Guilt Driven Reciprocity in a Psychological Signaling Game

Yuk-fai Fong, Chen-Ying Huang, Theo Offerman*

December 7, 2007

Abstract

We propose a theory of reciprocity according to which reciprocal behavior is driven by a donor's guilt. Through an experiment we show that subjects respond to factors which induce guilt but do not reflect allocative equity or intention. When the guilt inducing factor is privately observed by the donor, a psychological signaling game results. We solve for the separating and pooling equilibria. In a separating equilibrium, the donor distorts her gift to signal a low level of the guilt inducing factor. Consequently, conditional on the realization of a low level of guilt inducing factor, the donor gives a smaller gift when the guilt inducing factor is privately observed than when it is commonly known. Our experiment confirms this implication of the separating equilibrium.

*Fong: Management and Strategy Department, Northwestern University, 2001 Sheridan Rd., Evanston, IL 60208, USA. Huang: Department of Economics, National Taiwan University, 21 Hsu-Chow Road, Taipei 100, Taiwan. Offerman: Department of Economics, University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands. We would like to thank CREED-programmer Jos Theelen for programming the experiment. The research of Theo Offerman has been made possible by a fellowship of the Royal Netherlands Academy of Arts and Sciences.

1 Introduction

Imagine that you just paid the bill for a dinner, for which you included a 15% gratuity for your waitress's service. As you are walking out of the restaurant, the owner comes up to you and insists that you take back the gratuity. He explains to you that he has a policy of not allowing his waiting staff to accept tips. Should you be happy or upset about it? If your reciprocity is driven by fairness concern or by the waitress's kindness (as manifested by her good service), then the gratuity you paid is the outcome that maximizes your utility. The restaurant owner's rejection of your tips on behalf of the waitress imposes a constraint on your maximization problem and should lower your utility. By contrast, if your decision to reward the waitress is rather driven by guilt, then you might feel relieved that there is no expectation for gratuity and you would be glad to save some money. This example suggests that although in many situations fairness/kindness driven reciprocity and guilt driven reciprocity are observationally indistinguishable, the welfare implications of these two types of reciprocity can be rather different. Therefore, it is important to know the source of reciprocal behavior.

In many real-life settings like the example above, actions that exhibit kindness and equity concerns typically cause guilt at the same time. As a result, it is difficult to distinguish whether certain reciprocal behavior is motivated by kindness, equity, or guilt. There are, however, special situations where it is possible to distinguish among the sources of reciprocity. One such situation is sketched in the following story.

Michael courts Diana and gives her a richly ornate ring to show his affection for her. The ring is a family heirloom. At the moment that Michael gives the ring, neither Michael nor Diana knows whether the ring is an ingenious piece of kitsch or a true piece of art. After some time, Diana experiences some serious financial set-backs and finds herself in dire straits, she decides to sell the ring to a third person. Only *after* the sale of the ring its true value is revealed.

First consider the case where both Diana and Michael later learn the value of the ring. Suppose that after the dust is settled the ring turns out to be a true piece of art. Does this information make Diana feel more indebted to Michael compared to the case where the ring would have been a piece of kitsch? According to kindness and equity based theories it should not. Notice that the value of the ring does not reflect Michael's kindness to Diana, because at the time the ring was given to Diana, its value was unknown to Michael. By now the value of the ring is also immaterial to Diana's payoff and the allocative equity between Michael and her, because she already sold the ring before the value

was revealed to any of the players. Nevertheless, it seems quite natural that Diana feels more guilty to Michael once she learns that the ring is a piece of art instead of a piece of kitsch. This may cause her to return a larger gift to Michael when it is her turn to do so.

Next consider the situation where the value of the ring is revealed to Diana but not to Michael. If it is true that Diana cares about Michael's belief of the value of the ring, a psychological signaling game results in which Diana's gift to Michael may reflect an attempt to affect Michael's belief of the value of the ring.

In this paper, we theoretically and experimentally investigate the story of Diana and Michael in the setup where after Michael has taken his action both players are informed of the value of the ring as well as in the setup where only Diana is informed of the value of the ring. In the former setup, we tease apart the reciprocal behavior motivated by guilt from that motivated by kindness or equity concern. In the situation where both players are informed of the state of the world, subjects (in the role of Diana) reciprocate more when the realization is high (the ring is a piece of art) than when the realization is low (the ring is a piece of kitsch). This result underlines the independent role that guilt plays in reciprocal behavior. In fact, the proportion of subjects who exhibit some guilt-based reciprocity (40%) is in comparable magnitude to that of subjects who reciprocate only out of fairness concern or based on kindness (45%).

Given this result, it makes sense to analyze the situation where Michael is uninformed of the value of the ring as a psychological signaling game. This signaling game contains pooling as well as separating equilibria. In a pooling equilibrium, Diana sends back the same amount of money in both states of the world and thus hides the true value of the ring from Michael. There are also separating equilibria in which when Diana observes a low-valued ring, she returns a smaller gift than that under full information. The basic idea behind this finding is that, on top of the possibility that returning a gift may reduce guilt, giving a small gift allows Diana to signal the low value of the ring and doing so relieves Diana from the guilt caused by Michael's expectation. As in a standard separating equilibrium of a signaling game, such a distortion happens only when the realization is low. Therefore, when the ring realization is high, Diana acts identically as she does under full information. The experimental results are in line with these two central features of the separating equilibrium.

We mould the story between Michael and Diana in the following gift-exchange game between two players. The first mover (player X) has a coupon of which the value to player X could be either high

or low with equal probability. If player X decides to give up the coupon for the second mover (player Y), then player Y gets a fixed amount of euros *for sure* and observes the realization of the value of the coupon. In one treatment (the "Uninformed-X" treatment), player X never learns the realization if he gives up the coupon for player Y and in another treatment (the "Informed-X" treatment), he does learn the realization after he gives up the coupon for player Y.[1] Notice that if player Y reciprocates because player X is kind enough to give up the coupon for player Y, then since the realization of the value of the coupon is random and does not indicate the kindness of player X (nor does it affect the amount player Y receives), it should not affect the gift of player Y to player X. Alternatively, if player Y has concern for allocative equity, once player X gives up the coupon, then since player X no longer owns the coupon and the coupon realization does not affect player Y's payoff, the fairness of the allocation between player X and player Y is independent of the coupon realization. Hence the gift of player Y to player X should not be affected by the realization either. On the contrary, if player Y feels more guilty when player X gives up a coupon of higher value for her, then she may reciprocate more when the realization is higher. Hence the gift of player Y to player X may depend on the realization of the coupon.

Previous papers on reciprocity mainly focussed on one of two approaches. The first approach suggests that agents dislike unequal outcomes. The second approach suggests that players form a perception of the other players' intentions and reciprocate for kindness. That is, kind intentions are rewarded while unkind intentions are punished. Leading theoretical papers that adopt the first approach include Fehr and Schmidt (1999) and Bolton and Ockenfels (2000); those that adopt the second include Rabin (1993) and Dufwenberg and Kirchsteiger (2004). There are also some attempts to merge the two approaches (Levine (1998) and Falk and Fischbacher (2006)). A wealth of experimental papers investigates these approaches (among others, Charness and Rabin (2002), Offerman (2002), Falk, Fehr and Fischbacher (2003) and Fehr and Schmidt (2003)). Overall, the picture that emerges from these papers is that both intentions and inequality considerations are driving forces behind reciprocal behavior.[2]

Our paper contributes to the recent literature on the role of guilt aversion. Using psychological game

[1] In this paper, following the motivating example, we will refer to player X as a male and player Y as a female. In the experiment, of course, males and females were randomly assigned to each role.

[2] Because in our experimental setup the monetary allocation is independent of the state of the world, we also control for other allocative motives like the possibility that subjects pursue efficiency (Charness and Rabin (2002) and Engelmann and Strobel (2004)).

theory, originally developed in Geanakoplos, Pearce and Stacchetti (1989), Battigalli and Dufwenberg (2007) develop a general theory for guilt in games. In their model, a player's guilt depends on the degree to which he believes that he lets another player down. The second player is let down when the first player's choice causes the second player to get a lower monetary payoff than the second player has expected before the play starts.[3] This theoretical work is accompanied by a couple of experimental papers investigating the role of guilt aversion. Charness and Dufwenberg (2006) examine the impact of guilt aversion on the effectiveness of communication. Subjects make promises and promises become credible because people feel guilty if they do not live up to other people's expectations. Hopfensitz and Reuben (2006) show that social emotions like guilt need to be present for punishment to be an effective tool in deterring uncooperative behavior. Regner and Barria (2007) provide field evidence showing that customers of an online music store that offers unlimited pre-purchase listening make a higher voluntary payment of 8.20 dollar than the required minimum of 5 dollar. Their results suggest a role for guilt aversion in the field, though their data do not rule out other explanations.[4]

The remainder of the paper is organized in the following way. Section 2 presents the theoretical framework and provides the analysis of the psychological signaling game. Section 3 discusses the experimental design. Section 4 presents the experimental results and section 5 concludes.

2 Theoretical Framework

We consider a gift exchange game played between two players, the first mover (player X) and the second mover (player Y). The first mover has a coupon which he can either give up for player Y or keep for himself. Let $x_1 \in \{0, 1\}$ denote the decision of player X where 0 means he gives up the coupon for player Y and 1 means he keeps the coupon.[5] If player X decides to give up the coupon for player Y, then player Y receives a fixed amount *irrespective of the value of the coupon*. After player X decides, player Y can reciprocate and decide her gift for player X. Let $x_2 \in [0, m]$ denote the amount of player

[3] Other theoretical contributions that use psychological game theory to model guilt in games are provided by Ruffle (1999) and Battigalli and Dufwenberg (2005).

[4] In the psychological literature, there is evidence that shame and guilt have significant impact on many aspects of human functioning and interpersonal relationships (Baumeister, Stillwell and Heatherto (1994), Tangney (1995) and Tangney and Dearing (2002)). For instance, Ketelaar and Au (2003) find that feelings of guilt provoke individuals to cooperate in repeated social bargaining games.

[5] In the following, we use subscript 1 to denote the decision or belief about player X and subscript 2 that about player Y.

Y's gift for player X where m is the maximum of player Y's possible gift to player X.

As stated in the introduction, the realized value of the coupon can be either high (H) or low (L) to player X. To accommodate the possible guilty feeling of player Y, we need to specify player Y's beliefs, in particular her belief of player X's belief of the realization of the coupon. To this effect, denote $z_1 \in [0,1]$ player X's belief of the probability that the realization is high and $\tilde{z}_1 \in [0,1]$ player Y's belief of player X's belief of the probability that the realization is high. Also denote $z_2 \in [0,1]$ player Y's belief of the probability that the realization is high. Since player Y always observes the realization in our experiment, z_2 corresponds to the true realization. That is, $z_2 = 1$ when the realization is H and $z_2 = 0$ when the realization is L.

One can imagine that player Y's utility depends on X's decision, her own decision and her beliefs. Therefore, in a quite general form, player Y's utility can be expressed as $V(x_1, x_2, \tilde{z}_1, z_2)$. Hence player Y solves

$$\max_{x_2 \in [0,m]} V(x_1, x_2, \tilde{z}_1, z_2) \tag{1}$$

when it is her turn to make a decision. The utility of player Y is general enough to accommodate reciprocity based on kindness, equity, or guilt.[6] Since we focus on guilt driven reciprocity, we care most about the case when player X gives up the coupon for player Y (when $x_1 = 0$). Therefore, denote $U(x_2, \tilde{z}_1, z_2) \equiv V(0, x_2, \tilde{z}_1, z_2)$.

We model player Y's guilt as the loss in her utility and the increase in propensity to give a larger gift to player X which are caused by a *guilt inducing factor*. In our model, the guilt inducing factor is the coupon. Related beliefs on the coupon realization include $\tilde{z}_1$ and z_2. A high coupon realization leads to guilt either through player Y's direct observation of it, i.e., z_2, or player Y's belief of player X's

[6] We give examples for kindness driven reciprocity and equity driven reciprocity.

Example of kindness driven reciprocity: For instance, $V(x_1, x_2, \widetilde{z_1}, z_2) = x_2^{1-x_1}(m - x_2)$. In this case, when X does not give up the coupon for Y (when $x_1 = 1$), Y only cares about how much she can consume, $(m - x_2)$. On the other hand, when X gives up the coupon for her (when $x_1 = 0$), Y's utility is the product of her gift to X and how much she leaves for herself. Thus, when $x_1 = 1$, Y maximizes by giving up nothing for X. On the other hand, when $x_1 = 0$, Y maximizes by giving up a half of m for X.

Example of equity driven reciprocity: For instance, $V(x_1, x_2, \widetilde{z_1}, z_2) = \min\{x_1 + x_2, 1 - x_1 + m - x_2\}$. Suppose for simplicity that the expected value of the coupon to X is 1. In this case, when X does not give up the coupon for Y (when $x_1 = 1$), Y wants to equalize what X keeps for himself plus what Y gives up for X and what Y leaves for herself. Thus, Y maximizes by giving up $\frac{m-1}{2}$ for X. On the other hand, when X gives up the coupon for Y (when $x_1 = 0$), Y wants to equalize what she gives up for X and what X gives up for her plus what she leaves for herself. Thus, Y maximizes by giving up $\frac{m+1}{2}$ for X.

belief of it, i.e., $\tilde{z}_1$, if player X may update based on player Y's gift size. We now impose assumptions on $U(x_2, \tilde{z}_1, z_2)$ to draw experimentally testable implications on how player Y's gift to player X would respond to the coupon realization in the setting with commonly observed coupon realization and in the setting in which the coupon is privately observed by player Y.[7]

Assumptions:

(A1) *For any x_2, both $U(x_2, \tilde{z}_1, 0)$ and $U(x_2, \tilde{z}_1, 1)$ are strictly decreasing in $\tilde{z}_1$.* In words, conditional on the true realization z_2, if player Y believes player X believes that the probability of H is higher, then player Y might feel she needs to give more to player X to meet the higher expectation of player X. This results in more guilt if player Y still gives the same amount of gift x_2 to player X and thus reduces the utility of player Y.

(A2) *For any $\tilde{z}_1$ and z_2, $U(x_2, \tilde{z}_1, z_2)$ is continuous in x_2 and single-peaked.* This is merely a technical assumption.

Since $U(x_2, \tilde{z}_1, z_2)$ is single-peaked, denote

$$\begin{aligned} x_L^* &\equiv \arg \max_{x_2 \in [0,m]} U(x_2, 0, 0), \\ x_H^* &\equiv \arg \max_{x_2 \in [0,m]} U(x_2, 1, 1). \end{aligned}$$

(A3) *Assume that $x_L^* < x_H^*$.* In words, if player Y believes that player X believes that the realization

[7] One possible and intuitive set of assumptions which provides a foundation for and implies (A1) and (A3) is as follows. Holding player Y's gift to X fixed, a high realization of coupon causes player Y to feel more guilty and this results in a lower utility of player Y. Player Y also feels more guilty if she believes player X believes that the probability of a high realization is higher, and this again results in a lower utility of player Y. Therefore

$$\begin{aligned} U(x_2, \tilde{z}_1, 0) &> U(x_2, \tilde{z}_1, 1) \text{ and} \\ U(x_2, \tilde{z}_1', z_2) &> U(x_2, \tilde{z}_1, z_2), \text{ if } \tilde{z}_1' < \tilde{z}_1. \end{aligned}$$

Since the loss in utility caused by guilt may be partially recovered by treating X more favorably, therefore,

$$\begin{aligned} \arg \max_{x_2 \in [0,m]} U(x_2, \tilde{z}_1, 0) &< \arg \max_{x_2 \in [0,m]} U(x_2, \tilde{z}_1, 1) \text{ and} \\ \arg \max_{x_2 \in [0,m]} U(x_2, \tilde{z}_1', z_2) &< \arg \max_{x_2 \in [0,m]} U(x_2, \tilde{z}_1, z_2), \text{ if } \tilde{z}_1' < \tilde{z}_1. \end{aligned} \quad (*)$$

That is, when related beliefs ($\tilde{z}_1$ and z_2) are lower, Y does not feel as necessary to return a gift. Hence her optimal gift to X is lower. In particular, suppose we set $\tilde{z}_1 = 1$, $\tilde{z}_1' = 0$ and $z_2 = 0$ in (*), then we get

$$\begin{aligned} \arg \max_{x_2 \in [0,m]} U(x_2, 1, 0) &< \arg \max_{x_2 \in [0,m]} U(x_2, 1, 1) \text{ and} \\ \arg \max_{x_2 \in [0,m]} U(x_2, 0, 0) &< \arg \max_{x_2 \in [0,m]} U(x_2, 1, 0), \end{aligned}$$

and these imply assumption (A3).

is low than high (and player Y believes player X has correct beliefs as will happen when player X can observe the coupon realization), player Y does not feel as necessary to give a gift for player X to reduce her guilt. As a result, her optimal gift to player X is lower.

A higher $\tilde{z}_1$ implies that Y believes that X believes that the coupon is more valuable to X should X keep it. (A1) means that the increase of this belief will decrease the utility of Y. Moreover, this decrease of Y's utility can be partially recovered if she treats X more favorably and hence in (A3), we assume that when the belief is high rather than low, Y treats X more favorably.

Battigalli and Dufwenberg's (2007) framework is sufficiently general that one can view our model as an adaptation of theirs. In the context of their framework, the novelty of our model of guilt is that it allows player Y's belief of player X's belief of the coupon realization to affect player Y's guilt and thus her gift to player X. This can happen in a generalized framework of Battigalli and Dufwenberg's (2007) when play X believes that the coupon realization is high, he will have a higher *expectation about his material payoff*, causing him to expect a higher gift from Y. Player X's expectation about his material payoff in turn affects player Y's guilt and her gift to player X.

In the experiment we have two treatments, Informed-X and Uninformed-X, depending on whether player X observes the realization of the coupon after he has decided to give up the coupon for player Y. We can now work out the pure-strategy equilibria in both treatments if player X indeed gives up the coupon for player Y.

In the Informed-X treatment, since player X observes the realization as well, there is no room for player Y to signal whether the coupon is of high or low value to player X. Player Y will then optimally choose to give the gift of x_L^* to player X when the realization is low and give the gift of x_H^* to player X when the realization is high.

In the Uninformed-X treatment, we look for both the separating and the pooling equilibria.

Separating Equilibria

Suppose in a separating equilibrium, player Y chooses to give x_L to player X when the realization is low and to give x_H to player X when the realization is high. Moreover, for any gift x_2 from player Y to player X, player Y believes player X believes the realization is high with probability $\tilde{z}_1(x_2) \in [0, 1]$.

For these to be an equilibrium, we must have[8]

$$\begin{aligned} &\tilde{z}_1(x_L) = 0, \\ &\tilde{z}_1(x_H) = 1, \\ &U\left(x_L, 0, 0\right) \geq U\left(x_2, \tilde{z}_1(x_2), 0\right) \text{ for any } x_2, \\ &U\left(x_H, 1, 1\right) \geq U\left(x_2, \tilde{z}_1(x_2), 1\right) \text{ for any } x_2. \end{aligned} \tag{2}$$

We can now state some properties of the separating equilibria.

Proposition 1 *For any separating equilibrium in the Uninformed-X treatment,* $x_H = x_H^*$.

Proof. Suppose to the contrary that there is a separating equilibrium where $x_H \neq x_H^*$. Then when the realization is high, player Y can deviate to choose x_H^*. By doing so, her utility will be $U\left(x_H^*, \tilde{z}_1(x_H^*), 1\right)$. Since $U\left(x_H^*, 1, 1\right) > U\left(x_H, 1, 1\right)$ by (A2) and $U\left(x_H^*, \tilde{z}_1(x_H^*), 1\right) \geq U\left(x_H^*, 1, 1\right)$ by (A1), no matter what $\tilde{z}_1(x_H^*)$ is, this constitutes a profitable deviation, yielding a contradiction.[9] ■

We need more notations to proceed. Denote $x_{HL}^* \equiv \arg\max_{x_2 \in [0,m]} U(x_2, 1, 0)$. Let $A = \{x_2 \in [0, m] : U(x_2, 0, 0) \geq U(x_{HL}^*, 1, 0)\}$ and $B = \{x_2 \in [0, m] : U(x_H^*, 1, 1) \geq U(x_2, 0, 1)\}$.

Proposition 2 *For any separating equilibrium in the Uninformed-X treatment,* $x_L \in A \cap B$.[10]

Proof. To show that $x_L \in A$, suppose to the contrary that $x_L \notin A$; then $U(x_L, 0, 0) < U(x_{HL}^*, 1, 0)$. Then when the realization is low, Y can deviate to choose x_{HL}^*. By doing so, her utility will be $U\left(x_{HL}^*, \tilde{z}_1(x_{HL}^*), 0\right) \geq U(x_{HL}^*, 1, 0)$ by (A1). No matter what $\tilde{z}_1(x_{HL}^*)$ is, this constitutes a profitable deviation, yielding a contradiction.

To show that $x_L \in B$, suppose to the contrary that $x_L \notin B$, then $U(x_H^*, 1, 1) < U(x_L, 0, 1)$. When the realization is high, Y can deviate to choose x_L. Since in equilibrium $\tilde{z}_1(x_L) = 0$, her payoff will be $U(x_L, 0, 1)$, constituting a profitable deviation and yielding a contradiction. ■

Since for any x_2, $U(x_2, 0, 0) > U(x_2, 1, 0)$ by (A1) and $U(\cdot, 0, 0)$ is single-peaked, there must exist two points x_A^1, x_A^2 such that $A = [x_A^1, x_A^2]$ where $x_L^* \in A$. Similarly, since for any x_2, $U(x_2, 1, 1) < U(x_2, 0, 1)$

[8] Note that in equilibrium, beliefs correspond to reality (Geanakoplos, Pearce and Stacchetti (1989) and Rabin (1993)). Therefore, player Y's belief of player X's belief of the realization corresponds to player X's belief of the realization, which in turn corresponds to the true realization.

[9] It is quite standard in signaling games that there is no distortion for the unfavorable type.

[10] Note that for any $x_L \in A \cap B$, it is easy to check that as long as $\widetilde{z_1}(x_L) = 0$ and $\widetilde{z_1}(x_2) = 1$ for any $x_2 \neq x_L$, x_L and $x_H = x_H^*$ satisfy (2). In other words, as long as Y believes X has a high expectation off the equilibrium path, we can easily support x_L and $x_H = x_H^*$ as a separating equilibrium.

by (A1) and $U(\cdot, 0, 1)$ is single-peaked, when B is not empty, either there exists a point $x_B^1 < x_H^*$ such that $U(x_H^*, 1, 1) = U(x_B^1, 0, 1)$ or there exists a point $x_B^2 > x_H^*$ such that $U(x_H^*, 1, 1) = U(x_B^2, 0, 1)$ or both. Hence when B is not empty, either $B = [0, x_B^1]$, $B = [x_B^2, m]$ or $B = [0, x_B^1] \cup [x_B^2, m]$. This implies that when B is not empty, $A \cap B = [x_A^1, \min\{x_A^2, x_B^1\}]$, $[\max\{x_A^1, x_B^2\}, x_A^2]$ or $[x_A^1, \min\{x_A^2, x_B^1\}] \cup [\max\{x_A^1, x_B^2\}, x_A^2]$. To simplify the analysis, instead of going through all three possibilities, we work on the third possibility. The other two possibilities can be worked out analogously. (Figures 1 and 2 illustrate two of the possible ways in which the sets A and B may intersect. We use thickened segments along the x-axis to highlight $A \cap B$.)

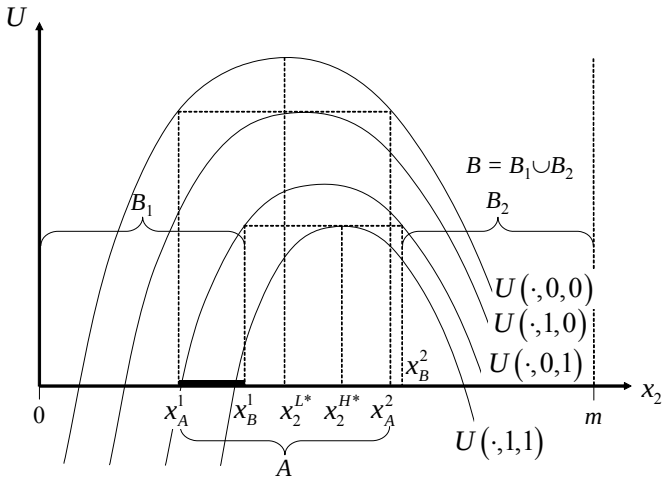

Figure 1

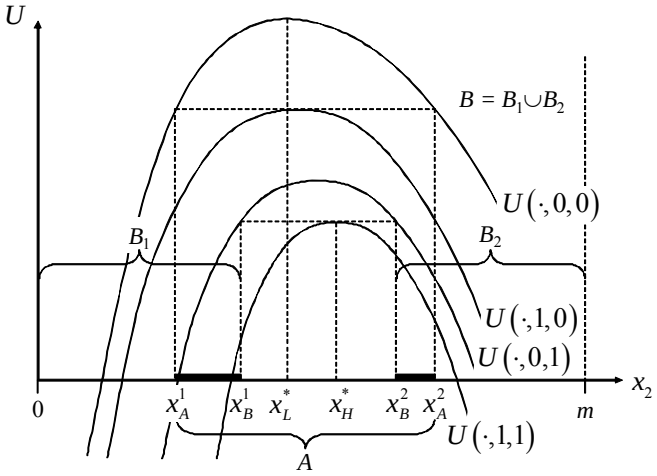

Figure 2

Note that without any further restriction, it is possible that $x_L^* \in A \cap B$.[11] That is, in the Uninformed-X treatment, player Y gives x_H^* to player X when the realization is high and gives x_L^* to player X when the realization is low. In other words, player Y behaves exactly the same way as she does in the Informed-X treatment. When this is the case, by (2), $U(x_H^*, 1, 1) \geq U(x_L^*, \tilde{z}_1(x_L^*), 1) = U(x_L^*, 0, 1)$. We now explore the possibility when on the contrary that $U(x_H^*, 1, 1) < U(x_L^*, 0, 1)$ holds.

When $U(x_H^*, 1, 1) < U(x_L^*, 0, 1)$, $x_L^* \notin B$. Since $B = [0, x_B^1] \cup [x_B^2, m]$, this implies that $x_B^1 < x_L^* < x_B^2$. Recall that $A \cap B = [x_A^1, \min\{x_A^2, x_B^1\}] \cup [\max\{x_A^1, x_B^2\}, x_A^2]$, depending on which interval x_L lies in, there are two cases. In the first case where $x_L \in [x_A^1, \min\{x_A^2, x_B^1\}]$, we have $x_L < x_L^*$. This is quite intuitive as when the coupon realization is low, player Y lowers her gift to player X (from the full information level x_L^*) to signal that the realization is indeed low. In the second case if $x_L \in [\max\{x_A^1, x_B^2\}, x_A^2]$, then $x_L \geq x_B^2 > x_L^*$. Since $x_B^2 > x_H^*$, in fact $x_L > x_H^*$. In other words, there is a type of perverse separating equilibria in which when the coupon realization is low, player Y signals that the coupon realization is low by giving player X more than x_H^*, her gift to player X in the Informed-X treatment when the coupon realization is high. Without specifying functional forms for the utility of player Y, this perverse signalling equilibrium cannot be ruled out. We now summarize these findings.

Corollary 3 *If $U(x_H^*, 1, 1) < U(x_L^*, 0, 1)$, then it is impossible that player Y in the Uninformed-X treatment behaves exactly the same way as she does in the Informed-X treatment. Specifically, $x_L < x_L^*$ or $x_L > x_H^*$. In words, when the realization is low, in the Uninformed-X treatment, player Y signals the realization is low either by giving less than what she would give in the Informed-X treatment with the low realization, or perversely by giving more than what she would give in the Informed-X treatment with the high realization.*

Figure 1 presents the natural case in which player Y may only signal a low coupon realization by giving player X less than what she would give in the Informed-X treatment with the low realization, i.e., $x_L < x_L^*$. Figure 2 presents the case in which perverse signaling, characterized by $x_L > x_H^*$, is possible. In both figures, the set of x_L with which player Y can convincingly signal that the coupon realization is low is thickened along the x-axis.

Pooling Equilibria

Suppose in a pooling equilibrium, player Y chooses to give x_p to player X no matter the realization

[11] Such a possibility is not illustrated in Figures 1 and 2.

is high or low. Moreover, for any gift x_2 from player Y to player X, player Y believes player X believes the realization is high with probability $\tilde{z}_1(x_2) \in [0,1]$. Denote ρ the prior probability that the realized coupon is of high value. For these to be an equilibrium, we must have

$$\begin{aligned} &\tilde{z}_1(x_p) = \rho, \\ &U\left(x_p, \rho, 0\right) \geq U\left(x_2, \tilde{z}_1(x_2), 0\right) \text{ for any } x_2, \\ &U\left(x_p, \rho, 1\right) \geq U\left(x_2, \tilde{z}_1(x_2), 1\right) \text{ for any } x_2. \end{aligned} \tag{3}$$

We need more notations to proceed. Let $\tilde{A} = \{x_2 \in [0,m] : U(x_2, \rho, 0) \geq U(x^*_{HL}, 1, 0)\}$ and $\tilde{B} = \{x_2 \in [0,m] : U(x_2, \rho, 1) \geq U(x^*_H, 1, 1)\}$.

Proposition 4 *For any pooling equilibrium in the Uninformed-X treatment,* $x_p \in \tilde{A} \cap \tilde{B}$.[12]

Proof. To show that $x_p \in \tilde{A}$, suppose to the contrary that $x_p \notin \tilde{A}$; then $U(x_p, \rho, 0) < U(x^*_{HL}, 1, 0)$. Then when the realization is low, player Y can deviate to choose x^*_{HL}. By doing so, her utility will be $U\left(x^*_{HL}, \tilde{z}_1(x^*_{HL}), 0\right) \geq U(x^*_{HL}, 1, 0)$ by (A1). No matter what $\tilde{z}_1(x^*_{HL})$ is, this constitutes a profitable deviation, yielding a contradiction.

To show that $x_p \in \tilde{B}$, suppose to the contrary that $x_p \notin \tilde{B}$; then $U(x_p, \rho, 1) < U(x^*_H, 1, 1)$. Then when the realization is high, player Y can deviate to choose x^*_H. By doing so, her utility will be $U\left(x^*_H, \tilde{z}_1(x^*_H), 1\right) \geq U(x^*_H, 1, 1)$ by (A1). No matter what $\tilde{z}_1(x^*_H)$ is, this constitutes a profitable deviation, yielding a contradiction. ■

Figure 3 depicts the range of pooling equilibrium x_2. As in Figures 1 and 2, we use thickened

[12]Note that for any $x_p \in \tilde{A} \cap \tilde{B}$, it is easy to check that as long as $\widetilde{z_1}(x_p) = \rho$ and $\widetilde{z_1}(x_2) = 1$ for any $x_2 \neq x_p$, x_p satisfies (3). In other words, as long as Y believes X has a high expectation off the equilibrium path, we can easily support x_p as a pooling equilibrium.

segments along the x-axis to highlight $\tilde{A} \cap \tilde{B}$.

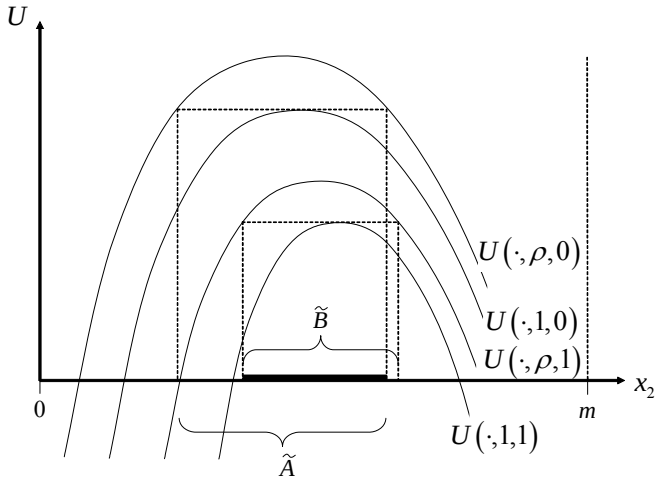

Figure 3

3 Experimental Setup

3.1 Description of the game

We now describe our game in the same terms as we did to the subjects. The game is played by two players, each of whom earns the sum of three amounts of money. Each player's first amount is a fixed starting capital of 5 euros. Each player's second amount is determined by the first mover, player X. Player X decides whether to keep a coupon or give it to the player with role Y. The coupon has an unknown value to player X. Its value to player X equals 0 or 4 euros, and both possibilities occur with equal probability. When player X makes his decision, neither player is informed of the value of the coupon to player X.

The choice of player X has the following consequences for the second amount to the players. If player X keeps the coupon, player Y receives 0 euro, while player X receives 0 euro in case the value of the coupon equals 0 or 4 euros in case the value of the coupon equals 4. If player X gives the coupon to player Y, player X receives 0 euro and player Y receives 6 euros, *irrespective of the value of the coupon* to player X.

The third amounts are determined by player Y, who decides whether to give 0, 1, 2, 3, or 4 euros to player X. The amount player X receives is two times the amount that player Y gives to player X. Player Y loses an amount equal to the amount that she gives to player X. Table 1 summarizes the game.

Table 1
The Game

Choice of player	value of coupon to X	probability	payoff to X	payoff to Y
X keeps the coupon	0	0.5	+0	+0
X keeps the coupon	4	0.5	+4	+0
X gives the coupon			+0	+6
X gives the coupon			+0	+6
Y gives 0 to X			+0	-0
Y gives 1 to X			+2	-1
Y gives 2 to X			+4	-2
Y gives 3 to X			+6	-3
Y gives 4 to X			+8	-4

Notes: Payoffs are in euros. The last two columns list what players earn in addition to their starting capital of 5 euros as a function of the choices of players X and Y and the value of the coupon. Player Y is informed of the choice of player X and the value of the coupon before she makes the choice how much to give to X.

We consider two variants of this game. In the first variant, player X remains uninformed of the value of the coupon throughout the game once he chooses to give the coupon to player Y. We refer to this variant as treatment "Uninformed-X". In the other variant, player X is informed of the value of the coupon after he gives the coupon to player Y. We refer to this variant as treatment "Informed-X".

3.2 Experimental design and procedures

The experiment was run at the University of Amsterdam. In total, 160 subjects participated. The experiment and its instructions were computerized. We explained that all subjects read the same instructions. Subjects could read the instructions at their own pace. At the end of the instructions, we informed subjects that half of them would participate in treatment Uninformed-X and half of them would participate in treatment Informed-X.[13] However, they would learn which treatment they were in only after all players X had made their decisions. This way we make sure that the behavior of players X is unaffected by the treatment and that a difference in the behavior of players Y between the two treatments cannot be attributed to players Y having different perceptions of player X's behavior.

[13] In the instructions, we used the labels treatment A and treatment B to refer to Uninformed-X and Informed-X, respectively.

After reading the instructions, subjects had to answer a set of control questions correctly before they could proceed. They received a handout with a summary of the instructions before the experiment started. Then subjects played the game described only once (to avoid repeated game effects). The experiment concluded with a questionnaire that asked subjects to rate their feelings about some situations that could occur in the experiment. We postpone the description of the questionnaire and the analysis of subjects' answers to the questionnaire to the next section. After completing the questionnaire, subjects were paid one by one in private. They could immediately leave the laboratory after they had been paid. The experiment lasted for about 45 minutes in which subjects earned on average 7.75 euros.[14]

In the instructions we explained that the subjects were randomly matched in pairs. Only after everybody had read the instructions, subjects learned whether they had the role of player X or player Y. At that time, we also informed each subject where the person s/he was matched with sat. We deliberately made this information available, because we think that people experience emotions like shame and guilt more realistically if they know that their partners know who they are. Subjects were not allowed to communicate, though.

After each player X decided whether to give up the coupon for the player Y he was paired up with, all players were informed of the treatments they were assigned to. Without informing player Y of player X's decision, we asked her to formulate a strategy. That is, we asked player Y to specify what she would choose in each of the three possibilities that might occur: (i) player X keeps the coupon; (ii) player X gives the coupon and its value is 0 euro to player X; (iii) player X gives the coupon and its value is 4 euros to player X. We explained that it was up to player Y to decide whether she wanted to give the same amount in all these three cases, or to give different amounts in different cases.

Only after player Y had constructed her plan, the choice of player X was communicated to her. The computer then drew a random value that determined whether the coupon had value 0 or 4 to player X. Player Y was informed of the value of the coupon to player X in the case player X had decided to give the coupon to her. If player X had kept the coupon, the value of the coupon to player X was not revealed to player Y.

The computer reminded player Y of the choice that she had planned for the particular case that actually happened. At this point, player Y was offered an opportunity to revise her previously planned

[14] Because the experiment was so short, we decided to run it after another experiment. The experiment that preceded this one dealt with a completely unrelated and anonymous market game.

choice at a small cost of 20 euro-cents. If player Y did not change her decision, then her planned choice became her final decision, and she did not incur extra costs.

The advantage of using the strategy method is that within-measures are obtained of how a person responds to different situations. This advantage is especially salient in experiments that investigate social preferences, because previous experiments showed that there is a lot of heterogeneity in subjects' social preferences (Offerman, Sonnemans and Schram (1996), for overviews see Fehr and Gächter (2000) and Burlando and Guyala (2005)). A within design "allows the effect of a treatment to shine through when the nuisance of individual difference is controlled for" (Camerer (2003), p. 41). In contrast, in a between design individual differences easily wash away a true treatment effect.

The danger of using the strategy method is that people may behave differently than they would have if the game were played naturally. For instance, when filling out a strategy, people may not be able to anticipate the emotions that they will experience when the particular situation actually occurs. Whether the strategy method unintentionally affects behavior is ultimately an empirical question, though. There are several studies that report only a limited impact of using the strategy method (Cason and Mui (1998), Brandts and Charness (2000), Oxoby and McLeish (2004), Bosch-Domènech and Silvestre (2005) and Falk and Kosfeld (2006)). A few studies find different behavior when the strategy method is used (Güth, Huck and Müller (2001), Brandts and Charness (2003) and Brosig, Weimann and Yang (2003)). We chose to implement a pragmatic solution that helps us to avoid the problem of the traditional strategy method. We used the strategy method, but we gave subjects the possibility to change their decision at a small cost if they wanted to. This way, those subjects who were not satisfied with how their planned choice worked out were given the possibility to revise their choice, allowing us to see how the strategy method had affected their choice.[15]

3.3 Testable hypotheses

The discussion in the section of theoretical framework leads to the following hypotheses that we will test in the next section. The first hypothesis deals with the question regarding whether player Y reacts to the lost opportunities of player X. Our hypothesis is that player Y will feel more guilty if she does not return anything after finding out that the coupon would have yielded 4 euros to player X instead

[15] We made subjects pay a small cost to change their decisions because we did not want subjects to change their decisions when they were indifferent.

of 0 euro. As a consequence, she sends more money back to player X in the former case. A direct test of this prediction is possible in treatment Informed-X, because here player Y's behavior is not affected by signaling. This corresponds to a test on assumption (A3).

Hypothesis 1: in treatment Informed-X, player Y sends back a lower amount of money when player X gives a coupon with a value of 0 than when he gives a coupon with a value of 4 euros.

The other hypotheses are related to signaling. For a separating equilibrium to occur in the Uninformed-X treatment, it is necessary that player Y sends a different amount when the value of the coupon is low than when it is high, while for pooling player Y should send the same amount irrespective of the value of the coupon. Hypothesis 2 distinguishes between pooling and separating equilibria.

Hypothesis 2: in treatment Uninformed-X, player Y's behavior does not change with the value of the coupon.

The null-hypothesis corresponds to the pooling equilibrium. Alternatively, player Y chooses to separate and send back a lower (or higher in the perverse case) amount when the value of the coupon is 0 than when it is 4. The next two hypotheses are conditional on the separating equilibrium being played. Proposition 1 shows that in any separating equilibrium player Y's reciprocal behavior is not affected by the treatment variable if the coupon has a high value. The next hypothesis is built upon this result.

Hypothesis 3: player Y sends back the same amount when player X is informed of the fact that he gave a coupon of 4 euros as when player X is not informed of this fact.

According to Corollary 3, in a separating equilibrium of treatment Uninformed-X, player Y may distort her gift in treatment Informed-X to signal that the coupon has a low value. This becomes the final hypothesis.

Hypothesis 4: player Y sends back a different amount when player X is informed of the fact that he gave a coupon of 0 euro than when player X is not informed of this fact.

4 Experimental Results

We present the results in two parts. In section 4.1 we will deal with the question of how well the model organizes the experimental data. In section 4.2 we will use the data collected in the post-experimental

questionnaire to shed light on the question concerning how reported guilt correlates with subjects' choices. In addition, we will briefly explore how gender affects trust and trustworthiness.

4.1 Hypotheses and data

Before we present the results on the hypotheses, we discuss the evidence we have on the question regarding whether the strategy method distorted what people chose. In total, 13 of the 80 players Y chose to revise their decision when we gave them the opportunity to do so. One of these persons did so by mistake, because after exercising the option to revise the planned decision, this person made the same decision as the one originally planned. So 12 out of 80 people or 15% revised their choice. This error rate is not unusual. It is for instance in the same ballpark as the error rates reported in experiments on individual decision making. Harless and Camerer (1994) mention that "direct estimates of error rates, derived by having subjects make the same choice twice without realizing it, suggests a natural error rate of 15-25%" (see also Camerer (1989) and Starmer and Sugden (1989)). More importantly, when subjects revised their choice they did not do so in a systematic direction; 6 subjects decreased their gifts while 6 increased their gifts. So the strategy method did not systematically bias subjects' choices. Therefore, we chose to use the data that we collected with the strategy method, with the only exception that for those people who revised their choices, we replaced the relevant entries in the strategies with their final choices.

Table 2 presents an overview of the choices of players Y conditional on the treatment and the choice of players X. In line with previous experimental results, we find that players Y responded reciprocally to the gifts of players X. Abstracting from the value of the coupon, in the Uninformed-X treatment players Y returned on average an amount of 1.43 when players X gave the coupon versus 0.18 when players X kept the coupon. According to a Wilcoxon rank test, players Y gave significantly more when they received a coupon than when they did not ($p = 0.00$).[16] Likewise, in the Informed-X treatment players Y returned significantly more when players X gave the coupon than when players X kept the coupon (on average 1.79 versus 0.20, respectively, $p = 0.00$). So our data convey the usual signs of intentionality driven reciprocity and inequality aversion driven reciprocity, but our experiment is not designed to distinguish between the two.

[16] All the tests reported in this paper are two-sided tests.

Table 2
Gifts of players Y

	X keeps coupon	X gives coupon=0	X gives coupon=4	All coupons
(1) Uninformed-X	0.18 (0.71)	1.13 (0.97)	1.73 (0.82)	1.43 (0.78)
(2) Informed-X	0.20 (0.76)	1.60 (0.98)	1.98 (0.92)	1.79 (0.86)
Mann-Whitney	0.98	0.03	0.10	0.04

Notes: The cells list the average amounts sent back by players Y (standard deviations in parentheses). The last row presents the p-values of the Mann-Whitney tests comparing the two rows above it.

Now we turn to the main hypotheses of our paper. Recall that hypothesis 1 is on how player Y reacted to the lost opportunities of player X when player X gave the coupon. The Informed-X treatment gives the best evidence for this question, because here player Y's choices are not distorted by the possibility of signaling. In accordance with our modeling of guilt aversion, players Y who received a high-valued coupon gave back an amount of on average 1.98 which exceeded the average reciprocated amount of 1.60 for the cases where they received a low-valued coupon. According to a Wilcoxon rank test, players Y gave significantly more when the value of the coupon was high than when it was low ($p = 0.01$). This rejection of the null hypothesis (that in treatment Informed-X player Y sends back the same amount of money when player X gives up a coupon with a value of 0 as when he gives up a coupon with a value of 4 euros) implies that our subjects' reciprocal behavior is not only motivated by intentions and inequality aversion. This finding is summarized in Result 1.

Result 1: In the Informed-X treatment, players Y reveal aversion to guilt resulting from lost opportunities of players X.

Now we turn our attention to hypothesis 2 concerning whether subjects play in accordance with the logic of a separating equilibrium or a pooling equilibrium in the Uninformed-X treatment. A necessary condition for a separating equilibrium is that player Y sends back a different amount when the value of the coupon is high than when it is low. Table 2 shows that in the Uninformed-X treatment, players Y send back a higher amount (1.73) when the value of the coupon is high than when it is low (1.13). The difference is significant according to a Wilcoxon rank test ($p = 0.00$). Notice that separation occurs in a way as one would naturally expect, that is, receivers of a low-valued coupon send back less money

than receivers of a high-valued coupon. This is summarized in Result 2.

Result 2: The necessary condition for the natural separating equilibrium is fulfilled. Specifically, in the Uninformed-X treatment, players Y send back a lower amount when the value of the coupon is low than when it is high.

Theoretically, as implied by Proposition 1, in a separating equilibrium of the Uninformed-X treatment, players Y who receive a high-valued coupon reciprocate to the same extent as they would when players X were informed of the value of the coupon. The data provide guarded support for this prediction. In fact, players Y who received a high-valued coupon send back somewhat less when player X was not informed than when he was informed (1.73 versus 1.98). The difference is rather small though. A Mann-Whitney test does not reject the null hypothesis players Y send back the same amount at the conventional level ($p = 0.10$). Because the test result comes close to significance, we also ran the corresponding t-test. The t-test produces a p-value of $p = 0.20$, further away from the threshold of significance. Although the data do not provide overwhelming support that there is no difference whatsoever in reciprocal behaviors of players Y with a high-valued coupon in the two treatments, the null hypothesis that player Y's behavior is unaffected by the treatment variable when she received a high-valued coupon is not rejected. This is the core of Result 3.

Result 3: In accordance with the separating equilibrium, the hypothesis that players Y who observe a high-valued coupon send back the same amounts in the two treatments is not rejected.

The fact that under private information of the coupon realization players Y who receive a high-valued coupon on average return a somewhat smaller gift than under full information is suggestive of the possibility that a fraction of recipients of the high-valued coupon attempt to mimic those who receive a low-valued coupon. The data are not sufficiently conclusive to draw a firm inference here, however.

According to the theoretical model, in a separating equilibrium signaling may change the behavior of players Y in the Uninformed-X treatment when the value of the coupon is low. Players Y who receive a low-valued coupon may distort their preferred allocation to underline the fact that they did receive a low-valued coupon instead of a high-valued coupon. Indeed, the data suggest that this actually happened in the experiment. Receivers of a low-valued coupon reciprocated less in the Uninformed-X treatment than in the Informed-X treatment (on average 1.13 versus 1.60). The difference is economically meaningful

and statistically significant according to a Mann-Whitney rank test at $p = 0.03$. (The corresponding t-test also rejects the null-hypothesis at $p = 0.03$). This finding is noted in result 4.

Result 4: In accordance with the separating equilibrium, players Y who observe a low-valued coupon reciprocate to a lesser extent in the Uninformed-X treatment than in the Informed-X treatment.

These results show that players Y were sensitive to the opportunities that players X lost when they gave up their coupon. It is not the case, however, that all players Y were affected in the same way by the lost opportunities of players X. In fact, like in other experiments that investigate social preferences, we observe substantial heterogeneity in how players Y behaved. We divided the players Y into different classes of social preferences. The results of this exercise are reported in Table 3.

Table 3
Types of players Y

	Uninformed-X	Informed-X	Pooled
Selfish	5%	10%	7.5%
Altruistic	0%	5%	2.5%
Reciprocal, intentionality/allocative	45%	45%	45%
Reciprocal, guilt	32.5%	12.5%	22.5%
Reciprocal, all factors	15.0%	20%	17.5%
Unclassified	2.5%	7.5%	5%

Notes: Types are defined as follows (notations used in the following should be obvious). Selfish: gift[keep] = gift[coupon=0] = gift[coupon=4] = 0; Altruistic: gift[keep] = gift[coupon=0] = gift[coupon=4] > 0; Reciprocal, intentionality/allocative: gift[keep] < gift[coupon=0] = gift[coupon=4]; Reciprocal, guilt: gift[keep] ≥ gift[coupon=0] and gift[coupon=0]< gift[coupon=4]; Reciprocal, all factors (intentionality/allocative and guilt): gift[keep] < gift[coupon=0] < gift[coupon=4];

Pooled across two treatments, 40% (22.5%+17.5%) of the players Y gave back more money when the coupon had a high value than when it had a low value to player X. This fraction is only slightly smaller than the fraction of subjects who behaved reciprocally without responding to the value of the coupon (45%). A remarkable result is that we have only few players Y classified as being selfish (7.5%). In their survey, Fehr and Gächter (2000) report that 34%-60% of subjects behave in a selfish way. Our conjecture is that having people observe with whom they are matched makes them act in a more social way. Our results may be more representative for real world situations where people are usually informed about with whom they interact.

Given the responses of player Y, player X would increase his material payoff by giving up the coupon. Players X who gave the coupon earned on average 8.23 euros while those who kept the coupon earned on average 6.86 euros. These earnings are pooled across treatments because players X did not yet know whether they would participate in Informed-X or Uninformed-X when they made their decisions. Players X tended to behave in agreement with maximizing their material payoffs. The fraction of players X who kept the coupon was exactly 35% in each of the two treatments.

4.2 Reported guilt and reciprocal behavior

In the post-experimental questionnaire, we asked subjects how they felt about certain situations that might or might not have occurred to them. For example, we asked players Y to state on a 7-point scale to what extent they agreed with the statements:

(1) If player X gives the coupon to player Y, it is only fair that player Y gives some money to player X.

(6) Consider the following case (that may or may not have happened to you). Player X gives her/his coupon to you. To what extent would you feel guilty if you do not give any money to player X?

(9) Consider the following case (that may or may not have happened to you). Player X gives her/his coupon to you. To what extent would you feel ashamed if you do not give any money to player X?

(12) Consider the following case (that may or may not have happened to you). Player X gives her/his coupon to you. To what extent do you think that player X will feel irritated if you do not give any money to player X?

Not surprisingly, it turns out that subjects' answers to these four questions are highly correlated. In fact, a factor analysis on the answers of these questions extracted only one factor with an eigenvalue greater than 1.[17] The extracted factor revealed the highest correlation with question 6 eliciting a subject's feelings of guilt, which confirms the intuition behind our model that guilt aversion is a driving force of reciprocal behavior of players Y. The result of the factor analysis motivated our choice to work with a variable that is simply the average of the subjects' reported answers to questions 1, 6, 9 and 12, and to call this variable "guilt".

We asked a similar set of questions for the case where the subject observed a coupon of value 0 and

[17] We apply the Kaiser criterion here, which prescribes to retain only factors with eigenvalues greater than 1. In essence this means that a factor is dropped unless it extracts at least as much as the equivalent of one original variable. This criterion is probably the one most widely used in factor analysis.

the case where the subject observed a coupon of value 4. For each case, factor analysis extracted only one factor with an eigenvalue larger than 1. Again, we constructed variables that equal the average of subjects' answers, and we refer to them as "guilt0" and "guilt4", respectively. Table 4 lists some aggregate guilt statistics for each treatment.

Table 4
Guilt reported by players Y

	guilt	guilt0	guilt4	Wilcoxon
(1) Uninformed-X	5.29 (1.42)	4.66 (1.64)	5.65 (1.38)	0.00
(2) Informed-X	5.18 (1.19)	4.55 (1.59)	5.41 (1.31)	0.00
Mann-Whitney	0.41	0.70	0.28	

Notes: The cells list the average guilt reports by players Y (standard deviations in parentheses). The last row presents the p-values of the Mann-Whitney tests comparing the two rows above it. The final column presents p-values of the Wilcoxon tests comparing the two columns before it.

The data confirm the role played by the factor of anticipated guilt. Subjects reported that they anticipated more guilt when they would not reciprocate a high-valued coupon than when they would not reciprocate a low-valued coupon (see the last column of Table 4). In each treatment this effect is significant according to a Wilcoxon rank test. Mann-Whitney rank tests do not reveal any difference between two treatments with respect to the reported feelings of anticipated guilt (see the last row of Table 4).

The important question is whether these feelings carried over to how players Y behaved when they made their choices about how much to give back to players X. To investigate this, we computed the difference between the gift that player Y chose for the case where she received a coupon of 4 euros minus the gift that she chose for the case where she received a coupon of 0 euro. In the bulk of cases, this difference equaled 0 or 1 (67 out of 80). Thus we construct a binary variable difgift04 with the value 0 if the difference was smaller than or equal to 0 and 1 if the difference was at least 1. Table 5 reports the results of a logistic regression in which we regress difgift04 on the difference in reported guilt for not reciprocating a coupon of 4 versus a coupon of 0 euros ("guilt4-guilt0"), a subject's major, the treatment and the gender. Notice that the coefficient of "guilt4-guilt0" is positive and significant. Indeed, people who felt more guilty for not reciprocating high-valued coupons than for low-valued coupons tended to be the ones that sent back higher amounts when they received a coupon of 4 euros instead of 0 euro.

In other words, it is consistent with our conjecture that guilt drives the reciprocating behaviors of Y.

Table 5
Logistic regression on difference in gift between high and low-valued coupons

	estimate (standard deviation)	p
guilt4-guilt0	0.53 (0.22)	0.02
econ	0.90 (0.52)	0.08
treat	0.65 (0.50)	0.19
gender	0.69 (0.58)	0.23
χ^2 (4); p=0.01		

Notes: dependent variable difgift04=0 if gift[coupon=4] – gift[coupon=0] ≤ 0; difgift04=1 if gift[coupon=4] – gift[coupon=0] > 0; guilt4-guilt0 represents the average reported guilt for a 4 euro coupon minus the average reported guilt for a 0 euro coupon; treat=0 for Informed-X and treat=1 for Uninformed-X; gender=0 for females and gender=1 for males.

We also elicited how players X felt about the case where player Y did not reciprocate anything. For example, we asked the question:

(6) Consider the following case (that may or may not have happened to you). You give the coupon to player Y. Player Y does not give any money to you. To what extent would you feel irritated in this case?

Notice that this question corresponds to question (12) of player Y. For the irritation question, players X reported an average of 5.45 while players Y anticipated an irritation level of on average 5.74. The difference is not significant according to a Wilcoxon rank test ($p = 0.26$). For none of the comparisons of the other questions we get a significant difference. Thus, players X did not report to experience stronger emotions than players Y anticipated.

Finally, we briefly explore to what extent gender differences play a role in the experiment. In an unobtrusive way our design allowed subjects to find out about the gender of the person they were matched with. We first discuss how gender affected trust. Table 6 reports the decisions to keep the coupon for each gender combination of player X and player Y.

Table 6
Trust by gender

	male Y	female Y	$\chi^2(1)$
male X	35.3% *34*	28.6% *14*	0.65
female X	45.5% *22*	20.0% *10*	0.17
$\chi^2(1)$	0.45	0.64	

Notes: The cells list the % of keep decisions by player X; in italics the number of pairs in the particular cell is listed. The last row (column) presents the p-values of the $\chi^2(1)$ tests comparing the two rows (columns) above (before) it.

A first glance at the table suggests that females tended to trust males less often than they trusted females. This difference in trust, although economically substantial, fails to reach a significant level. A $\chi^2(1)$ test does not reject the null hypothesis that the two relevant entries are equal ($p = 0.17$). Likewise, males did not distinguish between a female and a male when it came to trust. Neither was it the case that male (or female) players Y were trusted differently by male players X or female players Y. That is, the comparison of different rows in the table also fails to reach significance.

Neither did we find any gender effects in how trustworthy players Y were. Female players Y did not distinguish between male and female players X, and neither did male players Y. Female players X were not rewarded differently than male players X by any of the sexes.[18] The relevant statistics are presented in Table 7. These results agree with Camerer ((2003), p. 64) who surveys the literature and concludes

[18] Even though the gender differences are quite far from significant, some of them look fairly large. It is possible that with a larger sample some of the differences would become significant.

that "the mixed effects suggest that gender does not have a simple main effect on social preferences".[19]

Table 7
Trustworthiness by gender

	male X	female X	Mann-Whitney
male Y	1.76 (0.63) *34*	1.48 (0.96) *22*	0.19
female Y	1.64 (0.77) *14*	1.30 (1.21) *10*	0.38
Mann-Whitney	0.65	0.55	

Notes: The cells list average gift of player Y when player X gave the coupon (gift[coupon=0]+gift[coupon=4])/2); standard deviations in parentheses; in italics the number of pairs in the particular cell is listed. The last row (column) presents the p-values of the Mann-Whitney tests comparing the two rows (columns) above (before) it.

5 Conclusion

In this paper we identified an independent effect of guilt on reciprocity after controlling for the concern for allocative equity and first mover's intention. To isolate such effect, we investigated a gift exchange game with these properties: (i) the first mover decides his gift to the second mover without knowing the value of the gift to himself, (ii) the second mover's benefit from the gift is independent of its value to the first mover, and (iii) the second mover observes first mover's valuation of the gift before she chooses the size of the gift she returns to the first mover.

To avoid feeling guilty, second movers tend to reciprocate more when *ex post* the gift turns out to be more valuable to the first mover. When the first mover remains uninformed of the value of the gift to him, a psychological signaling game results where the second mover may attempt to influence the first mover's belief of the state of the world by sending a costly signal of the value of the gift. We characterized the pooling and separating equilibria of this game.

In a pooling equilibrium, the second mover sends back the same amount when the value of the gift to the first mover is low as when it is high. Thus, the state of the world remains unknown to the first mover. In a separating equilibrium, the second mover may choose to send back a lower than preferred amount when the value of the gift is low to convince the first mover of the true state of the world. In contrast, when the value of the gift to the first mover is high, the second mover sticks to the gift that she

[19] More recently, Bellemare and Kröger (2007) draw the same inference from a representative sample of the Dutch population and a survey of the literature.

would choose when the first mover were actually informed of the state of the world. The experimental results reported in this paper are in line with the predictions of the separating equilibrium.

References

[1] Battigalli, P. and M. Dufwenberg, 2005: "Dynamic Psychological Games", IGIER working paper 287, Bocconi University.

[2] Battigalli, P. and M. Dufwenberg, 2007: "Guilt in Games", *American Economic Review*, **97**, 170-176.

[3] Baumeister, R.F., A.M. Stillwell and T.F. Heatherton, 1994: "Guilt: An Interpersonal Approach", *Psychological Bulletin*, **115**, 243-267.

[4] Bellemare C. and S. Kröger, 2007: "On Representative Social Capital", *European Economic Review*, **51**, 183-202.

[5] Bolton, G. and A. Ockenfels, 2000: "ERC: A Theory of Equity, Reciprocity and Competition", *American Economic Review*, **90**, 166-193.

[6] Bosch-Domènech A., and J. Silvestre, 2005: "Elicitation in the Lab. Does Embedding Matter?", Discussion Paper.

[7] Brandts, J. and G. Charness, 2000: "Hot vs. Cold: Sequential Responses and Preference Stability in Experimental Games", *Experimental Economics*, **2**, 227-238.

[8] Brandts, J. and G. Charness, 2003: "Truth or Consequences: An Experiment", *Management Science*, **49**, 116–130.

[9] Brosig, J., Weimann, J. and C. Yang, 2003: "The Hot Versus Cold Effect in a Simple Bargaining Experiment", *Experimental Economics*, **6**, 75-90.

[10] Burlando, R.M. and F. Guyala, 2005: "Heterogeneous Agents in Public Goods Experiments", *Experimental Economics*, **8**, 35-54.

[11] Camerer, C.F., 1989: "An Experimental Test of Several Generalized Utility Theories", *Journal of Risk and Uncertainty*, **2**, 61-104.

[12] Camerer, C.F., 2003: *Behavioral Game Theory.* Princeton University Press, Princeton, New Jersey.

[13] Cason, T. and V. L. Mui, 1998: "Social Influence in the Sequential Dictator Game", *Journal of Mathematical Psychology*, **42**, 248-265.

[14] Charness, G., 2004: "Attribution and Reciprocity in an Experimental Labor Market", *Journal of Labor Economics*, **22**, 665-688.

[15] Charness, G. and M. Dufwenberg, 2006: "Promises and Partnership", *Econometrica*, **74**, 1579-1601.

[16] Charness, G. and M. Rabin, 2002: "Understanding Social Preferences With Simple Tests", *Quarterly Journal of Economics*, **117**, 817-869.

[17] Dufwenberg, M. and G. Kirchsteiger, 2004: "A Theory of Sequential Reciprocity", *Games and Economic Behavior*, **47**, 268-298.

[18] Engelmann, D. and M. Strobel, 2004: "Inequality Aversion, Efficiency and Maximin Preferences in Simple Distribution Experiments", *American Economic Review*, **94**, 857-869.

[19] Falk, A., E. Fehr and U. Fischbacher, 2003: "On the Nature of Fair Behavior", *Economic Inquiry*, **41**, 20-26.

[20] Falk, A. and U. Fischbacher, 2006: "A Theory of Reciprocity", *Games and Economic Behavior*, **54**, 293-315.

[21] Falk, A. and M. Kosfeld, 2006: "The Hidden Costs of Control", *American Economic Review*, **96**, 1611-1630.

[22] Fehr, E. and S. Gächter, 2000: "Fairness and Retaliation: the Economics of Reciprocity", *Journal of Economic Perspectives*, **14**, 159-181.

[23] Fehr, E. and K. Schmidt, 1999: "A Theory of Fairness, Competition, and Cooperation", *Quarterly Journal of Economics*, **114**, 817-868.

[24] Fehr, E. and K. Schmidt, 2003: "Theories of Fairness and Reciprocity - Evidence and Economic Applications", in: Dewatripont, M. et.al., *Advances in Economics and Econometrics*, Eighth World Congress of the Econometric Society, **1**, Cambridge: Cambridge University Press, 208-257.

[25] Geanakoplos, J., D. Pearce and E. Stacchetti, 1989: "Psychological Games and Sequential Rationality", *Games and Economic Behavior*, **1**, 60-79.

[26] Güth, W., Huck, S. and W. Müller, 2001: "The Relevance of Equal Splits in Ultimatum Games", *Games and Economic Behavior*, **37**, 161-169.

[27] Harless, David W. and Colin F. Camerer, 1994: "The Predictive Utility of Generalized Expected Utility Theories", *Econometrica*, **62** (2), 1251-1289.

[28] Hopfensitz, A. and E. Reuben, 2006: "The Importance of Emotions for the Effectiveness of Social Punishment", Discussion Paper.

[29] Ketelaar, T. and W.T. Au, 2003: "The Effects of Feelings of Guilt on the Behaviour of Uncooperative Individuals in Repeated Social Bargaining Games: An Affect-As-Information Interpretation of the Role of Emotion in Social Interaction", *Cognition and Emotion*, **17**, 429-453.

[30] Levine, D., 1998: "Modelling Altruism and Spitefulness in Experiments", *Review of Economic Dynamics,* **1**, 593-622.

[31] Offerman, T., 2002: "Hurting Hurts More Than Helping Helps", *European Economic Review,* **46**, 1423-1437.

[32] Offerman, T., Sonnemans, J. and A. Schram, 1996: "Value Orientations, Expectations, and Voluntary Contributions in Public Goods", *Economic Journal*, **106**, 817-845.

[33] Oxoby, R. and K. McLeish, 2004: "Sequential Decision Strategy Vector Methods in Ultimatum Bargaining: Evidence on the Strength of Other-Regarding Behavior", *Economics Letters*, **84**: 399-405.

[34] Rabin, M, 1993: "Incorporating Fairness into Game Theory and Economics", *American Economic Review*, **83**, 1281-1302.

[35] Regner, T. and J.A. Barria, 2007: "Do Consumers Pay Voluntarily? The Case of Online Music", *Discussion Paper.*

[36] Ruffle, B, 1999: "Gift Giving with Emotions", *Journal of Economic Behavior and Organization*, **39**, 399-420.

[37] Starmer, C., and R. Sugden, 1989: "Probability and Juxtaposition Effects: An Experimental Investigation of the Common Ratio Effect", *Journal of Risk and Uncertainty*, **2**, 159-178.

[38] Tangney, J.P., 1995: "Recent Advances in the Empirical Study of Shame and Guilt", *American Behavioral Scientist*, **38**, 1132-1145.

[39] Tangney, J.P. and R.L. Dearing, 2002: *Shame and Guilt.* The Guilford Press: New York.

Appendix

Instructions for the experiment

Please read the following instructions carefully. Everybody will read the same instructions. We will hand out a summary of the instructions on paper after everybody has finished reading the instructions and before the experiment starts. In this experiment, you will be randomly matched with another participant. There are two roles in this experiment: role X and role Y. One of the players of your pair will be assigned role X, the other role Y. When everybody has finished reading the instructions, you will learn whether you are player X or player Y. At that time, you will also be informed of the table behind which the person with whom you are matched is sitting, just like this person will be informed of the table behind which you are sitting. Each pair of players will play only once the game that is explained below. After this game, you will be asked to fill out a questionnaire. Then you will be paid one by one in private. You can leave the laboratory immediately after you have been paid.

Three sources for earning money

There are three sources for earning money. The first is a starting capital that does not depend on your choice. The second and the third depend on your own choice and the choice of the player with whom you are matched. (It is also possible that you lose some money, but this is in your own control). We now explain them one by one.

For the first source, you will be given a starting capital of 5 euros that you do not have to pay back to the experimenter at the end of the experiment.

The second amount that you earn is determined by the player with role X. The player with role X will receive a "coupon" of unknown value. The value of the coupon to player X is 0 euro with probability 0.5 and 4 euros with probability 0.5. The computer will randomly determine the value of the coupon, but the value remains unknown to the players at this stage. Player X decides whether to keep the

coupon or to give it to player Y. When player X makes this decision, neither player is informed of the value of the coupon to player X.

If player X gives the coupon to player Y, player Y will receive 6 euros, irrespective of the value of the coupon to player X. That is, if player X gives the coupon to player Y, player Y will receive 6 euros in case the value of the coupon to player X is 0 euro as well as in the case the value of the coupon to player X is 4 euros. If player X gives the coupon to player Y, s\he will no longer own the coupon. As a consequence, the second amount that s\he earns equals 0 euro.

If player X keeps the coupon, s\he will receive 0 euro (in case the coupon has a value of 0 euro to player X) or 4 euros (in case the coupon has a value of 4 euros to player X). Player Y will then receive 0 euro.

The third amount that you earn is determined by the player with role Y. Player Y has to decide whether to give 0, 1, 2, 3 or 4 euros to player X. The amount player X receives is double the amount player Y gives. This means that if player Y gives 0 euro, her or his payoff and the payoff of player X remain unchanged. If player Y gives 1 euro, her or his payoff decreases by 1 euro and the payoff of player X increases by 2 euros. If player Y gives 2 euros, her or his payoff decreases by 2 euros and the payoff of player X increases by 4 euros, etc.

Summary

To summarize, the total amount that you earn in this experiment is the sum of three amounts. The first amount is your starting capital of 5 euros. The second amount is determined by player X and the third amount by player Y. Below we list the possible choices of X and Y and their consequences on the second and third amount:

choice of X	value of coupon to X	probability	payoff to X	payoff to Y
Keep the coupon	0	0.5	+0	+0
Keep the coupon	4	0.5	+4	+0
Give the coupon	0	0.5	+0	+6
Give the coupon	4	0.5	+0	+6

Choice of Y	payoff to X	payoff to Y
Give 0 euro to X	+0 euro	-0 euro
Give 1 euro to X	+2 euros	-1 euro
Give 2 euros to X	+4 euros	-2 euros
Give 3 euros to X	+6 euros	-3 euros
Give 4 euros to X	+8 euros	-4 euros

Procedure

Before the choice of player X is communicated to player Y, player Y will be asked to construct a plan that specifies what s\he will do in each of the following three possibilities.

1. player X keeps the coupon
2. player X gives the coupon, and its value is 0 euro to player X
3. player X gives the coupon, and its value is 4 euros to player X

Player Y may decide for herself or himself whether to give the same amount in all these three cases, or to give different amounts in different cases.

Only after player Y has constructed this plan, the choice of player X will be communicated to player Y. In addition, player Y will be informed of the value of the coupon to player X in case player X gave the coupon to player Y. (If player X keeps the coupon, the value of the coupon to player X is not revealed to player Y.)

The computer will remind player Y of the choice that s\he had planned for the particular case that actually happened. At this point, player Y will be offered an opportunity to revise her or his previously planned choice at a cost. If player Y changes her or his decision, apart from that fact that the payoffs to players X and Y will be changed due to the change in player Y's action, player Y's payoff will be reduced by 20 euro-cents. If player Y does not change her or his decision, then her or his planned choice will be her or his final decision, and s\he will not incur extra costs.

One of two variants

We will run one of two different variants of the experiment described above. We refer to these variants as treatments A and B. We will only communicate which treatment applies after player X has made her or his choice. We implement each of the two treatments with equal probability. Both treatments use the rules described above, but they differ in the following aspects.

In treatment A, player X will NOT be informed of the value of the coupon to player X if s\he gives the coupon to player Y. In this case, s\he will only receive information about the amount that player Y

gives to her or him. As mentioned before, player Y does learn about the value of the coupon to player X if player X gives the coupon. If player X keeps the coupon, s\he will be the only one to learn the value of the coupon to her or him.

In treatment B, player X will be informed of the value of the coupon to her or him if s\he gives the coupon to player Y. That is, if player X gives the coupon to player Y, both players X and Y are informed of the value of the coupon to player X. If player X keeps the coupon, s\he will be the only one to learn about the value of the coupon to her or him.

Please answer the following questions.

(Note that subjects had to answer all questions correctly before they could proceed. If one or more answers were incorrect, this was communicated to them and they had to answer all questions again without knowing what answer(s) were wrong.)

How much does the payoff of player Y increase if player X gives the coupon to player Y, and the value of the coupon is 4 euros?

Answer: [6] euros.

How much does the payoff of player X increase if player X gives the coupon to player Y, the value of the coupon is 4 euros and player Y gives 4 euros to player X?

Answer: [8] euros.

How much does the payoff of player X increase if player X keeps the coupon, the value of the coupon is 4 euros and player Y gives 3 euros to player X?

Answer: [10] euros.

Suppose player X has given the coupon to player Y and the value of the coupon is 0 euro. Originally, player Y was planning to give 1 euro to player X in this case. However, when player Y is offered the opportunity to revise her or his decision at a cost, s\he chooses to increase the amount s\he gives to player X to 2 euros. What is player Y's total payoff from the game (excluding the starting capital)?

Answer: [3.80] euros.

Post Experiment Questionnaire for player X

Thank you for participating in the experiment. Please answer the following questions seriously. Your answers will help us understanding the finding in this study. The questionnaire is anonymous.

Unless otherwise specified, please answer the following questions on a seven-point scale where "1" indicates that you strongly disagree with the statement, "4" means neutral, and "7" means strongly agree.

☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
Strongly disagree			Neutral			Strongly agree

(All questions had this format.)

(1) I believed that my total payoff would be on average higher if I gave up my coupon for player Y than if I did not, taking into consideration player Y's response to my action.

(2) If player X gives the coupon to player Y, it is only fair that player Y gives some money to player X.

(3) In case player X gives the coupon to player Y, player Y is more likely to give some money to player X if the value of the coupon is 4 euros than if the value is 0 euro.

(4) In case player X gives the coupon to player Y, player Y is responsible in compensating player X.

(5) In case player X gives the coupon to player Y, player Y is responsible in compensating player X more if the value of the coupon is 4 euros than if the value is 0 euro.

(6) Consider the following case (that may or may not have happened to you). You give the coupon to player Y. Player Y does not give any money to you. To what extent would you feel irritation in this case?

(7) Consider the following case (that may or may not have happened to you). You give the coupon to player Y and the value of the coupon is 0 euro. Player Y does not give any money to you. To what extent would you feel irritation in this case?

(8) Consider the following case (that may or may not have happened to you). You give the coupon to player Y and the value of the coupon is 4 euros. Player Y does not give any money to you. To what extent would you feel irritation in this case?

(9) Consider the following case (that may or may not have happened to you). You give the coupon to player Y. Player Y does not give any money to you. To what extent would you feel contempt for player X in this case?

(10) Consider the following case (that may or may not have happened to you). You give the coupon to player Y and the value of the coupon is 0 euro. Player Y does not give any money to you. To what extent would you feel contempt for player Y in this case?

(11) Consider the following case (that may or may not have happened to you). You give the coupon to player Y and the value of the coupon is 4 euros. Player Y does not give any money to you. To what extent would you feel contempt for player Y in this case?

In the experiment, we informed you that there would be two different treatments. If player X gave the coupon, player Y only would be informed of its value in treatment A, while both players would be informed in treatment B.

(12) In comparing treatments (A) to (B), I believe that the high realization of the coupon value would have a larger impact on the amount player Y would give up if the realization is communicated to me as well as player Y.

+ standard background questions about study, gender etc.

Post Experiment Questionnaire for Player Y

Thank you for participating in the experiment. Please answer the following questions seriously. Your answers will help us understanding the findings in this study. The questionnaire is anonymous.

Unless otherwise specified, please answer the following questions on a seven-point scale where "1" indicates that you strongly disagree with the statement, "4" means neutral, and "7" means strongly agree.

☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	☐ 7
Strongly disagree			Neutral			Strongly agree

(All questions had this format.)

(1) If player X gives the coupon to player Y, it is only fair that player Y gives some money to player X.

(2) If player X gives the coupon to player Y and the value of the coupon is 0 euro, it is only fair that player Y gives some money to player X.

(3) If player X gives the coupon to player Y and the value of the coupon is 4 euros, it is only fair that player Y gives some money to player X.

(4) In case player X gives the coupon to player Y, player Y is responsible for compensating player X.

(5) In case player X gives the coupon to player Y, player Y is responsible for compensating player X more if the value of the coupon is 4 euros than if the value is 0 euro.

(6) Consider the following case (that may or may not have happened to you). Player X gives her/his coupon to you. To what extent would you feel guilty if you do not give any money to player X.

(7) Consider the following case (that may or may not have happened to you). Player X gives her/his coupon to you and the realization of the coupon value is 0 euro. To what extent would you feel guilty if you do not give any money to player X.

(8) Consider the following case (that may or may not have happened to you). Player X gives her/his coupon to you and the realization of the coupon value is 4 euros. To what extent would you feel guilty if you do not give any money to player X.

(9) Consider the following case (that may or may not have happened to you). Player X gives her/his coupon to you. To what extent would you feel ashamed if you do not give any money to player X.

(10) Consider the following case (that may or may not have happened to you). Player X gives her/his coupon to you and the realization of the coupon value is 0 euro. To what extent would you feel ashamed if you do not give any money to player X.

(11) Consider the following case (that may or may not have happened to you). Player X gives her/his coupon to you and the realization of the coupon value is 4 euros. To what extent would you feel ashamed if you do not give any money to player X.

(12) Consider the following case (that may or may not have happened to you). Player X gives her/his coupon to you. To what extent do you think that player X will feel irritated if you do not give any money to player X.

(13) Consider the following case (that may or may not have happened to you). Player X gives her/his coupon to you and the realization of the coupon value is 0 euro. To what extent do you think that player X will feel irritated if you do not give any money to player X.

(14) Consider the following case (that may or may not have happened to you). Player X gives her/his coupon to you and the realization of the coupon value is 4 euro. To what extent do you think that player X will feel irritated if you do not give any money to player X.

In the experiment, we informed you that there would be two different treatments. If player X gave the coupon, only player Y would be informed of its value in treatment A, while both players would be informed in treatment B.

(15) Consider the following case (that may or may not have happened to you). Player X gives her/his coupon to you. In comparing treatments (A) to (B), the high realization of the coupon value would have a larger impact on the amount I would give up for player X if the realization is communicated to me as well as player X.

+ standard background questions about study, gender etc.