Trust, Society, and Information: An Experiment

Aisling Scott
UC Berkeley

July 22, 2018

Abstract

Trust is an essential ingredient for unlocking economic surplus. However, consider the prisoner's dilemma—all parties gain from cooperation, yet each party has an incentive to deviate. How can we organize society to unlock the possible gains from trust in such situations? We've all had experiences that indicate it is possible. Studies have shown prosocial individuals are more trustworthy. We can take advantage of this fact and suggest pairing prosocial individuals with less prosocial individuals who will trust them if their type is known. In this case, it takes information, timing, and only one pro-social individual to unlock the trust surplus. Finally, we turn to a fundamental question: Will society organize itself so as to unlock surplus? We find that self-organization does just as well as imposed assignment.

1 Introduction

Trust serves as the glue that binds our society together. It is essential for unlocking economic and social surplus. Economists have long viewed trust as a calculative action taken only when the expected gain is positive (Williamson, 1993). We find numerous situations where potential economic gains to trust exist when all parties work together, yet each party has a financial incentive to act opportunistically. How do we solve this conundrum? Over time, societies have devised many solutions to this problem. For instance, the legal system serves as an important institution that can provide redress should one party act opportunistically at the expense of the others. However, contracting to preclude all of the paths for opportunism is difficult and expensive. Likewise, the court system itself is costly and works slowly. Another possible solution is reputation (Greif, 1993). Reputation works well when individuals interact often with each other or society and the threat of future losses sufficiently outweigh the gains of opportunism. Indeed game theorists have shown us that punishment and rewards mechanisms that can help cooperation become credible in repeated interaction.

Yet, in practice, many situations are not covered by these solutions. Contracts are too costly, interaction is too infrequent, anonymity is too great for our varied array of solutions to apply. Are we then doomed to forego the possible gains from cooperation in such situations? Introspection suggests that the answer is no —we have all experienced situations where formal and informal protections from opportunism were absent, yet trust did not break down. Even when incentives exist to free ride, researchers have found people cooperate (Dawes and Thaler, 1988).

Research indicates that prosocial preferences such as altruism and reciprocity lead people to trustworthy (Camerer, 2011; Ostrom and Walker, 2003). Many individuals exhibit prosocial preferences and we can take advantage of that fact. Consider when a less prosocial person has the opportunity to loan money to a prosocial individual. One can gain from loaning money because the recipient will return the money plus some. Here we need the participants to be in the right order and the sender to have information about their prosocial returner. To test this, we conducted a laboratory experiment using the classic Berg et al. (1996) trust game. Using survey instruments to identify pro-sociality, we placed together two individuals, only one of whom scores high on pro-

sociality. When types are publicly labeled, we compared trust under four scenarios: sequential pro-social first, sequential selfish first, simultaneous moves pro-social first, and simultaneous selfish first. We find that individuals trust less when they learn their partner is not prosocial. When their partner is prosocial, they trust the same amount as the control group with no labels. Our design is similar to Burks et al. (2003) in that subjects play both roles in the trust game. Burks et al. (2003) discover that when participants play both roles they give less on average, although we show in Table 2 ours do not differ from their control group.

Then, with the same pairing of two people with opposite prosocial types, we turn to a more fundamental question: Will society organize itself so as to unlock surplus? In these rounds, the subjects determine which of the trust games they wish to play using a random dictator. To our knowledge we are the first paper to investigate voting preferences and whether individuals can self-organize in the trust game. We find that overwhelmingly regardless of type the participants vote for the returner position in the sequential game, which gave them the highest payoff in the non-voting rounds. Additionally, voting seems to generate the same levels of trust as the non voting rounds. To recap, the main findings from our studies are:

- Result 1: In the treatment, prosocial senders, who learn their partner is less prosocial, trust less, while less prosocial senders, who learn their partner is prosocial, do not change their level of trust.
- Result 2: In the control, prosocial participants are more trustworthy than less prosocial participants.
- Result 3: In the treatment, prosocial senders update their expectations downward, while less prosocial senders update their expectations upward.
- Result 4: Voters prefer the sequential game and the returner position.
- Result 5: Voting (Self-Organization) generates as much trust as role assignment.

The remainder of the paper proceeds as follows: Section 2 offers an overview of the motivation, experimental design, and the experimental rationale. In section 3, we present our main findings. Finally, we conclude and offer some additional discussion concerning our findings.

2 Experiment

2.1 Motivation

The focal point of the experiment is that it takes one pro-social individual to gain the surplus from trust. In the standard trust game the Nash equilibrium is that the sender sends zero resulting in zero trust surplus being realized. Additionally, trust seems not to pay in that the return to trusting is about zero and a fairly robust result (Camerer, 2011). However, countless research experiments indicate when people play the trust game the sender sends money (Barr, 1999; Cox, 2004; Fahr and Irlenbusch, 2008; Glaeser et al., 2000) even when they are informed of the history that trust does not pay they trust (Ortmann et al., 2000). Ashraf et al. (2006) argue that individuals may trust even when it is not beneficial to them. This view contrasts with the classical economics view that trust is calculated (Williamson, 1993). While our experiment assumes the classical economics view, it is designed in such a way that we could possibly challenge that view.

Suppose that trust is calculative. With this assumption, one will trust when they know the likelihood of return is high. Some people are more likely to return the money than others. Now imagine a person less likely to return the money is in the position of trusting. The person may not trust if they know nothing about their partner; their priors may be that the other person is their same type and will not return the money. However, if we let the person know that their partner is a trustworthy type, they may trust since they will benefit from this transaction. Consequently, we hypothesize it would be beneficial to put the less trustworthy person in the position of trusting and the trustworthy person in the position of returning money. We test this by pairing opposite types and giving the sender information about their partner's type. Our experiment will test whether people respond to information that allows them to be calculative. Furthermore, we'll be able to understand if society can facilitate trust even in the presence of less prosocial individuals.

2.2 Experimental Design

218 UC Berkeley students participated in the experiment over 11 experimental sessions at XLAB, UC Berkeley. Subjects interacted exclusively through individual computer terminals. The computer terminals each had dividers so that subjects cannot see the screens of any other subject. The author used oTree online experimental software to design the experiment (Chen et al., 2016). The author

and a student helper oversaw all sessions. The experimental instructions were on the screen for subjects to read. Sessions were randomly assigned treatment or control. Additionally, the order of the trust game scenarios was randomly assigned in each session. Subjects filled out consent forms and started the experiment with the pro-social survey. Subjects were informed there are two parts to the experiment: a (non-incentivized) survey and an (incentivized) game. The survey results determined, in part, the pairings of subjects during the game. Subjects were informed that the survey would affect the game, but were not offered the detail as to its exact relationship.

Survey

At the start of the experiment, subjects were asked to complete a 5 question survey about attitudes. Falk et al. (2016) introduced and validated this instrument as a measure of pro-social traits such as trust, altruism, and positive reciprocity.[1] The survey took on average 6 minutes for subjects to complete.

The results of the survey created each subject's pro-social score.[2] These scores were ranked and, those whose scores were above the median for a session were assigned to the pro-social group while those below were assigned to the non pro-social group. At the end of the survey, subjects were told that the survey concerned pro-sociality. To ensure common knowledge about pro-sociality, subjects saw the definition of pro-sociality meaning voluntary behavior intended to benefit another. Additionally, subjects were informed their responses to the survey and the responses of others in the room were used to calculate their scores.

Trust Game. As referenced in the experimental motivation section, our trust game uses Berg et al. (1995) classic investment game. Participants were explained the instructions of the game. They confirmed their understanding of the game by answering a couple of questions.

The experiment uses a between subjects design to examine the effect of providing information about the pro-sociality of the returner to the sender (referred to as the information effect). It uses a with-in to investigate the impact of sequential versus simultaneous moves (timing effect). As the motivation stated, it is both the timing and the information that should matter.

In the treatment group for the information effect, the sender will get an information signal indicating whether the returner is above or below the median pro-social score. The language of

[1] Technically, we administered a streamline form of the survey, as suggested by Falk et al. (2016)

[2] See appendix for detailed calculations.

the treatment is shown in Figure **??**. Importantly, the sender does not learn about his/her own score, nor does the returner learn anything about the sender's score. Furthermore, senders stated their expectations about what percent of their money they expect to receive from the returner. In the information treatment, senders stated expectations before and after the pro-social score of their partner was revealed.

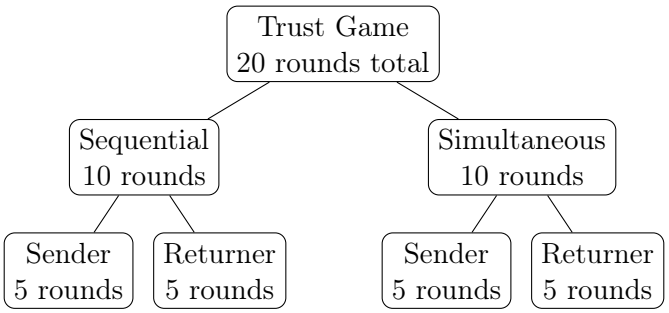


Figure 1: Experimental Design

The experiment uses a 2 X 2 X 2 design as demonstrated in Figure 1. Sessions are randomly assigned to treatment or control. In order to minimize confusion subjects play 10 rounds of either sequential or simultaneous and then switch to the other mode of play. Within those 10 rounds of sequential or simultaneous play, the subject will be both the sender for 5 rounds and the returner for 5 rounds. These orderings were randomly assigned such that half the sessions started with sequential play and half with simultaneous play. Subjects were made aware when they were switching from sequential to simultaneous. On each round, subjects were informed what timing they were playing as shown in Figure **??** and Figure **??**. Anonymity is kept throughout the experiment to avoid reputation effects. Subjects may be rematched with the same players in some rounds, but they will not know that.

Balance. We check whether our randomization achieved the goal of balanced groups across some individual covariates. At the end of the experiment, we asked participants some questions about their background and personality. We use these and the prosocial survey to

Voting. After 20 rounds of play, the subjects had 5 rounds to vote on what scenario to play. They chose if they would like to play sequentially or simultaneously and whether they would like

Table 1: Balance Table

Variable	Treat. Mean	Control Mean	Difference	P-value
Female	0.620	0.591	0.029	0.658
White	0.213	0.309	-0.096	0.107
Hispanic	0.130	0.182	-0.052	0.290
Asian	0.611	0.555	0.057	0.399
Black	0.056	0.009	0.046**	0.052
Number of Econ Courses	1.444	1.264	0.181	0.606
Extraversion	4.009	4.205	-0.195	0.187
Agreeable	4.667	4.523	0.144	0.199
Conscientious	5.000	5.205	-0.205	0.101
Emotional Stability	4.370	4.573	-0.202	0.134
Open	4.870	5.236	-0.366***	0.000
People Have Best Intentions	5.750	5.373	0.377	0.200
Reciprocate Favor	8.704	8.764	-0.060	0.748
Give without Expectations	7.370	7.664	-0.293	0.266
Donate	343.796	324.609	19.187	0.693
Give a Present	3.407	3.482	-0.074	0.742
Expectations in Trust	107.074	116.873	-9.799	0.248

to be the sender or the returner. The voting mechanism used is a Random Dictator (Dal Bó et al., 2017). The subjects learned that either their first choice or their partner's first choice would be implemented with 90 percent probability and 10 percent probability that their second choice would be implemented. The computer chose the outcome and then the subjects played the game. In the no information treatment subjects knew nothing about their type or their partner's type. In the information treatment subjects learned their partner's pro-social score.

Payment. At the end of the session, Xlab staff paid the subjects. One round was randomly chosen for real payoffs and subjects had knowledge of this at the start of the experiment. On average, participants made $16.78 and sessions lasted between 45 minutes and one hour.

2.3 Experimental Rationale

As the timing, information about pro-social preferences, and the order of pro-social participants plays an important role in trust. The 2X2X2 design allowed us to capture the main effects and the interaction effects of information, timing, and order. By including both simultaneous (SIM) and sequential (SEQ) play, we identified whether an individual will signal belief or use trust strategically.

In the sequential move, the sender's action serves as commitment device that signals their belief in returner returning a fair share (Servátka et al., 2008). In other words, the amount sent, is a costly message. Our main hypothesis that it takes only one to trust hinges on the fact that the sender must know the type of the returner and that he/she can signal to them that he/she holds a belief in their trustworthiness, otherwise trust will break down.
Participants were always paired with someone of the opposite type, but always a real participants since studies with confederates and computer simulations have been shown to change outcomes (Johnson and Mislin, 2011).

Ideally, both pro-social beliefs and the trust game would have be incentivized. Given monetary constraints and the advancement of research validating survey measures, this experiment implemented a non-incentivized survey. As stated above, the survey has been validated to correlate with incentivized measures. It was important subjects believed the pro-social signal regardless of their validation. We used expectations to observe whether subjects believed the information. In the information treatment group stated their expectations before and after observing the pro-sociality signal of their returner.

Since subjects played many rounds of the trust game, this study randomly selected one of the rounds for payment to the subjects, also know as the random lottery selection procedure? (Holt 1986a). Under this procedure, the subjects act truthfully each round. Azrieli et al. (2012) argue that in general games paying for one round is at least as good and sometimes better than paying for all rounds.

In order to see how subjects would self-organize, subjects were asked to vote on their preferences. It was important to elicit participants' true preferences absent of any coordination or strategy with their partner. We used the random dictator voting mechanism to achieve this goal.

3 Results

Of the 218 subjects 108 subjects were in the information treatment and 110 subjects were in the control. We found 5 main results:

- Result 1: In the treatment, prosocial senders, who learn their partner is less prosocial, trust less, while less prosocial senders, who learn their partner is prosocial, do not change their level of trust.
- Result 2: In the control, prosocial participants are more trustworthy than less prosocial participants.
- Result 3: In the treatment, prosocial senders update their expectations downward, while less prosocial senders update their expectations upward.
- Result 4: Voters prefer the sequential game and the returner position.
- Result 5: Voting (Self-Organization) generates as much trust as role assignment.

Overview. On average, participants sent \$6.09 and returned 33% of the amount tripled, in other words, they returned exactly the amount they got. Although our design has participants play both roles, our results align with the standard results from trust games. In a survey of experimental results Camerer (2011) reports that senders sent about 50 percent of their endowment and returners send back about one third of the tripled amount.

Table 2 presents comparisons of our data in comparison with Burks et al. (2003) denoted BCV and the classic trust game Berg et al. (1995) denoted BDMc. We compare to BCV since their design is most similar to ours. They also use the classic BDMc trust game with individuals playing both the sender and returner. Therefore, we believe BCV serves as a better check on our data due to the fact that our players also play both roles in the trust game. We use average dollars sent and returned as well as average fraction returned defined as the amount returned over the tripled amount sent. Another important metric is the amount returned over the amount sent or the return on trust. The largest return on trust in our experiment is in the sequential game with no information with 1.17 (\$8.06/6.89) percent return on \$1 sent, whereas the control in BCV the return is 1.31. Although

our study has a much lower return to trust than the control in BCV, it is higher than BDMc.

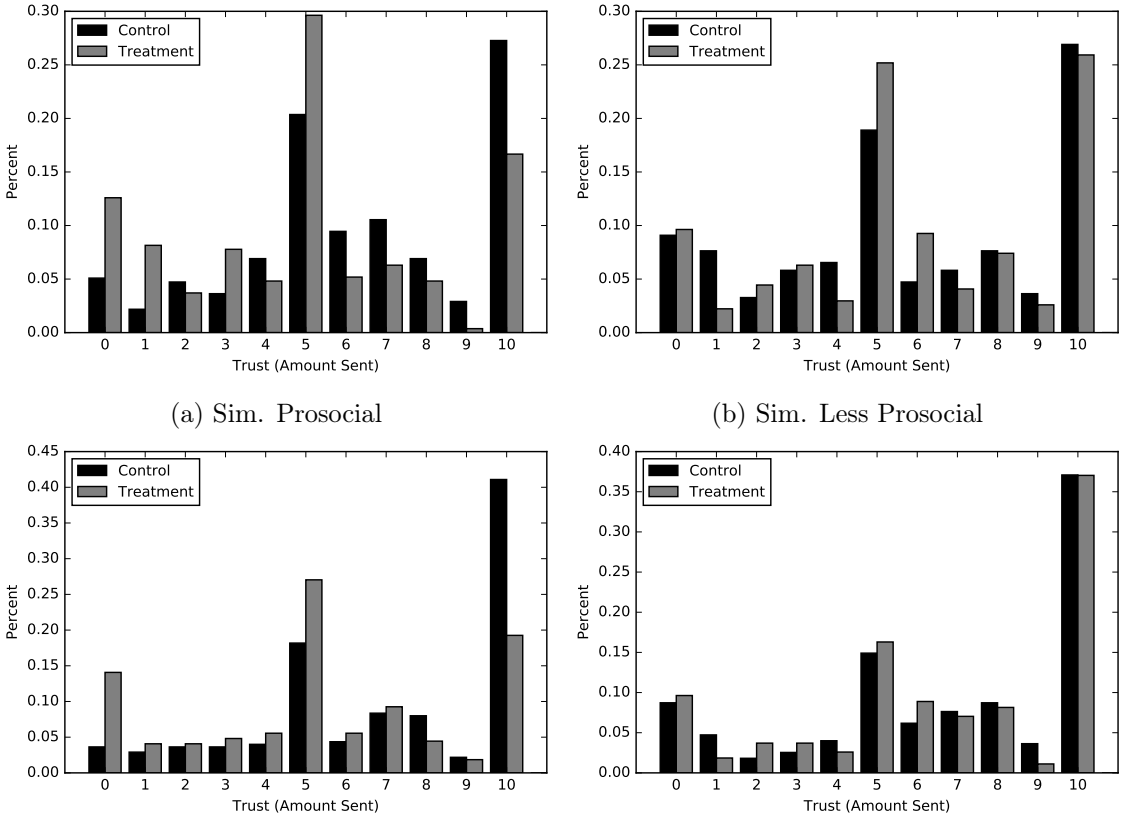
Table 2: Comparisons of Averages to BCV and BDMc

	Simultaneous		**Sequential**		**BCV**	**BDMc**
	Control	**Treatment**	**Control**	**Treatment**	**Control**	**All**
Sent	6.11 (3.21)	5.42 (3.24)	6.89 (3.22)	5.94 (3.37)	6.5 (3.61)	5.16 (2.94)
Return	6.7 (5.88)	5.02 (4.82)	8.06 (6.37)	6.18 (5.45)	8.5 (7.99)	4.66 (5.55)
Frac. Return	0.35	0.31	0.35	0.32	0.4	0.3

Notes: First two rows are average dollar amounts, last row is average fraction. Standard deviations are in parentheses.

Our results are close with those of BDMc and BCV. The differences between our data and BCV are not statistically significant. For amount sent, our sequential no information treatment mean has a t-statistic of t=0.48 and a p-value of p=0.63 when compared to the control in BCV. For amount returned, there is no statistically significant (p=0.80) difference between BCV and our sequential control. Finally, when we compare average fraction returned between the two groups we get a z-statistic of z=-0.45 and a p-value of p=0.65. We conclude that the behavior of our study control group with UC Berkeley undergraduates is similar to that of BCV's control group with Middlebury College students.

Result 1: In the treatment, prosocial senders, who learn their partner is less prosocial, trust less, while less prosocial senders, who learn their partner is prosocial, do not change their level of trust. Figure 3 presents histograms of the amount sent broken down by information treatment and control for each prosocial type and each type of play. In the control group, prosocial participants chose to give $10, the highest amount of trust, 41.09% of the time in the sequential play and 27.27% of the time in simultaneous play. While the prosocial participants in the treatment group, give $10 19.25% and 16.67% of the time in the sequential and simultaneous play respectively. The less prosocial participants look very similar in the treatment and control groups. In the sequential play, 37% of the time both groups give the full $10. In the simultaneous play, the treatment group gives $10 25.93% of the time contrasting to the control group who gives that amount 26.9% of the time.



(a) Sim. Prosocial

(b) Sim. Less Prosocial

(c) Seq. Prosocial

(d) Seq. Less Prosocial

Figure 2: Distribution of Trust

Figure 3 visually displays the amount of trust by each of the 8 scenarios. Prosocial senders in the sequential rounds sent more than the simultaneous rounds. Prosocial senders trusted significantly less after they are told their partners are less prosocial than the median player. The Wilcoxon rank-sum statistics for prosocial senders are z=-5.323 and z=-6.262 for simultaneous and sequential rounds respectfully. Consequently, the information treatment decreased trust for prosocial senders. On the other hand, the treatment seems not to influence less prosocial senders. The differences between control and treatment groups average amount sent by less prosocial senders were small and not statistically significant. For less prosocial senders, the Wilcoxon rank-sum statistic for simultaneous rounds is z=0.357 and a p-value of 0.360. During the sequential rounds, the Wilcoxon rank-sum statistic is z=-0.322 and a p-value of 0.370.

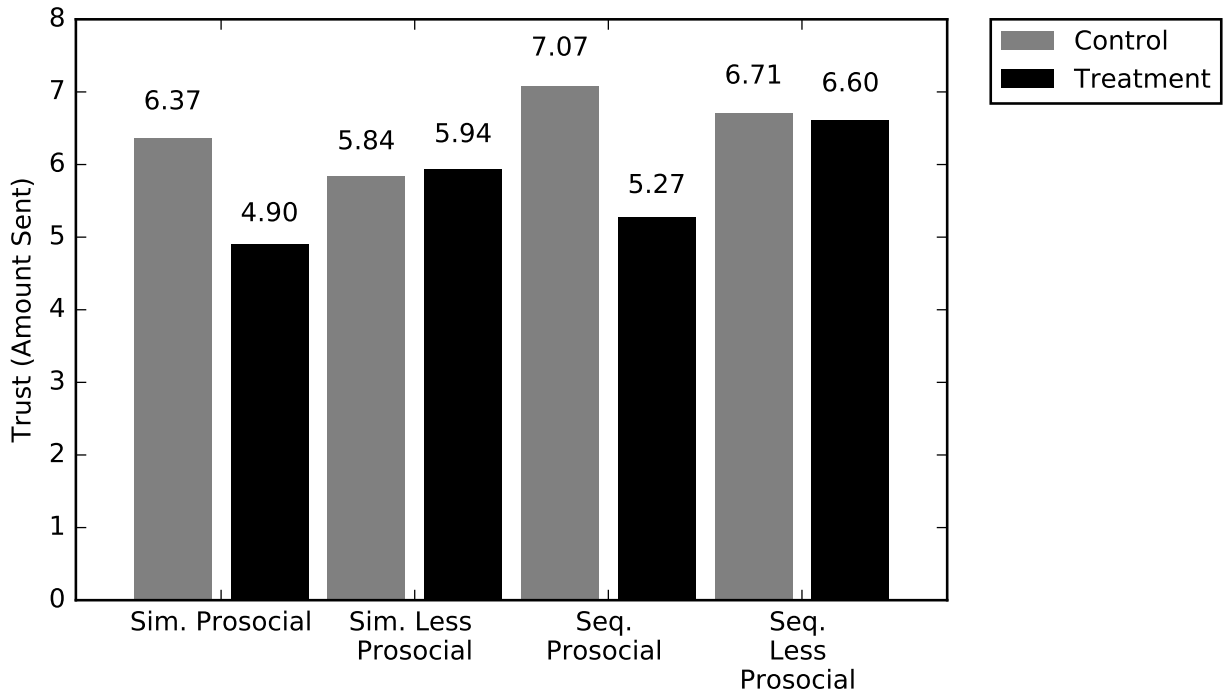


Figure 3: Trust in All Scenarios

We estimate the following full equation separately on simultaneous and sequential rounds to demonstrate the impact of the treatment on trust:

$$Y_{it} = \beta_0 + \beta_1 T + \beta_2 T * P + \beta_3 P + X_i + \gamma_t + \epsilon \tag{1}$$

where P is a dummy variable for whether the individual is above the pro-social median score and T is the treatment dummy. We include time fixed effects (γ_t) and the standard errors are robust. Table **??** reports the results. Our results confirm the non-parametric results from figure 3. The interaction between the information treatment and prosocial senders is statistically significant, meaning we can reject the null hypothesis that there is not statistical difference in the amount sent in the control group and treatment group by prosocial senders. Treatment and prosocial are not always statistically significant on their own suggesting what matters is the interaction between them. A prosocial sender in the treatment group sequential play sent \$1.04 (.31 standard deviations) less than in the control group. There is not statistical effect for less prosocial senders. The treatment seems to work in only one direction, that is downward, but does not increase trust for non-prosocial senders.

We aggregate the data to understand the effect of the treatment overall. We find that the overall amount given in the information treatment is $5.68 while in the control the average amount given is $6.50. The Wilcoxon rank-sum statistic for the control versus information treatment is z = -5.6290, with a p-value of 0.0, indicating that we can reject the null hypothesis that the amounts sent are equal. Since we have high sample size when we aggregate, we can also use the test statistic from a traditional t-test. Again, the difference is statistically significant at the 1% level.Overall, giving coarse information about the pro-sociality of the sender's partner decreases trust 13.8% in sequential play and 11.25% in simultaneous play.

Result 2: In the control, prosocial participants are more trustworthy than less prosocial participants. The percent returned patterns typically depend on the amount sent in the trust game, when the game is play sequentially (Ashraf et al., 2006). We expected the simultaneous rounds to have no relationship with the amount sent since the returners did not know that information when they made their decision. Figure 4 shows the median percent returned for each amount sent to the returner in the simultaneous rounds. The amount returned is fairly flat along the amount given as to be expected. The median returner closely matches the theoretical positive reciprocator, indicating that 50 percent of returners in the simultaneous rounds are willing to return at least the amount received. Figure 5 demonstrates median returners in the sequential rounds map closely to theoretical egalitarian predictions for low values of the amount given. In both the treatment and control the median returner does not give as much as predicted when the sender sends 7-9 dollars. When the sender sent $10 the median returner gave back 50 percent of the $30 they received.

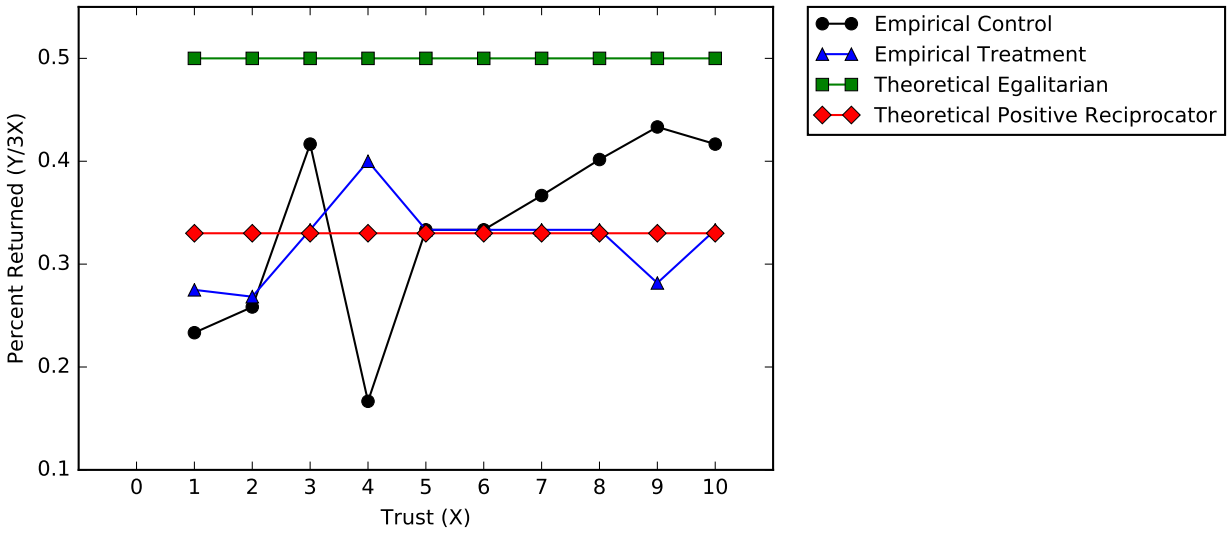


Figure 4: Fraction Returned in Simultaneous Rounds

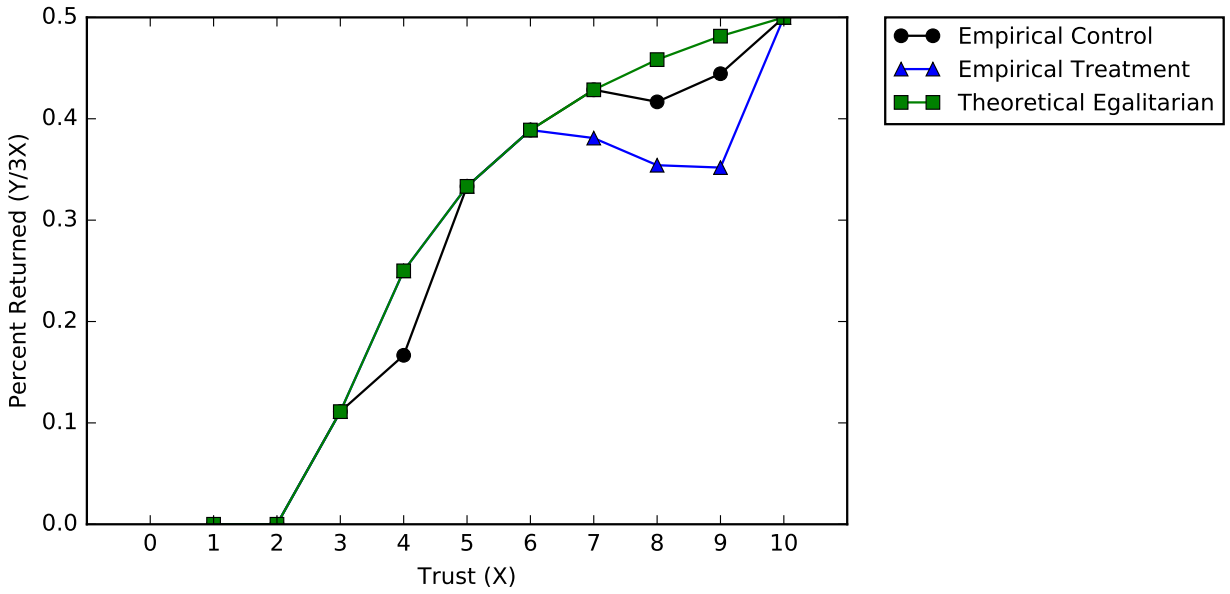


Figure 5: Fraction Returned in Sequential Rounds

We estimate the determinants of the percent returned in Table **??**. We lose some observations when the sender sends \$0 since the returner does not have a choice in their percent returned in that case. Columns Sim (1) and Seq (2) show the regressions without individual controls and time fixed effects. Column Sim (2) demonstrates that in the control prosocial returners do indeed return a higher percent in the simultaneous rounds when controlling for the amount given. In fact, a prosocial returner in the control simultaneous rounds sends .18 standard deviations more. The significance

of prosocial returners disappears in the Seq (2) and the magnitude is much smaller as well. In the sequential rounds a $1 increase in the amount given increases the percent returned by .16 standard deviations. While in the simultaneous rounds, the amount given increases the percent returned by .03 standard deviations.

Result 3: With the information treatment, prosocial senders update their expectations downward, while less prosocial senders update their expectations upward. Since the treatment decreased trust for prosocial senders but did not increase trust for less prosocial senders, we further examine whether participants actually updated their expectations based on the information given in the treatment group. Figure 6 shows the mean of the sender's expectations in the treatment pre and post they received information about their partners prosocial score. All differences are statistically significant at the 1% significance level using the Mann-Whitney test.

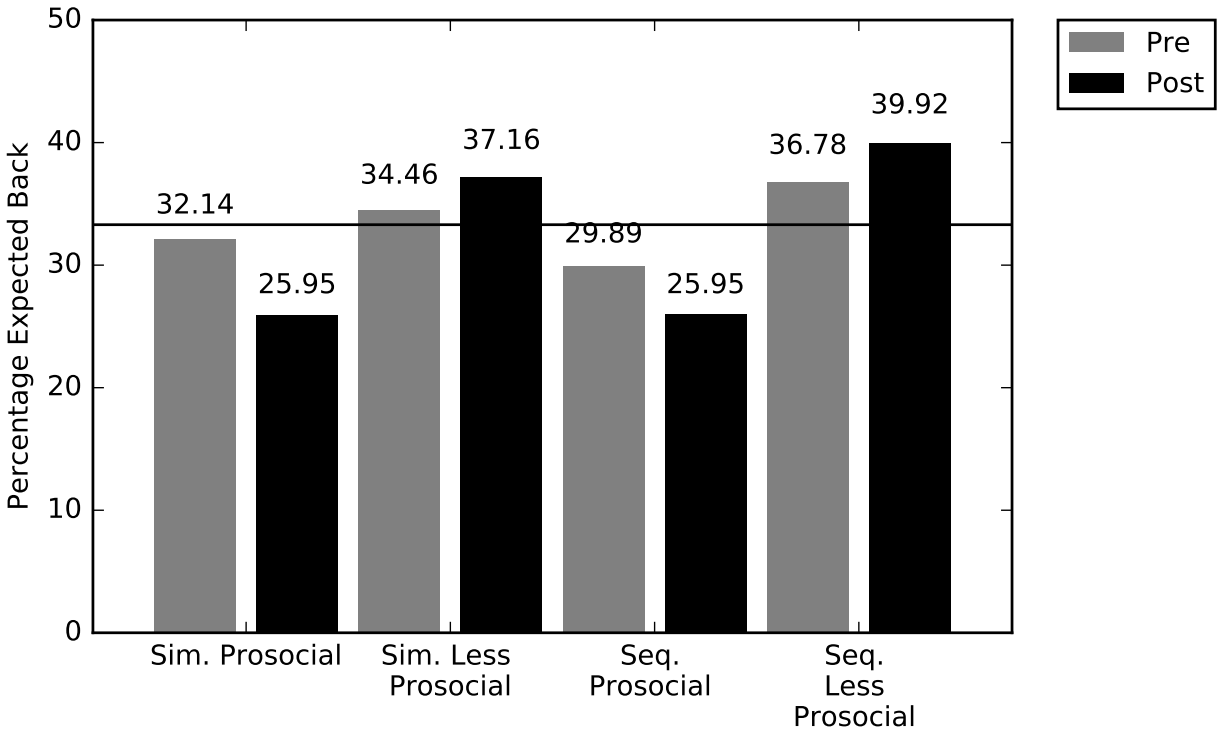


Figure 6: Expectations In Treatment Pre and Post Information About Partner

We see in Table **??** that expectations in round 1 before the treatment has learned any information are fairly balanced. The prosocial senders who start in the sequential play mode expect less in the treatment group than the control group. This difference is statistically significant at

the 5% significance level using the Mann-Whitney test for small samples. This initial difference could suggest why it appears the treatment impacted the prosocial participants. However, on a whole, the treatment and control groups are balanced in round 1 on expectations as shown in Table **??**. Furthermore, in round 1 we observe the treatment influence expectations as shown in **??**. All scenarios expect the sequential less prosocial senders significantly update their expectations. Given that the less prosocial senders who play the sequential rounds first, do not significantly increase their expectations when told their partners are prosocial in round 1, this could suggest that the less prosocial individuals may be less responsive to the treatment.

In table **??** we show the determinants of trust while controlling the expectations of treatment (post-treatment) and control. We break the regressions down by simultaneous and sequential as well as prosocial and less prosocial. Including expectations mitigates some of the main effect of the information treatment. We see that prosocial senders start at a higher level of trust with intercepts of 5.62 for simultaneous rounds and 4.33 for sequential rounds. The less prosocial participants have steeper slopes for their expectations than the prosocial participants, meaning they are more generous with using their expectations to give more to the returner. However, the treatment does make the less prosocial participants more stingy with their expectations. Expectations are a statistically significant determinant in trust and account for a fair amount of variation as the r-squared values increased from .10 to .38 in the sequential regression[3]. However, expectations do not influence trust in large magnitudes moving trust .02 standard deviations at the most. While the treatment influences the expectations and expectations do matter for trust, expectations alone do not dramatically move the dial on trust.

The control the median participant expects 35.83% back, while the median treatment participant expects 33.33% back, indicating that half believed they would receive at least what they have given. In the control group, 50.81% of the time participants believed they would actually make money — expecting more than 33.33%— compared to 39.53% of the time participants in the treatment group. The treatment group results are similar to Ashraf et al. (2006), where 39% of participants believed they would make money. In the treatment group, the prosocial individuals expect to make money 25% of the time, while the less prosocial individuals expect to make money 53.51% of the

[3]See table **??**

time. Clearly, the treatment impacted the expectations.

Result 4: Voters prefer the sequential game and the returner position. In the last 5 rounds of the experiment, individuals were able to vote on which game they would like to play sequential versus simultaneous as well as which order to play in sender versus returner. We asked the voting preferences questions in a way to elicit their un-coordinated beliefs.Over all our treatments and rounds only 27% of individuals' preferences actually coordinated, so in a world where there is no random dictator and individuals are forced to agree transactions would only occur 27% of the time. Overwhelmingly, subjects wanted to play the sequential game and 53% wanted to play the returner in the sequential game. While 22% of the votes wanted to play the sequential sender, only 9% wanted to be the simultaneous sender. Finally, 16% wanted to play the simultaneous returner.

Participants preference for the returner in the sequential game indicates their belief that the returner makes more money. Indeed, on average that is true. Table **??** displays the payoffs from the first 20 rounds of non-voting. Returners regardless of prosocial type or information treatment make more money, but with a higher variance. The sequential play scenario does yield the the highest payoff. Consequently, participants preferences in the voting game were well calibrated for making the most money. When asked why they chose their preference, one participant stated, "I wanted to steal the money at the end."

The participants that got their preference to play the returner in the sequential game made on average $14.63 in the control group and $13.65 in the treatment group. One might imagine that less prosocial participants chose to play the sequential returner so they "could steal the money at the end." However, there is an even split with prosocial and less prosocial participants listing the sequential returner as their top choice, with 50.4% of the votes being from prosocial participants. Surprisingly, participants did not steal all the money at the end. The percentage return median for the control was .38 and .33 for the treatment group. The median returner returned at least the amount they were given. Senders made $10.71 and $10.38 on average in the control and the treatment groups respectively.

Voters picked the returner sequential game as their preference and they did on average benefit financially from the vote if they were the chosen dictator.

Result 5: Voting (Self-Organization) generates as much trust as role assignment. We wanted to investigate whether participants would coordinate to gain the surplus from trust. Table 3 compares trust in the voting and non-voting rounds of the experiment. Using the Mann-Whitney test there are no statistically significant differences between trust in the voting and non-voting rounds and differences are less than $1.

Table 3: Mean Trust for Voting and Non-Voting Rounds

Scenario	Voting	Non-Voting	Difference	P-value
Control Sim	5.435	6.105	-0.670	0.257
Control Seq	6.808	6.889	-0.082	0.184
Treatment Sim	4.515	5.419	-0.904	0.474
Treatment Seq	6.183	5.937	0.246	0.298

4 Discussion

We were interested in the effect of information plus the order of prosocial types. However, if one wanted to fully test the effect of giving information about prosocial preferences, one would need to pair similar types as well. The point of this experiment was to pair opposite types, but the results indicate that we would be able to answer further questions with the pairing of all combinations of types.

Additionally, the information treatment sends a coarse signal to the senders. While we did our best to ensure individuals understood the meaning of prosocial and the coarse signal they received, participants may have reacted more to probabilities or more certain estimates that their partner was really trustworthy. As stated in the experimental rationale, its very difficult to get a predictive

measure for trustworthiness and we wanted to be honest with the participants so coarse measures were a good option in light of these concerns.

We saw in results 1-3 that trust for less prosocial types did not significantly increase with the information treatment while their expectations did. What could be some reasons for these puzzling results? Herrmann et al. (2008) find that some societies punish those that are cooperative. Applying this idea in our case, the less prosocial types could punish the prosocial types when they learn they are prosocial. While the less prosocial types do not increase the amount sent despite their increase in expectations, that is not necessarily punishing the prosocial types. Additionally, given the demographics of UC Berkeley students (more than 50% Asian in our sample) we do not believe that the results from Herrmann et al. (2008), who finds punishing strongest in Mediterranean countries, explain our finding that less prosocial types do not change their level of trust. Another hypothesis is that the less prosocial individuals could be less risk seeking than the prosocial individuals in the control. Eckel and Wilson (2004) show there is no correlation between trust and risk. Finally, perhaps the types differ in their ambiguity seeking behavior. Corcos et al. (2012) find a negative correlation between ambiguity aversion and trust, in other words, when people are more ambiguity averse they trust less. It is possible that less prosocial individuals are more ambiguity averse than prosocial individuals, however, the information signal could mitigate that ambiguity aversion. We can not rule out that perhaps the signal did not create any clarity for ambiguity aversion and that might be the reason for the difference.

Another hypothesis as to why less prosocial participants do not increase their trust of prosocial returners concerns other differences between the two groups. One can imagine that the treatment on less prosocial is weaker because of the differences in priors about the world for less prosocial individuals compared to prosocial individuals. For instance, perhaps less prosocial individuals believe most people are prosocial, consequently, receiving a signal that their partner is prosocial doesn't change their perspective on the world. Table **??** demonstrates that on average less prosocial senders had slightly higher expectations than their prosocial counterparts. From result 3, we do indeed see expectations changing and updating upward for less prosocial individuals paired with prosocial returners. However, we can not rule out the fact this effect may have been created by experimental demand since we asked them in a sequential order. Running the above mentioned experiment with

all combinations of pairs could clear up some of the questions the current experiment leaves us with.

5 Conclusion

When we pair prosocial types with less prosocial types, we find in the treatment, where the sender learns their partner's type, prosocial senders send less. Meanwhile, less prosocial senders do not change their behavior. Our evidence suggests that the treatment influenced expectations but not actions. Although the less prosocial senders increased their expectations that their paired prosocial returners would return more, they did not trust them more. While we may have created a weak treatment, regardless of that fact, our results align with those in behavioral economics showing that it can be hard to increase prosocial behavior (Gneezy et al., 2011).

The subjects may play the investment game, knowing they might not receive their money back and view it as a triple dictator game. Their view of the game could dramatically alter the amount of money they give. Camerer (2011) asserts that researchers might not know which game the participants actually believe they are playing. Consequently, it is important to gather beliefs as we did in this experiment. We find that subjects do believe the trust game has some rewards. Half of subjects seemed to believe that they could make money from trust with 53.51% of the time the less prosocial believed they could make money and 50.81% in the control group. This contrasts with results from Ashraf et al. (2006), who find 39% of subjects expect to make money. In our study, subjects with no information or positive information about prosocial partners believe they will have some return on their trusting behavior and this explains a significant amount of the variation in trust.

Expectations play a significant role in trust behavior, but they do not influence trust in a large magnitude. For instance, in the control sequential rounds increasing expectations by 10% increases trust by \$0.21, which is .06 standard deviations. Researchers have found that unconditional kindness and risk (Ashraf et al., 2006) influence trust, but again expectations explained more of the variation. Expectations seem to play an important role in trust, but are not the whole story. We believe there is a lot of room for further research to explore expectations in how it relates to trust.

References

Ashraf, N., Bohnet, I., Piankov, N., 2006. Decomposing trust and trustworthiness. Experimental economics 9, 193–208.

Azrieli, Y., Chambers, C.P., Healy, P.J., 2012. Incentives in experiments: A theoretical analysis. Journal of Political Economy .

Barr, A., 1999. Familiarity and trust: An experimental investigation. The Centre for the Study of African Economies Working Paper Series , 107.

Berg, J., Dickhaut, J., McCabe, K., 1995. Trust, reciprocity, and social history. Games and economic behavior 10, 122–142.

Burks, S.V., Carpenter, J.P., Verhoogen, E., 2003. Playing both roles in the trust game. Journal of Economic Behavior & Organization 51, 195–216.

Camerer, C.F., 2011. Behavioral game theory: Experiments in strategic interaction. Princeton University Press.

Chen, D.L., Schonger, M., Wickens, C., 2016. otree?an open-source platform for laboratory, online, and field experiments. Journal of Behavioral and Experimental Finance 9, 88–97.

Corcos, A., Pannequin, F., Bourgeois-Gironde, S., 2012. Is trust an ambiguous rather than a risky decision. Economics Bulletin 32, 2255–2266.

Cox, J.C., 2004. How to identify trust and reciprocity. Games and economic behavior 46, 260–281.

Dal Bó, E., Dal Bó, P., Eyster, E., 2017. The demand for bad policy when voters underappreciate equilibrium effects. The Review of Economic Studies 85, 964–998.

Dawes, R.M., Thaler, R.H., 1988. Anomalies: cooperation. Journal of economic perspectives 2, 187–197.

Eckel, C.C., Wilson, R.K., 2004. Is trust a risky decision? Journal of Economic Behavior & Organization 55, 447–465.

Fahr, R., Irlenbusch, B., 2008. Identifying personality traits to enhance trust between organisations: An experimental approach. Managerial and Decision Economics 29, 469–487.

Falk, A., Becker, A., Dohmen, T., Huffman, D., Sunde, U., 2016. The preference survey module: A validated instrument for measuring risk, time, and social preferences .

Glaeser, E.L., Laibson, D.I., Scheinkman, J.A., Soutter, C.L., 2000. Measuring trust. The quarterly journal of economics 115, 811–846.

Gneezy, U., Meier, S., Rey-Biel, P., 2011. When and why incentives (don't) work to modify behavior. Journal of Economic Perspectives 25, 191–210.

Greif, A., 1993. Contract enforceability and economic institutions in early trade: The maghribi traders' coalition. The American economic review , 525–548.

Herrmann, B., Thöni, C., Gächter, S., 2008. Antisocial punishment across societies. Science 319, 1362–1367.

Johnson, N.D., Mislin, A.A., 2011. Trust games: A meta-analysis. Journal of Economic Psychology 32, 865–889.

Ortmann, A., Fitzgerald, J., Boeing, C., 2000. Trust, reciprocity, and social history: A re-examination. Experimental Economics 3, 81–100.

Ostrom, E., Walker, J., 2003. Trust and reciprocity: Interdisciplinary lessons for experimental research. Russell Sage Foundation.

Servátka, M., Tucker, S.J., Vadovic, R., 2008. Strategic use of trust .

Williamson, O.E., 1993. Calculativeness, trust, and economic organization. The Journal of Law and Economics 36, 453–486.