

Separating Gratitude from Guilt in the Laboratory*

Hong (Hannah) Lin¹ and David Ong²

ABSTRACT

In contrast to guilt based reciprocity, which hypothesizes that reciprocity is an increasing function of the 2nd order expectation of trustor's expectation for reciprocation, we test for reciprocity which is a decreasing function of trustees 2nd order expectations, i.e., that people can reciprocate out of gratitude. To unambiguously decrease 2nd order expectations in our treatment, we broke up a standard trust game into a two stage dictator game where the 1st round dictator was not informed about the possibility of a 2nd round. Furthermore, the 2nd dictator could "silently exit" by not sending anything to the 1st round dictator. We found a significant increase in both the amount of reciprocation and the number of people reciprocating as compared to our standard trust and dictator games controls. The strength of the positive reciprocity observed rules out inequality aversion as a cause. We found support for our hypothesis in the prior data of others who tested for guilt based reciprocity. Our result also seems to reconcile conflicting results in that literature. To our knowledge, this is the first paper which shows that kindness distinct from guilt, shame, efficiency and inequity aversion could be a motive for reciprocity. Our strong positive reciprocity result also suggests why it has been difficult to find in the past.

JEL Codes: C91, C72, D64

Keywords: reciprocity, guilt, gratitude, trust game, double-blind, silent exiting

¹ Email: hlin05@gmail.com, Tel: +86-136-3263-7246, Fax: +86-755-2603-5344, Address: N1007, PKU, Univ.Town, Shenzhen, P.R.C, 518055.

² Corresponding author email: dvdong@gmail.com, Tel: +86-755-2603-2655, Fax: +86-755-2603-5344, Address: C402, PKU, Univ.Town, Shenzhen, P.R.C, 518055.

* This research was funded by the Dean's Fund No.2009039, Peking University Shenzhen Graduate School. The author gratefully acknowledge the helpful comments of Luz Marina Arias, Zhuoqiong (Charlie) Chen, Zemin (Zachary) Zhong, Tore Ellingsen, Ernesto Reuben, Wooyong Lim, Alexander Smith, the participants in our session at the Economic Science Association North American Meeting, Chicago, July 2011.

I. INTRODUCTION

The importance of reciprocity has long been recognized in economics, particularly in contracting when institutions of legal enforcement are underdeveloped or absent, e.g., in emerging markets, or in social settings where contracts are implicit and informal. Prior experimental literature has largely demonstrated that people are not entirely selfish (in the colloquial sense of the word). However, as Charness and Rabin (2002) pointed out, evidence for positive reciprocity, where recipients of favor respond more generously than dictators with the same amount of endowment, has been elusive. They regarded McCabe et al., (2003) as a notable exception in the literature. However, even in that experiment, inequity aversion and psychological motives like guilt and shame, which are not obviously forms of positive reciprocity, are still possible confounds. These psychological confounds could also explain why positive reciprocity has been difficult to find in other experiments where beliefs had generally not been the focus of the experiment and hence, were not well controlled. In standard gift exchange or trust experiments, recipients could believe that the sender gave in order to incite reciprocation. In that case, the recipients might have doubts about the senders' generous motives and hence, might feel that positive reciprocity was not warranted. Thus, we believe that a test of positive reciprocity would have to control the 2nd order beliefs of recipients. We focused upon reciprocating behavior that is a decreasing function of those 2nd order expectations. This is our interpretation of what is commonly meant by "gratitude"³ or "thankfulness." In contrast, the current literature on belief based reciprocity has largely focused on "guilt," which has

³ We are grateful to Luz Marina Arias for pointing out that some people are more likely to reciprocate if they knew that a gift giver had not been expecting anything in return.

been modeled as an increasing function of the beliefs of the receiver/trustee in gift exchange/ trust games.

In this somewhat extended introduction and literature review, we discuss the problems and anomalies in this literature first in order to motivate both our hypothesis of gratitude and our somewhat radical methodology for measuring it. We argue that these largely center around the difficulties of controlling 2nd order beliefs when players are fully aware of the game structure. We also show that though prior work tested for an alternative hypothesis, we are able to find support for our hypothesis in their data. We then discuss the implications of our study for why positive reciprocity has been difficult to find.

The guilt based reciprocity literature hypothesizes that people reciprocate out of disappointment aversion. Charness and Dufwenberg (2006; hereafter, CD) showed that the making of cheap talk promises prior to trust games were correlated with the average giving of trustors, and the average reciprocation by trustees. Furthermore, average trustees' 2nd order belief in reciprocation is correlated with average reciprocation by trustees. However, CD did not reveal the actual beliefs of trustors to the trustees with whom they were paired. Thus, they could not establish causation. Ellingsen et al. (2010) in an attempt to establish causation, revealed the actual belief of each trustor to a paired trustee. They found that the trustor's belief had no significant effect on the paired trustee's reciprocation, seemingly refuting the guilt based reciprocity hypothesis. Particularly interesting for our gratitude hypothesis was the fact that the majority of trustees who faced the most pessimistic beliefs of trustors reciprocated the most. This suggests a counter-guilt factor consistent with our gratitude based reciprocation hypothesis.

Reuben et al. (2009) tested the guilt based reciprocation hypothesis with a different design. To get trustors to truthfully reveal their expectations to trustees, Ellingsen et al. (2010) did not tell trustors that their beliefs would be revealed to trustees. To get trustees to believe that trustors had been truthful, they told the fact of the omission to trustees. Reuben et al. (2009) did not use asymmetric instructions, which might cause trustees to question whether the experimenter had been forthright with them as well. Subjects played twice in each role of sender and receiver. Reuben et al. (2009) elicited subjects' beliefs when they were senders during a round when beliefs were unused, and then used them in the 2nd round. Reuben et al. (2009) also used significantly higher payments for both belief elicitation and for actual play. They did find a significant correlation between revealed beliefs of trustors and the reciprocating behavior of the paired trustees. However, even here when beliefs were well controlled, about 10% of subjects went in the opposite direction predicted by guilt in each of their pessimistic and optimistic belief treatments.

Stanca et al.'s (2007) study of intrinsic motivation suppressed 2nd order beliefs in a non-equilibrium setting. In their setup, the 1st player was not informed of the possibility of reciprocation in the 2nd stage. Hence, their 2nd player could not discount the kindness intent of 1st player giving as being merely done in anticipation of reciprocation. They found significantly increased positive reciprocation as compared to their full information control for their strategy elicitation (though not for their decision method elicitation). However, with respect our hypothesis, this design does not indentify gratitude since guilt is still a possible confound. They controlled for guilt from ex-ante but not ex-post expectations. The 1st dictator was informed of the game structure in the 2nd stage. Guilt of the 2nd player was still possible from the disappointment of *ex-post expectations*. Hence, aversion to possible disappointment could still drive reciprocation. Furthermore,

unrelated to their hypothesis, but relevant to our gratitude hypothesis, people could reciprocate from shame. There was no “silent exiting” (Dana et al., 2006). Subjects could not give 0 or low and not be noticed. The design was also not double-blind. Subjects might be motivated by shame from experimenter or other subjects knowing about non-reciprocation (Hoffman et al., 1994, 1996).

We adopted a design similar to Stanca et al., but had more controls for the above mentioned confounds, and found strong evidence for positive reciprocity. In contrast to Stanca et al., the proportion of subjects who reciprocated, as well as the proportion of endowment given back by our recipients who received a positive transfer, was significantly higher than both our dictator and standard trust games control.

Furthermore, our results for this apparent positive reciprocity were less likely to have been motivated by guilt as might have been the case in Stanca et al. or inequity aversion, efficiency, or shame, as might have been the case in (McCabe et al., 2003). As in Stanca et al., we controlled for guilt from anticipation of disappointing ex-post expectation by also not informing the 1st round dictator of the possibility of a 2nd round. However, we also further diminished the possibility of guilt from ex-post beliefs by permitting implicit silent exiting with our payment procedure. Inequity aversion as the motive for reciprocation was ruled out in the results.

The fact that trustees gave significantly more when they were less sure that they were expected to reciprocate suggests that guilt crowds out gratitude. This could help explain the conflicting results in Ellingsen et al. 2010 and Reuben et al. 2009, as well as the scarcity of results on positive reciprocity in prior experiments mentioned by Charness and Rabin (2002). First, the opposition between these belief based incentives of guilt and

gratitude could make guilt experiments very sensitive to the beliefs of subjects. Similar experiments are likely to yield conflicting results for even slightly different setups. Secondly, if guilt can crowd out gratitude, trustees could discount the generosity of trustors giving as motivated by anticipation of reciprocation. This could help explain the paucity of evidence for positive reciprocity.

II. EXPERIMENTAL DESIGN

We highlight the novel part of the design first before going into details. We used the decision method of elicitation. To avoid strategic behavior, we gave our subjects instructions step-by-step. We broke the trust game into a two stage dictator game. The 1st stage dictator (1st dictator) was not informed of the 2nd stage to prevent the 1st dictator from giving to incite guilt, and also to prevent the 2nd dictator believing that it was the 1st dictator's motive in giving. The 2nd stage dictator (2nd dictator) was informed of the game structure and the 1st dictator's lack of knowledge of the 2nd round, and thus also of the possibility of reciprocation by the 2nd dictator. To avoid guilt from ex-post expectations, the 2nd dictator was told that the 1st dictator would not be informed about the source of the money, which we sent by electronic bank transfer with the money which they did not give to the 2nd dictator.

The following is a summary of the procedures. Please refer to the game tree in Figure 1.

In Stage 1, subjects were told that they could divide 20 CNY with another player and that the remainder would be sent to their bank account in 3-4 days. Hence, dictator 1's payoff was $20 - x$, where x was the amount sent to dictator 2. As mentioned, the

dictator 1s' were not informed of possibility that the 2nd dictators could reciprocate. The payment was through a non-descript electronic deposit which the 1st dictators might not even notice as including an extra payment.

In Stage 2, the 2nd dictators received 3*transfer from the 1st dictator + show-up fee of 5 CNY. Then the 2nd dictators could decide to give something back. Hence, dictator 2s got $3x+5-y$, where y was the amount sent to dictator 1s. Dictator 1's final payment was $20-x+y$.

Dictator 2s were informed that the 1st stage dictators were not informed about possibility of reciprocation. Reciprocated money was sent with the stage 1 payment to dictator 1. The 2nd dictators took what money they did not give back upon leaving the experiment.

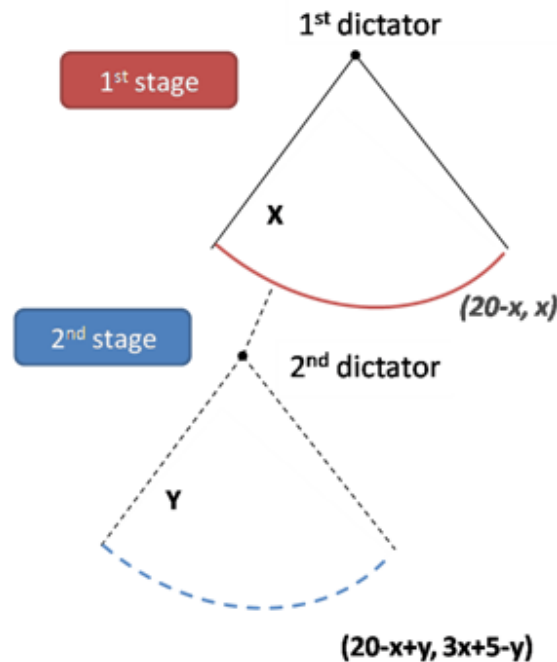


FIGURE 1: GAME TREE OF THE EXPERIMENT

The experiment involved 90 subjects recruited from graduate students at Shenzhen University Town. For the main treatment, 70 subjects were put into 35 pairs in the two stage dictator game. “Control 1” is the decision of the 1st dictators (while the main treatment was the decision of the 2nd dictator). The other 20 subjects were in the standard trust game. “Control 2” is the decision of these trustees. Figure 2 shows the main forms that subjects filled out. We elicited the *factual* 2nd order beliefs of the trustees in control 2 as to the expectations of the trustor. We also elicited the counterfactual 2nd order beliefs of the 2nd dictators about the first dictators expectations for reciprocation *had they known about the 2nd stage* in the treatment. See appendix A for the translation of the instructions.

The figure displays three forms used in the experiment, each with English and Chinese text. The first form is a registration form where participants provide their bank account and name, and write their university ID card number. The second form is a dictator form where the dictator decides how much of a 20 CNY endowment to give to player 2. The third form is a trustee form where the trustee decides how much of an endowment to give to player 1.

Form 1: Registration

Please write your bank account and name here. We will send your payoff to it in 3-4 days.

请将实验内容保密两周。请将你的校园卡卡号填于此。我们将在3-4天内将实验收入转到你的账上。

卡号 _____ 姓名 _____

现在你已经完成了所有的问题，感谢你的合作。请在离开教室前将问卷投入教室前高的纸箱中。如果你在领取实验报酬上有任何问题，请在实验结束后单独与实验人员联系。

Form 2: Dictator

Dear player 2, I give you _____ CNY from 20 CNY.

致参与者 2:

我从 20 元钱里分给你 _____ 元。

Form 3: Trustee

I give _____ CNY to player 1 from my _____ CNY. Player 2

我从 _____ 元钱中给参与者 1 _____ 元钱。

参与者 2

FIGURE 2. FORMS USED IN THE EXPERIMENT

III. MAIN RESULTS

The summary of our main results is as follows. In our control 1, 22/35=63% of our 1st dictators gave a positive amount. This is the standard result. The mean giving (ratio of the average) of these subjects was 4.4/20 CNY =22% (same as the average of the ratio). In control 2, 8/10=80% of these trustees in this standard trust game reciprocated. Mean giving was 6.3/21.5 CNY=29%, with the average of the ratio being 31%. However, in the treatment, 22/22=100% of the 2nd dictators who received positive transfers from the 1st dictator reciprocated⁴. Mean giving: 10.8/26 CNY=41%, with the average of the ratio being 39%. The difference in mean giving to the trust game control was significant at the 5% level, while that to the 1st stage dictator game was significant at the 1%. Inequity aversion was also ruled out by the result. 2/35=5.7% of the 1st dictator gave more than half of their endowment. (Only these 1st dictators could have thought they were poorer than their 2nd dictator pair, given that they did not know their transfer would be multiplied by 3 and have 5 CNY added.) However, 17/22=77% of the 2nd dictators who received a positive amount from 1st dictators were poorer than their 1st dictator pair, after giving. These 2nd dictators knew the true payoffs of the 1st dictators. This rules out inequity aversion as their motive for giving. The ratios of transfers and the average ratios of each

⁴ There was one 1st dictator who transferred 0.1 CNY. Since 0.1 is much closer to 0 than any other positive transfers, we did not include the 2nd dictator who received this 0.1 into the data of 2nd dictators who received amounts greater than 0.

treatment group are shown in Figure 3.

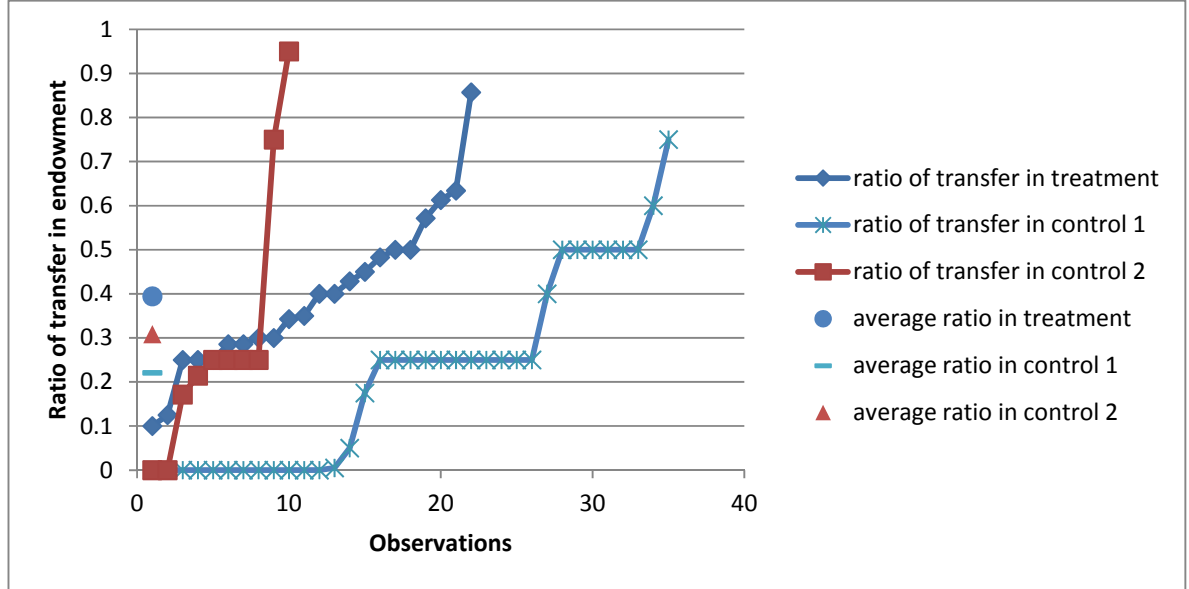


FIGURE 3. RATE OF TRANSFER

We also tested for the guilt sensitivity of all subjects using a standard test from psychology, the TOSCA-3. See Tangney et al., 2002 for details. We gave them the test in part as a distracter since they might have found it strange to get “money for nothing”. We found no significant correlation between guilt sensitivity either in 1st dictator giving, or in 2nd dictators reciprocation, or trustees reciprocation in the standard trust game. The lack of significance of the 2nd dictator reciprocation would be predicted if gratitude is something distinct from guilt. The lack of significance with 1st dictator giving could be due to irrelevance of guilt as a motive. The lack of significance with trustee giving in the standard trust game could be due to small sample size, which in this case was 10 subjects.

IV. DISCUSSION

Our results are significantly stronger than Stanca et al. (2007)’s results. There, identical with us, 1st dictator’s transferred 22% of their endowment in treatment.

However, their 2nd dictator reciprocation decreased to 12% for strategy method elicitation and 10% for decision method elicitation. That is considerably less than our 41% decision method elicitation. We chose the decision method because prior studies had shown that the strategy method may not be a reliable tool to measure reciprocity. Fehr et al. (2003) concluded that the strategy method is less emotionally arousing than the game method. Casari and Cason (2009) concluded that measured trustworthiness is lower using strategy method elicitation. Our results could also be stronger because people in China are less accustomed to giving by strangers. Hence, the surprise might cause them to reciprocate more.

Lazear et al. (2011) also had a two stage dictator game with an uninformed 1st dictator and a 2nd dictator who could silently exit. They found significant silent exiting by the 2nd dictator (31%), where we found none. However, they restricted the choice set of 1st dictators to either (7,5) or (6,6), which meant that 1st dictators could only give 14% of their endowment. This would have made an important difference in our experiment where only one of our 22 1st dictators transferred less than 14%. Limiting the amount of giving by 1st dictators could induce a norm/reference point for the 2nd dictator to follow. More importantly, such a ceiling on giving would limit the gratitude that the 2nd dictators could feel. Furthermore, if “type preference” as introduced by Gul and Pesendorfer 2010 was a motive for the 2nd dictators’ giving, then 1st dictators’ signal of the generosity of their type would be hampered. Under such a restricted choice, a specific 2nd dictator could only know that their paired 1st dictator was not mean. That might warrant only a similar lack of meanness from the 2nd dictator – silent exiting. Furthermore, the lowered ex post belief of the 1st dictator, which could have been anticipated by the 2nd dictator should the 2nd dictator not silently exit, would induce less guilt aversion in the 2nd dictator, decreasing

the amount of giving when 2nd dictator does actually not silently exit. Small amounts of money may also have counter intuitive incentive effects as suggested by Gneezy and Rustichini (2000).

There are several possible limitations to our study. First, we *assumed* like Stanca et al. (2007) but did not prove that by not informing 1st dictators about the possibility of the 2nd dictator reciprocating, we had diminished the ex ante 2nd order belief of the 2nd dictator that the first dictator was expecting reciprocation. We did survey for the counter-factual beliefs of the 2nd dictator of about 1st dictator beliefs *had the 1st dictator known* about the 2nd round. Unlike Ellingsen et al. 2010, this was highly significantly correlated with actual 2nd dictator reciprocation. This is evidence that ex-post beliefs mattered, and therefore, could drive guilt driven reciprocity, as we suggested was a limitation of Stanca et al. (2007) with respect to testing our gratitude hypothesis.

This significant correlation, though would not be additional evidence for guilt based reciprocity. 2nd dictator giving and 2nd order counter-factual beliefs could have been driven by the amount and the implied beliefs of 1st dictator giving. In other words, the 2nd dictator gave more when he had more to give, not only because he believed that 1st dictator expected more. We could not separate the independent effects of endowment and implied beliefs on 2nd dictator giving. 2nd order beliefs and the endowment had a correlation of 0.93. The consequent multi-collinearity yielded no significant regression. What the correlation does suggest is that the 2nd order beliefs would have been a confound for the apparent generosity of 1st dictator giving had the counter-factual world been true, supporting our and Stanca et al. (2007)'s results that people might reciprocate more when they do not believe that the reciprocation had been expected. The 2nd order

beliefs of trustees in the control group confirmed results from CD 2006. These were correlated with trustee giving at less than 1% level of significance when regressed on both 2nd order beliefs and the endowment from the trustor.

Second, our silent exiting for 2nd dictators was implicit. 2nd dictators were told that 1st dictators were uninformed about the existence of the 2nd round. Furthermore, they were told that the 1st dictators would not be told that the money was from them, the 2nd players, should they choose to reciprocate. To see that this implies silent exiting, note that the statement which rules out silent exiting, “1st dictator will be told that you did not give back anything” would be inconsistent with the spirit and the letter of this instruction⁵. In any case, the effect of guilt should be weaker than in either standard trust games or in Stanca et al.’s design, if not eliminated. Third, the double-blindness in our experiment was also implicit and unannounced. We simply had no identifiers for 2nd dictators’ identity on the envelopes they submitted for giving. We chose not to make our double-blindness more salient in order to avoid possible demand effects (George Loewenstein, 1999). Furthermore, Barmettler et al. (2011) found no difference between implicit double-blindness and explicit double-blindness. If the double-blindness was not salient, then subjects might still feel some shame from “appearing” ungrateful before experimenter, in which case, it would be a 2nd order reaction to gratitude. The use of implicit silent exiting and double-blindness should in any case have decreased both shame and guilt and hence should not have led to significantly increased rates of reciprocation in our treatment, if shame and guilt were the cause.

⁵ The exact statement was “They were not told about the existence of this round. That means, when player 1 made their choice (to give to you), Player 1s were not told the amount they transferred will be multiplied by 3 and plus 5. They were also not told that you can pay them back in this round. They will not be told that the money is given by you.”

Third, we also had a small number of observations. However, our results should be read as a partial replication of Stanca's et al. As was suggested in the introduction, our hypothesis was corroborated as the driver of the contra-guilt data in Ellingsen et al. (2010) and in Reuben et al. (2009), and in the apparent conflict between their main results.

Fourth, we leave for future work how to reconcile gratitude and guilt within a general model of reciprocity. As mentioned in the introduction, our strong result when guilt was suppressed suggests that gratitude is not merely a substitute for guilt, but can be "crowded out" by guilt. How this crowding out occurs and what determines whether a person feels gratitude or guilt is beyond the scope of this paper. However, one possible reason for our results is that part of gratitude is motivated by "type preference." 2nd dictators might like to give more to those 1st dictators who gave them more because 2nd dictators prefer to give more to people who have shown themselves to be more altruistic. To test this, Lin and Ong 2011 introduced a 3rd dictator (30 subjects) who observed 1st dictator giving to 2nd dictator, and could give to 1st dictator from 3 exogenous levels of endowment: 6, 20, 32 CNY. They also had implicit and explicit double-blindness and silent exiting treatments. They found that 3rd player giving correlated with 1st dictator giving but not with endowment when regressed on both in all cases except for the 6 CNY explicit double-blindness and silent exiting treatment. In that case the correlation was inverted. They suggest that this could be a demand effect interacting with the small endowment. They also found no difference between explicit and implicit double-blindness except for the 6 CNY. Their result supports our result, and suggests that type preferences contributed to reciprocation in our experiment. Thus, their result seems to address the limitations we mentioned above in regards to separating the belief from

endowment effect, and the saliency of implicit double-blindness and silent exiting, and sample size.

In summary, our results suggest that guilt and inequity aversion are unnecessary for reciprocation. Gratitude can also drive reciprocation. We found greater reciprocation when guilt was ruled out, suggesting that guilt might crowd out gratitude. Our result could help explain opposing results in Ellingsen et al. (2010), Reuben et al. (2009). The opposing belief dependent incentives would make experiment sensitive to setup. Our strong positive reciprocity result also suggests why it has been difficult to find in the past.

REFERENCES

- Barmettler, F.; Fehr, E. & Zehnder, C. (2011), 'Big Experimenter Is Watching You! Anonymity and Prosocial Behavior in the Laboratory'(5925), Technical report, Institute for the Study of Labor (IZA).
- Casari, M. & Cason, T. N. (2009), 'The strategy method lowers measured trustworthy behavior', *Economics Letters* **103**(3), 157 - 159.
- Charness, G. & Dufwenberg, M. (2006), 'Promises and Partnership', *Econometrica* **74**(6), 1579--1601.
- Charness, G. & Rabin, M. (2002), 'Understanding Social Preferences With Simple Tests', *The Quarterly Journal of Economics* **117**(3), 817-869.
- Dana, J.; Cain, D. M. & Dawes, R. M. (2006), 'What you don't know won't hurt me: Costly (but quiet) exit in dictator games', *Organizational Behavior and Human Decision Processes* **100**(2), 193 - 201.
- Elizabeth, H.; Kevin, M.; Keith, S. & Vernon, S. (1994), 'Preferences, Property Rights, and Anonymity in Bargaining Games', *Games and Economic Behavior* **7**(3), 346-380.
- Ellingsen, T.; Johannesson, M.; Tjøtta, S. & Torsvik, G. (2010), 'Testing guilt aversion', *Games and Economic Behavior* **68**(1), 95 - 107.
- Fehr, E.; Fischbacher, U.; von Rosenbladt, B.; Schupp, J. & Wagner, G. G. (), 'A Nation-Wide Laboratory: Examining trust and trustworthiness by integrating behavioral experiments into representative surveys'(iewwp141), Technical report, Institute for

Empirical Research in Economics - University of Zurich.

Gul, F. & Pesendorfer, W. (2010), 'Interdependent Preference Models as a Theory of Intentions', Technical report, Working paper.

Hoffman, E.; McCabe, K. & Smith, V. L. (1996), 'Social Distance and Other-Regarding Behavior in Dictator Games', *American Economic Review* **86**(3), 653-60.

Lazear, E.; Malmendier, U. & Weber, R. (2011), Sorting in experiments with application to social preferences, *American Economic Journal: Applied Economics*, forthcoming..

Loewenstein, G. (1999), 'Experimental Economics from the Vantage-Point of Behavioural Economics', *Economic Journal* **109**(453), F23-34.

Lin, H. & Ong, D. (2011), 'Deserving Altruism: Type Preferences in the Laboratory', *unpublished*.

McCabe, K. A.; Rigdon, M. L. & Smith, V. L. (2003), 'Positive reciprocity and intentions in trust games', *Journal of Economic Behavior & Organization* **52**(2), 267-275.

Reuben, E.; Sapienza, P. & Zingales, L. (2009), 'Is mistrust self-fulfilling?', *Economics Letters* **104**(2), 89-91.

Stanca, L.; Bruni, L. & Corazzini, L. (2009), 'Testing theories of reciprocity: Do motivations matter?', *Journal of Economic Behavior & Organization* **71**(2), 233-245.

Tangney, J. P. & Dearing, R. L. (2003), *Shame and Guilt (Emotions And Social Behavior)*, The Guilford Press.