although the descriptive statistics suggested this was the case, as role C players chose IN fewer times in the TT, the analysis proved that this difference is not statistically significant. Most of the time in the TT, when role A players chose IN, so did role B players. A possible explanation of why we did not find an effect may be due to the fact that C's kindness judgments toward A and B may offset each other, which might make it difficult to gauge their influence relative to one another. Role C players may be torn between deciding in favor of A or in favor of B. Although we did not find a difference in B and C's kindness evaluation of IN choices, we did find that *perceived kindness of A is evaluated as less kind with round number*. In our analyses, we took the kindness attributions each player made after being informed of the actions of each of the other participants as dependent variables (see Tables A.1, A.2 and A.3 in the Appendix A). We estimated a normal response model using iterative generalized least squares models. In the first step, we took as a dependent variable how kind or unkind B and C found A in the TT. The model for the kindness attribution made by B of A's actions represents an improvement with respect to the empty model (χ^2 = 5.21, p = 0.02; see Model 1 in Table A.1). Note that the coefficient parameter for round 0.062 (SE = 0.027) has a negative sign. Therefore, *A is found to be progressively less kind by role B players*. The same is true for Role C players (χ^2 = 6.20, p = 0.02, see Model 1 in Table A.2).

Since subjects play the game repeatedly, the question remains whether they learn throughout the experiment. Note that levels of A's kindness worsened with round number. This effect is brought about not only by role C players' attributions, but also those of B players. It may be the case that with round number, cooperation between players B and C made A less optimistic about trusting. Consequently, fewer A players chose to pass the play to B, which in turn lowered A's score on the kindness scale.

In a nutshell, the results confirm that there is no main effect of our manipulation on the proportion of IN choices. However, C players chose to play IN fewer times. This suggests that being in charge of the final output (their role being pivotal) may enhance peoples' sense of responsibility, and therefore induce them to refrain from behaving selfishly. In addition, the inclusion of a measurement of participants' kindness attributions showed that their responses in the different experimental conditions are partly moderated by their kindness judgments. Thus, individual behavior seems responsive to the way participants consider others' kindness.

5. Conclusion

The objective of this study was to evaluate if partners to a corrupt transaction would refrain from infringing externalities on a third party that had

previously relied on them. With this goal in mind, we considered a three-player corruption game in which players A could signal trust. In addition, we used a coarse measure of kindness (attributions) as a proxy of trust. We considered a three-player corruption game in which players A could signal fairness intentions. In addition, we measured kindness attributions using a scale. According to intention-based theories of reciprocity, we would have expected a lessening of corruption (e.g., fewer IN choices by B and C) in the Trust treatment relative to the baseline. However, our main contention, that the presence of an active player suffering from negative externalities would make players B and C refrain from corruption, thus stated, is not borne out by the data. The presence of a kind first player does not have a significant impact on the decisions of other players to cooperate or defect. This result seems to concur with the literature on experimental corruption (see Abbink et al., 2002; Barr and Serra, 2009) in that transaction partners hardly take into consideration harm done to others, but goes against the predictions of intention-based reciprocity models. Even more puzzling is the fact that second movers (players B) chose the self-maximizing option as they considered that A was kinder. The strong point of our study is that we do find that behavior affects kindness attributions, so "the first step in the mechanism" seems to work. However, these kindness attributions either do not affect behavior or affect it in an unexpected way.

Our results bring out some aspects of kindness-driven reciprocity that have not received much attention, and that require further scrutiny. For instance, passing on the play even when it invariably increases others' earnings might not be the only element triggering a reciprocal response. An action may have other contextual features that could make its appraisal difficult in terms of kindness. As our study shows, being put in a moral dilemma or in a situation that enhances one's responsibility may also be a factor modulating reciprocal responses. When they get their turn, C players bear a heavier burden than B players, since they are ultimately responsible for final payoffs. This is manifest in the fact that unlike B players' selfish actions (which could be reversed by C choosing OUT), those of C players are not reversible.

Contrary to claims made in previous studies (Abbink et al., 2002; Barr and Serra, 2009), which have stated that reciprocity is a main mechanism through which corruption is sustained, kindness attributions seem orthogonal to reciprocation in our experiment, as implied by our data. Although subjects were not indifferent to the degree of kindness of the other players, their behavior does not support a reciprocity mechanism. This result points to the need to amend models of kindness-based reciprocity with more fine-grained utility functions that might take into consideration those aspects, such as pivotality or responsibility, which arose as key drivers of reciprocity (of the lack of it) in our study. To be sure, the lack of support for a *reciprocity mechanism* between first and second mover in our experiment does not preclude that this mechanism might have a stronger impact in a repeated environment. It seems possible that corruption may be sustained by reciprocity but as a sub-product of repeated

game incentives (see, for instance, Abbink et al., 2002), while corruption may be fueled by self-interest in one-shot scenarios. Another explanation for the lack of effect of a reciprocity mechanism between A and B could be that at least a fraction of subjects could have perceived A's actions as strategically motivated, since by passing the play to B, A players' behavior may be driven by the expectancy of higher payoffs. It has been shown in trust games that reciprocity is stronger when this type of strategic considerations can be ruled out (see Stanca et al., 2009). Future studies should attempt to rule out strategic considerations in order to disentangle the effect of kindness from other potential confounders. As in the study of Stanca et al., this could be achieved by manipulating the information sets of the players.

In line with recent research (Gneezy et al., 2018), albeit in a different setting, our results also suggest that corruption is driven mainly by greed or self-interest. Gneezy et al. (2018) recently stated that it is important from a public policy perspective whether the main motivation for bribery is reciprocity or greed (in the sense of payoff-maximization) because if "greed drives bribery, policy interventions should focus on preventing bribes that are contingent on a certain outcome. On the other hand, if reciprocity motivates bribery, policy interventions should focus on making reciprocity more difficult, for example, by decreasing personal contact through anonymity and staff rotation."

From the perspective of nudging away people from corrupt practices, our results seem to suggest that policy efforts should be directed towards enhancing public officials' awareness of their pivotality. One interesting avenue to explore is the use of normative "nudges" to prime prescribed/proscribed behaviors. It has been shown that focusing people's attention on norms helps to tilt the balance towards pro-social behavior (Cialdini et al., 1990; Krupka and Weber, 2009). In the context of corruption, this may well be done by using a meaningful context (see Alekseev et al., 2017; Senci et al., 2019), by eliciting appropriate norms prior to game playing (see Krupka and Weber, 2009) or by introducing explicit norms (see, for example, Karakostas and Zizzo, 2016).

Our results also show that signaling kindness is not as straightforward as it may seem, and that other stronger motivations may play a counteracting role. Identifying the mechanisms responsible for corruption to emerge may be more intricate, because corruption could be the result of the aggregate working of multiple mechanisms. Nevertheless, the blatant contradictions between our results and what is hypothesized by social preferences theory in such a simple setting like ours is striking.

Funding

Bibliographic references

Abbink, Klaus (2004) "Staff rotation as an anti-corruption policy: an experimental study". *European Journal of Political Economy*, 20 (4), 887–906.
<https://doi.org/10.1016/j.ejpoleco.2003.10.008>

Abbink, Klaus; Irlenbusch, Bernd and Renner, Elke (2002). "An Experimental Bribery Game". *Journal of Law, Economics and Organization*, 18 (2), 428–454.
<https://doi.org/10.1093/jleo/18.2.428>

Alekseev, Aleksandr; Charness, Gary and Gneezy, Uri (2017). "Experimental methods: When and why contextual instructions are important". *Journal of Economic Behavior & Organization*, 134, 48–59.
<https://doi.org/10.1016/j.jebo.2016.12.005>

Barr, Abigail and Serra, Danila (2009). "The effects of externalities and framing on bribery in a petty corruption experiment". *Experimental Economics*, 12 (4), 488–503.
<https://doi.org/10.1007/s10683-009-9225-9>

Bartling, Björn and Fischbacher, Urs (2012). "Shifting the Blame: On Delegation and Responsibility". *Review of Economic Studies*, 79 (1), 67–87.
<https://doi.org/10.1093/restud/rdr023>

Berg, Joyce; Dickhaut, John and McCabe, Kevin (1995). "Trust, Reciprocity, and Social History". *Games and Economic Behavior*, 10 (123), 122–142.
<https://doi.org/10.1006/game.1995.1027>

Cialdini, Robert; Reno, Raymond and Kallgren, Carl (1990). "A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places". *Journal of Personality and Social Psychology* 58, 1015–1026.
<https://doi.org/10.1037/0022-514.58.6.1015>

Coffman, Lucas, (2011). "Intermediation reduces punishment and reward". *American Economic Journal: Microeconomics*, 3, 77–106.
<https://doi.org/10.1257/mic.3.4.77>

Drugov, Mikhail; Hamman, John and Serra, Danila (2014). "Intermediaries in corruption: an experiment". *Experimental Economics*, 17, 78–99.
<https://doi.org/10.1007/s10683-013-9358-8>

Dufwenberg, Martin and Kirchsteiger, Georg (2000). "Reciprocity and wage undercutting".*European Economic Review*, 44(4), 1069–1078.
<https://doi.org/10.1016/s0014-2921(99)00047-1>

Falk, Armin; Fehr, Ernst and Fischbacher, Urs (2008). "Testing Theories of Fairness—Intentions Matter". *Games and Economic Behavior*, 62, 287–303.
<https://doi.org/10.1016/j.geb.2007.06.001>

Falk, Armin and Fischbacher, Urs (2006). "A theory of reciprocity." *Games and Economic Behavior*, 54, 293–315.
<https://doi.org/10.1016/j.geb.2005.03.001>

Falk, Armin and Szech, Nora (2016). "Diffusion of Being Pivotal and Immoral Outcomes". *Human Capital and Economic Opportunity Global Working Group working paper series*, Working Paper 2016-013.

Fehr, Ernst and Gächter, Simon (2000)."Cooperation and Punishment in Public Goods Experiments". *The American Economic Review*, 90(4), 980–994.
<https://doi.org/10.1257/aer.90.4.980>

Fischbacher, Urs (2007). "z-Tree. Zurich Toolbox for Ready-made Economic Experiments". *Experimental Economics*, 10, 171–178.
<https://doi.org/10.1007/s10683-006-9159-4>

GNEEZY, Uri; SACCARDO, Silvia and VAN VELDHUIZEN, Roel (2018). "Bribery: Behavioral Drivers of Distorted Decisions". *Journal of the European Economic Association*, 17(3), 917–946.
<https://doi.org/10.1093/jeea/jvy043>

GREINER, Ben (2015). "Subject pool recruitment procedures: organizing experiments with ORSEE". *Journal of the Economic Science Association* 1: 114.
<https://doi.org/10.1007/s40881-015-0004-4>

HOX, Joop J. (1995). *Applied Multilevel Analysis*. Amsterdam: TT-Publikaties.

JACQUEMET, Nicholas (2012). "Corruption as betrayal: Experimental Evidence". Work in progress. Retrieved from http://www.nicolasjacquemet.com/Corruption.pdf. Accessed on May 10, 2017.

KARAKOSTAS, Alex and ZIZZO, Daniel John (2016). "Compliance and the power of authority". *Journal of Economic Behavior & Organization*, 124, 67–80.
<https://doi.org/10.1016/j.jebo.2015.09.016>

KOLLOCK, Peter (1998). "Social dilemmas: the anatomy of cooperation". *Annual Review of Sociology*, 24, 183–214.
<https://doi.org/10.1146/annurev.soc.24.1.183>

KRUPKA, Erin and WEBER, Roberto A. (2009). "The focusing and informational effects of norms on pro-social behavior". *Journal of Economic Psychology*, 30, 307–320.
<https://doi.org/10.1016/j.joep.2008.11.005>

LAMBSDORFF, Johann G. and FRANK, Björn (2010). "Bribing versus Gift-Giving: An Experiment". *Journal of Economic Psychology*, 31 (3), 347–357.
<https://doi.org/10.1016/j.joep.2010.01.004>

LECKIE, Georg and CHARLTON, Chris (2012). "runmlwin: A program to run the MlwinN Multilevel Modeling Software from within Stata". *Journal of Statistical Software*, 52(11), 1–40.
<https://doi.org/10.18637/jss.v052.i11>

LEVINE, David K. (1998). "Modeling Altruism and Spitefulness in Experiments". *Review of Economic Dynamics*, 1, 593–622.
<https://doi.org/10.1006/redy.1998.0023>

LÓPEZ-PÉREZ, Raúl (2008). "Aversion to norm breaking: a model". *Games and Economic Behavior*, 64 (1), 237–267.
<https://doi.org/10.1016/j.geb.2007.10.009>

MCCABE, Kevin, A.; RIGDON, Mary L. and SMITH, Vernon L. (2003). "Positive reciprocity and intentions in trust games". *Journal of Economic Behavior and Organization*, 52, 267–275.
<https://doi.org/10.1016/s0167-2681(03)00003-9>

OSTROM, Elinor; GARDNER, Roy and WALKER, James (1994). *Rules, Games, and Common-Pool Resources*. Ann Arbor: Univ. of Michigan Press.
<https://doi.org/10.3998/mpub.9739>

RABIN, Matthew (1993). "Incorporating Fairness into Game Theory and Economics". *American Economic Review,* 83, 1281–1302.

RASBASH, John; CHARLTON, C.; BROWNE, W.; HEALY, M. and CAMERON, B. (2009). "MLwiN Version 2.1." Centre for Multilevel Modelling, University of Bristol: Bristol, UK.

SENCI, Carlos M.; HASRUN, Hipólito M.; MORO, Rodrigo and FREIDIN, Esteban (2019). "The influence of prescriptive norms and negative externalities on bribery decisions in the lab". *Rationality and Society*.
<https://doi.org/10.1177/1043463119853893>

Snijders, Tom A. B. and Bosker, Roel J. (2012). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling.* London: Sage Publishers.

Sommet, Nicolas and Morselli, Davide (2017). "Keep Calm and Learn Multilevel Logistic Modeling: A Simplified Three-Step Procedure Using Stata, R, Mplus, and SPSS". *International Review of Social Psychology*, 30 (1), 203–218.
<https://doi.org/10.5334/irsp.90>

Stanca, Luca; Bruni, Luigino and Corazzini, Luca (2009). "Testing Theories of Reciprocity: Do Motivations Matter?" *Journal of Economic Behavior and Organization*, 71 (2), 233–245.
<https://doi.org/10.1016/j.jebo.2009.04.009>

Transparency International (2012). "What Is the Corruption Perceptions Index?" Available online at www.transparency.org/cpi2011/in_detail (accessed 12 August 2018).

Weisel, Ori and Shalvi, Saul (2015). "The collaborative roots of corruption". *PNAS* 112 (34), 10651–10656.
<https://doi.org/10.1073/pnas.1423035112>

Appendix A. Supplementary Results

Table A.1. MlwiN's IGLS (Iterative generalized least squares) estimation of regression coefficients for the kindness judgments made by B toward A's action

	Dependent variable: kindness of A		
	Model 0	Model 1	Model 2
Intercept	4.500 (0.191)	4.499 (0.197)	2.520 (0.149)
Round Center		-0.062 (0.027)*	0.010 (0.013)
Action A			3.689 (0.154)*
Diff. in deviance		**5.219**	**255.293**
Random Part			
Btw. subject variance	0.319 (0.273)	0.396 (0.287)	0.268 (0.114)
Level-1 variance	4.361 (0.495)	4.180 (0.475)	0.938 (0.107)
No. of obs.	180	180	180

Note: Difference in deviance in bold characters indicates that the model is an improvement over the previous model. * Indicates $p < 0.05$ (two-tailed test).

Source: Own elaboration.

Table A.2. MlwiN's IGLS (iterative generalized least squares) estimation of regression coefficients. Kindness judgments made by C toward A is our dependent variable.

	Dependent variable: kindness of A		
	Empty Model	Model 1	Model 2
Intercept	4.527	4.521	2.421
	(0.165)	(0.171)	(0.135)
Round Center		-0.071	-0.001
		(0.028)*	(0.011)
Action A			3.900
			(0.127)*
Diff. in deviance		**6.208**	**317.862**
Random Part			
Btw. subject variance	0.009	0.095	0.266
	(0.208)	(0.221)	(0.101)
Level-1 variance	4.818	4.574	0.657
	(.548)	(0.521)	(0.075)
No. of obs.	180	180	180

Note: Difference in deviance in bold characters indicates that the model is an improvement over the previous model. * Indicates $p < 0.05$ (two-tailed test).

Source: Own elaboration.

Table A.3. MlwiN's IGLS (Iterative generalized least squares) coefficient estimates. Kindness judgments made by C toward B is our dependent variable.

	Dependent variable: kindness of B				
	Empty Model	Model 1	Model 2	Model 3	Model 4
Intercept	5.19	5.192	3.792	3.748	3.896
	(0.140)*	(0.139)*	(0.195)*	(0.227)*	(0.242)*
Round Center		0.002	-0.004	-0.004	-0.002
		(0.012)	(0.010)	(0.010)	(0.010)
Action B			1.662	1.661	1.484
			(0.168)*	(0.168)*	(0.200)*
Treatment				0.100	-0.409
				(0.269)	(0.412)
Action B*Treat					0.606
					(0.376)
Diff. in deviance		0.022	**82.931**	0.139	2.558
Random Part					
Btw. subject variance	0.806	0.808	0.806	0.805	0.769
	(0.204)	(0.204)	(0.190)	(0.190)	(0.183)
Level-1 variance	1.277	1.277	0.919	0.919	0.917
	(0.117)	(0.117)	(0.084)	(0.084)	(0.084)
No. of obs.	296	296	296	296	296

Note: Difference in deviance in bold characters indicates that the model is an improvement over the previous model. * Indicates $p < 0.05$ (two-tailed test).

Source: Own elaboration.

Appendix B. Instructions

1. Instructions for the No_Trust Treatment

Instructions

You are participating in a decision-making experiment. Please read the following instructions carefully. These instructions state everything you need to know to participate in the experiment.

If you have questions, please raise your hand. One of the experimenters will approach you to answer your questions.

In this experiment you will earn money: You will earn points during the experiment. The number of points that you earn depends on your own choices and the choices of other participants. The number of points others earn, depends on their own choices and your choice, as well. At the end of the experiment, the total number of points you earn during the experiment will be exchanged at an exchange rate of:

5 points = 10 euro cent

The money you receive will be paid anonymously and in cash at the end of the experiment. The other participants will not learn what you receive and you will not learn what the others have received, either. Further instructions will follow on the next screens.

During the experiment you are not allowed to communicate with other participants and you are not allowed to use your cell phones.

[Please click to continue.]

Overview of the experiment

In today's experiment you can earn money. In every round of the experiment groups of three (3) participants are formed. The participants are labeled A, B and C. They will take turns. First participant B will make a decision. Then participant C. Participant A does not make any decision. Participants, when given their turn, choose between two actions: IN or OUT. At the end, all participants are informed of the decisions made in his/her group and the points earned by each member. The instructions below explain the experiment in detail.

Formation of groups

At the beginning of each round of the experiment all participants will be randomly divided into groups of three (3) people. Note that groups change after every round.

Assigning IDs

In each round, after the groups are formed, the computer randomly assigns an ID to each participant in a group. There will be 3 IDs: Participant A, Participant B and Participant C. These IDs will also be reassigned in each round.

You will participate in 20 rounds of decision making and then you will be asked to answer a short questionnaire.

[Please click to continue.]

Decisions

In every round, you will see a screen as shown below. You will find a graphic representation of the sequence of decisions that are made. As you can see in the figure, decisions are made sequentially: first participant B and then participant C.

As shown in the figure, participant A cannot make decisions in the experiment. Notice that the points are presented in the following order [Points for A, Points for B, Points for C]

Participant B chooses between actions IN or OUT. In case OUT is chosen the decision making ends and the points shown in the figure will be assigned to each participant. In case B chooses IN, participant C continues. Participant C chooses between IN or OUT, and the points are assigned as shown in the figure.

To better understand the points you can earn depending on your decisions and the decisions of the other participants, some examples are shown on the next screen.

[Please click to continue.]

Points

In every round, after the decision making is over, all the participants will see a screen as shown below. Three elements are displayed on the screen: a graphic representation of the decisions made (left), a table summarizing the choices of each participant and the points earned (right) and a box in which you have to answer a question. Notice that participants who were not given the chance to make a choice are represented in the table by a dash (-).

To better understand the points you can earn in each situation, please look at the examples by clicking on the buttons below. Each example shows one of the four possible decision-making outcomes.

Once you have looked at the four examples, please click to continue.

Examples

Example 1 (B plays OUT)
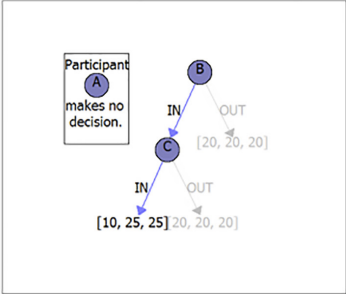
Example 2 (C plays IN)

Example 3 (C plays OUT)

The instructions are over!

You will now start the decision making. Remember that:

— You will make decisions in 20 rounds.
— In each round new groups will be formed randomly
— In each round you will be randomly assigned a new ID: Participant A, B or C.
— At the top of the first screen you will see your ID for the current round.
— A screen summarizing all the decisions made and the points earned will be shown at the end of each round.

[Please click to continue.]

2. Instructions for the Trust_Treatment

Instructions

You are participating in a decision-making experiment. Please read the following instructions carefully. These instructions state everything you need to know to participate in the experiment.

If you have any questions, please raise your hand. One of the experimenters will come to answer your questions.

In this experiment you will earn money: You will earn points during the experiment. The number of points that you earn depends on your own choices and the choices of other participants. The number of points others earn depends on their own choices and your choice as well. At the end of the experiment, the total number of points you earn during the experiment will be exchanged at an exchange rate of:

5 points = 10 cents

The money you receive will be paid anonymously and in cash at the end of the experiment. The other participants will not know what you receive and you will not know what the others have received either. Further instructions will follow on the next screens.

During the experiment you are not allowed to communicate with the other participants and you are not allowed to use your cell phones.

[Please click to continue.]

Overview of the experiment

In today's experiment you can earn money. In every round of the experiment groups of three (3) participants are formed. The participants are labeled A, B and C. They will take turns. First, participant A will make a decision. Then participant B. Then participant C. When given their turn, participants choose between two actions: IN or OUT. At the end, all participants are informed of the decisions made in his/her group and the points earned by each member. The instructions below explain the experiment in detail.

Formation of groups

At the beginning of each round of the experiment all participants will be randomly divided into groups of three (3) people. Note that the groups change after every round.

Assigning IDs

In each round, after the groups are formed, the computer randomly assigns an ID to each participant in a group. There will be 3 IDs: Participant A, Participant B and Participant C. These IDs will also be reassigned in each round.

You will participate in 20 decision-making rounds and then you will be asked to answer a short questionnaire.

[Please click to continue.]

Decisions

In each round, you will see a screen as shown below. You will find a graphic representation of the sequence of decisions that are made. As you can see in the figure, decisions are made sequentially, starting with participant A, then B and then C.

As shown in the figure, Participant A can choose action IN or action OUT. If A chooses action OUT the decision making ends (B and C do not make a decision) and each participant will earn the points shown in the box below the OUT arrow in the image. Notice that the points are presented in the following order [Points for A, Points for B, Points for C]

If participant A chooses option IN, participant B will be allowed to continue the decision making. Participant B chooses between actions IN or OUT. In case OUT is chosen the decision making ends, and the points shown in the figure will be assigned to each participant. In case B chooses IN, participant C continues. Participant C chooses between IN or OUT, and the points are as illustrated in the graphic.

To better understand the points you can earn depending on the decisions you make and the decisions of the other participants, some examples are shown on the next screen.

Figure for Decision Screen.
[Please click to continue.]

Points

In each round, after the decision making is over, all the participants will see a screen like the ones hown below. Three elements are displayed on the screen: a graphic representation of the decisions made (left), a table summarizing the choices of each participant and the points earned (right), and a box in which you have to answer a question. Notice that participants who were not given the chance to make a choice are represented in the table by a dash (-).

To better understand the points you can earn in each situation, please go through the examples by clicking on the buttons below. Each example shows one of the four possible decision-making outcomes.

Once you have looked at the four examples, please click to continue.
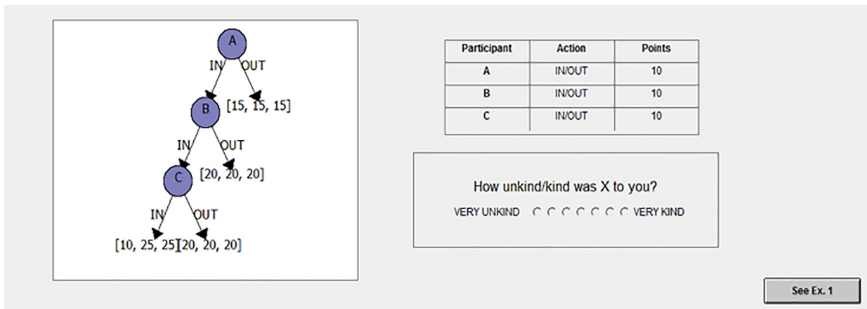
Figure for the Points screen.

Examples

Example 1 (A plays OUT)

Example 2 (B plays OUT)

Example 3 (C plays IN)

Example 4 (C plays OUT)

The instructions are over!
You will now start the decision making. Remember that:

— You will make decisions in 20 rounds.
— In each round new groups will be formed randomly.
— In each round you will be randomly assigned a new ID: Participant A, B or C.
— At the top of the first screen you will see your ID for the current round.
— A screen summarizing all the decisions made and the points earned will be shown at the end of each round.

[Please click to continue.]