

# IS MISTRUST SELF-FULFILLING?\*

*ERNESTO REUBEN*

Northwestern University

*PAOLA SAPIENZA*

Northwestern University, NBER, and CEPR

*LUIGI ZINGALES*

University of Chicago, NBER, and CEPR

## ABSTRACT

We study experimentally the effect of expectations on whether trust is repaid. Subjects respond with untrustworthy behavior if they see that little is expected of them. This suggests that guilt aversion plays an important role in the repayment of trust.

Date of first version: May 2008

Date of this version: March 2009

Keywords: trust; trustworthiness; reciprocity; guilt aversion; experiment

JEL classification: C72; C91; Z13

---

\* We thank the Templeton Foundation for financial support.

*Note:* This is the author's version of a work that was accepted for publication in *Economics Letters*. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in <http://dx.doi.org/10.1016/j.econlet.2009.04.007>

## 1. Introduction

The degree to which you expect others to behave in a trustworthy manner is crucial for your willingness to trust. An interesting question, however, is whether your expectation (if manifested) can affect the likelihood that your trust is repaid. The answer to this question has important implications for the stability of trust (or lack thereof).

To observe whether expectations impact the repayment of trust, we ran an experiment based on a variation of the trust game of Berg et al. (1995) called the lost wallet game (Dufwenberg and Gneezy, 2000). In the game, a *sender* is endowed with \$50, which he can keep or send to a *receiver*. If the sender sends (i.e., trusts), the receiver receives \$150 of which she can return any amount to the sender. Subjects play the game twice as receivers. The first time, they do not know the amount the sender expects back. The second time, they are informed of the sender's expectation. Half of them are matched with senders with high expectations and the rest with senders with low expectations. We test whether observing a high/low expectation increases/decreases the amount returned.

In this context, it is useful to consider the theory of guilt aversion (Battigalli and Dufwenberg, 2007) as it assumes a direct impact of the sender's expectations on the receiver's returned amount. Specifically, receivers who return money do so in order to avoid feeling guilty for disappointing the sender. Hence, since senders with low expectations are not disappointed by low returns, facing such senders motivates receivers to return less.

## 2. Design

The experiment was conducted in October 2007 with MBA students from Northwestern University. Average earnings in the game were \$47.39. The detailed experimental procedures are found in Reuben et al. (2008).

To convey the senders' expectations in an incentive-compatible way, subjects are told they will make four decisions, one of which will be paid at random. Moreover, to avoid subjects from influencing each other, they are not informed in-between decisions of the behavior of others, and they are given the instructions for each decision just before they make it.

The first three decisions are the same for all subjects. Decision 1 consists of playing the lost wallet game as a sender. For decision 2, subjects are first explained that receivers make their choice using the strategy method—that is, they choose their return conditional on receiving the \$150—and then, they are asked to predict the amount their receiver returns.<sup>1</sup> Decision 3 consists of playing the same game as receivers using the abovementioned contingent choice. For decision 4, subjects play once again the lost wallet game. Some of them play as senders and the rest as receivers. Senders are told that their receiver will be informed of their expected return (elicited in decision 2) and receivers are told their sender's expectation. The role of sender is given at random to subjects with low<sup>2</sup> or high expectations (less than \$10 or more than \$70).<sup>3</sup>

### 3. Related work

There are a few papers that study the effect of expectations on the repayment of trust. Dufwenberg and Gneezy (2000), Charness and Dufwenberg (2006), and Bacharach et al. (2007) study variations of trust games and measure the receivers' second-order beliefs (i.e., what they believe the sender expects them to return). They find that receivers with high second-order beliefs return higher amounts.

---

<sup>1</sup> Subjects are paid according to their accuracy: they get  $\max\{\$75 - |\text{actual amount returned} - \text{predicted amount returned}|, \$0\}$ .

<sup>2</sup> One might worry that the receivers' choice is irrelevant for senders with low expectations. However, given the various motivations for sending (e.g. altruism), we consider there is always a positive probability that the sender sends. In the experiment, 25% of those expecting back \$10 or less send money.

<sup>3</sup> Subjects are unaware that decision 2 affects future roles. They are simply informed of the role they are assigned. A similar procedure is used by, for example, Gächter and Thöni (2005).

The study that comes closest to ours is Ellingsen et al. (2008), hereafter EJTT. The authors use variations of trust games in which the senders' expectations are communicated to the receivers.<sup>4</sup> They do not find a significant correlation between revealed expectations and amounts returned, and suggest that the abovementioned result of second-order beliefs is driven by the "false consensus effect".

We should point out that both EJTT's design and ours have common limitations. First, we cannot be certain that senders are revealing their true expectation. In particular, if senders foresee that their expectation will be revealed, they might submit an untruthful expectation in order to manipulate the behavior of receivers.<sup>5</sup> Second, receivers might think that senders are submitting an untruthful expectation. Both cases work against finding support for guilt aversion.

Compared to EJTT, we take the following steps to lessen any effects introduced by these limitations. First, we pay senders considerable amounts for the accuracy of their expectation, which makes it unlikely that they can increase their earnings by manipulating receivers.<sup>6</sup> Second, we give all subjects the same instructions and sequence of tasks to avoid triggering skepticism by receivers—that is, subjects that reveal their true expectation have little reason to believe others do not. In contrast, EJTT tell receivers that senders are uninformed that their expectation will be revealed precisely when they are in fact revealing it. This procedure makes this omission salient, which can easily make receivers wonder if senders anticipated it. Third, we use within-subject comparisons, and therefore, we can be confident that receivers who *change* their return must believe true expectations are at least partially revealed (receivers that think senders are manipulating expectations should simply ignore them).

---

<sup>4</sup> Senders are not informed that receivers see their expectation.

<sup>5</sup> In EJTT expectations might also be affected by hedging, which we avoid by paying only one decision.

<sup>6</sup> For manipulation to work, senders must believe that a dollar deviation in their expectation will increase the amount returned by more than a dollar.

## 4. Results

For the statistical analysis, we concentrate on the 52 subjects who were receivers in decisions 3 and 4. In the latter, 24 (28) subjects faced a sender with a low (high) expectation. Overall, 67.3% of the subjects send money as senders. Their mean expected return is \$56.67, which is close to the actual mean return in decision 3: \$52.88.

The senders' expectations—observed in decision 4—produce a stark difference in the amount returned. On average, receivers who observe a low expectation decrease their return from \$53.75 to \$40.63. The decrease is significant with a Wilcoxon-signed-ranks test ( $p = 0.030$ ). A majority of them, 54%, decrease the amount returned, 33% keep it constant, and 13% increase it.<sup>7</sup>

In contrast, receivers who observe a high expectation increase their return from \$52.14 to \$58.29. The increase, however, is not significant with a two-sided Wilcoxon-signed-ranks test ( $p = 0.125$ ). In this case, 61% of the receivers keep their return constant, 28% increase it, and 11% decrease it.

The difference in amounts returned is clearly seen in Figure 1, which reports the cumulative distribution functions of the returned amounts in the three situations (the “unknown expectations” line is the CDF of decision 3). It can be seen that the CDF under low expectations first-order stochastically dominates the CDF under unknown expectations, which in turn first-order stochastically dominates the CDF under high expectations.

---

<sup>7</sup> This effect seems to be due to the sender's expectation and not to receivers typing the number on the screen. Of those who change their amount returned, only 30% return the observed expectation. Thus, a considerable majority are clearly thinking of how to adjust their decision and are not blindly complying with the sender's expectation or using it as an anchor. EJTT directly test the anchoring hypothesis and find no effect.

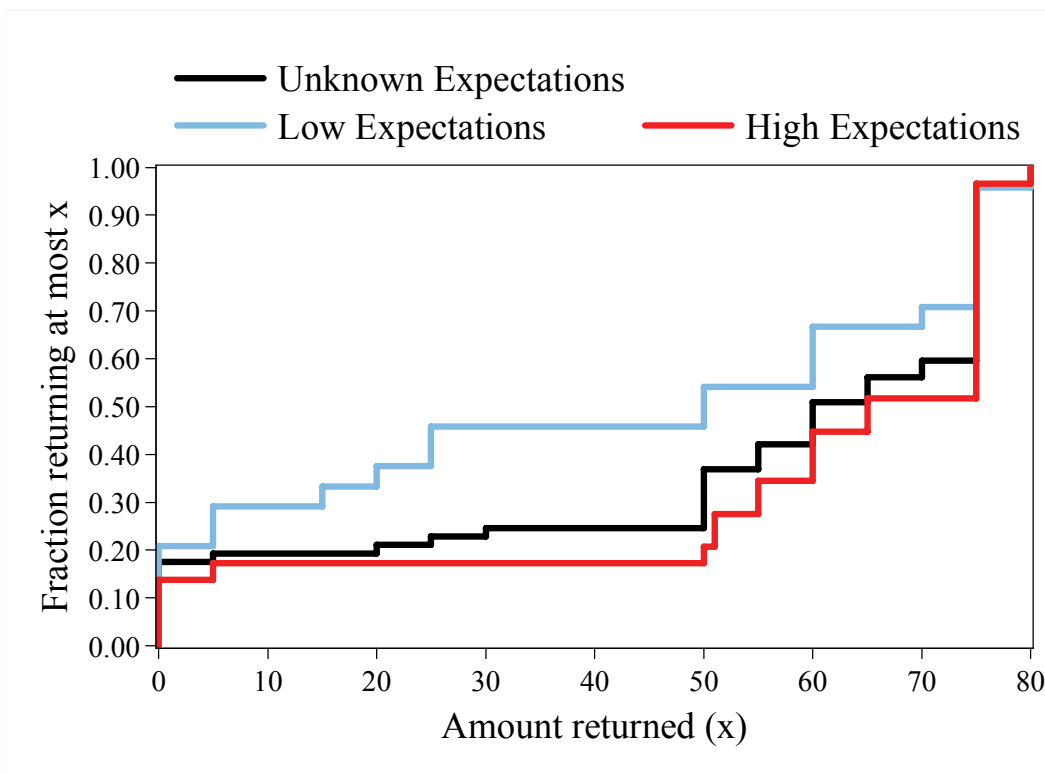


FIGURE 1 – CUMULATIVE DISTRIBUTIONS OF RETURNED AMOUNTS

To observe the effect of the senders' expectations whilst controlling for other variables, we run Tobit regressions with the amount returned in decision 4 as the dependent variable (see Table 1). We censor at \$0 since negative returns are not possible and at \$75 since most receivers seem to consider this as an upper bound to the amount returned (36.5% return exactly \$75, 59.6% return less, and only 3.9% return more).

In column I we use four independent variables: a dummy variable indicating whether the subject sees a low expectation, the amount returned by the subject under unknown expectations, a dummy variable indicating whether the subject sent money in the sender role, and the subject's expectation of the amount returned by others.

Compared to a sender with high expectations, playing with a mistrusting sender is associated with a \$31.77 lower return—a striking 43% difference. The subjects' expectation of the amount returned by others is also significant: the more they expect others to return, the more they return themselves. Finally, the amount returned under unknown expectations has a weakly positive effect.

TABLE 1 – ESTIMATION OF TRUSTWORTHINESS

	I	II
Low expectations	-31.77** (13.66)	26.09 (27.91)
Amount returned with unknown expectations	0.51* (0.29)	0.97*** (0.35)
Sent money	2.96 (16.82)	3.59 (15.92)
Expected money returned	1.13** (0.49)	1.19** (0.48)
Amount returned with unknown expectations × low expectations		-1.11** (0.49)
Constant	-17.1 (23.96)	-44.74 (27.65)

*Note:* Standard errors in parentheses, \*, \*\*, \*\*\* indicate statistical significance at 10%, 5%, and 1%.

In column II we run the same regression including the only significant interaction term: the interaction between observing a low expectation and the subject's return under unknown expectations. In this regression, the dummy variable for low expectations loses significance, which is picked up by the interaction variable. The regression results suggest that the effect of observing a low expectation is stronger the higher the amount returned under unknown expectations. This can be seen as consistent with the theory of guilt aversion in the sense that, once we control for the subjects' second-order beliefs,<sup>8</sup> the amount returned serves as a proxy for their guilt sensitivity, and the more sensitive to guilt subjects are, the more they should react to a the sender's low expectation.

<sup>8</sup> For this interpretation we must assume that a subject's expected amount returned strongly correlates with their belief of the expected amount returned of other's.

## 5. Conclusions

In this paper we study the effect of the senders' expectations on the amount returned in a variation of a trust game. In line with the theory of guilt aversion, we show that most receivers respond with untrustworthy behavior when they face a sender with low expectations. Hence, motivations such as reciprocity and altruism are insufficient to explain returns. Receivers are also driven by a desire to comply with the sender's expectation.

These results suggest that expecting others to behave untrustworthily (if manifested) causes untrustworthy behavior, which further confirms the initially low expectation. This makes mistrust self-fulfilling.

## References

- Berg, J., Dickhaut, J., and McCabe, K. (1995). Trust, Reciprocity and Social History. *Games and Economic Behavior* 10: 122-142.
- Bacharach, M., Guerra, G., and Zizzo, D.J. (2007). The Self-Fulfilling Property of Trust: An Experimental Study. *Theory and Decision* 63: 349-388.
- Battigalli, P. and Dufwenberg, M. (2007). Guilt in Games. *American Economic Review Papers & Proceedings* 97: 170-176.
- Charness, G., and Dufwenberg, M. (2006). Promises and Partnership. *Econometrica* 74: 1579-1601.
- Dufwenberg, M. and Gneezy, U. (2000). Measuring Beliefs in an Experimental Lost Wallet Game. *Games and Economic Behavior* 30: 163-182.
- Ellingsen, T., Johannesson, M., Tjøtta, S., and Torsvik, G. (2008). Testing Guilt Aversion. SSE/EFI Working Paper.
- Gächter, S. and Thöni, C. (2005). Social Learning and Voluntary Cooperation Among Like-Minded People. *Journal of the European Economic Association* 3: 303-314.
- Reuben, E., Sapienza, P., and Zingales, L. (2008). Is Mistrust Self-fulfilling? Supplementary Material. Mimeo.